Technische Universität München
Fakultät für Physik

Quantum Information Methods in Many-Body Physics

Jiří Guth Jarkovský

Vollständiger Abdruck der von der Fakultät für Physik
der Technischen Universität München
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigten Dissertation.

Vorsitzender: Prof. Dr. Rudolf Gross

Prüfer der Dissertation: 1. Hon.-Prof. Dr. J. Ignacio Cirac
2. Prof. Dr. Michael Knap

Die Dissertation wurde am 17.08.2022 bei der Technischen Universität München
eingereicht und durch die Fakultät für Physik am 16.09.2022 angenommen.

Abstract

This thesis explores the connection between two important fields in quantum physics — quantum information theory (QIT) and quantum many-body physics. We utilise methods and approaches from the former to solve difficult problems in the latter.

One of the main results of the thesis is a criterion for the approximability of mixed quantum states by matrix product operators (MPO), a type of variational family called *tensor networks*. We have proven that if a quantity called *Rényi entanglement of purification* between different parts of a multi-partite mixed state scales slowly enough, the state can be approximated efficiently by MPO.

Another topic of the thesis is the detection of spontaneous symmetry breaking and the search for the ground states of symmetric Hamiltonians despite its presence. We find that regardless of spontaneous symmetry breaking, there always exists a symmetric purification of the ground state. Furthermore, the entanglement of this purification indicates the presence of spontaneous symmetry breaking and its properties allow us to find out which symmetry is broken (in case of multiple symmetries).

We apply this method to completely solve a toy model of permutationally-invariant Hubbard model with off-site interactions. We derive the ground state phase diagram and characterize each phase by which of the original symmetries has been spontaneously broken.

The thesis also contains some research conducted in parallel to the research described above. We show an alternative approach to solving the Hubbard model on a cubic lattice in infinite dimensions and we provide some analytical results for the *Rényi mutual information* of pure and classically-correlated states.

Zusammenfassung

Diese Dissertation erforscht die Verbindung zwischen zwei wichtigen Gebieten der Quantenphysik — der Quanteninformationstheorie (QIT) und der Quantenvielteilchenphysik. Es werden Methoden und Ansätze aus der Quanteninformationstheorie verwendet, um quantenphysikalische Probleme zu lösen.

Eines der wichtigsten Ergebnisse dieser Dissertation ist ein Kriterium für die Approximierbarkeit von gemischten Quantenzuständen durch Matrixproduktoperatoren (MPO), eine Art Variationsfamilie, die *Tensornetzwerke* genannt wird. Wir haben bewiesen, dass der Zustand effizient durch MPO approximiert werden kann, wenn eine Größe namens *Rényi-Verschränkung der Reinigung* zwischen verschiedenen Teilen eines gemischten Mehrteilchenzustands langsam genug skaliert.

Ein weiteres Thema der Dissertation ist der Nachweis spontaner Symmetriebrechungen und die Suche nach Grundzuständen symmetrischer Hamiltonianer trotz ihrer Anwesenheit. Wir stellen fest, dass unabhängig von spontanen Symmetriebrechungen immer eine symmetrische Reinigung des Grundzustandes existiert. Darüber hinaus zeigt die Verschränkung dieser Reinigung das Vorhandensein von spontanem Symmetriebruch an, und ihre Eigenschaften erlauben es uns, herauszufinden, welche Symmetrie gebrochen ist (im Falle von mehreren Symmetrien).

Wir wenden diese Methode an, um ein Spielzeugmodell des permutationsinvarianten Hubbard-Modells mit Off-Site-Wechselwirkungen vollständig zu lösen. Wir leiten das Phasendiagramm des Grundzustands ab und charakterisieren jede Phase dadurch, welche der ursprünglichen Symmetrien spontan gebrochen wurde.

Die Dissertation enthält auch einige kürzere Forschungsarbeiten, die parallel zu den oben beschriebenen Arbeiten durchgeführt wurden. Wir zeigen einen alternativen Ansatz zur Lösung des Hubbard-Modells auf einem kubischen Gitter in unendlichen Dimensionen und liefern einige analytische Ergebnisse für die *gegenseitige Rényi-Information* von reinen und klassisch korrelierten Zuständen.

Acknowledgements

This thesis would never have been possible without the support of many wonderful people who I have met along the way.

Firstly, I would like to thank my supervisor Ignacio Cirac for his leadership throughout my PhD. You helped me explore vast parts of the ever-vaster world of quantum physics. From every discussion we had, I left knowing much about topics which I had not even known existed before and I already looked forward to our next meeting to learn even more.

I am very thankful to everyone else from the theory group, for making it such a lively and inclusive environment to work in. I am particularly thankful to the organizers of the group workshops and the many topical seminars, but I am also thankful to everybody else for always being available for scientific discussion on all of the different topics of quantum physics.

I am especially grateful to Sonya Gzyl for all her help as the IMPRS-QST coordinator. The IMPRS-QST programme enriched my PhD in many different ways — from broadening my scientific overview and expanding my soft skills, to growing my professional networks and providing many social activities.

A big thank you also goes to the theory group secretaries Andrea Kluth, Regina Jasny and Elena Wiggert for their incredible helpfulness with any issues I encountered during my PhD, going far and beyond in every way to provide assistance with not only administrative matters.

I am hugely grateful to my office mates Caroline de Groot, Albert Gasull and David Stephen for all their support during my PhD, for many scientific, cultural and political discussions we have had over the years and for the deep friendships we have developed by sharing our common workplace. Special thanks go to Albert, Caroline, Benjamin Schiffer and Arthur Christianen for proofreading parts of this thesis.

Last but not least, I am grateful for my family, my parents, sister and brother, for their love and food. Finally, I would like to express my gratitude to my loving wife Sofi Maytesyan for all her care and support throughout my journey.

Contents

Chapter 1

Introduction

In 1948 Claude Shannon introduced a new mathematical theory of communication [1]. His article *A Mathematical Theory of Communication* established a rigorous mathematical description of information, randomness and communication. Near the end of the 20th century, significant attention was focused on generalizations of Shannon's theory to quantum systems [2–4]. On top of that, several phenomena unique to the quantum setting were also discovered, such as quantum teleportation [5] or super-dense coding [6]. Thus quantum information theory (QIT) was born.

One of the main features that makes quantum systems distinct from classical systems is entanglement [7, 8]. This "spooky action at a distance", as Albert Einstein called it [9], plays countless important roles all over quantum physics and related disciplines. Entanglement is the backbone of most known quantum computing algorithms [10, 11], in quantum metrology it enhances the resolution of microscopes [12] and in quantum communication it enables quantum teleportation, which may one day pave the way for quantum internet [5, 13, 14].

And old proverb says "Fire is a good servant, but a bad master" [15]. The same could be stated about entanglement. While it is helpful in many information-theoretic tasks, in quantum many-body physics it is responsible for the large complexity of quantum states. Because of entanglement, the computational resources required to exactly represent a generic quantum state of N particles grow exponentially in N, making it nearly impossible to classically describe even modestly large systems ($N \sim 50$). One of the main goals of quantum many-body physics is to develop methods to overcome this obstacle. The main questions asked are "Which systems can be described efficiently?" and "Which methods can be used for this description?".

One possible answer to the first question is "systems with little entanglement". Our method of choice to describe such systems are *tensor networks* (TN). Tensor networks are used to represent complicated objects (such as state wavefunctions, operators or their expected values) as contractions of many easy-to-describe tensors. There exist various TN architectures designed to represent different kinds of systems. In 1 dimension, *matrix product states* (MPS) are particularly suitable for describing the ground states of local, gapped Hamiltonians [16]. Together with a numerical method called the density-matrix renormalization group (DMRG) [17], one-dimensional tensor networks have seen great success in expressing thermal

states [18–27] and simulating (short-)time dynamics [27–36].

A different group of quantum systems that can be described efficiently are those with "extreme dimensionality". Some many-body systems can be solved analytically on a lattice in 1 or 2 spatial dimensions [37–43], often employing the Bethe ansatz [44]. Conversely, *mean field theories* are a class of methods that describes well physical systems in high dimensions [45, 46]. However, mean field theories have also been used widely as approximations for low-dimensional systems [47–50] and even outside of physics [51–53].

Another important feature that makes quantum systems easier to describe is the presence of symmetries, as symmetries greatly reduce the number of relevant degrees of freedom of a many-body system. Nowhere does this stand out more prominently than in systems with permutation symmetry. As such systems are not very common in nature, they are often studied from a very theoretical point of view by the mathematical physics community. In quantum physics, the tool to describe permutationally-invariant states is the quantum de Finetti theorem [54–58], which originates from a similar theorem in probability theory [59]. In QIT, the quantum de Finetti theorem has found uses in quantum cryptography [56] and in entanglement detection [60].

This thesis seeks to deepen the connection between quantum information theory and quantum many-body physics. We will apply the tools of QIT, such as various entanglement measures, purifications and the de Finetti theorem to problems in quantum many-body physics, such as approximability by tensor networks and the search for ground states of Hamiltonians in infinite dimensions or with permutation symmetry.

1.1 Structure

The thesis is divided into 8 chapters (including the introduction and the conclusion) and 3 appendices. The first chapter **Introduction** serves to introduce the reader into the topic of the thesis and to provide a detailed overview of the structure of the thesis.

The second chapter **Background on Quantum Information Theory** covers some basic concepts of quantum information theory. The most important of these are *entanglement* and *purifications*, but the chapter contains a lot of other minor topics needed throughout the thesis, from *majorization* to the various *norms* used in quantum information theory.

The third chapter of the thesis is titled **Background on Quantum Many-Body Physics**. The field of quantum many-body physics is very broad and deep, so this chapter contains only the ingredients necessary in later chapters of the thesis. The most important topic introduced in this chapter is *tensor networks*.

In the fourth chapter **Approximating 1D Mixed States with Matrix Product Operators** we combine the knowledge from chapters 2 and 3. First an older result is presented, using entanglement entropy as a criterion for approximability of pure states by tensor networks. After that, we use a quantity called *entanglement of purification* to construct a criterion for approximabilty of *mixed*

states by similar tensor networks.

The fifth chapter **Symmetries and Purifications** describes a novel method of connecting spontaneous symmetry breaking in ground states of symmetric Hamiltonians to entanglement in purifications of said ground states. This approach is directly applied in chapter 6, but it is first presented as a stand-alone chapter because its applications are potentially much broader.

In the sixth chapter **Hubbard Model with Permutation Symmetry**, we investigate a particular quantum many-body model. The model is a version of the famous *Hubbard model* with permutation symmetry of all its sites. The chapter showcases a particular use of the method described in chapter 5 and shows the appearance of some interesting ground state phases for particular configurations of the model Hamiltonian.

The seventh chapter **Hubbard Model in Infinite Dimensions** describes some research done into the *Hubbard model on infinite-dimensional cubic lattice* utilising methods described earlier in the thesis. In particular, different variational families of quantum states are used to approximate the true ground state energy.

The thesis is supported by three appendices. Appendix A **Hubbard Model Phase Diagram** contains the lengthier calculations involved in finding the ground state of the Hubbard model in chapter 6 and plotting the associated phase diagram. Appendix B **Selected Topics from Functional Analysis** contains information about some topics from functional analysis needed for the rigorous mathematical foundations of the results in this thesis. And finally appendix C **Rényi Mutual Information** contains some unpublished and relatively inconsequential original research into Rényi mutual information.

The relations between the various parts of the thesis are summarized in the figure 1.1. While the two main projects covered in this thesis tackle very different problems of quantum many-body physics (approximability by tensor networks and spontaneous symmetry breaking in the permutationally-invariant Hubbard model), the approach they use is similar (see figure 1.2). In both cases we have a physical many-body system, which we purify into an ancillary subspace. In both cases we are interested in the entanglement of this purification.

Publications

This thesis contains results from one article published during the PhD and one article currently in preparation.

- Efficient Description of Many-Body Systems with Matrix Product Density Operators. **Jiří Guth Jarkovský**, András Molnár, Norbert Schuch and J. Ignacio Cirac. PRX Quantum, 1, 010304, (2020) [62]

  The content of this article is covered in chapter 4.

- Detecting Spontaneous Symmetry Breaking with Purifications. **Jiří Guth Jarkovský**, Lorenzo Piroli and J. Ignacio Cirac. *In preparation* [63]

  The content of this article is covered in chapters 5 and 6.

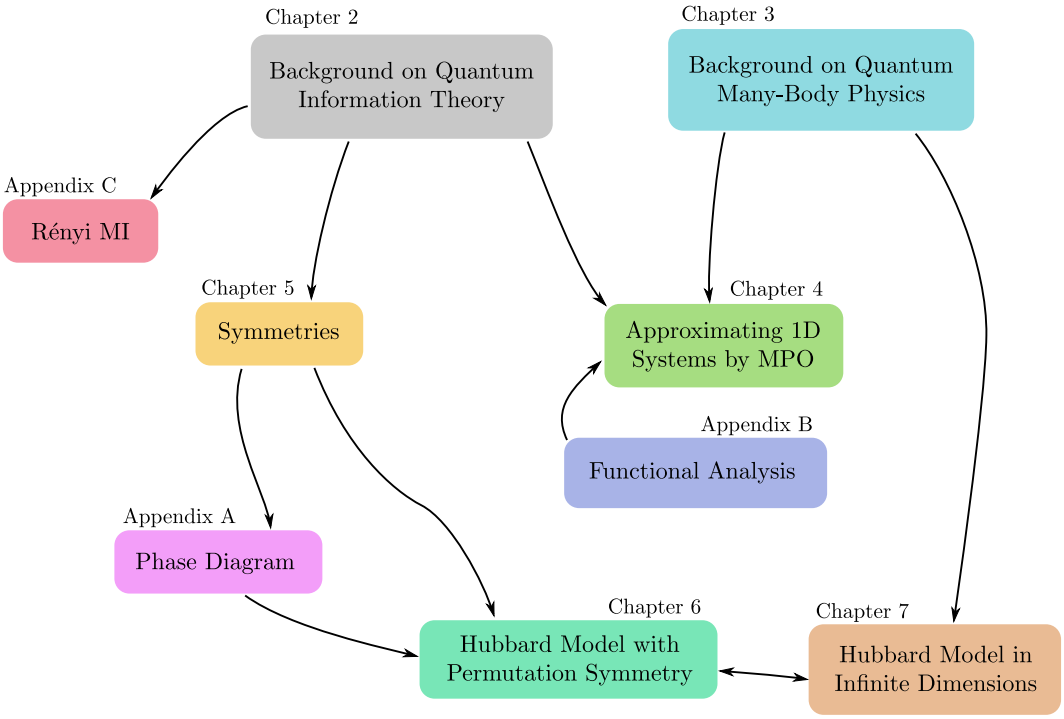


Figure 1.1: The main relations between the various topics within the thesis visualized. Figure inspired by a similar figure of de Groot [61].

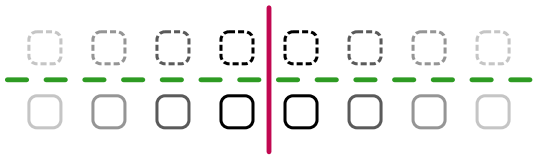


Figure 1.2: Both projects covered in this thesis use the entanglement of purifications. In the approximability project we are interested in entanglement between two parts of the system (solid red line), whereas in the symmetry-breaking project we use the entanglement between the ancillary subspace and the physical space (dashed green line).

Chapter 2

Background on Quantum Information Theory

In this chapter we review some necessary background on quantum information theory to serve as a reference for the rest of the thesis. We will go over the basic definitions of quantum states (section 2.1), entanglement and the various quantities used to quantify it (section 2.2). We will also review some important topics needed for the later chapters — purifications of quantum states (section 2.3), majorization (section 2.4), norms (section 2.5), fermionic systems (section 2.6) and some methods of quantifying the correlations present in mixed states (subsections 2.2.3 and 2.3.1). For a much deeper look into quantum information theory, there are excellent textbooks by Wilde [64], Watrous [65] or Nielsen and Chuang [66].

The chapter assumes knowledge of linear algebra and of the basic principles of quantum physics but most required concepts from quantum information theory will be developed from the ground up.

2.1 Quantum States

In this section we define basic quantum states and describe some of their properties. A quantum system corresponds to a Hilbert space $\mathcal{H}$. We denote $\mathcal{B}(\mathcal{H})$ the set of all bounded operators acting on $\mathcal{H}$. We may label one orthonormal basis of $\mathcal{H}$ the *computational basis* and label its elements $\{|i\rangle\}_{i=0}^{\dim \mathcal{H}-1}$. Then all vectors in the Hilbert space may be expressed as complex vectors of coefficients with respect to this basis (and all operators in $\mathcal{B}(\mathcal{H})$ may be expressed as matrices).

2.1.1 Pure and Mixed States

Generally speaking, quantum states are described by a density matrix ρ. A density matrix is an operator acting on the Hilbert space $\rho \in \mathcal{B}(\mathcal{H})$. In order to describe a valid quantum state, the density matrix has to be postive-semidefinite $\rho \geq 0$ and normalized $\mathrm{Tr}\{\rho\} = 1$. We will refer to the set of all valid density matrices over the Hilbert space $\mathcal{H}$ as $\mathcal{S}(\mathcal{H})$.

As a positive-semidefinite matrix, ρ has non-negative eigenvalues $\lambda_i \geq 0$. The normalization condition ensures that the sum of all the eigenvalues is equal to one

$\sum_i \lambda_i = 1$.

If ρ has one eigenvector $|\phi\rangle \in \mathcal{H}$ with an eigenvalue 1, then the quantum state is called *pure*. In that case the density matrix can be written as an outer product of the vector $|\phi\rangle$ with itself $\rho = |\phi\rangle\langle\phi|$. When talking about pure states, it is common to only refer to the vector $|\phi\rangle$ (as opposed to the corresponding density matrix). If a quantum state is not pure, it is *mixed*. The density matrix of a mixed state is a convex combination of density matrices of pure states and can be interpreted as a statistical ensemble of multiple different pure states.

Definition 1 (Maximally mixed state)**.** *Let $\mathcal{H}$ be a d-dimensional Hilbert space. The state*

$$\rho = \frac{\mathbb{1}}{d} \in \mathcal{S}(\mathcal{H}) \tag{2.1}$$

is called the maximally mixed state*.*

The maximally mixed state corresponds to having no information about the state of the system (other than the Hilbert space dimension). It can be interpreted as an equal mixture of all the (pure) basis states of $\mathcal{H}$.

2.1.2 Multipartite States

Two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ may be multiplied to create a larger Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. This corresponds to a larger quantum system being composed of two subsystems. In quantum information theory, a d-dimensional Hilbert space without any extra structure is called a *qudit*. Except for section 2.6, throughout this chapter we assume that the Hilbert space $\mathcal{H}$ is constructed as a tensor product of multiple qudit Hilbert spaces.

Definition 2 (Partial Trace)**.** *Let $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be a bipartite quantum state.* A partial trace *over the subspace $\mathcal{H}_B$ is an operation*

$$\mathrm{Tr}_B\{\cdot\} : \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{S}(\mathcal{H}_A). \tag{2.2}$$

Let $\rho_A = \mathrm{Tr}_B\{\rho_{AB}\}$. Then the partial trace satisfies

$$\forall O \in \mathcal{B}(\mathcal{H}_A) : \mathrm{Tr}\{\rho_A O\} = \mathrm{Tr}\{\rho_{AB}(O \otimes \mathbb{1})\}. \tag{2.3}$$

We refer to ρ_A as the reduced state of ρ_{AB} to subsystem A*.*

The physical interpretation is that ρ_A describes the state of ρ_{AB} on subsystem A only. The partial trace corresponds to an operation of discarding (or hiding) the subsystem B. Equality (2.3) ensures that ρ_A has the same expectation values as ρ_{AB} with any observable acting on system A only.

2.2 Entropies and Entanglement

In classical thermodynamics, *entropy* is described as a measure of disorder in a system. In classical information theory, it is the measure of information (or lack

thereof). Quantum information theory takes inspiration from the interpretation of classical information theory and applies it to quantum systems. Because there are many different ideas of how to measure (the lack of) information, there are many different entropies. In classical information theory, the most common one is the Shannon entropy [1]. Because of the connections between classical information theory and quantum information theory, all logarithms throughout this thesis are base 2.

Definition 3 (Shannon entropy)**.** *For a random variable X with possible outcomes $x_1, x_2, ..., x_n$, we define the Shannon entropy of X as*

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log \left[p(x_i)\right], \tag{2.4}$$

where $p(x_i)$ is the probability of the outcome x_i and we use the convention that $p(x_i) \log \left[p(x_i)\right] = 0$ if $p(x_i) = 0$.

In quantum information theory, we calculate the entropy of quantum states instead of distributions of random variables.

Definition 4 (von Neumann entropy)**.** *For a density matrix ρ with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$, we define the von Neumann entropy of ρ as*

$$S(\rho) = -\operatorname{Tr}\{\rho \log \rho\} = -\sum_{i=1}^{n} \lambda_i \log \left[\lambda_i\right], \tag{2.5}$$

where we use the convention that $\lambda_i \log \left[\lambda_i\right] = 0$ if $\lambda_i = 0$.

From the formula (2.5) we can see that the von Neumann entropy of a pure state is equal to 0 (analogous to the Shannon entropy of a distribution with just one outcome).

2.2.1 Rényi Entropies

There are many different ideas of what it means to measure information, and many different entropies capture those ideas. One particular family of entropies are the *Rényi entropies* [67]. For brevity we will only state the definition for quantum states, as the formula looks very similar whether applied to classical random variables or quantum states.

Definition 5 (Rényi entropies)**.** *Let ρ be a quantum state density matrix and $\alpha \in (0, 1) \cap (1, +\infty)$ be a real parameter. The α-Rényi entropy of ρ is defined as*

$$S_\alpha(\rho) = \frac{\log(\operatorname{Tr}\{\rho^\alpha\})}{1 - \alpha}. \tag{2.6}$$

The limits $\lim_{\alpha \to 1^+} S_\alpha(\rho)$ and $\lim_{\alpha \to 1^-} S_\alpha(\rho)$ are both well defined and they are equal to the von Neumann entropy $S(\rho)$ (which can also be labelled as $S_1(\alpha)$ for consistency). Two other entropies (which measure information in yet another way) are the min- and max- entropies.

Definition 6 (Min- and max-entropies)**.** *The min-entropy of a quantum state ρ with eigenvalues $\lambda_1, \lambda_2, ...\lambda_n$ is*

$$S_\infty(\rho) = -\log \max_i \{\lambda_i\}. \tag{2.7}$$

The max-entropy [68] of a quantum state ρ is

$$S_0(\rho) = \log \operatorname{rank} \rho. \tag{2.8}$$

As the notation suggests, the limits of the Rényi entropies $\lim_{\alpha\to 0^+} S_\alpha(\rho)$ and $\lim_{\alpha\to+\infty} S_\alpha(\rho)$ are both well defined and they are equal to the min- and max-entropies respectively.

Example 1 (The effect of α in Rényi entropies)
Roughly speaking, as α decreases, the corresponding Rényi entropies give more weight to the number of eigenvalues of ρ and less weight to their value.

Consider the following two density matrices

$$\rho_1 = 0.8 \left|0\right\rangle \left\langle 0\right| + 0.1 \left|1\right\rangle \left\langle 1\right| + 0.1 \left|2\right\rangle \left\langle 2\right|, \tag{2.9}$$
$$\rho_2 = 0.5 \left|0\right\rangle \left\langle 0\right| + 0.5 \left|1\right\rangle \left\langle 1\right|. \tag{2.10}$$

These have the following Rényi entropies

$$S_{1/2}(\rho_1) = 1.2212..., \qquad S_2(\rho_1) = 0.5995..., \tag{2.11}$$
$$S_{1/2}(\rho_2) = 1, \qquad S_2(\rho_2) = 1. \tag{2.12}$$

As we can see, according to 1/2-Rényi entropy, the density matrix ρ_1 is more "mixed" than density matrix ρ_2, but according to 2-Rényi entropy, it is the other way around.

Example 2 (Calculations of Rényi entropies)
In this example we demonstrate the application of the formula (2.6) to a few special example states.

First consider a pure state $\rho = \left|\phi\right\rangle \left\langle\phi\right|$. This density matrix taken to any power α remains the same: $\rho^\alpha = \left|\phi\right\rangle \left\langle\phi\right|^\alpha = \left|\phi\right\rangle \left\langle\phi\right|$. The trace of this is 1 and the logarithm of 1 is 0. Therefore all Rényi entropies of this state are equal to 0 (and by applying the limits of $\alpha \to 0^+, 1, +\infty$ also the min-, von Neumann and max-entropies).

Now consider the maximally mixed state $\rho = \frac{\mathbb{1}}{d}$. The eigenvalues of this density matrix are all equal to $\frac{1}{d}$. After taking the αth power, the eigenvalues will transform to $\frac{1}{d^\alpha}$. The trace of ρ^α will then be $\frac{d}{d^\alpha} = d^{1-\alpha}$. The logarithm of that is $(1-\alpha)\log d$. All Rényi entropies of this state will therefore be $\log d$ (and by applying the limits of $\alpha \to 0^+, 1, +\infty$ also the von Neumann, min- and max-entropies).

Lemma 1. *Rényi entropies are non-increasing in α, i.e.*

$$\forall \alpha, \beta \in (0, \infty) : \alpha \geq \beta \implies S_\alpha(\rho) \leq S_\beta(\rho). \tag{2.13}$$

As can be seen in example 2, the Rényi entropies may not be strictly decreasing in α. In fact, they can be all the same.

2.2.2 Schmidt Decomposition

The Schmidt decomposition is a particular way to write a bipartite pure state as a sum of product states.

Definition 7 (Schmidt Decomposition)**.** *Let $|\Psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a bipartite pure state. Let d_A be the dimension of the subspace $\mathcal{H}_A$ and d_B the dimension of the subspace $\mathcal{H}_B$. The* Schmidt Decomposition *of $|\Psi\rangle_{AB}$ is the following expression*

$$|\Psi\rangle_{AB} = \sum_{i=1}^{\min(d_A,d_B)} \sqrt{\lambda_i}\, |\chi_i\rangle_A \otimes |\xi_i\rangle_B \,, \tag{2.14}$$

where $\sqrt{\lambda_i}$ are the real Schmidt coefficients *satisfying $\sum_i \lambda_i = 1$ and the states $|\chi_i\rangle_A$ and $|\xi_i\rangle_B$ are elements of orthonormal bases in their respective Hilbert spaces.*

While the Schmidt decomposition is not unique, the Schmidt coefficients are (up to reordering). The Schmidt decomposition is closely tied to entanglement of pure states.

Definition 8 (Entanglement of pure states)**.** *A bipartite pure state is called a* product state *if it has only one non-zero Schmidt coefficient (which is equal to* 1*). Otherwise it is* entangled*.*

A pure product state is named as such because it can be written as a tensor product of two pure states

$$|\Psi\rangle_{AB} = |\chi\rangle_A \otimes |\xi\rangle_B \,. \tag{2.15}$$

Equation (2.15) is technically a Schmidt decomposition with only a single non-zero Schmidt coefficient.

If we take a partial trace of the state $|\Psi\rangle_{AB}$ from equation (2.14) and plug in the Schmidt decomposition, i.e.

$$\mathrm{Tr}_A \left\{ |\Psi\rangle \langle\Psi|_{AB} \right\} = \sum_{i=1}^{\min(d_A,d_B)} \lambda_i \, |\xi_i\rangle \langle\xi_i|_B \,, \tag{2.16}$$

$$\mathrm{Tr}_B \left\{ |\Psi\rangle \langle\Psi|_{AB} \right\} = \sum_{i=1}^{\min(d_A,d_B)} \lambda_i \, |\chi_i\rangle \langle\chi_i|_A \,, \tag{2.17}$$

we learn of another meaning of the Schmidt coefficients. We see that they are the square roots of the eigenvalues of either of the reduced density matrices we obtain by tracing out one part of the system. This implies that the reduced state of a product pure state is again a pure state. Conversely, the reduced state of an entangled pure state is always mixed. Furthermore, we can quantify entanglement by the entropy of the reduced density matrix. As can be seen from equations (2.16) and (2.17), the two reduced density matrices have the same non-zero eigenvalues and therefore

they have the same von Neumann and Rényi entropies. The literature usually refers to the *entropy of entanglement* or *entanglement entropy* when talking about the von Neumann entropy of the reduced states as a measurement of entanglement of the pure bipartite state.

Example 3 (Entropy of entanglement)
Consider the following three bipartite states:

$$|\phi_1\rangle = |00\rangle\,, \tag{2.18}$$
$$|\phi_2\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \tag{2.19}$$
$$|\phi_3\rangle = \frac{|00\rangle + \sqrt{2}\,|11\rangle}{\sqrt{3}}. \tag{2.20}$$

Their reduced density matrices on the first subsystem are, in order:

$$\rho_1 = |0\rangle\langle 0|\,, \tag{2.21}$$
$$\rho_2 = \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}, \tag{2.22}$$
$$\rho_3 = \frac{|0\rangle\langle 0| + 2\,|1\rangle\langle 1|}{3}. \tag{2.23}$$

The von Neumann entropies of these three density matrices are:

$$S(\rho_1) = 0, \tag{2.24}$$
$$S(\rho_2) = \log(2) = 1, \tag{2.25}$$
$$S(\rho_3) = -\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3} \approx 0.9183... \tag{2.26}$$

We have $S(\rho_2) > S(\rho_3) > S(\rho_1)$, which implies that the state $|\phi_2\rangle$ is more entangled than the state $|\phi_3\rangle$, which in turn is more entangled than the state $|\phi_1\rangle$ (which is a product state).

For completeness we also provide the definition of entanglement of mixed states.

Definition 9 (Entanglement of mixed states)**.** *A bipartite mixed state $\rho_{AB} \in \mathcal{B}(\mathcal{H}_{AB})$ is called a* product state *if it can be written as a tensor product of two states*

$$\rho_{AB} = \sigma_A \otimes \theta_B \tag{2.27}$$

for some density matrices $\sigma_A \in \mathcal{B}(\mathcal{H}_A)$ and $\theta_B \in \mathcal{B}(\mathcal{H}_B)$. A bipartite mixed state is called a separable state *if it can be written as a convex combination of product states. Otherwise it is an* entangled state*.*

Unfortunately, for mixed states there is no connection between the entropy of the reduced states and the entanglement present therein.

Example 4 (Un-entangled mixed states)
Consider the bipartite state $\rho_{AB} = \frac{\mathbb{1}}{d_A} \otimes \frac{\mathbb{1}}{d_B}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ (with dimensions d_A and d_B respectively). This corresponds to the maximally mixed state on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Its reduced states on the subsystems A and B are $\rho_A = \frac{\mathbb{1}}{d_A}$ and $\rho_B = \frac{\mathbb{1}}{d_B}$ respectively. These two states have their von Neumann entropies $S(\rho_A) = \log d_A$ and $S(\rho_B) = \log d_B$ respectively. However, the state ρ_{AB} is not entangled at all. It is in fact a product state! As we can see, the von Neumann entropy of reduced states is not a good quantity to quantify entanglement of mixed states.

The Schmidt decomposition can also be defined for operators, but it does not have the same properties as the Schmidt decomposition for pure states:

Definition 10 (Operator-Schmidt decomposition)**.** *The Schmidt decomposition of a bipartite operator is*

$$\rho_{AB} = \sum_{i=1}^{\min[(d_A)^2,(d_B)^2]} \gamma_i \left(\Sigma_i\right)_A \otimes \left(\Pi_i\right)_B , \tag{2.28}$$

where $\gamma_i \in \mathbb{R}$ *are the* operator-Schmidt coefficients*. The operators* Σ_i, Π_i *are elements of* $\mathcal{B}(\mathcal{H}_A)$ *and* $\mathcal{B}(\mathcal{H}_B)$ *respectively, orthonormal with respect to the Hilbert Schmidt inner product, i.e.*

$$\mathrm{Tr}\Big\{\Sigma_i^\dagger \Sigma_j\Big\} = \mathrm{Tr}\Big\{\Pi_i^\dagger \Pi_j\Big\} = \delta_{ij}. \tag{2.29}$$

2.2.3 Mutual Information and Relative Entropy

Mixed states may contain some correlations that are not entanglement. Those are usually referred to as *classical correlations*. A quantity that allows us to quantify all correlations present in a mixed quantum state is *mutual information*. The mutual information of a bipartite density matrix ρ_{AB} is defined as

$$I(A:B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{2.30}$$

Alternatively, the mutual information may be defined in a number of equivalent ways [64]:

$$\begin{aligned} I(A:B)_\rho &= D(\rho_{AB}||\rho_A \otimes \rho_B) && (2.31)\\ &= \min_{\sigma_B} D(\rho_{AB}||\rho_A \otimes \sigma_B) && (2.32)\\ &= \min_{\sigma_A} D(\rho_{AB}||\sigma_A \otimes \rho_B) && (2.33)\\ &= \min_{\sigma_A,\sigma_B} D(\rho_{AB}||\sigma_A \otimes \sigma_B) && (2.34) \end{aligned}$$

where the minimizations run over valid density matrices and $D(\cdot||\cdot)$ is the quantum relative entropy.

Definition 11 (Quantum relative entropy)**.** *Quantum relative entropy between two density matrices ρ and σ is*

$$D(\rho||\sigma) = \mathrm{Tr}\{\rho\,(\log\rho - \log\sigma)\}. \tag{2.35}$$

The mutual information has several desirable properties. It is zero for product states and non-zero for entangled states and classically-correlated separable states. For illustration of mutual information calculations, see example 5.

There are several ways to generalize the mutual information into Rényi mutual information. These generalizations and some original research into their properties can be found in appendix C

Similar to mutual information is a quantity called the *conditional mutual information*. This quantity is defined for tripartite states as

$$I(A:B|C)_\rho = S(\rho_{AC}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_C). \tag{2.36}$$

As the name suggests, conditional mutual information gives us the expected value of the mutual information between two subsystems, given knowledge of the state on a third subsystem. The non-negativity of conditional mutual information $I(A:B|C)_\rho \geq 0$ is called the *strong subadditivity of von Neumann entropy* [69].

2.3 Purifications

In this section we introduce the concept of *purifications*. Purifications play a significant role in chapters 4, 5 and 6.

Definition 12 (Purification)**.** *For a density matrix ρ_A on a Hilbert space $\mathcal{H}_A$ a purification is a pure state $|\Psi\rangle_{AB}$ on a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B = \mathcal{H}_{AB}$ such that*

$$\mathrm{Tr}_B\left\{|\Psi\rangle\langle\Psi|_{AB}\right\} = \rho_A. \tag{2.37}$$

Note that a purification is not unique. In particular, for any isometry U_B, the state $\mathbb{1}_A \otimes U_B\,|\Psi\rangle_{AB}$ is also a purification of ρ_A.

2.3.1 Entanglement of Purification

Entanglement of purification [70] is a measure of correlations in a mixed state. We use entanglement of purification as a criterion of approximability of mixed states in chapter 4.

Definition 13 (Entanglement of purification)**.** *For a bipartite density matrix ρ_{AB}, the* entanglement of purification *is defined as*

$$E^p(\rho_{AB}) = \min_{|\Psi\rangle_{AA'BB'}} S\left(\rho_{AA'}\right), \tag{2.38}$$

where the minimization runs over all purifications $|\Psi\rangle_{AA'BB'}$ of ρ_{AB} that use two ancillary subspaces A' and B'. The density matrix $\rho_{AA'} = \mathrm{Tr}_{BB'}\,|\Psi\rangle\langle\Psi|_{AA'BB'}$ is this purification with both the subsystems B and B' traced out.

The von Neumann entropy of $\rho_{AA'}$ can be interpreted as the entanglement of the purification $|\Psi\rangle_{AA'BB'}$, hence its name. The entanglement of purification captures both classical and quantum correlations, similarly to mutual information.

The entanglement of purification has several interesting properties. In the following we will denote by $A : BC$ when a tripartite state is interpreted as a bipartite state with the subsystems B and C together being treated as one. The entanglement of purification is bounded from above by the von Neumann entropy of the reduced density matrices

$$E^p(\rho_{AB}) \leq \min[S(\rho_A), S(\rho_B)], \tag{2.39}$$

it is monotonic with respect to the partial trace

$$E^p(\rho_{A:BC}) \geq E^p(\rho_{AB}), \tag{2.40}$$

and it is bounded from below by half the mutual information

$$E^p(\rho_{AB}) \geq \frac{I(A:B)_\rho}{2}. \tag{2.41}$$

For an overview of the entanglement of purification, see Nguyen et. al. [71].

Recently, there has been a new surge of interest in entanglement of purification [72–78]. This has been sparked by a seminal paper [79] suggesting that a quantity called the *minimal cross-section of the entanglement wedge* could be its holographic counterpart in the setting of the *anti-de Sitter / conformal field theory* (AdS/CFT) correspondence [80]. The AdS/CFT correspondence is an important conjecture linking string theory to quantum field theories, and is therefore receiving a lot of attention in the theoretical physics community [81–83]. The recent wave of research is focused mostly on the AdS/CFT interpretation of entanglement of purification and is therefore not particularly relevant for the topic of this thesis.

In comparison to quantum mutual information, the entanglement of purification gives more weight to classical correlations, as can be seen in example 5.

Example 5 (Entanglement of purification vs. mutual information)
Consider the following two states:

$$|\Phi\rangle_{AB} = \sum_i \sqrt{\lambda_i}\, |i\rangle_A \otimes |i\rangle_B\,, \tag{2.42}$$

$$\rho_{AB} = \sum_i \lambda_i\, |i\rangle\langle i|_A \otimes |i\rangle\langle i|_B\,. \tag{2.43}$$

Both these states have the same reduced states on the subsystems A and B, $\rho_{A/B} = \sum_i \lambda_i\, |i\rangle\langle i|_{A/B}$. However, the first state is pure and entangled whereas the second state is mixed and classically correlated. The mutual information of the pure state is

$$I(A:B)_\Phi = S(\rho_A) + S(\rho_B) - S(|\Phi\rangle\langle\Phi|_{AB}) = 2S(\rho_A). \tag{2.44}$$

In contrast the mutual information of the mixed state is

$$I(A:B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = S(\rho_A), \tag{2.45}$$

which is half the mutual information of the pure state.

Now consider the entanglement of purification. Since $|\Phi\rangle_{AB}$ is pure, any of its purifications will be a product state

$$|\Phi\rangle_{AA'BB'} = |\Phi\rangle_{AB} \otimes |\Phi'\rangle_{A'B'}\,, \tag{2.46}$$

whose reduced state on subsystems AA' looks like

$$\rho_{AA'} = \rho_A \otimes \rho'_{A'}. \tag{2.47}$$

The entropy of $\rho_{AA'}$ is be minimized if $\rho'_{A'}$ is pure, in which case it equals

$$E^p(|\Phi\rangle\langle\Phi|_{AB}) = S(\rho_{AA'}) = S(\rho_A). \tag{2.48}$$

One possible purification of ρ_{AB} is

$$|\Psi\rangle_{AA'BB'} = \sum_i \sqrt{\lambda_i}\,|i\rangle_A \otimes |i\rangle_B \otimes |0\rangle_{A'} \otimes |i\rangle_{B'}\,. \tag{2.49}$$

with entanglement

$$E^p(\rho_{AB}) = S(\rho_{AA'}) = S(\rho_A). \tag{2.50}$$

We skip the proof that this is optimal and instead refer the reader to Nguyen et al. [71].

Both states have the same entanglement of purification, but they have different mutual information. This can be interpreted as the mutual information giving less "weight" to classical correlation whereas entanglement of purification treating classical correlations on equal footing with entanglement.

The entanglement of purification can be easily generalized to the *Rényi entanglement of purification* by simply replacing the von Neumann entropy in the definition (2.38) with a desired Rényi entropy,

$$E^p_\alpha(\rho_{AB}) = \min_{|\Psi\rangle_{AA'BB'}} S_\alpha(\rho_{AA'}). \tag{2.51}$$

The Rényi entanglement of purification retains most of the important properties of (von Neumann) entanglement of purification. For instance the inequalities (2.39) and (2.40) still hold. Even an analogy of inequality (2.41) holds, but to properly state it and prove it, we first need to define Rényi mutual information. For that reason, the Rényi version of (2.41) is only stated and proven in section C.3 of appendix C.

2.4 Majorization

Now we summarize a bit of information about majorization, an important concept in quantum information theory (see theorem 1). In this thesis we will only use it once as a part of a proof in chapter 4. In statistical mathematics, majorization provides the notion of a preorder on probability distributions.

Definition 14. *We say that a discrete probability distribution* $\{p_i\}_{i=1}^d$ majorizes *a different probability distribution* $\{q_i\}_{i=1}^d$ *(symbolically* $p \succ q$*) if*

$$\forall k \in \{1, 2, ...d\} : \quad \sum_{i=1}^{k} p_i^{\downarrow} \geq \sum_{i=1}^{k} q_i^{\downarrow}. \tag{2.52}$$

Here $\{p_i^{\downarrow}\}_{i=1}^d$ *and* $\{q_i^{\downarrow}\}_{i=1}^d$ *are the distributions obtained from* $\{p_i\}_{i=1}^d$ *and* $\{q_i\}_{i=1}^d$ *by reordering their values from largest to lowest.*

Majorization has several uses throughout quantum information theory. Namely, it governs the allowed transformations of bipartite pure states via local operations and classical communication (LOCC).

Theorem 1 (Nielsen's theorem [84])**.** *A bipartite pure state* $|\phi_1\rangle_{AB} = \sum_{i=1}^d \sqrt{\lambda_i} |i\rangle_A |i\rangle_B$ *can be deterministically transformed to* $|\phi_2\rangle_{AB} = \sum_{i=1}^d \sqrt{\sigma_i} |i\rangle_A |i\rangle_B$ *via local operations assisted by classical communication if and only if the distribution* $\{\lambda_i\}_{i=1}^d$ *is majorized by the distribution* $\{\sigma_i\}_{i=1}^d$*.*

In this thesis, we will use majorization in context of Schur-convex functions.

Definition 15 (Schur-convexity)**.** *A multi-variate function* f *is called* Schur-convex *if and only if* $f(p) \leq f(q)$ *for any probability distributions* p, q *satisfying* $p \prec q$*. It is called* Schur-concave *if its negative* $-f$ *is* Schur-convex*.*

Importantly for our work, all Rényi entropies are Schur-concave, where the probability distribution is given by the eigenvalues of the input density matrix [85].

2.5 Norms

In this section we go over some commonly used norms in quantum information theory and beyond. Knowledge of various norms and their relations is crucial to be able to prove efficient approximations of quantum states by tensor networks.

Any Hilbert space has a norm induced by the inner product. The norm of a pure state is an example of one such norm, defined as

$$\| |\phi\rangle \| = \sqrt{\langle \phi | \phi \rangle}. \tag{2.53}$$

Operators acting on a Hilbert space have a norm induced by this vector norm. This is called the *operator norm* $\|\cdot\|_\infty$ and it is defined as

$$\|O\|_\infty = \sup_{|\phi\rangle \in \mathcal{H}} \frac{\|O |\phi\rangle\|}{\|\phi\|}. \tag{2.54}$$

Generally, for operators we can introduce a family of norms called the *Schatten norms*. They are defined for $p \in [1, +\infty)$ as

$$\|O\|_p = (\mathrm{Tr}\{|O|^p\})^{\frac{1}{p}}. \tag{2.55}$$

The special case of Schatten norm with $p = 1$ is called the *trace norm*. The limit of $p \to +\infty$ corresponds to the operator norm $\|\cdot\|_\infty$.

The Shatten norms have multiple useful properties [65]:

1. They are non-increasing in p. For any $p, q \in [0, +\infty]$ and any operator A, we have
$$p \leq q \Rightarrow \|A\|_p \geq \|A\|_q. \tag{2.56}$$

2. There exists a duality between pairs of the norms. For any $p, q \in [1, +\infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$ we have
$$\|O\|_p = \sup_T \frac{\left|\mathrm{Tr}\left\{O^\dagger T\right\}\right|}{\|T\|_q}. \tag{2.57}$$
Relaxing the supremum, this turns into a useful inequality
$$\|O\|_p \|T\|_q \geq \left|\mathrm{Tr}\left\{O^\dagger T\right\}\right|. \tag{2.58}$$
This is an analogy of Hölder's inequality for L^p norms of integrable functions [86].

3. For any $p \in [0, +\infty]$, Schatten norms satisfy
$$\|ABC\|_p \leq \|A\|_\infty \|B\|_p \|C\|_\infty. \tag{2.59}$$

4. Despite the non-increasing property of equation (2.56), there exists an upper bound on $\|\cdot\|_1$ in terms of $\|\cdot\|_2$. For any A, we have
$$\|A\|_1 \leq \sqrt{d}\|A\|_2, \tag{2.60}$$
where d is the dimension of the Hilbert space that the operator A acts on.

The *trace distance* is a frequently used distance measure between two density matrices, defined as half of the trace norm of the difference of the two density matrices $\frac{1}{2}\|\rho - \sigma\|_1$. Because of the inequality (2.58), the trace distance may be used to bound the difference in expectation values of a bounded observable O in two states ρ and σ,

$$\mathrm{Tr}\{\rho O\} - \mathrm{Tr}\{\sigma O\} \leq \|O\|_\infty \|\rho - \sigma\|_1. \tag{2.61}$$

For pure states, the trace distance reduces to [66]

$$\frac{1}{2}\||\phi\rangle\langle\phi| - |\psi\rangle\langle\psi|\|_1 = \sqrt{1 - |\langle\phi|\psi\rangle|^2}. \tag{2.62}$$

2.6 Fermionic States

All the previous sections in this chapter discussed systems of qubits in which the information is stored in an effective spin, e.g. an inner degree of freedom of an atom, a polarization angle of a photon or the actual spin of an electron. However, in chapters 6 and 7 we want to describe different versions of the Hubbard model, which is a model of fermions. Investigating the states of particles following the Fermi-Dirac statistics brings with it a new layer of intricacy, which requires different mathematical structures to be introduced.

To describe fermionic Hilbert spaces, we first need to define the fermionic annihilation and creation operators a_i and $a_i^\dagger$. These operators follow the *canonical anti-commutation relations* (CAR):

$$\{a_i, a_j\} = 0, \tag{2.63}$$

$$\{a_i^\dagger, a_j^\dagger\} = 0, \tag{2.64}$$

$$\{a_i, a_j^\dagger\} = \delta_{ij}. \tag{2.65}$$

Here the subindex i, j refers to the species of the fermion. In practice the species index can be split further to represent some other labelling of the fermions, e.g. their position on a lattice or their momentum. The physical interpretation is that the creation operator $a_i^\dagger$ adds a fermion of species i to the system, and the annihilation operator a_i removes a fermion of species i from the system. The anti-commutation relations imply that there can only be one (or none) fermion of species i in the system.

To build a Hilbert space of discrete fermionic systems, we begin with the vacuum state $|\Omega\rangle$. This state corresponds to a system without any particles in it. The annihilation operators are defined to destroy the vacuum state: $a_i |\Omega\rangle = 0$, but we can construct physical states by acting with creation operators on the vacuum state $a_i^\dagger |\Omega\rangle$. Because of the anti-commutation relation (2.64), acting with the same creation operator twice will destroy the state. Therefore the dimension of the Hilbert space is 2^K where K is the cardinality of the set of possible subindices i of the creation operators.

Because of the anti-commutation of the creation operators, the Hilbert space of a fermionic many-body system does not have the structure of a tensor product of individual Hilbert spaces, as was the case for effective spin qubits. We can still use most of the concepts introduced earlier in this chapter, but we need to be more careful with some of the definitions.

We may divide the fermionic Hilbert space into subspaces based on the fermionic species present therein. This allows us to define a reduced state.

Definition 16 (Fermionic reduced state)**.** *Let $\mathcal{J}$ and $\mathcal{I}$ be two sets of fermionic species satisfying $\mathcal{J} \subset \mathcal{I}$. Let $\mathcal{H}_\mathcal{I}, \mathcal{H}_\mathcal{J}$ be the Hilbert spaces of fermionic pure states containing only the fermionic species from the corresponding set. Let $\rho \in \mathcal{B}(\mathcal{H}_\mathcal{I})$. We call $\sigma \in \mathcal{B}(\mathcal{H}_\mathcal{J})$ the reduced state of ρ onto the subspace $\mathcal{H}_\mathcal{J}$ if*

$$\forall O \in \mathcal{B}(\mathcal{H}_\mathcal{J}) : \mathrm{Tr}\{\rho O\} = \mathrm{Tr}\{\sigma O\}. \tag{2.66}$$

This allows us to use the define pure entangled / product states:

Definition 17. *A pure fermionic state is entangled / product across a bipartition of fermionic species if its reduced state is mixed / pure.*

Chapter 3

Background on Quantum Many-Body Physics

Quantum many-body physics is an area of research which studies quantum systems with many subsystems. The tensor-product nature of quantum system Hilbert spaces means that the dimensionality of the entire system usually grows exponentially fast with the number of subsystems. As an example, a system consisting of N d-dimensional qudits has a Hilbert space of dimensionality d^N. The large dimensionality of the Hilbert space is problematic for the study of such systems. Even just accurately describing the wavefunction of an arbitrary state in many-body Hilbert space requires computational resources exponentially large in the system size N.

Many different approaches have been developed to overcome this problem. One way to simplify the situation is to artificially restrict the set states to consider. Due to e.g. symmetries or entanglement structure, we may choose to simply limit the solution of a many-body problem to a relatively small set of states, which itself is easy to describe accurately. This set of states is typically characterized by a relatively small number of parameters (typically scaling polynomially in the system size) and it is called a *variational family* of states. In this chapter we will discuss two common variational families of states: *Gaussian states* (section 3.1) and *tensor networks* (section 3.2). These two families of states have become commonplace throughout quantum many-body physics.

Quantum many-body physics is a rich field with a plethora of problems, theorems, methods and approaches. In this chapter, we will focus on the topics particularly important for the content of this thesis and therefore we will omit a lot of other interesting topics. For a more in-depth look into quantum many-body physics, see e.g. Coleman [87].

3.1 Gaussian States

Gaussian states are a variational family of many-body states with a lot of useful properties. One of them is Wick's theorem [88] — the expectation values of higher-order products of creation and annihilation operators with Gaussian states can be reduced to products of expectation values of at most two creation and

annihilation operators [89]. Furthermore, Gaussian states are the exact ground states of quadratic (non-interacting) Hamiltonians. However, they also give very good results when used as variational family for interacting systems. That is what we will do in subsection 7.5.2 to tackle the infinite-dimensional Hubbard model.

Gaussian states exist for both bosonic and fermionic systems, but in this thesis we work exclusively with fermionic Gaussian states, so we will adapt all definitions and theorems to that. Gaussian states are obtained by acting on the vacuum $|\Omega\rangle$ with an operator, which is the exponential of an anti-Hermitian operator quadratic in the creation and annihilation operators, e.g. $|\psi\rangle = \mathrm{e}^{\mathrm{i}a^\dagger a}|\Omega\rangle$[1]. One way to parameterize Gaussian states is by parameterizing this unitary (or the anti-Hermitian exponent).

However, we will choose a different parametrization given by the covariance matrix Γ. Before properly defining Γ, we switch from the creation and annihilation operators to *Majorana fermionic* operators p and q:

$$a_i^\dagger = \frac{q_i - \mathrm{i}p_i}{\sqrt{2}}, \tag{3.1}$$

$$a_i = \frac{q_i + \mathrm{i}p_i}{\sqrt{2}}. \tag{3.2}$$

As a reminder, here the index i labels the site of the lattice on which those operators act. Those operators are Hermitian and they follow the following anti-commutation relations

$$\{q_i, q_j\} = \delta_{i,j}, \tag{3.3}$$

$$\{p_i, p_j\} = \delta_{i,j}, \tag{3.4}$$

$$\{q_i, p_j\} = 0. \tag{3.5}$$

with other anti-commutation relations being zero. With this we can define the covariance matrix of a Gaussian state $|\psi\rangle$

$$\Gamma_{xy} = -2i\,\langle\psi|\,xy\,|\psi\rangle + \mathrm{i}\delta_{x,y}, \tag{3.6}$$

where x, y can stand for any of $\{p_i, q_j\}$, labelling all of the fermionic indices. The letter p, q labels the "type" of the Majorana fermion and the sub-index labels the species of the original fermion. By construction, the Γ matrix is anti-symmetric. Due to Wick's theorem, any expectation value can be computed from two-point expectation values. The covariance matrix determines all possible two-point expectation values, so it determines all expectation values. This way, the covariance matrix determines the Gaussian quantum state while only having number of entries quadratically proportional to the number of fermionic species.

The covariance matrix Γ squares to negative identity

$$\Gamma^2 = \sum_y \Gamma_{xy}\Gamma_{yz} = -\mathbb{1}. \tag{3.7}$$

There exist many additional variational families built upon Gaussian states [90], and developing new ones better suited for specific applications is an active area of research [91, 92].

[1] That is where the name *Gaussian* comes from

3.1.1 Imaginary Time Evolution

A commonly used methods to minimize the energy of Hamiltonians is *imaginary time evolution*. We will only use it in connection to Gaussian states, so we mention it here. The idea of imaginary time evolution is simple. First, recall that (ordinary) time evolution operator of an isolated system with Hamiltonian H is

$$U(t) = \mathrm{e}^{\mathrm{i}tH}. \tag{3.8}$$

Similarly, the thermal state of a system with Hamiltonian H at temperature T is equal to

$$\rho_T = \frac{\mathrm{e}^{\frac{1}{T}H}}{Z}, \tag{3.9}$$

where $Z = \mathrm{Tr}\left\{\mathrm{e}^{\frac{1}{T}H}\right\}$ is the partition function. The ground state corresponds to $T = 0$, so its density matrix can in principle be obtained as $\rho_{\mathrm{GS}} = \lim_{T\to 0^+} \rho_T$. Note the similarities between the right-hand sides of equations (3.8) and (3.9). The main difference is the imaginary unit i in the (ordinary) time evolution. Interpreting the inverse temperature as imaginary time, we can try to find the ground state by evolving along this imaginary time far enough (reaching very low temperature T). Naturally, we need to normalize the state along the way, as the imaginary time evolution is not unitary.

If we perform imaginary time evolution within the manifold of Gaussian states, the covariance matrix evolves as [89]

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\Gamma = -4\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma} + \Gamma\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\Gamma\right). \tag{3.10}$$

3.2 Tensor Networks

In this section we provide the basic information about tensor networks and describe two simple TN architectures. But before doing that we need to clarify the nomenclature. Throughout this work, a tensor simply refers to a generalization of a matrix (an array of numbers, often complex numbers). Tensors are characterized by their indices: a tensor with one index corresponds to an ordinary vector in $\mathbb{C}^n$, a tensor with two indices corresponds to a matrix, etc. Throughout this work (and much of the tensor network literature), we make no distinction between covariant and contravariant indices.

The term "tensor network" refers to objects made up of multiple tensors contracted over some of their indices. As tensor networks can be quite complicated, we use the following graphical notation:

- A tensor is represented by a box[2].
- The indices of a tensor are represented by legs sticking out of the box.
- When the indices of two tensors are contracted, the corresponding legs are connected to form a link between the two boxes.

[2]Some authors use a circle

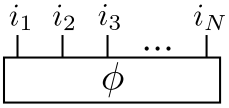
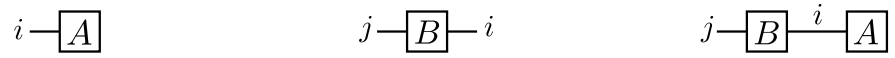


Figure 3.2: The tensor $\phi_{i_1,i_2,\,\ldots\,i_N}$ from equation (3.11) in the tensor network notation.

By contracting the indices of many tensors, we create a tensor network.

Example 6 (Simple tensor networks)
The tensor network notation can be used even for relatively simple objects. The following figure depicts some simple objects from linear algebra depicted using the notation of tensor networks.



(a) Vector A_i (b) Matrix B_{ji} (c) $\sum_i B_{ji}A_i$

Figure 3.1: A vector, a matrix and their product in the tensor network notation

Note that the product $\sum_i B_{ji}A_i$ in the tensor network notation has one free leg, which means that as a whole it is again a vector.

In quantum many-body physics tensor networks are used to represent complicated state vectors, operators or their expectation values. As will be shown below, the tensor network representation can provide very efficient ways to describe the desired objects. The question of which states can be represented (approximated) by tensor networks is discussed in chapter 4.

For a more in-depth overview of tensor networks and their methods see [93–96]. In the remainder of this section we will discuss in detail linear tensor networks — matrix product states and matrix product operators. In subsection 3.2.4 we briefly comment on projected entangled pair states (PEPS) — a tensor network architecture which generalizes matrix product states to 2D.

3.2.1 Matrix Product States and Matrix Product Operators

A generic pure state $|\phi\rangle$ of a system of N d-dimensional qudits can be described as

$$|\phi\rangle = \sum_{i_1=0}^{d-1} \sum_{i_2=0}^{d-1} \ldots \sum_{i_N=0}^{d-1} \phi_{i_1,i_2,\,\ldots\,,i_N} \, |i_1 i_2 \ldots i_N\rangle \,, \tag{3.11}$$

where $|i_1 i_2 ... i_N\rangle$ are the basis states of the Hilbert space written in the computational basis. The state is defined by its coefficients $\phi_{i_1,i_2,\ldots i_N}$. The object $\phi_{i_1,i_2,\ldots i_N}$ itself (depicted in figure 3.2) is a tensor containing d^N components, making it difficult to describe exactly.

Definition 18 (Matrix Product State)**.** *The matrix product state [97] (MPS) is a pure state of the following form:*

$$|\phi\rangle = \sum_{i_1=0}^{d-1}\sum_{i_2=0}^{d-1}\ldots\sum_{i_N=0}^{d-1} \; \overset{i_1}{\boxed{A^{[1]}}}\!\!-\!\!\overset{i_2}{\boxed{A^{[2]}}}\!\!-\;\ldots\;-\!\!\overset{i_N}{\boxed{A^{[N]}}}\; |i_1 i_2 \ldots i_N\rangle. \tag{3.12}$$

Here the objects $A^{[j]}$ (for $j \notin \{1, N\}$) are tensors with dimensions $d \times D_{j-1} \times D_j$. The tensors $A^{[1]}$ and $A^{[N]}$ have dimensions $d \times D_1$ and $d \times D_{N-1}$ respectively. The free indices are called the physical indices *and their dimension d is called the* physical dimension*. In contrast, the contracted indices are called* virtual indices *and their dimensions D_j are called the* bond dimensions*.*

If we fix the physical index of the tensors $A^{[j]}$, they become matrices and the tensor network in equation (3.12) becomes a simple matrix product, which is where the name comes from.

Example 7 (MPS, adapted from [98])
Consider a system of many qubits (i.e. $d = 2$). We want to describe an MPS on such a system. To simplify notation, we split the tensors $A^{[j]}$ into two parts — one corresponding to the physical index being equal to 0 and another corresponding to it being equal to 1. Consider the following set of tensors (for $j \in \{2, 3, ..., N-1\}$:

$$A_0^{[1]} = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad A_0^{[j]} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad A_0^{[N]} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3.13}$$

$$A_1^{[1]} = \begin{pmatrix} 0 & 1 \end{pmatrix}, \qquad A_1^{[j]} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A_1^{[N]} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.14}$$

If we plug this set of tensors into equation (3.12), we obtain the GHZ state $|\phi\rangle = \frac{1}{\sqrt{2}}(|000...00\rangle + |111...11\rangle)$.

Example 8 (Schmidt decomposition as an MPS)
The MPS formalism can be applied to Schmidt decompositions, defined in equation (2.14). Consider a generic bipartite Schmidt-decomposed state $|\Psi\rangle_{AB}$

$$|\Psi\rangle_{AB} = \sum_{i=1}^{D} \sqrt{\lambda_i}\, |\chi_i\rangle_A \otimes |\xi_i\rangle_B\,, \tag{3.15}$$

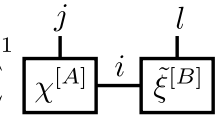
where we substituted $D = \min(d_A, d_B)$. The states $|\phi_i\rangle_A$ and $|\xi_i\rangle_B$ can be expanded in the computational bases of their respective subsystems as

$$|\chi_i\rangle_A = \sum_{j=0}^{d_A-1} \chi_{ji}^{[A]}\, |j\rangle_A \qquad \text{and} \qquad |\xi_i\rangle_A = \sum_{l=0}^{d_B-1} \xi_{li}^{[B]}\, |l\rangle_B\,. \tag{3.16}$$

Furthermore, we simplify the expression by substituting $\tilde{\xi}_{ji}^{[A]} = \sqrt{\lambda_i}\xi_{ji}^{[A]}$. This

leads to

$$|\Psi\rangle_{AB} = \sum_{i=1}^{D} \sum_{j=0}^{d_A-1} \sum_{l=0}^{d_B-1} \chi^{[A]}_{ji} \tilde{\xi}^{[B]}_{li} |jl\rangle_{AB} = \sum_{j=0}^{d_A-1} \sum_{l=0}^{d_B-1}$$



$$|jl\rangle_{AB}. \tag{3.17}$$

The j-th tensor of an MPS has $dD_{j-1}D_j$ components, other then the first and last, which have dD_1 and dD_{N-1} components respectively. Therefore, to fully describe an MPS, we need

$$d\left[D_1 + \sum_{j=2}^{N-1} (D_{j-1}D_j) + D_{N-1}\right] \tag{3.18}$$

components. This can be upper-bounded by NdD^2 for $D = \max_j D_j$.

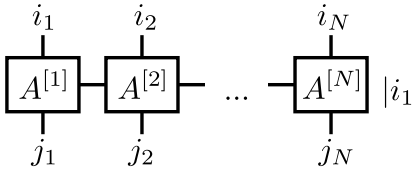
A very similar approach can be taken to describe a generic (density) operator acting on N d-dimensional qudits,

$$\rho = \sum_{i_1=0}^{d-1} \sum_{i_2=0}^{d-1} ... \sum_{i_N=0}^{d-1} \sum_{j_1=0}^{d-1} \sum_{j_2=0}^{d-1} ... \sum_{j_N=0}^{d-1} \rho_{i_1,i_2,...i_N,j_1,j_2,...,j_N} |i_1 i_2 ... i_N\rangle \langle j_1 j_2 ... j_N| . \tag{3.19}$$

The object $\rho_{i_1,i_2,...i_N,j_1,j_2,...,j_N}$ is a tensor containing d^{2N} components.

Definition 19 (Matrix Product Operator)**.** *A matrix product operator (MPO) is an operator of the following form:*

$$\rho = \sum_{i_1=0}^{d-1} ... \sum_{i_N=0}^{d-1} \sum_{j_1=0}^{d-1} ... \sum_{j_N=0}^{d-1}$$



$$|i_1 i_2 ... i_N\rangle\langle j_1 j_2 ... j_N|. \tag{3.20}$$

An MPO can describe both a density operator and any other operator acting on the Hilbert space (for example a quantum circuit [99]). When used to describe a density operator, the object is sometimes called *matrix product density operator* (MPDO). However, this name implies that the operator is positive, which is something that we can not guarantee for the operators that we will encounter in this thesis. Therefore, we will continue to call our tensor networks MPO, despite the fact that in this work we try to approximate exclusively density operators.

The j-th tensor of an MPO has $d^2D_{j-1}D_j$ components, other then the first and last, which have d^2D_1 and d^2D_{N-1} components respectively. Therefore, to fully describe an MPO we need

$$d^2\left[D_1 + \sum_{j=2}^{N-1} (D_{j-1}D_j) + D_{N-1}\right] \tag{3.21}$$

components. This can be upper-bounded for by Nd^2D^2 for $D = \max_j D_j$.

To simplify the notation of more complicated expressions, we sometimes omit the sums and the computational basis elements. Instead we represent states and operators only by their tensor networks, especially later in the thesis, e.g. in figure 4.3.

3.2.2 Bond Dimension and Efficiency

In the quantum many-body physics community, a task is considered to be done efficiently if the number of resources required scales at most polynomially in the size of the input[3]. In our case, the size of the input denotes the number of sites of the quantum system in question. In this language, exactly representing a generic pure quantum state on N sites (equation (3.11)) is not efficient, because it requires the knowledge of d^N complex coefficients (d^{2N} for a generic mixed state).

Describing an N-site system using an MPS (MPO) with a fixed bond dimension D requires the knowledge of NdD^2 (Nd^2D^2) complex numbers. This could be efficient, depending on our choice of D. Allowing D to grow at most polynomially with N keeps the MPS (MPO) description efficient, but it prevents us from being able to represent arbitrary pure (mixed) states exactly. The question of which states can be efficiently and accurately represented with MPS (MPO) is extremely important for quantum many-body physics research. The chapter 4 discusses sufficient conditions on pure (mixed) states to be approximable efficiently and accurately by MPS (MPO).

Throughout this work, when discussing how one variable scales with respect to some other variable, it is implied that we are interested in the limiting behaviour, i.e. $N \to +\infty$. This limit is called the *thermodynamic limit* and is discussed in more detail in section 3.3.

3.2.3 Entanglement Area Law

In this subsection we look at the entanglement present in an MPS. Take the MPS from equation (3.12). Consider a bipartition of the system into first k sites and the remaining $N-k$ sites. Consider the following pair of substitutions:

$$\left|\psi_l^{[k]}\right\rangle = \sum_{i_1,\ldots,i_k=0}^{d-1} \left(A_{i_1}^{[1]} A_{i_2}^{[2]} \ldots A_{i_k}^{[k]}\right)_l |i_1 \ldots i_k\rangle \,, \tag{3.22}$$

$$\left|\xi_l^{[k]}\right\rangle = \sum_{i_1,\ldots,i_k=0}^{d-1} \left(A_{i_{k+1}}^{[k+1]} A_{i_{k+2}}^{[k+2]} \ldots A_{i_k}^{[N]}\right)_l |i_{k+1} \ldots i_N\rangle \,. \tag{3.23}$$

We can write the MPS $|\phi\rangle$ as

$$|\phi\rangle = \sum_{l=1}^{D_k} \left|\psi_l^{[k]}\right\rangle \otimes \left|\xi_l^{[k]}\right\rangle . \tag{3.24}$$

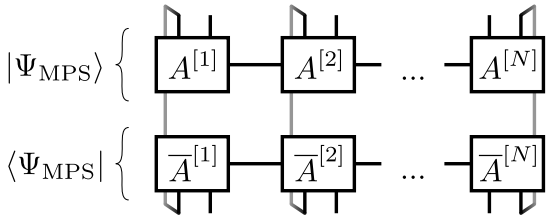
This indicates that $|\phi\rangle$ has Schmidt rank at most D_k across the bipartition, which is in turn bounded by $D = \max_k(D_k)$. The max-entanglement-entropy of a state is equal to the logarithm of its Schmidt rank, so it is upper-bounded by $\log(D)$. And because of the monotonicity of Rényi entropies (2.13), they are all also be bounded by $\log(D)$ for all $\alpha \in (0, +\infty)$. Importantly, the entanglement of the MPS depends only on the bond dimension, not on k. This is a special case of what is traditionally referred to as *entanglement area law*.

[3]Sometimes quasi-polynomial scaling is considered efficient too.

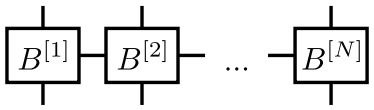
The *entanglement area law* is a property of some special quantum many-body states. It means that after bipartitioning the system into two sections, the entanglement between the two sections scales linearly with the *area* of the boundary of the two sections, as opposed to their *volume* which is the situation for a generic quantum state [100]. In one dimension, the area of a section corresponds only to the two sites at the edge of the section. The size of the area does not change (regardless of the size of the section). In a system obeying the entanglement area law, we would expect the same for entanglement, as is the case for a finite-bond MPS.

The situation with MPOs is more complicated. Mixed states contain classical correlations on top of entanglement, so it makes sense to talk about area law for some more general measure of correlations. In 2008 Wolf et. al. [101] noted that the mutual information between two parts of an MPO is bounded by the logarithm of the bond dimension of its *local purification* — an MPS which yields the MPO once its alcillary space is traced out. In other words, an MPO which is in a locally-purified form obeys the mutual information area law. The result was later generalized for Rényi mutual information [102]. However, it has also been demonstrated that there exist MPOs with fixed bond dimension and with a local purification MPS whose bond dimension grows logarithmically with the system size [103]. For this counter-example the mutual information bound does not imply area law because it grows with the system size, albeit slowly. Therefore, the question of whether all MPOs with a fixed bond dimension obey the area law for the mutual information remains open.

This question is partially addressed in appendix C for classically correlated MPOs (for Rényi mutual information).



(a) MPO in a locally purified form, obtained by tracing out the ancillary subspace of an MPS $|\Psi_{\mathrm{MPS}}\rangle$.

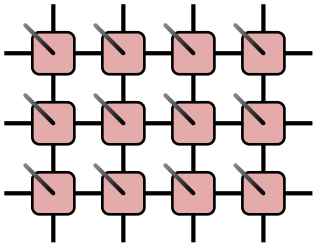


(b) MPO obtained by merging the tensors $A^{[k]}$ with their complex conjuagte $\overline{A}^{[k]}$.

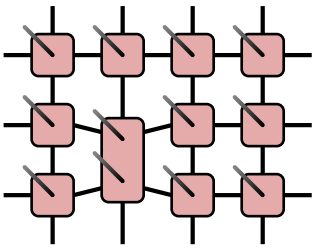
Figure 3.3: Locally-purified form explained with tensor network diagrams. Figure adapted from [103]

3.2.4 Projected Entangled Pair States

In this subsection, we briefly comment on projected entangled pair states (PEPS). This tensor network architecture is not studied anywhere else in this thesis, so we include the information here solely for completeness. A simple way to generalize MPS/MPO is to arrange the tensors not in a line, but in a 2D lattice. Such tensor networks are usually referred to as *projected entangled pair states / operators* (PEPS / PEPO).



(a) A square lattice PEPS tensor network.



(b) Contracting two tensors together in a PEPS produces a tensor with 6 virtual legs.

Figure 3.4: A diagrammatic depiction of a PEPS tensor network and its state after contracting two of the tensors. The legs originating in the middle of the tensors are the physical legs and they are pointing outside of the plane.

Most commonly PEPS are constructed on a square lattice, so that each of the tensors has 4 virtual legs connecting it to the other tensors (see figure 3.4a). Despite their similarity to MPS, the 2D nature of PEPS brings with it several fundamental limitations. It can be easily seen from the diagram that even contracting a PEPS tensor network (which is necessary to do in order to calculate expectation values) gets complicated as contracting two neighbouring tensors gives a tensor with more virtual legs (see figure 3.4b). Contracting more and more tensors will lead to larger and larger tensors, negating the speed-up of using tensor networks in the first place. In fact, Schuch [104] showed that the task of contracting PEPS lies in the complexity class of $\#P$-complete, i.e. it is at least as hard as NP-complete problems [105]. Furthermore, it has been shown that several questions about PEPS (such as invariance under symmetries or gap of parent Hamiltonians) are fundamentally undecidable [106].

That being said, PEPS still approximate ground states of local gapped Hamiltonians well [104] (just like MPS) and they are used for studying ground states of both gapped and critical systems [107–110] and also some non-quantum problems [111].

3.3 Thermodynamic Limit

The thermodynamic limit of a system is usually defined as the situation when the number of particles N gets very large, or symbolically $N \to +\infty$. The thermodynamic limit is important for two reasons. First, most systems systems encountered in nature contain very large numbers of particles $N \sim 10^{23}$, so it makes sense physically to investigate what happens when N is very large. Secondly, in some cases the math describing large systems actually becomes simpler when taking the limit $N \to +\infty$, allowing us to get better results than for finite (small) N.

A mathematically rigorous way to describe systems in the thermodynamic limit is by using the formalism of C^*-algebras. However, in practice in physics it is done in different ways. Usually when interested in the large N limit, we do not actually

talk about states or operators (e.g. the Hamiltonian), but *families of states* and *families of operators* which are defined for an arbitrary N. For example the 1D transverse-field Ising Hamiltonian

$$H = -J \sum_{i=1}^{N} \sigma_i^x \sigma_{i+1}^x - h \sum_{i=1}^{N} \sigma_i^z \tag{3.25}$$

represents a family of Hamiltonians (for various values of N), each acting on a different system and each having a different ground state. Similarly, on N qubits, we can have a family of states which are all a product of qubits being in the state $|0\rangle$: $\phi = |0\rangle^{\otimes N}$. In examples like these, it is simple to intuitively take the thermodynamic limit $N \to +\infty$. For a more rigorous treatment of the thermodynamic limit, see section B.1.

In most literature, the distinction between states / operators and *families of states / operators* is not explicitly made. In order to remain consistent, we will also sometimes write about *states* or *operators* when we mean *families of states / operators*.

Chapter 4

Approximating 1D Mixed States with Matrix Product Operators

In the previous chapter, we described tensor networks and outlined some of their most important uses. This chapter focuses on an essential question regarding 1D tensor networks (MPS, MPO): Which quantum states can be accurately described/approximated by such tensor networks? There are a few ways to answer this question. The main focus of this chapter is criteria based on entanglement and correlations present in the state (see section 4.4 for a brief overview of other criteria of approximability).

First, in section 4.1 we provide a short motivation, explaining why we are interested in this particular result. Then in section 4.2 we present an older result which proves that a linear pure state is approximable by MPS if its Rényi entanglement scales only logarithmically with system size (see theorem 2 for a more specific formulation). This is a seminal result, which has a strong influence on our approach as well. In section 4.3 we will present our own result on the approximability of mixed states by MPO. In that case Rényi entanglement is replaced by Rényi entanglement of purification and the scaling required is slightly different (see theorem 3). At the end in section 4.4 we discuss other criteria used for approximability of states by tensor networks.

4.1 Motivation

As was shown in chapter 3, MPS follow by construction the entanglement area law. This suggests that they are well suited to describe quantum states that follow the entanglement area law. It wa shown in [112] that there is indeed a connection between the entanglement in a pure state and its approximability by MPS. This has been an important result for quantum many-body physics and for tensor networks. However, a weakness of this result is that it only applies to pure states. For years there has been no analogous result quantifying the approximability of mixed states by MPO.

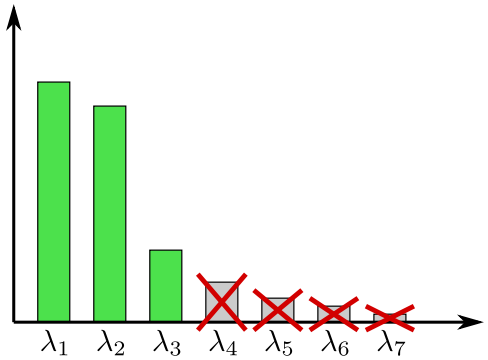


Figure 4.1: If the eigenvalues λ_i decay rapidly, we make only a small error by truncating them.

4.2 Approximating 1D Pure States with Matrix Product States

In this section we review the result of [112], which our own result is built upon. Most of the statements in this section come from that source. The main result states that a 1D pure state can be efficiently approximated by a matrix product state (MPS) if there exists $\alpha < 1$, such that the corresponding Rényi entropy of entanglement S_α across every bipartition scales at most logarithmically with the number of sites. We will go through the proof of the statement in this section because it will be illustrative for the proof of our own statement.

4.2.1 Schmidt Decomposition Truncation

The idea of constructing the MPS is built on truncating the Schmidt decomposition of the pure state across a bipartition. If the Rényi entropy of entanglement across a bipartition is small, it implies that the Schmidt coefficients decay rapidly and therefore we make only a small error by truncating (see figure 4.1). This idea is quantified in the following lemma:

Lemma 2. *Let ρ be a density matrix with eigenvalues λ_i labelled in descending order (i.e. $i \geq j \Rightarrow \lambda_i \leq \lambda_j$). Define $\varepsilon_D = \sum_{i=D+1}^{d} \lambda_i$ to be the sum of all but first D eigenvalues ($d = \dim \rho$). For any $0 < \alpha < 1$ we have*

$$\log \varepsilon_D \leq \frac{1-\alpha}{\alpha} \left[S_\alpha(\rho) - \log \frac{D}{1-\alpha} \right]. \tag{4.1}$$

Proof. The proof is done in several steps. First we seek to lower-bound the Rényi entropy of an arbitrary classical distribution $\{p_i\}_{i=1}^{d}$ as a function of its "tail" $\varepsilon_D^p = \sum_{i=D+1}^{d} p_i$.

Consider the following family of distributions (with h satisfying $0 < h < \frac{1-\varepsilon_D^p}{D}$

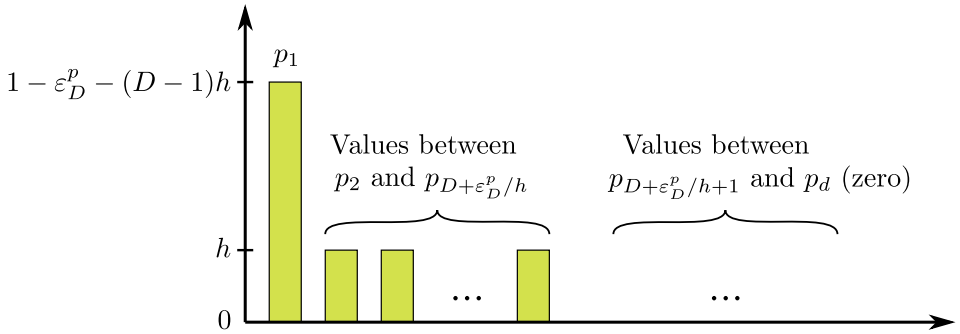


Figure 4.2: The distribution described in equations (4.2)–(4.4).

serving as a free parameter at this moment, see figure 4.2):

$$p_1 = 1-\varepsilon_D^p-(D-1)h, \tag{4.2}$$

$$p_2 = p_3 = ... = p_{D+\varepsilon_D^p/h} = h, \tag{4.3}$$

$$p_{D+\varepsilon_D^p/h+1} = p_{D+\varepsilon_D^p/h+2} = ... = p_d = 0. \tag{4.4}$$

These distributions majorize all other distributions with the same ε_D^p and p_{D+1}, therefore they have the lowest Rényi entropy among them (because of the Schur-concavity of Rényi entropies explained in section 2.4). The range of h allows to cover all possible values of p_{D+1}, so a distribution of this form will minimize the Rényi entropy $S_\alpha\left(\{p_i\}_{i=1}^d\right)$ among all distributions with fixed ε_D^p.

For the distribution described in equations (4.2)–(4.4) the following inequality holds:

$$\sum_{i=1}^d p_i^\alpha = [1-\varepsilon_D^p-(D-1)h]^\alpha + (D-1+\varepsilon_D^p/h)h^\alpha \geq Dh^\alpha + \varepsilon_D^p h^{\alpha-1}. \tag{4.5}$$

We minimize the right-hand side expression with respect to h to obtain

$$\sum_{i=1}^d p_i^\alpha \geq \frac{D^{1-\alpha}(\varepsilon_D^p)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha}. \tag{4.6}$$

Next, we plug this into the formula for the Rényi entropy to obtain

$$S_\alpha(\{p_i\}_{i=1}^d) \geq \frac{1}{1-\alpha}\log\left[\frac{D^{1-\alpha}(\varepsilon_D^p)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha}\right]. \tag{4.7}$$

As was said above, any other distribution with the same ε_D^p will have larger Rényi entropy. That includes the eigenvalues of ρ, $\{\lambda_i\}$ (for which $\varepsilon_D^p = \varepsilon_D$). We therefore have

$$S_\alpha(\rho) \geq \frac{1}{1-\alpha}\log\left[\frac{D^{1-\alpha}(\varepsilon_D)^\alpha}{(1-\alpha)^{1-\alpha}\alpha^\alpha}\right]. \tag{4.8}$$

And finally we express ε_D as

$$\log\varepsilon_D \leq \frac{1-\alpha}{\alpha}\left[S_\alpha(\rho) - \log\frac{D}{1-\alpha}\right] + \log\alpha. \tag{4.9}$$

The equation (4.1) is obtained by omitting the negative term $\log\alpha$. □

Lemma 2 gives us information about the rate of decay of the Schmidt coefficients based on the Rényi entropy of entanglement. It can be used to bound the error made by truncating the Schmidt decomposition of pure bipartite states. Let $|\phi\rangle_{AB}$ be a bipartite pure state with the Schmidt decomposition

$$|\phi\rangle_{AB} = \sum_{i=1}^{d} \sqrt{\lambda_i} \, |\chi_i\rangle_A \otimes |\xi_i\rangle_B \tag{4.10}$$

with $d = \min(\dim A, \dim B)$. For $D \leq d$ we define the (unnormalized) state $|\phi_D\rangle_{AB}$ obtained by truncating the Schmidt decomposition of $|\phi\rangle_{AB}$ after the first D terms:

$$|\phi_D\rangle_{AB} = \sum_{i=1}^{D} \sqrt{\lambda_i} \, |\chi_i\rangle_A \otimes |\xi_i\rangle_B \, . \tag{4.11}$$

The state $|\phi_D\rangle_{AB}$ approximates $|\phi\rangle_{AB}$ with the error $\varepsilon = \| |\phi\rangle_{AB} - |\phi_D\rangle_{AB} \|^2 = \sum_{i=D+1}^{d} \lambda_i$. This ε can be identified with ε_D from Lemma 2. Therefore, this error can be bounded by a function of the Rényi entropy of $\rho_B = \mathrm{Tr}_A \, |\phi\rangle \langle\phi|_{AB}$ and the bond dimension, according to equation (4.1). Changing the truncating cut-off D allows us to manipulate the upper bound on the error ε.

4.2.2 Approximating Multipartite States

Having quantified the error obtained by truncating a bipartite Schmidt decomposition, we are now ready to formulate and proof the approximability theorem for multipartite pure states. We will first state precisely the theorem, then give a brief sketch of the general idea of the proof and then go through the proof in detail. The result comes from Verstraete et al. [112], but the proof presented here is original.

Theorem 2 (Approximability of 1D Pure States)**.** *Let $|\phi\rangle$ be a pure state on a linear lattice with N sites containing qudits. Let $\rho_k = \mathrm{Tr}_{N \setminus k} \, |\phi\rangle \langle\phi|$ be its reduced density matrix on the first k sites. If there exists $\alpha \in [0, 1)$ and $c \in \mathbb{R}$ (independent of N) such that $\forall k$*

$$S_\alpha(\rho_k) \leq c \log N \tag{4.12}$$

then the state can be efficiently approximated by an MPS.

The proof of theorem 2 can be split into the following steps:

1. We take the Schmidt decomposition of the state $|\phi\rangle$ across every possible bipartition of the system into the first k sites and the remaining $N - k$ sites.

2. We truncate each of these decompositions after the first D terms. Because of the scaling of the Rényi entropy, we can apply lemma 2 to upper-bound the error created by this truncation. This way we create a set of $N - 1$ states, each with Schmidt rank D across a different bipartition.

3. We construct projectors onto these states. Applying those projectors enforces Schmidt rank D across the corresponding bipertition. Applying all of them to the original state $|\phi\rangle$ creates the resulting MPS, without accumulating too much error.

The state $|\phi\rangle$ can be exactly described by a d^N dimensional tensor $\phi_{i_1,i_2,\,\ldots\,,i_N}$, as seen in equation (3.11):

$$|\phi\rangle = \sum_{i_1=0}^{d-1} \sum_{i_2=0}^{d-1} \ldots \sum_{i_N=0}^{d-1} \phi_{i_1,i_2,\,\ldots\,,i_N} |i_1 i_2 \ldots i_N\rangle \,. \tag{4.13}$$

For $k \in \{1, 2, \ldots N-1\}$, consider the bipartition of the system into the first k sites and the remaining $N-k$ sites (hereinafter called the k-bipartition). Across the k-bipartition, the state $|\phi\rangle$ has the Schmidt decomposition:

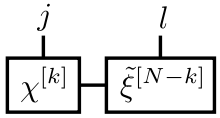
$$|\phi\rangle = \sum_{i=1}^{d^{\min(k,N-k)}} \sqrt{\lambda_{k,i}}\, |\chi_{k,i}\rangle \otimes |\xi_{k,i}\rangle \tag{4.14}$$

Let $|\phi_k\rangle$ be the state obtained from $|\phi\rangle$ by truncating its Schmidt decomposition across the k-bipartition after D terms[1], i.e.

$$|\phi_k\rangle = \sum_{i=1}^{D} \sqrt{\lambda_{k,i}}\, |\chi_{k,i}\rangle \otimes |\xi_{k,i}\rangle \,. \tag{4.15}$$

Following example 8 from section 4.2 we can write the states $|\phi_k\rangle$ as MPS-like tensor networks

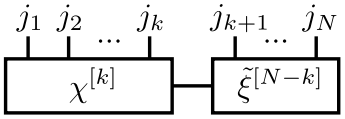
$$|\phi_k\rangle = \sum_{j=0}^{d^k-1} \sum_{l=0}^{d^{N-k}-1}$$



$$|j\rangle_k \otimes |l\rangle_{N-k}. \tag{4.16}$$

Here the states $|j\rangle_k$ form the computational basis on the first k qudits and $|l\rangle_{N-k}$ form the computational basis on the last $N-k$ qudits. To simplify notation, the Schmidt coefficient $\sqrt{\lambda_{k,i}}$ was absorbed into the tensor $\tilde{\xi}^{[N-k]}$. The bond dimension is naturally D. The indices j and l run from 0 to d^k-1 or $d^{N-k}-1$ and enumerate all computational basis states on the two parts of the state. We can split them into indices $\{j_1, j_2, \ldots, j_N\}$ running from 0 to $d-1$ which enumerate computational basis states on each site individually.

$$|\phi_k\rangle = \sum_{j_1=0}^{d-1} \sum_{j_2=0}^{d-1} \ldots \sum_{j_N=0}^{d-1}$$



$$|j_1\, j_2 \ldots j_N\rangle. \tag{4.17}$$

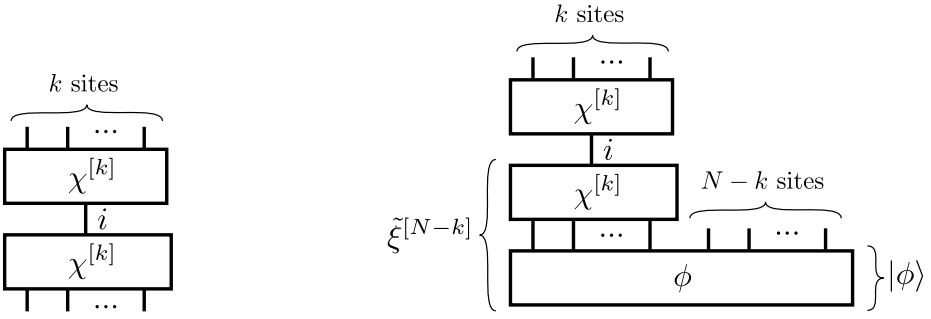
In equation (4.17) we can already start seeing the desired MPS structure emerging. Note that the equation describes many different states $|\phi_k\rangle$ for all possible values of k. Each of these states approximates the original state $|\phi\rangle$ with an error $\varepsilon_k = \| |\phi\rangle - |\phi_k\rangle \|^2$, which we can control by changing the bond dimension D (using

[1]This is only possible if $D \leq d^{\min(k,N-k)}$. That might not be true for k close to 1 or N. In those cases the sum in equation (4.14) does not even have D terms, so truncating it to D terms is ill-defined. To avoid badly defined expressions in those two cases we may additionally define $|\chi_{k,i}\rangle = 0$ and $\lambda_{k,i} = 0$ for $i \geq d^k + 1$ and similarly $|\xi_{k,i}\rangle = 0$ and $\lambda_{k,i} = 0$ for $i \geq d^{N-k} + 1$.

the result of lemma 2). The next step is to "merge" the states $|\phi_k\rangle$ together to create a state that has low bond dimension across every bipartition and which still accurately approximates the original state. To that end, we define the set of projectors $P_k = \sum_{i=1}^{D} |\chi_{k,i}\rangle \langle\chi_{k,i}|$ (depicted as tensor networks in figure 4.3a). We can see that the projectors P_k produce the states $|\phi_k\rangle$ when applied to the original state $|\phi\rangle$ (figure 4.3b):

$$P_k |\phi\rangle = \sum_{j=1}^{D} |\chi_{k,j}\rangle \langle\chi_{k,j}| \sum_{i=1}^{d^{\min(k,N-k)}} \sqrt{\lambda_{k,i}} |\chi_{k,i}\rangle \otimes |\xi_{k,i}\rangle = |\phi_k\rangle , \tag{4.18}$$

where we used the Schmidt decomposition (4.14) and the orthonormality property $\langle\chi_{k,j}|\chi_{k,i}\rangle = \delta_{ij}$.



(a) The projector P_k (b) The projector P_k applied to $|\psi\rangle$

Figure 4.3: The tensor networks corresponding to the projector P_k and its application to $|\phi\rangle$, creating $|\phi_k\rangle$. Compare the tensor network in 4.3b with the tensor network in equation (4.17).

As we can see in equation (4.18) and figure 4.3, applying the projectors P_k introduces the bond dimension D into the state. Consider what happens when we apply multiple of the projectors P_k in sequence. For $l > k$ consider

$$P_k P_l |\phi\rangle = P_k |\phi_l\rangle = \sum_{j=1}^{D} |\chi_{k,j}\rangle \langle\chi_{k,j}| \sum_{i=1}^{D} \sqrt{\lambda_{l,i}} |\chi_{l,i}\rangle \otimes |\xi_{l,i}\rangle \tag{4.19}$$

$$= \sum_{j=1}^{D} \sum_{i=1}^{D} |\chi_{k,j}\rangle \otimes |\zeta_{k,l,i,j}\rangle \otimes \sqrt{\lambda_{l,i}} |\xi_{l,i}\rangle \equiv |\phi_{kl}\rangle , \tag{4.20}$$

where we defined $\langle\chi_{k,i}|\chi_{l,j}\rangle \equiv |\zeta_{k,l,i,j}\rangle$. Remember that the states $|\chi_{l,j}\rangle$ and $|\chi_{k,i}\rangle$ live on a different number of sites, so their product $|\zeta_{k,l,i,j}\rangle$ is a state on $l - k$ sites (as can be seen in figure 4.4). The state $|\phi_{kl}\rangle$ obtained in equation (4.20) can be interpreted as the "fusion" of $|\phi_l\rangle$ and $|\phi_k\rangle$. This state now has bond dimension D across two bipartitions simultaneously — after the first k sites and after the first l sites.

Motivated by the result of consecutive application of the projectors P_k, we now apply all of them to the original state, one after another. We get

$$|\phi_{\text{MPS}}\rangle = \prod_{k=1}^{N-1} P_k |\phi\rangle , \tag{4.21}$$

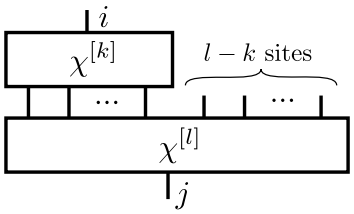


Figure 4.4: The tensor network representing $|\zeta_{k,l,i,j}\rangle = \langle\chi_{k,i}|\chi_{l,j}\rangle$.

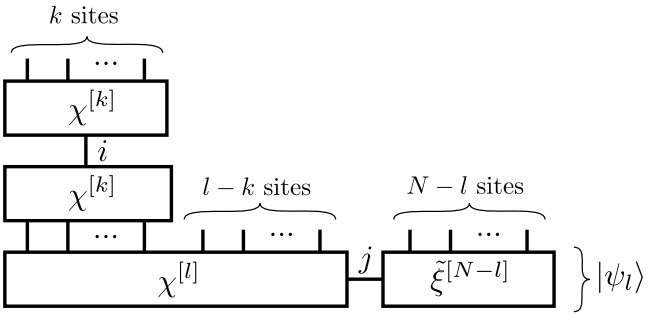


Figure 4.5: The tensor network representing the application of P_k to $|\phi_l\rangle$ (for $k < l$), equivalent to applying $P_k P_l$ to $|\phi\rangle$.

where the product is done in sequence (starting from the left), i.e. $P_1 P_2 ... P_{N-1}$, which means that P_{N-1} is applied to $|\phi\rangle$ first. The projectors do not commute with each other, so their order is important. See figure 4.6 for the tensor network representing equation (4.21).

The product of the projectors P_k in equation (4.21) produces single-site terms like

$$\left\langle \chi_{k,i_k} \middle| \chi_{k+1,i_{k+1}} \right\rangle = \left| \zeta_{k,k+1,i_k,i_{k+1}} \right\rangle . \tag{4.22}$$

The resulting state $|\phi_{\text{MPS}}\rangle$ can then be written as

$$|\phi_{\text{MPS}}\rangle = \sum_{i_1,i_2,\ldots,i_{N-1}=1}^{D} |\chi_{1,i_1}\rangle \otimes \left(\bigotimes_{k=1}^{N-2} \left| \zeta_{k,k+1,i_k,i_{k+1}} \right\rangle \right) \otimes \left| \xi_{N-1,i_{N-1}} \right\rangle \sqrt{\lambda_{i_{N-1}}}. \tag{4.23}$$

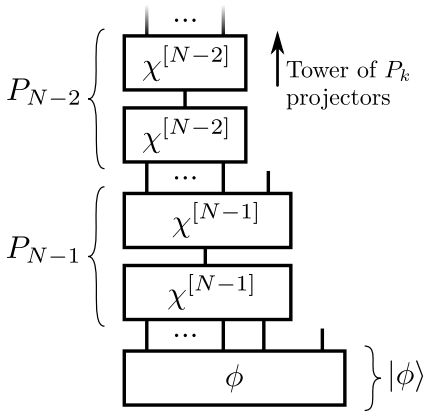


Figure 4.6: The tensor network representing the right-hand side of equation (4.21)

This state is an MPS, as can be seen by expressing the single-site states $|\chi_{1,i_1}\rangle$, $\left|\zeta_{k,k+1,i_k,i_{k+1}}\right\rangle$ and $\left|\xi_{N-1,i_{N-1}}\right\rangle$ in the computational basis:

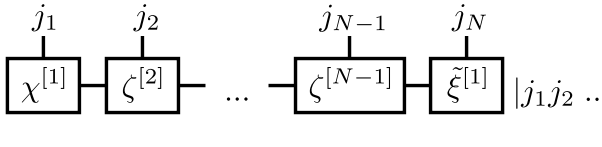
$$|\chi_{1,i_1}\rangle = \sum_{j=1}^{d} \chi^{[1]}_{ji_1} |j\rangle\,, \tag{4.24}$$

$$\left|\zeta_{k,k+1,i_k,i_{k+1}}\right\rangle = \sum_{j=1}^{d} \zeta^{[k]}_{ji_ki_{k+1}} |j\rangle\,, \tag{4.25}$$

$$\left|\xi_{N-1,i_{N-1}}\right\rangle = \sum_{j=1}^{d} \xi^{[1]}_{ji_{N-1}} |j\rangle\,. \tag{4.26}$$

This allows us to rewrite the state $|\phi_{\mathrm{MPS}}\rangle$ using tensor network notation:

$$|\phi_{\mathrm{MPS}}\rangle = \sum_{j_1=0}^{d-1}\sum_{j_2=0}^{d-1}\cdots\sum_{j_N=0}^{d-1}$$



$$|j_1 j_2 \ldots j_N\rangle. \tag{4.27}$$

The bonds connecting the tensors together correspond to the indices i_k (for $k \in \{1, 2, ..., N-1\}$) in equation (4.23). All these indices run from 1 to D. This implies that the state $|\phi_{\mathrm{MPS}}\rangle$ has exactly the form of an MPS with bond dimension D.

Now we examine how far this state is from the original state $|\phi\rangle$, i.e. what is the approximation error

$$\||\phi_{\mathrm{MPS}}\rangle - |\phi\rangle\|^2 = \left\|\prod_{k=1}^{N-1} P_k |\phi\rangle - |\phi\rangle\right\|^2. \tag{4.28}$$

To upper-bound the error, we repeatedly use the triangle inequality

$$\left\|\prod_{k=1}^{n} P_k |\phi\rangle - |\phi\rangle\right\| \leq \left\|\prod_{k=1}^{n} P_k |\phi\rangle - \prod_{k=1}^{n-1} P_k |\phi\rangle\right\| + \left\|\prod_{k=1}^{n-1} P_k |\phi\rangle - |\phi\rangle\right\| \tag{4.29}$$

valid for any $n \in \{2, 3, ..., N-1\}$. We apply this inequality recursively to reduce n all the way from $N-1$ down to 1. This generates terms $\left\|\prod_{k=1}^{n} P_k |\phi\rangle - \prod_{k=1}^{n-1} P_k |\phi\rangle\right\|$ for $n \in \{2, 3, ...N-1\}$. Each of these terms can be bounded by using $\|P_k\|_\infty = 1$, by the virtue of P_k being a projector:

$$\left\|\prod_{k=1}^{n} P_k |\phi\rangle - \prod_{k=1}^{n-1} P_k |\phi\rangle\right\| = \left\|\prod_{k=1}^{n-1} P_k \left(P_n |\phi\rangle - |\phi\rangle\right)\right\| \tag{4.30}$$

$$\leq \left\|\prod_{k=1}^{n-1} P_k\right\|_\infty \|P_n |\phi\rangle - |\phi\rangle\| \tag{4.31}$$

$$\leq \prod_{k=1}^{n-1} \|P_k\|_\infty \|P_n |\phi\rangle - |\phi\rangle\| \tag{4.32}$$

$$= \|P_n |\phi\rangle - |\phi\rangle\| \tag{4.33}$$

$$= \||\phi_n\rangle - |\phi\rangle\| = \sqrt{\varepsilon_n}. \tag{4.34}$$

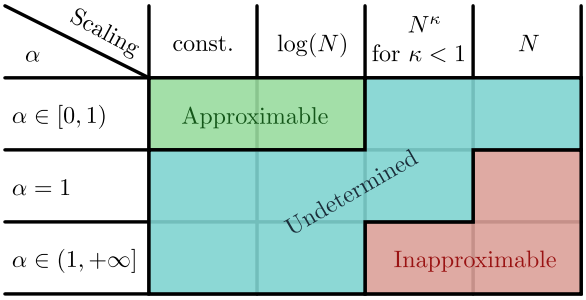


Figure 4.7: A table of approximability and inapproximability regions of 1D systems based on their scaling of various Rényi entropies of entanglement. The rows correspond to different values of $\alpha \in [0, +\infty]$ and the columns correspond to different possible scaling of the corresponding Rényi entropy of entanglement S_α (S_1 being the von Neumann entropy). Figure adapted from [113]

Summing all those terms together, we get $\| |\phi\rangle - |\phi_{\text{MPS}}\rangle \| \leq \sum_{k=1}^{N-1} \sqrt{\varepsilon_k}$. Using lemma 2 gives us

$$\log \varepsilon_k \leq \frac{1-\alpha}{\alpha} \left[c \log N - \log \frac{D}{1-\alpha} \right]. \tag{4.35}$$

For a given target error ε, we can set

$$D = (1-\alpha) \frac{N^{\frac{2\alpha}{1-\alpha}+c}}{\varepsilon^{\frac{\alpha}{1-\alpha}}}, \tag{4.36}$$

which ensures that $\varepsilon_k \leq \frac{\varepsilon}{N^2}$ and therefore $\| |\phi\rangle - |\phi_{\text{MPS}}\rangle \|^2 \leq \varepsilon$. The bond dimension from equation (4.36) is polynomial in N (and in ε). The MPS approximation is therefore efficient.

4.2.3 Inapproximability of 1D Pure States with Matrix Product States

Shortly after theorem 2 was published [112], the result was expanded on by Schuch et al. [113]. They provided entropy-based conditions implying the *impossibility* of efficient MPS approximation as well as examples to demonstrate that those conditions are tight. In this subsection we briefly summarize their results.

Schuch et al. proved that if the von Neumann entropy of entanglement scales linearly with the system size, the state can not be efficiently approximated with an MPS. Furthermore, if for any $\alpha > 1$, the Rényi entropy S_α grows sublinearly with the system size (i.e. proportional to N^κ for some $0 < \kappa < 1$), the same is true. These two situations correspond to the red region in figure 4.7. The green region corresponds to approximability based on the scaling of Rényi entropies for $\alpha < 1$, which was proven in subsection 4.2.2. The blue region corresponds to situation when (in)approximability by MPS can not be determined from the scaling of the entropies alone. In the blue region, Schuch et al. found examples of states which can be approximated efficiently as well as examples of states which can not.

Possibly the most interesting part of the diagram in figure 4.7 is the area corresponding to constant von Neumann entropy. This describes states whose von Neumann entanglement obeys the area law. Yet, that on its own is not a sufficient condition for MPS approximability, as an example by Schuch et al. demonstrates. However, this example is not translationally invariant, unlike all other provided examples. The question of whether any translationally invariant pure state obeying the von Neumann entanglement area law can be efficiently approximated by an MPS remains open.

4.3 Approximating 1D Mixed States with Matrix Product Operators

In this section we show our own research generalizing the result of section 4.2. The first question that we should ask ourselves is which quantity will play the role of the Rényi entropy to determine whether a mixed state is approximable. We cannot use the Rényi entropy of the reduced state anymore as it no longer has the meaning of entanglement of the original state, as was shown in example 4 in chapter 2. In fact, we cannot use any measure of entanglement alone. There exist separable states (with zero entanglement) which require large bond dimension to approximate accurately. For example the state $\rho_{AB} = \frac{1}{D}\sum_{i=1}^{D} |i\rangle\langle i|_A \otimes |i\rangle\langle i|_B$ for $D > 1$ has zero entanglement and requires bond dimension D to represent exactly. Therefore we need a quantity that captures both classical and quantum correlations. A natural candidate would be the mutual information (or rather its Rényi generalizations). We did some research into Rényi mutual information, some of which is presented in appendix C, in hopes of using it as approximability criterion. Unfortunately, we did not find a rigorous connection between the mutual information and approximability by MPO. Instead we used the Rényi entanglement of purification (see subsection 2.3.1 for definition), as this allows us to take advantage of Theorem 2.

Theorem 3 (Approximability of 1D Mixed States)**.** *Let ρ be a mixed state on a linear lattice with N sites. Consider the bipartition of the system into the first k sites and the remaining $N-k$ sites. If there exists $c > 0$ (independent of N) such that for any k the Rényi entanglement of purification across this bipartition follows*

$$E_\alpha^p(\rho) \leq c \log N \text{ for } \alpha = \frac{1}{\log N} \tag{4.37}$$

then the state can be efficiently approximated by an MPO

As before, we will prove this theorem in steps.

1. First we show that small entanglement of purification across a bipartition guarantees that we can approximate the mixed state by a low-rank approximation (analogous to the Schmidt decomposition truncation).

2. Then we will devise a scheme to merge two such bipartite approximations together in a way that preserves the low bond dimension without increasing the approximation error too much.

3. Finally, we devise a scheme to repeat this process iteratively to construct the final MPO.

Bounded entanglement of purification across a bipartition $E^p_\alpha(\rho_{AB}) \leq c \log N$ implies that there exists a purification $|\Psi\rangle_{AA'BB'}$ with an equally bounded Rényi entanglement across the same bipartition $S_\alpha(\rho_{AA'}) = E^p_\alpha(\rho_{AB}) \leq c \log N$. This purification can therefore be approximated by a low-rank approximation obtained by truncating the Schmidt decomposition $|\Psi\rangle_{AA'BB'} = \sum_{i=1}^{D} \sqrt{\lambda_i} \, |\chi_i\rangle_{AA'} \otimes |\xi_i\rangle_{BB'}$. For this approximation, we have $\| |\Psi\rangle - |\Psi_D\rangle \|^2 = \varepsilon_D$ with ε_D bounded by an equation analogous to (4.1),

$$\log \varepsilon_D \leq \frac{1-\alpha}{\alpha} \left[E^p_\alpha(\rho_{AB}) - \log \frac{D}{1-\alpha} \right]. \tag{4.38}$$

The error ε_D is made in the standard vector norm, i.e.

$$\| |\Psi\rangle - |\Psi_D\rangle \|^2 = \varepsilon_D = \sum_{i=D+1}^{d} \lambda_i = 1 - \sum_{i=1}^{D} \lambda_i = 1 - \langle \Psi_D | \Psi_D \rangle \,. \tag{4.39}$$

However, this time we are interested in the trace distance between the state and its approximation. To calculate that, we first normalize the approximating state $|\Psi'_D\rangle = \frac{|\Psi_D\rangle}{\sqrt{\langle \Psi_D | \Psi_D \rangle}}$. The trace distance between the state $|\Psi\rangle$ and its normalized approximation $|\Psi'_D\rangle$ is

$$\frac{1}{2} || |\Psi\rangle \langle \Psi| - |\Psi'_D\rangle \langle \Psi'_D| ||_1 = \sqrt{1 - |\langle \Psi | \Psi'_D \rangle|^2} = \sqrt{1 - \langle \Psi_D | \Psi_D \rangle} = \sqrt{\varepsilon_D}. \tag{4.40}$$

We can now apply a partial trace to the left side of (4.40). Partial trace is contractive with respect to the trace distance. Hence we obtain

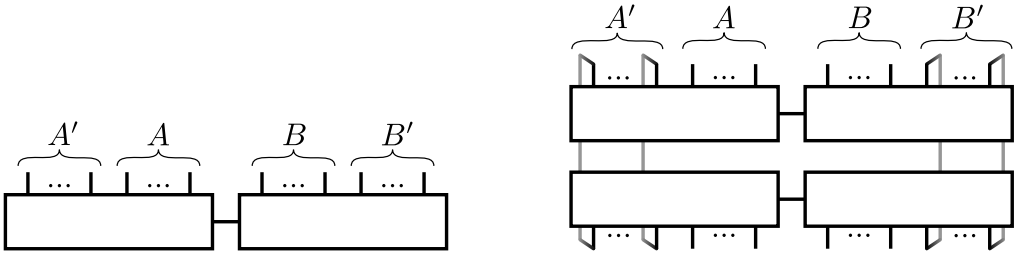
$$\left\| \rho_{AB} - \rho^D_{AB} \right\|_1 \leq 2\sqrt{\varepsilon_D}, \tag{4.41}$$

where ρ^D_{AB} is the density matrix obtained by tracing out the A' and B' subsystems from $|\Psi'_D\rangle$. This density matrix has the form

$$\rho^D_{AB} = \mathrm{Tr}_{A'B'} \, |\Psi'_D\rangle \langle \Psi'_D|_{AA'BB'} \tag{4.42}$$

$$= \sum_{i=1}^{D} \sum_{j=1}^{D} \sqrt{\lambda_i \lambda_j} \left[\mathrm{Tr}_{A'} \, |\chi_i\rangle \langle \chi_j|_{AA'} \right] \otimes \left[\mathrm{Tr}_{B'} \, |\xi_i\rangle \langle \xi_j|_{BB'} \right]. \tag{4.43}$$

Importantly, the density matrix ρ^D_{AB} has operator-Schmidt rank D^2.



(a) The pure state $|\Psi'_D\rangle_{AA'BB'}$ (b) The mixed state ρ^D_{AB}

Figure 4.8: The tensor networks corresponding to the purification approximation $|\Psi'_D\rangle_{AA'BB'}$ and its partial trace of subsystems A' and B' (which gives ρ^D_{AB}).

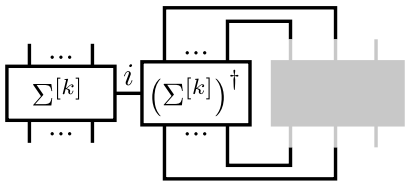


Figure 4.9: The superoperator $\mathcal{P}_k(\cdot)$. Input operators go in the shaded area.

4.3.1 Merging of Two Bipartite Approximations

As was shown above, if the mixed state has low entanglement of purification across one bipartition, we can approximate it accurately by a mixed state with low operator-Schmidt rank. Now the task is to merge those approximations into a tensor network with low bond dimension. In the case of pure states, a similar task was done by applying norm-non-increasing projectors onto the subspaces spanned by halves of the bipartite approximations in equation (4.21). We may attempt the same approach for mixed states. Assume that we have two approximations to ρ, with operator-Schmidt rank D^2 across different bipartitions (and $k < l$). They have the operator-Schmidt decompositions

$$\rho_k = \sum_{i=1}^{D^2} \gamma_i^{[k]} \Sigma_i^{[k]} \otimes \Pi_i^{[N-k]} \tag{4.44}$$

$$, \rho_l = \sum_{i=1}^{D^2} \gamma_i^{[l]} \Sigma_i^{[l]} \otimes \Pi_i^{[N-l]} \tag{4.45}$$

and they approximate ρ with error ε_k and ε_l respectively. The subspaces spanned by $\Sigma_i^{[k]}$ and $\Sigma_i^{[l]}$ are subspaces of operators. The projectors onto these subspaces are therefore superoperators (for any k)

$$\mathcal{P}_k(\cdot) = \sum_{i=1}^{D^2} \Sigma_i^{[k]} \operatorname{Tr}_{[k]} \left\{ \left(\Sigma_i^{[k]}\right)^\dagger \cdot \right\}. \tag{4.46}$$

The projectors $\mathcal{P}_k$ work similar to projectors P_k, i.e. $\mathcal{P}_k(\rho) = \rho_k$ and the operator obtained by their consecutive application

$$\rho'_{kl} \equiv \mathcal{P}_k(\rho_l) = \sum_{i_k=1}^{D^2} \sum_{i_l=1}^{D^2} \Sigma_{i_k}^{[k]} \otimes \operatorname{Tr}_{[k]} \left\{ \left(\Sigma_{i_k}^{[k]}\right)^\dagger \Sigma_{i_l}^{[l]} \right\} \otimes \gamma_{i_l}^{[l]} \Pi_{i_l}^{[N-l]} \tag{4.47}$$

has operator-Schmidt rank D^2 across two cuts, as can be seen in figure 4.10.

However, this approach fails because the projectors $\mathcal{P}_k$ are generally not norm-non-increasing with respect to the trace norm. The approximation error of ρ'_{kl} from equation (4.47) can be bounded as

$$\begin{aligned} \|\rho'_{kl} - \rho\|_1 &\leq \|\rho'_{kl} - \rho_k\|_1 + \|\rho_k - \rho\|_1 && (4.48)\\ &= \|\mathcal{P}_k(\rho_l - \rho)\|_1 + \varepsilon_k && (4.49)\\ &= \left\| \sum_{i=1}^{D^2} \Sigma_i^{[k]} \otimes \operatorname{Tr}_{[k]} \left\{ \left(\Sigma_i^{[k]}\right)^\dagger (\rho_l - \rho) \right\} \right\|_1 + \varepsilon_k. && (4.50) \end{aligned}$$

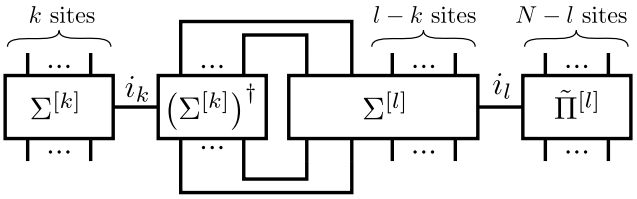


Figure 4.10: The operator ρ'_{kl} obtained by acting on ρ_l with $\mathcal{P}_k$. As we did before with states, we absorbed the operator-Schmidt coefficient $\gamma_{i_l}^{[l]}$ into $\Pi_{i_l}^{[N-l]}$, forming $\tilde{\Pi}_{i_l}^{[N-l]}$.

We know that $\|\rho_l - \rho\|_1 = \varepsilon_l$. Furthermore, we know that

$$\left\|\left(\Sigma_i^{[k]}\right)^\dagger\right\|_2 = 1 \;\Rightarrow\; \left\|\left(\Sigma_i^{[k]}\right)^\dagger\right\|_\infty = 1. \tag{4.51}$$

These together imply that

$$\left\|\mathrm{Tr}_{[k]}\left\{\left(\Sigma_i^{[k]}\right)^\dagger(\rho_l - \rho)\right\}\right\|_1 \geq \left\|\left(\Sigma_i^{[k]}\right)^\dagger\right\|_\infty \|\rho_l - \rho\|_1 \varepsilon_l. \tag{4.52}$$

However, the problem is with $\Sigma_i^{[k]}$. The trace norm $\left\|\Sigma_i^{[k]}\right\|_1$ is upper-bounded by $k\left\|\Sigma_i^{[k]}\right\|_2$, but this bound is too weak to be efficient. Therefore we cannot replicate the approach used for pure states.

These problems with norms do not arise if we attempt to bound the approximation error in the 2-norm $\|\cdot\|_2$:

$$\begin{aligned}
\|\rho'_{kl} - \rho\|_2 &\leq \|\rho'_{kl} - \rho_k\|_2 + \|\rho_k - \rho\|_1 && (4.53)\\
&= \|\mathcal{P}_k(\rho_l - \rho)\|_2 + \varepsilon_k && (4.54)\\
&\leq \|\rho_l - \rho\|_2 + \varepsilon_l && (4.55)\\
&\leq \varepsilon_k + \varepsilon_l. && (4.56)
\end{aligned}$$

The operators $\mathcal{P}_k$ are orthogonal projections on the space of operators, therefore, they do not increase the norm induced by the inner product. On the space of operators, the inner product is the Hilbert-Schmidt inner product and the norm induced by it is the 2-norm $\|\cdot\|_2$. If we are interested in an MPO approximation in the 2-norm, we can simply replicate approach described in section 4.2. The final MPO would be obtained by consecutively applying the projector $\mathcal{P}_k$ for all $k \in \{1, ..., N-1\}$ and the final error would be $\varepsilon = \sum_{k=1}^{N-1} \varepsilon_k$. It is possible to do this efficiently, analogously to how the pure state case was done. However, we aim for an MPO approximation in the trace norm, which means that we can not take this approach.

The approach that we developed works similarly — we also use projections, but different ones. We label our new projections $\mathcal{Q}_k$. We require that these project onto the span of $\Sigma_i^{[k]}$, so we have $\forall i : \mathcal{Q}_k(\Sigma_i^{[k]}) = \Sigma_i^{[k]}$ and consequently $\mathcal{Q}_k(\rho) = \rho_k$. To build these projections we consider a collection of linearly independent matrices $\Theta_i^{[k]}$ which span the same subspace as the matrices $\Sigma_i^{[k]}$. A possible projection onto

the subspace span($\Sigma_i^{[k]}$) is

$$\mathcal{Q}_k = \sum_{i=1}^{D^2} \Theta_i^{[k]} \otimes \mathrm{Tr}_{[k]} \left\{ \hat{\Theta}_i^{[k]\dagger} \cdot \right\}. \tag{4.57}$$

where $\hat{\Theta}_i^{[k]}$ are the dual matrices to $\Theta_i^{[k]}$ so that $\mathrm{Tr}\left\{\hat{\Theta}_i^{[k]\dagger}\Theta_j^{[k]}\right\} = \delta_{ij}$. Note that there is a lot of choices of the matrices $\Theta_i^{[k]}$ (and the corresponding $\hat{\Theta}_i^{[k]}$) that fulfill our requirements so far. Obviously, one such choice is $\Theta_i^{[k]} = \hat{\Theta}_i^{[k]} = \Sigma_i^{[k]}$, but as we saw earlier, that is not desirable.

When we apply this more general $\mathcal{Q}_k$ onto ρ_l we get

$$\rho_{kl} \equiv \mathcal{Q}_k(\rho_l) = \sum_{i_k=1}^{D^2} \sum_{i_l=1}^{D^2} \Theta_{i_k}^{[k]} \otimes \mathrm{Tr}_{[k]} \left\{ \left(\hat{\Theta}_{i_k}^{[k]}\right)^\dagger \Sigma_{i_l}^{[l]} \right\} \otimes \gamma_{i_l}^{[l]} \Pi_{i_l}^{[N-l]}. \tag{4.58}$$

This is not the same as the operator ρ'_{kl} which we got from equation (4.47). The tensor network structure is the same (see figure 4.10 for reference), but the tensors themselves are different.

Again, we ask how far this state is from ρ in the trace norm, i.e.

$$\begin{aligned} \|\mathcal{Q}_k(\rho_l) - \rho\|_1 &\leq \|\mathcal{Q}_k(\rho_l) - \rho_k\|_1 + \|\rho_k - \rho\|_1 && (4.59)\\ &= \|\mathcal{Q}_k(\rho_l) - \mathcal{Q}_k(\rho)\|_1 + \varepsilon_1 && (4.60)\\ &= \|\mathcal{Q}_k(\rho_l - \rho)\|_1 + \varepsilon_1. && (4.61) \end{aligned}$$

Now we examine the first term specifically, i.e.

$$\begin{aligned} \|\mathcal{Q}_k(\rho_l - \rho)\|_1 &= \left\| \sum_{i=1}^{D^2} \Theta_i^{[k]} \otimes \mathrm{Tr}_{[k]} \left\{ \hat{\Theta}_i^{[k]\dagger} (\rho_l - \rho) \right\} \right\|_1 && (4.62)\\ &\leq \sum_{i=1}^{D^2} \left\|\Theta_i^{[k]}\right\|_1 \left\| \mathrm{Tr}_{[k]} \left\{ \hat{\Theta}_i^{[k]\dagger} (\rho_l - \rho) \right\} \right\|_1 && (4.63)\\ &\leq \sum_{i=1}^{D^2} \left\|\Theta_i^{[k]}\right\|_1 \left\| \hat{\Theta}_i^{[k]\dagger} (\rho_l - \rho) \right\|_1 && (4.64)\\ &\leq \sum_{i=1}^{D^2} \left\|\Theta_i^{[k]}\right\|_1 \left\|\hat{\Theta}_i^{[k]\dagger}\right\|_\infty \|\rho_l - \rho\|_1. && (4.65) \end{aligned}$$

To keep this quantity as low as possible, we want to minimize $\left\|\Theta_i^{[k]}\right\|_1 \left\|\hat{\Theta}_i^{[k]\dagger}\right\|_\infty \geq \left|\mathrm{Tr}\{\Theta_i^{[k]}\hat{\Theta}_i^{[k]\dagger}\}\right| = |\delta_{ii}| = 1$. It turns out that it is possible to choose $\Theta_i^{[k]}$ and $\hat{\Theta}_i^{[k]}$ such that $\left\|\Theta_i^{[k]}\right\|_1 = \left\|\hat{\Theta}_i^{[k]\dagger}\right\|_\infty = 1$. To show that, we use the following theorem from functional analysis (for proof see appendix B):

Theorem 4 (Auerbach's Lemma)**.** *Let $V, |\cdot|$ be an n-dimensional normed vector space and V^* its dual space (with a dual norm $\|\cdot\|$). Then there exists a basis $\{e_1, e_2, ..., e_n\} \in V$ with a dual basis $\{f_1, f_2, ..., f_n\} \in V^*$ ($f_i(e_j) = \delta_{ij}$), such that $|e_i| = \|f_1\| = 1$.*

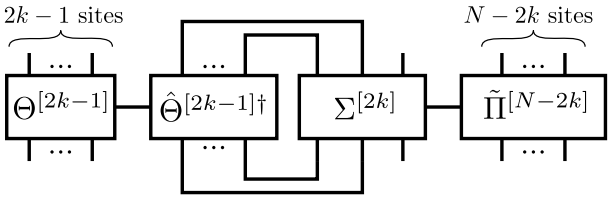


Figure 4.11: The operator $\rho_{2k-1,2k}$ obtained by merging ρ_{2k-1} with ρ_{2k}, i.e. by acting with $\mathcal{Q}_{2k-1}$ on ρ_{2k}.

The space $\mathrm{span}(\Sigma_i^{[k]})$ equipped with the trace norm $\|\cdot\|_1$ can be identified with V from Auerbach's lemma. Its basis will then be the matrices $\Theta_i^{[k]}$. For now we will denote the dual basis elements $\hat{\theta}_i \in V^*$. The space $V = \mathrm{span}(\Sigma_i^{[k]})$ together with its norm $\|\cdot\|_1$ is just a subspace of $\mathcal{B}(\mathcal{H}_{[k]})$ with the same norm. However, the same cannot be said for the dual spaces. Fortunately, we can use another tool of functional analysis — The Hahn-Banach theorem (for full formulation see appendix B). It guarantees that any dual element $\hat{\theta}_i \in V^*$ can be extended as $\tilde{\theta}_i \in \mathcal{B}^*(\mathcal{H}_{[k]})$ such that $\left\|\hat{\theta}_i\right\| = \left\|\tilde{\theta}_i\right\|$ (here the norms are the dual norms on the corresponding dual spaces). As an element of $\mathcal{B}^*(\mathcal{H}_{[k]})$, $\tilde{\theta}_i$ can be written as

$$\tilde{\theta}_i(\,\cdot\,) = \mathrm{Tr}\Big\{\hat{\Theta}_i^{[k]\dagger}\,\cdot\,\Big\}, \tag{4.66}$$

which implies that $\left\|\tilde{\theta}_i\right\| = \left\|\hat{\Theta}_i^{[k]}\right\|_\infty$.

This means that we can always find appropriate $\Theta_i^{[k]}$ (and corresponding $\hat{\Theta}_i^{[k]}$) such that $\left\|\Theta_i^{[k]}\right\|_1 = \left\|\hat{\Theta}_i^{[k]\dagger}\right\|_\infty = 1$ and therefore

$$\|\mathcal{Q}_k(\rho_l) - \rho\|_1 \leq D^2\|\rho_l - \rho\|_1 + \varepsilon_k \leq D^2\varepsilon_l + \varepsilon_k. \tag{4.67}$$

If for simplicity we label $\varepsilon = \max_i(\varepsilon_i)$, we get $\|\mathcal{Q}_k(\rho_l) - \rho\|_1 \leq (D^2+1)\varepsilon$, i.e. by merging two approximations together, we multiply the error by $D^2 + 1$.

4.3.2 Merging of Multiple Low-Rank Approximations

The fact that merging two approximations together multiplies the error by $D^2 + 1$ makes it impossible to merge the low-rank approximations sequentially as was done for the pure states. The overall error would grow exponentially with the number of sites, which would be too fast. Instead we merge the approximations in a tree-like fashion.

First assume that the number of sites is one more than a power of two: $N = 2^p+1$. We will begin by approximating the state ρ by a set of states $\{\rho_1, \rho_2, ...\rho_{N-1}\}$, such that the state ρ_i has low operator-Schmidt rank across the bipartition between the i-th and $(i+1)$th site and it approximates ρ with error ε_i. These states have the form of the states in equations (4.44) and (4.45). This is possible because of the scaling of the entanglement of purification. For simplicity define $\varepsilon = \max_i \varepsilon_i$

In the next step we merge together approximations ρ_{2k-1} and ρ_{2k} (for $k \in \{1, 2, ..., \frac{N-1}{2}\}$). The resulting approximations $\rho_{2k-1,2k}$ have low operator-Schmidt

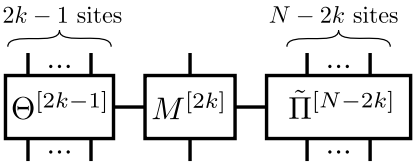


Figure 4.12: The operator $\rho_{2k-1,2k}$ after contracting $\mathrm{Tr}_{[2k-1]}\left\{\hat{\Theta}^{[2k-1]\dagger}_{i_{2k-1}}\Sigma^{[2k]}_{i_{2k}}\right\}$ into a single-site tensor $M^{[2k]}_{i_{2k-1},i_{2k}}$. Compare with the tensor network in figure 4.11.

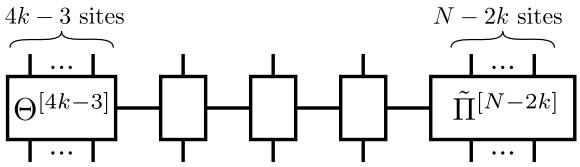


Figure 4.13: The operator $\rho_{4k-3,4k-2,4k-1,4k}$.

rank across two bipartitions. One of them is between the $(2k-1)$st site and $2k$-th site and the other is between the $2k$-th site and the $(2k+1)$st site. The states $\rho_{2k-1,2k}$ approximate the original state ρ with an error at most $(D^2+1)\varepsilon$.

In the next step, we want to merge pairs of these approximations, i.e. we merge $\rho_{4k-3,4k-2}$ with $\rho_{4k-1,4k}$ (for $k \in \{1,2,...,\frac{N-1}{4}\}$) to obtain $\rho_{4k-3,4k-2,4k-1,4k}$. For this, we have to construct new projections $\mathcal{Q}_{4k-3,4k-2}$.

The approach is the same as earlier, but there is one last thing we need to pay attention to. The projection $\mathcal{Q}_{4k-3,4k-2}$ is constructed to project onto the span of

$$\sum_{i_{4k-3}=1}^{D^2} \Theta^{[4k-3]}_{i_{4k-3}} \otimes M^{[4k-2]}_{i_{4k-3},i_{4k-2}}. \tag{4.68}$$

These operators (as a whole object) have operator-Schmidt rank D^2 across the bipartition between the $(4k-3)$rd site and the $(4k-2)$nd site. Any linear combination of them will also have the same operator-Schmidt rank across the same bipartition (or lower). This means that when $\mathcal{Q}_{4k-3,4k-2}$ is applied to $\rho_{4k-1,4k}$, the resulting object $\rho_{4k-3,4k-2,4k-1,4k}$ will have low operator-Schmidt coefficient across 4 different bipartitions (see figure 4.13). These $\rho_{4k-3,4k-2,4k-1,4k}$ approximate ρ with an error bounded by $(D^2+1)^2\varepsilon$.

We continue with this process until we reach $\rho_{1,...,N-1}$. This state is the sought MPO approximation and it approximates ρ with an error bounded by $(D^2+1)^{\log(N-1)}\varepsilon = (D^2+1)^p\varepsilon$ because $\log(N-1)$ is the number of steps of merging we needed to take to get $\rho_{1,...,N-1}$ (each step after the first doubles the number of bonds with low operator-Schmidt rank).

In case that N is not one more than a power of two, we simply append the original state ρ with product states, e.g.

$$\rho' = \rho \otimes \left(\bigotimes_{i=N+1}^{2^p+1} |0\rangle\langle 0|\right), \tag{4.69}$$

where $p = \lceil \log(N-1) \rceil$. This definitely does not change the scaling of the Rényi entanglement of purification. We then apply the above algorithm to ρ' to find its

MPO approximation. Finally, we trace out the appended system to get an MPO approximation of ρ.

The final error we get is $\varepsilon_{\text{final}} = (D^2+1)^{\lceil \log(N-1)\rceil}\varepsilon$ when here ε is the error obtained by truncation of the Schmidt decomposition of the purification across one cut, related to the entanglement of purification by (4.38). Putting both equations together, we get the expression for the final error,

$$\varepsilon_{\text{final}} = (D^2+1)^{\lceil \log(N-1)\rceil} \exp\left\{ \frac{1-\alpha}{\alpha} \left[E_\alpha^p(\rho_{AB}) - \log \frac{D}{1-\alpha} \right] \right\}. \tag{4.70}$$

Given a target error $\varepsilon_{\text{final}}$, we are interested in what scaling of the Rényi entanglement of purification is necessary to be able to reach the target error with the bond dimension growing at most polynomially $D \sim \text{poly}(N)$. First note the factor in front of the exponential $(D^2+1)^{\lceil \log(N-1)\rceil}$. This factor will grow as we increase the bond dimension, which means that the decay of the exponential has to offset this.

The expression for the final error can be manipulated to

$$\varepsilon_{\text{final}} \leq \exp\left\{ \log\left(D^2+1\right) \lceil \log(N-1)\rceil + \frac{1-\alpha}{\alpha} \left[E_\alpha^p(\rho_{AB}) - \log \frac{D}{1-\alpha} \right] \right\}. \tag{4.71}$$

For a fixed α, the first term in the exponential $\log (D^2+1)\lceil \log(N-1)\rceil$ will always outgrow the last term $-\frac{1-\alpha}{\alpha}\log\frac{D}{1-\alpha}$ for N large enough. This implies that the exponent as a whole will be positive and the error will grow with D. However, this does not happen if we allow α to change with N, specifically $\alpha = \frac{1}{3\log N}$ is enough. To see what effect this will have on $\varepsilon_{\text{final}}$, we plug in the bound on the entanglement of purification $E_\alpha^p(\rho_{AB}) \leq c \log N$, the allowed polynomial scaling of the bond dimension $D = N^\kappa$, and we omit some sub-leading terms in the calculation. This gives us the bound

$$\begin{aligned}
\varepsilon_{\text{final}} &\leq \exp\left\{ \log\left(D^2+1\right)(\log N + 1) + \frac{1-\alpha}{\alpha} \left[E_\alpha^p(\rho_{AB}) - \log \frac{D}{1-\alpha} \right] \right\} && (4.72)\\
&\lesssim \exp\left\{ \log\left(D^2\right) \log N + \frac{1}{\alpha} \left[E_\alpha^p(\rho_{AB}) - \log D \right] \right\} && (4.73)\\
&\leq \exp\left\{ \log\left(N^{2\kappa}\right) \log N + 3 \log N \left[c \log N - \log N^\kappa \right] \right\} && (4.74)\\
&\leq \exp\{(3c-\kappa)\log N \log N\}. && (4.75)
\end{aligned}$$

We can see that we just require $\kappa > 3c$ in order for the error to decay super-polynomially with N. Given that c does not depend on N, this is an efficient scaling of the bond dimension.

4.4 Other Criteria of Approximability

Aside from entanglement or purification and Rényi entropies, there are other criteria for approximability of pure (or mixed) states by an MPS (MPO). In this section we will briefly comment on those.

4.4.1 Decay of Conditional Mutual Information

One way to construct an MPO is via recovery maps. This was developed in the Bachelor's thesis of Sieber [114], which was based on a breakthrough result by Fawzi and Renner [115]. Sieber describes the following method to approximate a mixed state by an MPO:

1. First we trace out all but the first two sites of a mixed state.
2. Then we construct a recovery map which acts on the right-most site of the system and reconstructs the system to extend on one more site.
3. Repeat step 2 until the entire system is re-constructed.

By construction, this approach yields an MPO with bond dimension proportional to the local physical dimension (which can be increased by blocking sites together). The accuracy of this reconstruction comes from Fawzi and Renner's result and it is given by the conditional mutual information between two next-nearest neighbour sites, conditioned on the site in between. The conditional mutual information depends on blocking the sites together. If it decays exponentially in the number of sites blocked, then we can block the sites in such a way that the bond dimension grows polynomially with the system size.

A natural question to ask is when does the conditional mutual information decay exponentially with the block size. One answer apparent from the definition of conditional mutual information (2.36) is when the state is pure and the von Neumann entanglement saturates the area law exponentially fast. As was discussed in subsection 4.2.3, a von Neumann entanglement area law does not on its own imply efficient approximability by MPS. The result by Sieber [114] proves that a von Neumann entanglement area law is enough, if the constant entanglement entropy is saturated fast enough (with the size of the entangled region). Recently Kuwahara [116] discovered that the conditional mutual information decays quickly enough when the mixed state is a Gibbs state above a certain threshold temperature.

4.4.2 Ground States of Gapped Hamiltonians

Another criterion of approximability of pure states is whether they are a ground state of a gapped one-dimensional Hamiltonians. This result came together by a combination of several different works. The flow of implications is outlined in figure 4.14.

Hastings [16] (together with Koma [117]) proved that being a ground state of a gapped Hamiltonian implies that a pure state has exponentially decaying correlations and its Rényi entanglement entropies obey the area law (for some $\alpha < 1$). The result by Verstraete and Cirac [112] discussed in section 4.2 connects this result to approximability by MPS. Later in 2013, Arad, Vazirani, Landau and Kitaev [118] proved directly that ground states of 1D gapped Hamiltonians can be efficiently approximated by MPS, bypassing the area law result. This allowed them to get better efficiency in the approximation. Note that the connection between between exponentially decaying correlations and entanglement area law was rigorously proven only in 2015 by Brandão and Horodecki [119].

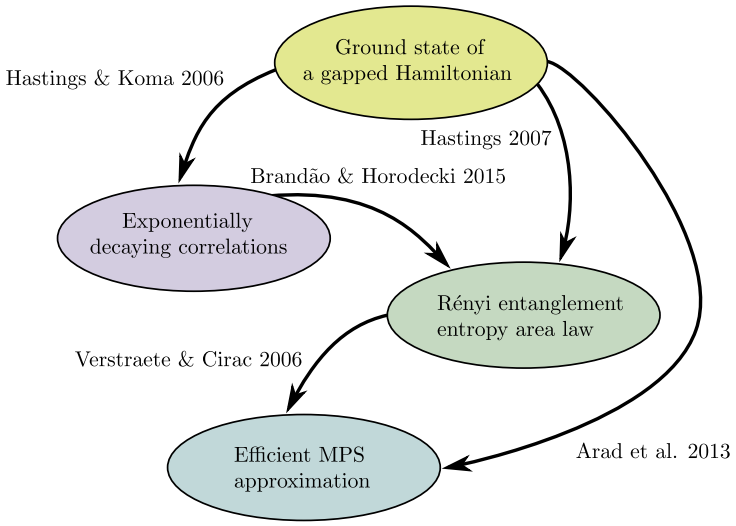


Figure 4.14: A diagram of the various results connecting approximability by MPS with the state being a GS of a gapped 1D Hamiltonian.

Chapter 5

Symmetries and Purifications

The laws of physics are rich with symmetries. So are most of the models studied in quantum many-body physics. Symmetries can be useful to us in a lot of ways. Knowledge of the symmetries of a system allows us to greatly simplify the its description in terms of only the crucial / relevant degrees of freedom.

Entire textbooks have been written about symmetries and symmetry breaking [120, 121] (not only in physics [122]), so in this chapter we will not go into too much detail. We investigate a method of simplifying the search for the ground states of symmetric quantum many-body Hamiltonians. As part of the method, we will use purifications (see section 2.3). This chapter is written without placing any unnecessary assumptions on the system such as the symmetry group or the statistics of the constituents. The method described in this chapter will turn out to be crucial in solving the permutationally-invariant Hubbard model in chapter 6.

5.1 Symmetries

In mathematics a *symmetry* is usually defined as an operation that leaves a certain object invariant. In quantum physics those operations will be represented by unitaries and the objects of interest are quantum states / operators.

Definition 20 (Symmetries)**.** *A pure quantum state $|\phi\rangle$ is considered symmetric with respect to a symmetry represented by a unitary matrix U if and only if*

$$U|\phi\rangle = |\phi\rangle . \tag{5.1}$$

We say that a pure quantum is symmetric up to a phase if and only if

$$U|\phi\rangle = \mathrm{e}^{\mathrm{i}\alpha}|\phi\rangle . \tag{5.2}$$

A quantum operator O is considered symmetric with respect to a symmetry represented by a unitary matrix U if and only if

$$UOU^{\dagger} = O. \tag{5.3}$$

We distinguish between *symmetric* pure states and pure states which are *symmetric up to a phase*. Usually, symmetry up to a phase is sufficient for pure

states, as the global phase does not have a physical meaning. However, in this work we will often use the more strict definition of symmetry which requires the global phase of the state to be preserved. The reason for that will become apparent in section 5.3. One of the advantages of this convention is that a superposition of symmetric states is always symmetric.

The unitaries representing a symmetry (of a state / operator) always form a group. The symmetry is often labelled by the name of the underlying group (**SU(2)**, **U(1)**, etc. ...).

5.2 Spontaneous Symmetry Breaking

Typically, the full symmetry group $\mathcal{G}$ of a quantum many-body Hamiltonian contains several different subgroups. Consider a symmetry subgroup $\mathcal{G}' \subset \mathcal{G}$ represented by unitaries U_g for $g \in \mathcal{G}$, i.e. $U_g H U_g^\dagger = H$. Let $|\phi_0\rangle$ be a ground state of a Hamiltonian H. It is possible that the state $|\phi_0\rangle$ does not have the symmetry U_g(for some $g \in \mathcal{G}'$), not even up to a phase, i.e.

$$U_g |\phi\rangle = |\phi_g\rangle \not\propto |\phi\rangle . \tag{5.4}$$

This is usually referred to as *spontaneous symmetry breaking* (although precise definitions differ [120, 123]).

Spontaneous symmetry breaking complicates the search for ground states. As described in chapter 3, the Hilbert spaces corresponding to quantum many-body systems are huge, so searching them for the ground state is very difficult. The search can be made easier by restricting it only to a symmetric subspace, i.e. the subspace of vectors that have the symmetries of the Hamiltonian. However, by enforcing this restriction, we will not be able to find symmetry-broken states.

One might claim that there always exists a ground state, which is symmetric with respect to the symmetry subgroup $\mathcal{G}'$. Starting from an arbitrary ground state $|\phi\rangle$, we can define the pure symmetrization

$$|\phi_{\text{sym}}\rangle = \int_{\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'} U_g |\phi\rangle . \tag{5.5}$$

Here the integration over the group uses the Haar measure $\mu_{\mathcal{G}'}$ on the group $\mathcal{G}'$. The vector $|\phi_{\text{sym}}\rangle$ constructed this way is symmetric by construction $U_g |\phi_{\text{sym}}\rangle = |\phi_{\text{sym}}\rangle$, but in some cases it is equal to the 0-vector (which has all the possible symmetries, but does not represent any physical quantum state).

A different symmetrization is

$$\rho_{\text{sym}} = \int_{\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'} U_g |\phi\rangle \langle\phi| U_g^\dagger . \tag{5.6}$$

This ρ_{sym} is always a valid quantum state (positive semi-definite and normalized). If the starting state $|\phi\rangle$ is symmetric up to a phase, then equation (5.6) gives $\rho_{\text{sym}} = |\phi\rangle \langle\phi|$. If $|\phi\rangle$ is not symmetric, even up to a phase, then ρ_{sym} is a mixed state. Note also that we get the same ρ_{sym} whether we start from $|\phi\rangle$ or from any $U_g |\phi\rangle$ for $g \in \mathcal{G}'$.

The energy of the state ρ_{sym} is

$$\begin{aligned}
\operatorname{Tr}\{H\rho_{\text{sym}}\} &= \operatorname{Tr}\Big\{H \int_{\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'} U_g \left|\phi\right\rangle \left\langle\phi\right| U_g^\dagger\Big\} && (5.7)\\
&= \int_{\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'} \operatorname{Tr}\Big\{U_g^\dagger H U_g \left|\phi\right\rangle \left\langle\phi\right|\Big\} && (5.8)\\
&= \int_{\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'} \operatorname{Tr}\{H \left|\phi\right\rangle \left\langle\phi\right|\} && (5.9)\\
&= \left\langle\phi\right| H \left|\phi\right\rangle . && (5.10)
\end{aligned}$$

As we can see, regardless of whether spontaneous symmetry breaking occurs or not, there always exists a symmetric density operator ρ_{sym} with the ground state energy.

5.3 Purified Ground States

As discussed in section 5.2 above, there always exists a symmetric state ρ_{sym} with the ground state energy. This ρ_{sym} may be a pure state, but it may also be a mixed state. Either way, we may purify it. We construct the *standard purification*

$$\left|\Psi\right\rangle_{AB} = \sqrt{\rho_{\text{sym}}}_A \left|\omega\right\rangle_{AB} . \tag{5.11}$$

where $\left|\omega\right\rangle$ is the maximally-entangled state between the subspaces A and B. In particular

$$\left|\omega\right\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} \left|i\right\rangle_A \otimes \left|i\right\rangle_B \qquad \text{for qudits,} \tag{5.12}$$

$$\left|\omega\right\rangle_{AB} = \prod_{i=1}^{d} \left(\frac{\mathbb{1}_i + a_i^\dagger b_i^\dagger}{\sqrt{2}}\right) \left|\Omega\right\rangle \qquad \text{for fermions.} \tag{5.13}$$

For fermions, use b_i and $b_i^\dagger$ as the annihilation and creation operators for the ancillary fermionic modes. The pure state $\left|\Psi\right\rangle_{AB}$ is a purification of ρ_{sym}, with the subspace B being ancillary. Since the matrix ρ_{sym} is symmetric with respect to U_g, so is its square root

$$\forall g \in \mathcal{G}' : U_g \sqrt{\rho_{\text{sym}}} U_g^\dagger = \sqrt{\rho_{\text{sym}}} \tag{5.14}$$

Now we plug equation (5.14) into equation (5.11), specifying the subspace that the operators act on with an additional subindices A and B,

$$\begin{aligned}
\left|\Psi\right\rangle_{AB} &= (U_g)_A \sqrt{\rho_{\text{sym}}}_A \left(U_g^\dagger\right)_A \left|\omega\right\rangle_{AB} && (5.15)\\
&= (U_g)_A \sqrt{\rho_{\text{sym}}}_A \left(\tilde{U}_g\right)_B \left|\omega\right\rangle_{AB} && (5.16)\\
&= (U_g)_A \left(\tilde{U}_g\right)_B \sqrt{\rho_{\text{sym}}}_A \left|\omega\right\rangle_{AB} && (5.17)\\
&= (U_g)_A \left(\tilde{U}_g\right)_B \left|\Psi\right\rangle_{AB} . && (5.18)
\end{aligned}$$

Here the form of $\tilde{U}_g$ depends on the nature of the quantum system again. For qudits any operator O obeys

$$O_A \left|\omega\right\rangle_{AB} = O_B^T \left|\omega\right\rangle_{AB} , \tag{5.19}$$

implying that $\tilde{U}_g = U_g^*$. For fermions we have

$$a_i \left|\omega\right\rangle_{AB} = b_i^\dagger \left|\omega\right\rangle_{AB}, \tag{5.20}$$

$$a_i^\dagger \left|\omega\right\rangle_{AB} = -b_i \left|\omega\right\rangle_{AB}, \tag{5.21}$$

so we construct $\tilde{U}_g$ from U_g by expressing it using the mode operators and using the equalities (5.20) and (5.21).

Either way, $(U_g)_A \left(\tilde{U}_g\right)_B$ is a unitary and the purification $\left|\Psi\right\rangle_{AB}$ is symmetric with respect to the action of this unitary.

The purification $\left|\Psi\right\rangle_{AB}$ has the same energy as $\left|\phi\right\rangle$ and ρ_{sym}, i.e.

$$\left\langle\Phi_0\right| H_A \left|\Phi_0\right\rangle_{AB} = \text{Tr}\{\rho_A H_A\} = \left\langle\phi_0\right| H_A \left|\phi_0\right\rangle_A. \tag{5.22}$$

Therefore, the problem is shifted to minimizing $\left\langle\Phi\right| H_A \left|\Phi\right\rangle_{AB}$ with respect to purifications $\left|\Phi\right\rangle_{AB}$ symmetric under the action of $(U_g)_A \left(\tilde{U}_g\right)_B$ (for all $g \in \mathcal{G}'$).

The calculations above show that regardless of whether the ground state spontaneously breaks the symmetry U_g or not, there always exists a pure state living in a Hilbert space appended by the ancillary subspace, which has the symmetry $(U_g)_A \left(\tilde{U}_g\right)_B$. If we want to find the ground state of H, we can proceed in the following steps:

1. Identify the symmetries of H and their respective unitaries U_g.
2. Append the Hilbert space with an ancillary subspace.
3. Construct $(U_g)_A \left(\tilde{U}_g\right)_B$ (for all $g \in \mathcal{G}$).
4. Search the enlarged Hilbert space for a ground state. Restrict the search to states symmetric under $(U_g)_A \left(\tilde{U}_g\right)_B$.

This approach allows us to restrict the search to symmetric states without having to worry about spontaneous symmetry breaking. On the other hand, the dimensionality of the Hilbert space that we need to search is larger now (square of the dimensionality of the physical Hilbert space alone). Simply restricting the Hilbert space to symmetric purifications will not be able to offset this dimensionality growth. As we will see in the examples in section 5.4 and chapter 6, we usually have to utilize some other approach that amplifies the advantage of symmetries of the purifications (e.g. the quantum de Finetti theorem or tensor networks).

5.3.1 Detecting Spontaneous Symmetry Breaking

In this subsection, we investigate the properties of $\left|\Psi\right\rangle$ and ρ_{sym} and how are they related.

A purification of a density matrix ρ is entangled (between the physical part and the ancillary part) if and only if ρ is mixed. The calculation above demonstrated that if a ground state $\left|\phi\right\rangle$ breaks the symmetry $\mathcal{G}'$ of the Hamiltonian, then the state ρ_{sym} is mixed and therefore its purification is entangled. But the logic can

also go the other way. If the purification ground state is a product state, then the corresponding ground state $\rho_{\text{sym}} = |\phi\rangle\langle\phi|$ is pure and symmetric. If the purification ground state is entangled, then the corresponding ground state ρ is mixed and therefore degenerate. Ground state degeneracy means that *some* symmetry of the Hamiltonian is spontaneously broken. We can always define a unitary permuting the various ground states. This is itself a symmetry of the Hamiltonian (according to the definition 20), but it might not be part of $\mathcal{G}'$ — the symmetry group that we're interested in.

It is generally advised to pick $\mathcal{G}'$ as large as possible. The larger the group $\mathcal{G}'$ is, the more restricted the space of symmetric purifications is. And the more restricted the space of symmetric purifications is, the easier it is to find the ground state purification. The examples in subsection 5.4.2 and chapter 6 are permutationally-invariant systems because the permutation group is large (exponential in the system size), so it restricts the Hilbert space a lot.

On the other hand, some of the symmetries of the Hamiltonian might not be very interesting and/or physical, so their spontaneous breaking is not relevant. As was mentioned above, a degenerate ground state automatically breaks a symmetry of the Hamiltonian which permutes the different ground states. Furthermore, depending on the method used to find the ground state purification, it might sometimes be beneficial to not include a known (and physically relevant) symmetry in the group $\mathcal{G}'$. A specific case of this will be discussed in subsection 5.4.2.

In the case that $\mathcal{G}'$ itself further decomposes into smaller subgroups, we can find out which of them was broken.

Theorem 5. *Let $\mathcal{G}''$ be a subgroup of $\mathcal{G}'$. We claim that the symmetry corresponding to $\mathcal{G}''$ is not spontaneously broken if and only if the purification ground state fulfills*

$$\forall g \in \mathcal{G}'' : (U_g)_A |\Phi\rangle_{AB} = \mathrm{e}^{\mathrm{i}\alpha(g)} |\Phi\rangle_{AB} . \tag{5.23}$$

Proof. For one direction of the equivalence, split the situation in two cases: If the ground state is not degenerate, then the ground state has the symmetry U_g for $g \in \mathcal{G}''$ up to a phase and the equation (5.23) is obvious.

If the ground state is degenerate then no spontaneous symmetry breaking implies that all of the ground states have to be symmetric with respect to U_g *up to the same phase*. If different ground states gave different phase, then a superposition of these would be a ground state spontaneously breaking the U_g symmetry. All ground states being symmetric up to the same phase $\alpha(g)$ implies $U_g \rho_{\text{sym}} = \mathrm{e}^{\mathrm{i}\alpha(g)} \rho_{\text{sym}}$, which immediatelly gives equation (5.23).

For the other direction of the equivalence, start from the equation (5.23). After decomposing the purification ground state as $|\Phi\rangle_{AB} = \sqrt{\rho_{\text{sym}}} |\omega\rangle_{AB}$ we get an equality between two operators acting on the maximally entangled state. It can therefore be reformulated as equality between two operators:

$$U_g \sqrt{\rho_{\text{sym}}} = \mathrm{e}^{\mathrm{i}\alpha(g)} \sqrt{\rho_{\text{sym}}}. \tag{5.24}$$

Next, we multiply both sides by $\sqrt{\rho_{\text{sym}}}$ from the right and apply trace to both sides:

$$\mathrm{Tr}\{U_g \rho_{\text{sym}}\} = \mathrm{e}^{\mathrm{i}\alpha(g)}. \tag{5.25}$$

Now, we express ρ_{sym} as the integral from equation (5.6) and take the integration outside of the trace:

$$\int_{f\in\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'}\, \mathrm{Tr}\Big\{U_g U_f \left|\phi\right\rangle \left\langle\phi\right| U_f^\dagger\Big\} = \mathrm{e}^{\mathrm{i}\alpha(g)}. \tag{5.26}$$

Next, we rearrange the trace to a bra-ket expectation value and multiply both sides of the equation by $\mathrm{e}^{-\mathrm{i}\alpha(g)}$:

$$\int_{f\in\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'}\, \left\langle\phi\right| U_f^\dagger U_g U_f \left|\phi\right\rangle \mathrm{e}^{-\mathrm{i}\alpha(g)} = 1. \tag{5.27}$$

Now, we take the real part of both sides of the equation. The operation of taking the real part of a complex number commutes with the integration:

$$\int_{f\in\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'}\, \mathrm{Re}\left[\left\langle\phi\right| U_f^\dagger U_g U_f \left|\phi\right\rangle \mathrm{e}^{-\mathrm{i}\alpha(g)}\right] = 1. \tag{5.28}$$

As the last step, we simply subtract 1 from both sides and take it inside the integral (utilising the property of the Haar measure $\int_{f\in\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'} = 1$)

$$\int_{f\in\mathcal{G}'} \mathrm{d}\mu_{\mathcal{G}'}\, \Big\{\mathrm{Re}\left[\left\langle\phi\right| U_f^\dagger U_g U_f \left|\phi\right\rangle \mathrm{e}^{-\mathrm{i}\alpha(g)}\right] - 1\Big\} = 0. \tag{5.29}$$

The absolute value of $\left\langle\phi\right| U_f^\dagger U_g U_f \left|\phi\right\rangle \mathrm{e}^{-\mathrm{i}\alpha(g)}$ is less or equal to 1. Its real part is thus also less or equal to 1. Therefore we have an integral of non-positive function that is equal to 0. The only possibility is that the function is 0 *almost everywhere*. That implies

$$\left\langle\phi\right| U_f^\dagger U_g U_f \left|\phi\right\rangle = \mathrm{e}^{\mathrm{i}\alpha(g)}. \tag{5.30}$$

In other words, almost all of the ground states $U_f \left|\phi\right\rangle$ have the symmetry U_g, up to the same phase $\mathrm{e}^{\mathrm{i}\alpha(g)}$. □

5.4 Examples

In this section we present a couple of simple examples to demonstrate the use of the method described about. We will start with an elementary example of two fermions. Then we will look at hardcore bosons on a permutationally-invariant lattice with all-to-all hopping and interactions. And the last example will be the Majumdar-Ghosh model. In chapter 6 we apply our method to yet another example — The Hubbard model with on a permutationally-invariant lattice with all-to-all hopping and interactions. However, that problem is relatively complex, so it deserves its own chapter.

5.4.1 Elementary Example — Two Fermions

In this example we have a Hilbert space of two fermions (a_1 and a_2) with the Hamiltonian

$$H = a_1^\dagger a_1.$$

We can clearly see that the ground state is either $|\Omega\rangle$ or $a_2^\dagger |\Omega\rangle$ or any superposition of the two, for example

$$\sin(\theta) |\Omega\rangle + \cos(\theta) a_2^\dagger |\Omega\rangle . \tag{5.31}$$

for a free parameter $\theta \in [0, \pi/2]$. As we can see, this ground state is degenerate.

The Hamiltonian has plenty of symmetries, for example:

- $\mathbf{U(1)}$ for the 2nd species with parameter ϕ: $a_2 \to \mathrm{e}^{\mathrm{i}\phi} a_2$
- $\mathbf{U(1)}$ for the 1st species with parameter φ: $a_1 \to \mathrm{e}^{\mathrm{i}\varphi} a_1$
- Particle-hole symmetry for the 2nd species: $a_2 \longleftrightarrow a_2^\dagger$

Looking back at the general ground state (5.31), we can see that it always breaks at least some of the above symmetries. The particle-hole symmetry is broken unless $\theta = \pi/4$. And the $\mathbf{U(1)}$ symmetry is broken unless $\theta \in \{0, \pi/2\}$.

However, a mixed state

$$\frac{|\Omega\rangle \langle\Omega|}{2} + \frac{a_2^\dagger |\Omega\rangle \langle\Omega| a_2}{2}$$

has all of the above symmetries. And it has a purification

$$|\Psi\rangle = \frac{1 + a_2^\dagger b_2^\dagger}{\sqrt{2}} |\Omega\rangle ,$$

which has all of the above symmetries and is entangled, indicating symmetry breaking for the physical ground state.

5.4.2 Hardcore Bosons with All-to-All Hopping and Interactions

For this example we consider a permutationally-invariant system of hardcore bosons living on N sites. Hardcore bosons commute when they are on different sites ($j \neq i$), but anticommute on-site:

$$[a_j, a_i] = 0, \tag{5.32}$$
$$[a_j, a_i^\dagger] = 0, \tag{5.33}$$
$$\{a_i, a_i^\dagger\} = 1, \tag{5.34}$$
$$\{a_i, a_i\} = 0. \tag{5.35}$$

The Hamiltonian

$$H = \frac{t}{N} \sum_{i \neq j} \left(a_i^\dagger a_j + \text{h.c.} \right) + \mu \sum_i a_i^\dagger a_i + \frac{V}{N} \sum_{i \neq j} a_i^\dagger a_i a_j^\dagger a_j \tag{5.36}$$

contains three terms: hopping, chemical potential and a two-body interaction. These three terms each have corresponding real parameters t, μ and V to tune their strength. The hopping and interaction terms are scaled so that all parts of

the Hamiltonian scale linearly with N (and therefore remain relevant for large N). Note that the sums in the hopping and interaction terms run over all pairs of sites. In fact, the Hamiltonian does not contain any information about underlying lattice structure at all. This is compatible with the permutation symmetry of the system. The Hamiltonian also has the $\mathbf{U(1)}$ particle-number symmetry (seen already in the previous example). The permutation symmetry will be examined in more detail in chapter 6.

As explained in section 5.3, there exists a purification ground state $|\Psi\rangle$ with the permutation symmetry and the $\mathbf{U(1)}$ symmetry. To find it, we can use the quantum de Finetti theorem [54–57]. It guarantees that for large N, any expectation values of local observables may be reproduced by a product state $|\psi_{\text{prod}}\rangle = \bigotimes_{i=1}^{N} |\psi_i\rangle$ (with error scaling as $1/N$). However, the quantum de Finetti theorem says nothing about preserving the symmetries of the original state. So while $|\psi_{\text{prod}}\rangle$ is permutationally-invariant by construction, it might not have the $\mathbf{U(1)}$ symmetry. If we force the variational family to be a product of $\mathbf{U(1)}$-symmetric states, it will not achieve the ground state energy (in some parts of the phase diagram). For this reason, we choose $\mathcal{G}'$ to only include the permutation symmetry and we will address the $\mathbf{U(1)}$ symmetry later.

In the purified Hilbert space, there will be another species of hardcore bosons, annihilated and created by $b_i, b_i^\dagger$ respectively. Most generally, the single-site state $|\psi_i\rangle$ has the following form:

$$|\psi_i\rangle = \left(\theta_0 + \theta_1 a_i^\dagger + \theta_2 b_i^\dagger + \theta_3 a_i^\dagger b_i^\dagger\right)|0_i\rangle\,. \tag{5.37}$$

with the four parameters $\theta_0, \theta_1, \theta_2, \theta_3$ subjected to the normalizition restriction $|\theta_0|^2 + |\theta_1|^2 + |\theta_2|^2 + |\theta_3|^2 = 1$. In terms of these parameters, the energy density of the state is

$$\frac{E}{N} = \frac{\langle\psi_{\text{prod}}|\, H\, |\psi_{\text{prod}}\rangle}{N} \tag{5.38}$$

$$= \mu\,\langle\psi_i|\, a_i^\dagger a_i\, |\psi_i\rangle + t\,\langle\psi_i|\,\langle\psi_j|\left(a_i^\dagger a_j + a_j^\dagger a_i\right)|\psi_i\rangle\,|\psi_j\rangle + V\left(\langle\psi_i|\, a_i^\dagger a_i\, |\psi_i\rangle\right)^2 \tag{5.39}$$

$$= \mu\left(|\theta_1|^2 + |\theta_3|^2\right) + 2t|\theta_0^*\theta_1 + \theta_2^*\theta_3|^2 + V\left(|\theta_1|^2 + |\theta_3|^2\right)^2. \tag{5.40}$$

In the second equality we used the fact that $\langle\psi_{\text{prod}}|$ is permutationally invariant to replace the sums in the Hamiltonian H with just one of its summands.

The expression for the energy density can be minimized by hand. We will split the calculation in two cases based on the sign of t.

First consider $t > 0$. In this case, we want to minimize the expression $|\theta_0^*\theta_1 + \theta_2^*\theta_3|$. This can be achieved by setting $\theta_2 = \theta_1 = 0$ without restricting the possible values of $|\theta_1|^2 + |\theta_3|^2$ (because of the freedom in setting θ_3). Finding the ground state energy density then reduces to minimizing $\mu|\theta_3|^2 + V|\theta_3|^4$ (for allowed range of $|\theta_3| \in [0, 1]$).

We can easily see that the ground state corresponds to $|\theta_3| = 0$ for $V \geq -\mu \leq 0$, $|\theta_3| = 1$ for $-\mu \geq 2V \leq -2\mu$ and to $|\theta_3| = \sqrt{-\frac{\mu}{2V}}$ in the remaining area ($2V \geq -\mu \geq 0$), see figure 5.1. Because of the normalization, changes in $|\theta_3|$ will change $|\theta_0|$, however this has no effect on the energy (after we forced $\theta_2 = \theta_1 = 0$).

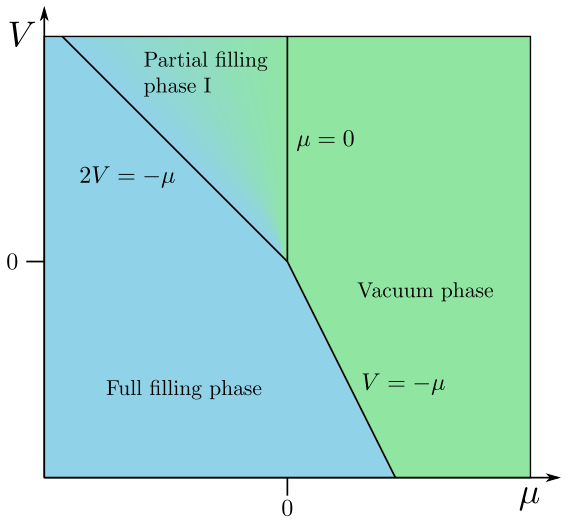


Figure 5.1: Ground state phase diagram for hardcore bosons for $t > 0$.

Now consider $t < 0$. In this case, we want to maximize $|\theta_0^* \theta_1 + \theta_2^* \theta_3|$. When minimizing the energy, we can assume all θ's to be real and non-negative. This is because their individual phases will not have any effect on the μ and V terms, while the t term is maximized if $\theta_0^* \theta_1$ and $\theta_2^* \theta_3$ are in phase. By substituting

$$\theta_0 = \sqrt{c_1} A, \tag{5.41}$$
$$\theta_2 = \sqrt{1 - c_1} A, \tag{5.42}$$
$$\theta_1 = \sqrt{c_2} B, \tag{5.43}$$
$$\theta_3 = \sqrt{1 - c_2} B, \tag{5.44}$$

for some real A, B, c_1, c_2 satisfying $c_1^2 + c_2^2 = A^2 + B^2 = 1$, we find the energy density to be

$$\frac{E}{N} = \mu B^2 + 2tA^2 B^2 \left(\sqrt{c_1 c_2} + \sqrt{(1 - c_1)(1 - c_2)} \right)^2 + V B^4. \tag{5.45}$$

The term in the parentheses is maximized by $c_1 = c_2$ (in which case it is just $A^2 B^2$). Setting $c_1 = c_2$ simplifies the energy density to

$$\frac{E}{N} = \mu B^2 + 2tA^2 B^2 + V B^4 \tag{5.46}$$

and the purification ground state ansatz to

$$|\psi_i\rangle = \left(A + B a_i^\dagger \right) \left(\sqrt{c_1} + \sqrt{1 - c_1} b_i^\dagger \right) |0_i\rangle \,. \tag{5.47}$$

The energy density (5.46) can be minimized using the Lagrange multipliers method. There are three possible solutions, depending on the paramteres V, μ and t:

- For $\mu > -t$ and $V > -\mu + t$, the energy density is minimized by $A = 1, B = 0$.

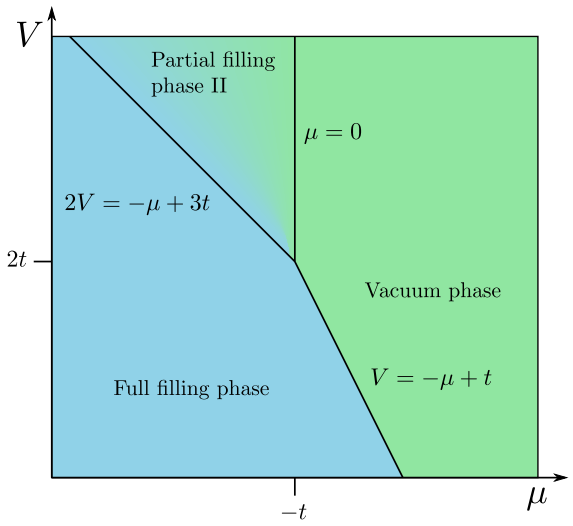


Figure 5.2: Ground state phase diagram for hardcore bosons for $t < 0$.

- For $V < -\mu + t$ and $2V < -\mu + 3t$, the energy density is minimized by $A = 0, B = 1$.
- For $\mu < -t$ and $2V > -\mu + 3t$, the energy density is minimized by

$$A = \sqrt{\frac{2V - 2t + \mu}{2V - 4t}}, \tag{5.48}$$

$$B = \sqrt{\frac{-2t - \mu}{2V - 4t}}. \tag{5.49}$$

Regardless of A and B, the resulting state is a product state between the ancillary subspace and the physical subspace.

The phase diagram looks similar to the $t > 0$ case. One difference is that the position of the critical point is changed, although the slope of the phase transition lines is not. Another difference is that now the partially-filled phase is different from what it was previously.

Symmetry breaking

Now, we analyze the phases from the perspective of spontaneous symmetry breaking. The vacuum and full filling phases are obviously product states and they do not break any of the symmetries of the Hamiltonian.

There are two partial filling phases. In the phase diagrams we labelled them as *Partial filling phase I / II*. The partial filling phase I for $t > 0$ corresponds to $(\theta_0, \theta_1, \theta_2, \theta_3) = (\sqrt{1 + \frac{\mu}{V}}, 0, 0, \sqrt{-\frac{\mu}{V}})$. The ground state in this phase is entangled, signaling spontaneous breaking of the permutation symmetry. On the other hand, the partial filling phase II for $t < 0$ is a product state (between the physical and ancillary degrees of freedom), as can be clearly seen in equation (5.47). The state breaks the $\mathbf{U(1)}$ symmetry and it is degenerate. When we were looking for the ground state, we restricted ourselves to real θ's. Lifting this restriction reveals the

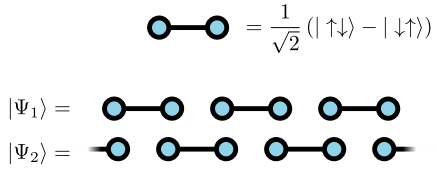


Figure 5.3: The two ground states of the Majumdar-Ghosh model

degeneracy — adding any complex phase to B does not change the ground state energy. Neither does any change of the ancillary part of the state (i.e. changing $c_1 = c_2$).

Physically, the difference between the two phases is the following: In phase I, some sites will be filled and some will be empty (breaking permutation symmetry), while in phase II, all sites will be in the same superposition of being filled and empty (preserving permutation symmetry, but breaking the **U(1)** symmetry).

5.4.3 Majumdar–Ghosh Model

The Majmundar-Ghosh model [124] is a simple model of a spin chain with a nearest-neighbour interaction as well as a next-nearest-neighbour interaction (of half strength). The Hamiltonian is

$$H = \sum_i \left(\vec{S}_i \cdot \vec{S}_{i+1} + \frac{1}{2} \vec{S}_i \cdot \vec{S}_{i+2} \right). \tag{5.50}$$

It is a special case of a more general class of $J_1 - J_2$ models (which can have any relative strength of the two interaction terms). The exact solution to this model is known [125]. The ground state corresponds to pairs of neighbouring spins forming singlets $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$. The ground state is 2-fold degenerate, depending on which pairs of sites are entangled into the singlets (see figure 5.3).

The symmetries of the model are translations (by any number of sites) and global rotations of all the spins simultaneously. The symmetry group of translations is much smaller than the symmetry group of permutations, so we do not expect to be able to solve this ground state problem by hand. Instead of attempting to find the solution from scratch by using the symmetric purification approach, we will take into account that we already know the solution and just present a symmetric purification ground state which reduces to a mixture of the two degenerate ground states.

First, note that the two ground states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ fulfill the global spin rotation symmetry. The symmetry broken is the translation symmetry (when translated by an odd number of sites).

The ground states are product states of blocks of two sites and can therefore be represented as matrix product states (MPS) with bond dimension two (see figure 5.4b). As operators, the ground states will be matrix product operators of bond dimension four, depicted in figure 5.4a.

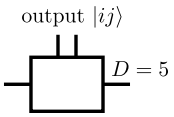
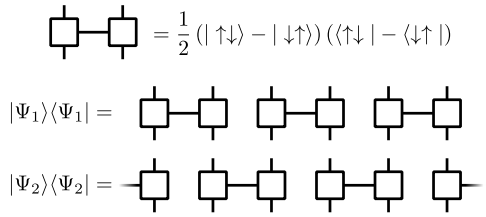


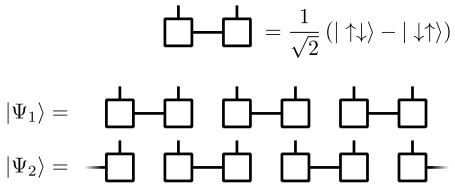
Figure 5.5: The MPS tensor corresponding to the purification ground state. For the tensor values see table 5.1.

Left / right bond index	1	2	3	4	5
1	0	0	0	0	$\|\uparrow\uparrow\rangle$
2	0	0	0	0	$i\,\|\uparrow\downarrow\rangle$
3	0	0	0	0	$i\,\|\downarrow\uparrow\rangle$
4	0	0	0	0	$\|\downarrow\downarrow\rangle$
5	$\|\downarrow\downarrow\rangle$	$i\,\|\downarrow\uparrow\rangle$	$i\,\|\uparrow\downarrow\rangle$	$\|\uparrow\uparrow\rangle$	0

Table 5.1: The components of the tensor constructing the purification ground state MPS. The table contains the output of the tensor, given the left and right bond indices (ranging from 1 to 5). The second spin in each pair is the ancillary spin.



(a) The two ground states of the Majumdar-Ghosh model as an MPO of bond dimension four.



(b) The two ground states of the Majumdar-Ghosh model as an MPS of bond dimension two.

Figure 5.4: The ground states of the Majumdar-Ghosh model depicted as tensor networks.

In order to get the ground state purification, we first need to mix those two states uniformly, corresponding to equation (5.6). This will by construction create a translationally invariant MPO with bond dimension five. From it we create a translationally invariant purification, analogously to equation (5.11). The tensor of the purification is diagramatically shown in figure 5.5 and its values are described in table 5.1. The ground state purification is a translationally-invariant MPS, so it is defined by only one tensor.

By construction, the MPS purification is invariant with respect to translations or global rotations of the spins. However, as was explained in section 5.3.1, when we apply translation to only the physical spins, the state is changed. This comes from the fact that the original ground states break the translation symmetry.

Chapter 6

Hubbard Model with Permutation Symmetry

In this chapter we will describe our approach towards solving the Hubbard model with permutation symmetry, taking advantage of the method described in chapter 5. We will start with a description of the Hubbard model in section 6.1, which includes the identification of the symmetries of its Hamiltonian. Then in section 6.2, we discuss extending those symmetries to the ancillary subspace and constructing a symmetric variational family on the extended Hilbert space. Finally in section 6.3 we describe the ground state phase diagram, including a description of all its phases. The calculation of the phase diagram is included in appendix A.

6.1 Hubbard Model

In this section we will describe the object of our study — the Hubbard model with permutation symmetry. In subsection 6.1.1 we will give a general description of models generally referred to as *Hubbard models*, their features and some known results. In subsection 6.1.2 we will describe our particular model. Lastly in subsection 6.1.3 we will identify and describe the symmetries of our model, in particular, the unitaries representing those symmetries.

6.1.1 Hubbard Model Generally

The Hubbard model [126], sometimes called the Fermi-Hubbard model to distinguish it from the Bose-Hubbard model [127], is a standard model in quantum many-body physics. It describes systems with two species of fermions on a lattice. The Hubbard Hamiltonian contains two important terms — the hopping term and the interaction term.

The hopping term serves to allow the hopping of fermions of both species between two neighbouring sites:

$$t \sum_{<i,j>} \left(a_{i,1}^\dagger a_{j,1} + a_{i,2}^\dagger a_{j,2} + \text{h.c.} \right), \tag{6.1}$$

where the sum is taken over all pairs of neighbouring sites i and j. The real parameter t is used to tune the strength of this term.

The interaction term represents the interaction that happens when two fermions (of different species) end up on the same site:

$$U \sum_i a_{i,1}^\dagger a_{i,1} a_{i,2}^\dagger a_{i,2}. \tag{6.2}$$

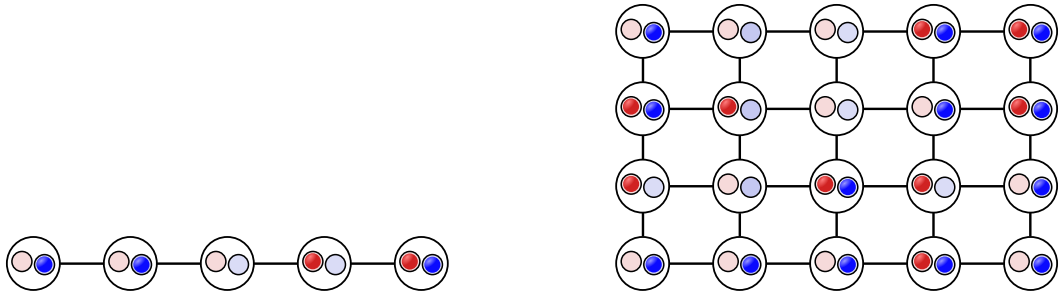
Here the sum is taken over all sites of the system. Again, this term has a real parameter U associated with it, whose sign determines whether this interaction is repulsive or attractive.

Usually, the Hubbard model is solved for a fixed number (or density) of particles [128, 129]. In this thesis, we focus on a different approach in which we let the number of particles change freely and add one more term to the Hamiltonian, representing the chemical potential,

$$\mu \sum_i \left(a_{i,1}^\dagger a_{i,1} + a_{i,2}^\dagger a_{i,2} \right). \tag{6.3}$$

This term simply adds energy μ for each particle in the system, regardless of which site it is on. Tuning the parameter μ of this term allows us (in principle) to control the density of particles in the system by punishing (or rewarding) adding extra particles.

The Hubbard Hamiltonian is made up of the three aforementioned terms. The real parameters t, U and μ change the relative importance of the terms of the Hamiltonian, which in turn influences the ground state. For $t \to +\infty$, the model describes a Fermi gas of freely moving (non-interacting) particles [130]. On the other hand for $t = 0$, the model reduces to a single-site problem [130]. There are many different versions of the Hubbard model based on the geometry of the underlying lattice [131–134]. The lattice geometry determines which sites are considered nearest neighbours for the sum in equation (6.1). At first sight, this might seem like a minor detail, but it can have profound effect on the ground state, or even our capability to find one. Alternative generalizations of the Hamiltonian stem from tweaking some of its terms (e.g. adding anisotropy) or adding other terms (e.g. off-site interaction).



(a) 1D linear lattice (b) 2D square lattice

Figure 6.1: Depiction of possible states of the Hubbard model on two simple lattices. There are two fermionic species: depicted as red and blue. Each site has space for one fermion of each species. The straight lines connect neighbouring sites.

Despite its simplicity, solving the Hubbard model has proven to be difficult. To date, only a handful of special cases or toy versions of the model have been solved. One of such is the analytical solution of the Hubbard model on a 1D chain [42, 135, 136] (see figure 6.1a) found by employing the Bethe ansatz [44]. Another is a numerical solution of the Hubbard model on an infinite-dimensional cubic lattice obtained by using *dynamical mean field theory* [137]. In this chapter we solve another toy version of the model — permutationally-invariant Hubbard model. In chapter 7 we provide some of our own investigation into the infinite-dimensional cubic lattice Hubbard model.

6.1.2 Permutationally-Invariant Hubbard Model with Off-Site Interactions

In our case we are going to solve the Hubbard model for a permutationally-invariant system. Here it should be mentioned that only after our work was finished did it come to our attention that some research into this model (and similar models) had already been done in the past [129, 138–142]. Some of the results overlap, but our approach is unique. Furthermore, we also add a permutationally-invariant off-site interaction term to the Hamiltonian, which is entirely new.

We will label by N the number of sites in our system and we are interested in the limit of very large N. The permutation invariance implies that there is no way of distinguish any specific pair of sites, contrary to what happens for usual lattices where we can identify nearest neighbours. In a permutationally-invariant system, every site is nearest neighbour with every other site. This means that in our case, the hopping term looks like

$$H_t = t \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \left(a_{i,1}^\dagger a_{j,1} + a_{i,2}^\dagger a_{j,2} \right). \tag{6.4}$$

This sum runs over all pairs of sites twice, however it only contains terms hopping *from site j to site i*, so no term is repeated.

The on-site interaction and the chemical-potential terms look like

$$H_U = U \sum_{i=1}^{N} a_{i,1}^\dagger a_{i,1} a_{i,2}^\dagger a_{i,2}, \tag{6.5}$$

$$H_\mu = \mu \sum_{i=1}^{N} \left(a_{i,1}^\dagger a_{i,1} + a_{i,2}^\dagger a_{i,2} \right). \tag{6.6}$$

On top of these three terms, we add an extra off-site interaction term to the Hamiltonian,

$$H_V = \frac{V}{2} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \left(a_{i,1}^\dagger a_{i,1} + a_{i,2}^\dagger a_{i,2} \right) \left(a_{j,1}^\dagger a_{j,1} + a_{j,2}^\dagger a_{j,2} \right). \tag{6.7}$$

The $\frac{1}{2}$ factor in the off-site interaction term is included to cancel double counting (because the sums will count every pair of sites twice). Note that unlike the on-site

interaction, the off-site interaction acts also between two fermions of the same species, which may sit on different sites. We could have different same-species and off-species off-site interactions, but that would *explicitly* break the **SU(2)** symmetry of the Hamiltonian. The overall Hamiltonian will consist of the sum of the above four terms (6.4)–(6.7).

In order for all of the terms of the Hamiltonian to be relevant even in the thermodynamic limit (number of sites $N \to +\infty$), we need to pay attention to how they scale with N. We use the result of Kraus et al. [143], who described the scaling of permutationally-invariant fermionic Hamiltonians with at most two-site terms. Terms which are sums of single-site operators, such as H_U and H_μ, scale linearly with N. For sums of two-site operators, there are two possibilities. The deciding factor is whether the two-site operators act with even or odd number of mode operators on each of the two sites. The off-site interaction term H_V belongs in the *even* category and therefore scales like N^2. The hopping term belongs in the *odd* category and therefore it scales linearly with N. This scaling behaviour is typical of fermionic operators and does not apply to permutationally-invariant bosonic systems (compare with the Hamiltonian for hardcore bosons (5.36)).

For simplicity, we scale the Hamiltonian so that the energy is of the scale $\mathcal{O}(1)$ as $N \to +\infty$,

$$H = \frac{H_t}{N} + \frac{H_U}{N} + \frac{H_\mu}{N} + \frac{H_V}{N^2}. \tag{6.8}$$

6.1.3 Symmetries of Hubbard Model with Permutation Invariance

To be able to apply the approach from chapter 5, we first need to identify the symmetries of the Hamiltonian from equation (6.8).

Permutation symmetry

The first symmetry (obvious from the construction) is the permutation symmetry. The corresponding group is called the *symmetric group* and it is labelled S_N. The permutation is an operation on the system sites i, transforming them into different system sites

$$\pi : i \to \pi(i). \tag{6.9}$$

The correspond unitary acts on the the annihilation operators as[1]

$$a_{i,1} \to a_{\pi(i),1}, \tag{6.10}$$
$$a_{i,2} \to a_{\pi(i),2}. \tag{6.11}$$

Note that we have to apply the same permutation to both species of the fermions. This makes sense physically because we permute only the sites and a site can host both species of the fermions. Any permutation can be constructed from multiple

[1] As the action of the symmetry is a unitary transformation, it looks exactly the same when annihilation operators are turned into creation operators by the Hermitian adjoint †.

swaps (permutations of only two sites). A unitary corresponding to the swap of sites $i \longleftrightarrow j$ looks like

$$U_{(ij)} = \prod_{k=1}^{2} \left(a_{i,k} a_{i,k}^\dagger a_{j,k} a_{j,k}^\dagger + a_{i,k}^\dagger a_{i,k} a_{j,k}^\dagger a_{j,k} + a_{i,k}^\dagger a_{j,k} + a_{j,k}^\dagger a_{i,k} \right). \tag{6.12}$$

Any more complicated permutation will be composed of multiple swaps, so the overall unitary will be a product of multiple swap unitaries.

Particle number symmetry

Another symmetry of the Hamiltonian is the **U(1)** particle number symmetry. The corresponding unitary representation will add a complex phase to all of the annihilation operators (and the opposite phase to all of the creation operators):

$$a_{i,1} \to \mathrm{e}^{\mathrm{i}\alpha} a_{i,1}, \tag{6.13}$$
$$a_{i,2} \to \mathrm{e}^{\mathrm{i}\alpha} a_{i,2}. \tag{6.14}$$

Here α is a real parameter. We can see that this is a symmetry of the Hamiltonian because each of its terms consists of products of an equal number of creation and annihilation operators. The corresponding unitary looks like

$$U_{\mathbf{U(1)}}(\alpha) = \prod_{i=1}^{N} \prod_{k=1}^{2} \left(a_{i,k} a_{i,k}^\dagger + \mathrm{e}^{-\mathrm{i}\alpha} a_{i,k}^\dagger a_{i,k} \right). \tag{6.15}$$

SU(2) rotation symmetry

The last symmetry that we will present here is the **SU(2)** spin symmetry. This symmetry corresponds to the following transformation:

$$a_{i,1} \to \beta a_{i,1} - \gamma^* a_{i,2}, \tag{6.16}$$
$$a_{i,2} \to \gamma a_{i,1} + \beta^* a_{i,2}. \tag{6.17}$$

Here β and γ are complex parameters satisfying $|\beta|^2 + |\gamma|^2 = 1$ (so that the transformation in equations (6.16) and (6.17) is a special unitary transformation). To see that the Hamiltonian has this symmetry, we look at each term individually:

$$a_{i,1}^\dagger a_{j,1} \to \left[\beta^* a_{i,1}^\dagger - \gamma a_{i,2}^\dagger \right] \left[\beta a_{j,1} - \gamma^* a_{j,2} \right] \tag{6.18}$$
$$= |\beta|^2 a_{i,1}^\dagger a_{j,1} - \beta^* \gamma^* a_{i,1}^\dagger a_{j,2} - \beta\gamma a_{i,2}^\dagger a_{j,1} + |\gamma|^2 a_{i,2}^\dagger a_{j,2}, \tag{6.19}$$

$$a_{i,2}^\dagger a_{j,2} \to \left[\gamma^* a_{i,1}^\dagger + \beta a_{i,2}^\dagger \right] \left[\gamma a_{j,1} + \beta^* a_{j,2} \right] \tag{6.20}$$
$$= |\gamma|^2 a_{i,1}^\dagger a_{j,1} + \beta^* \gamma^* a_{i,1}^\dagger a_{j,2} + \beta\gamma a_{i,2}^\dagger a_{j,1} + |\beta|^2 a_{i,2}^\dagger a_{j,2}. \tag{6.21}$$

If we add together the terms (6.19) and (6.21), we obtain $a_{i,1}^\dagger a_{j,1} + a_{i,2}^\dagger a_{j,2}$, proving that the hopping term has this symmetry. The same applies when we set $j = i$, proving that the chemical-potential and off-site interaction terms also have this

symmetry. To see what the on-site interaction term gets transformed into, we look at the transformation of $a_{i,1}a_{i,2}$:

$$a_{i,1}a_{i,2} \to [\beta a_{j,1} - \gamma^* a_{j,2}]\,[\gamma a_{j,1} + \beta^* a_{j,2}] \tag{6.22}$$
$$= |\beta|^2 a_{i,1}a_{i,2} - |\gamma|^2 a_{i,2}a_{i,1} = a_{i,1}a_{i,2}. \tag{6.23}$$

For the same reason, we get $a^\dagger_{i,1}a^\dagger_{i,2} \to a^\dagger_{i,1}a^\dagger_{i,2}$. Therefore

$$a^\dagger_{i,1}a_{i,1}a^\dagger_{i,2}a_{i,2} = -a^\dagger_{i,1}a^\dagger_{i,2}a_{i,1}a_{i,2} \to -a^\dagger_{i,1}a^\dagger_{i,2}a_{i,1}a_{i,2} = a^\dagger_{i,1}a_{i,1}a^\dagger_{i,2}a_{i,2}, \tag{6.24}$$

which proves that the interaction term also has the **SU(2)** symmetry. The corresponding unitary is

$$U_{\mathbf{SU(2)}}(\beta,\gamma) = \prod_{i=1}^{N} \Big(a_{i,1}a^\dagger_{i,1}a_{i,2}a^\dagger_{i,2} + a^\dagger_{i,1}a_{i,1}a^\dagger_{i,2}a_{i,2} \tag{6.25}$$
$$+ \beta a^\dagger_{i,1}a_{i,1}a_{i,2}a^\dagger_{i,2} + \beta^* a_{i,1}a^\dagger_{i,1}a^\dagger_{i,2}a_{i,2} \tag{6.26}$$
$$- \gamma^* a^\dagger_{i,2}a_{i,1} + \gamma a^\dagger_{i,1}a_{i,2}\Big). \tag{6.27}$$

Note that as it stands, the Hubbard Hamiltonian from equation (6.8) *does not* have a particle-hole symmetry. It can have it for some specific values of the parameters t, U, μ and V, but it does not have it generally.

6.2 Symmetries and Variational Family

In this section we will go through the process described in chapter 5 to extend the symmetries described in subsection 6.1.3 to act onto a larger Hilbert space of purifications. We will then describe the states within this space that have the required symmetries and which will therefore be used to find the ground state.

6.2.1 Extending Symmetries to the Ancillary Subspace

Before we start looking for the ground state purification, we need to extend the action of the symmetries discussed in subsection 6.1.3 to the ancillary subspace. The way to do that is shown in section 5.3. We let the Hermitian adjoint of the unitaries act on the physical half of a maximally entangled state (5.13) and then use the formulas (5.20) and (5.21) to transform the unitaries into different ones acting on the ancillary subspace instead.

We begin with the **U(1)** unitary from equation (6.15). We go slowly here

whereas for the other unitaries we skip over some steps for brevity:

$$\left[U^\dagger_{\mathbf{U(1)}}(\alpha)\right]_A |\omega\rangle_{AB} = \prod_{i=1}^{N}\prod_{k=1}^{2}\left(a_{i,k}a^\dagger_{i,k} + \mathrm{e}^{-\mathrm{i}\alpha}a^\dagger_{i,k}a_{i,k}\right)^\dagger |\omega\rangle_{AB} \tag{6.28}$$

$$= \prod_{i=1}^{N}\prod_{k=1}^{2}\left(a_{i,k}a^\dagger_{i,k} + \mathrm{e}^{\mathrm{i}\alpha}a^\dagger_{i,k}a_{i,k}\right) |\omega\rangle_{AB} \tag{6.29}$$

$$= \prod_{i=1}^{N}\prod_{k=1}^{2}\left(-a_{i,k}b_{i,k} + \mathrm{e}^{\mathrm{i}\alpha}a^\dagger_{i,k}b^\dagger_{i,k}\right) |\omega\rangle_{AB} \tag{6.30}$$

$$= \prod_{i=1}^{N}\prod_{k=1}^{2}\left(b_{i,k}a_{i,k} - \mathrm{e}^{\mathrm{i}\alpha}b^\dagger_{i,k}a^\dagger_{i,k}\right) |\omega\rangle_{AB} \tag{6.31}$$

$$= \prod_{i=1}^{N}\prod_{k=1}^{2}\left(b_{i,k}b^\dagger_{i,k} + \mathrm{e}^{\mathrm{i}\alpha}b^\dagger_{i,k}b_{i,k}\right) |\omega\rangle_{AB} \tag{6.32}$$

$$= \left[\tilde{U}_{\mathbf{U(1)}}(\alpha)\right]_B |\omega\rangle_{AB}\,. \tag{6.33}$$

The parentheses each act on a different site/species and contain an even number of fermionic operators, so they all commute with each other. That is why we can do the operations in parallel. We can see that the unitary $\left[\tilde{U}_{\mathbf{U(1)}}(\alpha)\right]_B$ looks almost exactly like the unitary $\left[U_{\mathbf{U(1)}}(\alpha)\right]_A$, but with the opposite phase (and acting on the ancillary fermions instead of the physical fermions). Taken as a single unitary $\left[U_{\mathbf{U(1)}}(\alpha)\tilde{U}_{\mathbf{U(1)}}(\alpha)\right]_{AB}$ acting on the entire Hilbert space, the $\mathbf{U(1)}$ symmetry acts like

$$a_{i,1} \to \mathrm{e}^{\mathrm{i}\alpha}a_{i,1}, \tag{6.34}$$

$$a_{i,2} \to \mathrm{e}^{\mathrm{i}\alpha}a_{i,2}, \tag{6.35}$$

$$b_{i,1} \to \mathrm{e}^{-\mathrm{i}\alpha}b_{i,1}, \tag{6.36}$$

$$b_{i,2} \to \mathrm{e}^{-\mathrm{i}\alpha}b_{i,2}. \tag{6.37}$$

Applying the same process, we find that the unitary $\left[\tilde{U}_{(ij)}\right]_B$ corresponding to the swapping of two sites extended onto the ancillary degrees of freedom looks like

$$\left[\tilde{U}_{(ij)}\right]_B = \prod_{k=1}^{2}\left(b_{i,k}b^\dagger_{i,k}b_{i,k}b^\dagger_{i,k} + b^\dagger_{i,k}b_{i,k}b^\dagger_{i,k}b_{i,k} + b^\dagger_{i,k}b_{j,k} + b^\dagger_{j,k}b_{i,k}\right). \tag{6.38}$$

This looks identical to $\left[U_{(ij)}\right]_A$ (other than acting on the ancillary fermions instead of the physical fermions). Therefore permutation acts in the same way on the physical and the ancillary fermions:

$$a_{i,1} \to a_{\pi(i),1}, \tag{6.39}$$

$$a_{i,2} \to a_{\pi(i),2}, \tag{6.40}$$

$$b_{i,1} \to b_{\pi(i),1}, \tag{6.41}$$

$$b_{i,2} \to b_{\pi(i),2}. \tag{6.42}$$

Similarly, we extend the action of the **SU(2)** symmetry to the ancillary fermions. The corresponding unitary is

$$\left[\tilde{U}_{\mathbf{SU(2)}}(\beta,\gamma)\right] = \prod_{i=1}^{N} \Big(b_{i,1}b_{i,1}^\dagger b_{i,2}b_{i,2}^\dagger + b_{i,1}^\dagger b_{i,1} b_{i,2}^\dagger b_{i,2} \tag{6.43}$$

$$+ \beta^* b_{i,1}^\dagger b_{i,1} b_{i,2} b_{i,2}^\dagger + \beta b_{i,1} b_{i,1}^\dagger b_{i,2}^\dagger b_{i,2} \tag{6.44}$$

$$- \gamma b_{i,2}^\dagger b_{i,1} + \gamma^* b_{i,1}^\dagger b_{i,2} \Big). \tag{6.45}$$

This looks like the **SU(2)** unitary for the physical fermions, except the coefficients β and γ are complex-conjugated. The **SU(2)** symmetry extended on the entire Hilbert space acts like

$$a_{i,1} \to \beta a_{i,1} - \gamma^* a_{i,2}, \tag{6.46}$$

$$a_{i,2} \to \gamma a_{i,1} + \beta^* a_{i,2}, \tag{6.47}$$

$$b_{i,1} \to \beta^* b_{i,1} - \gamma b_{i,2}, \tag{6.48}$$

$$b_{i,2} \to \gamma^* b_{i,1} + \beta b_{i,2}. \tag{6.49}$$

6.2.2 Fermionic de Finetti Theorem

Now we start restricting our state space to states compatible with the symmetries described in the previous subsection. We will start with the permutation symmetry. Bosonic states with permutation symmetry obey the quantum de Finetti theorem, which says that for $N \to +\infty$, they can be approximated well by product states (or mixtures thereof) [54–57].

An analogous result for fermionic systems was recently discovered by Krumnow et al. [58], called the *fermionic de Finetti theorem.*

Theorem 6 (Fermionic de Finetti)**.** *Let ρ be a permutationally invariant fermionic state on a system with $N \geq 6$ sites with p fermionic species per site. Then there exist single-site fermionic states $|\phi_l\rangle$ respecting the fermionic parity superselection rule and coefficients $a_l \geq 0$ such that $\sum_l a_l = 1$, satisfying*

$$\left\| \mathrm{Tr}_{N\setminus k}\, \rho - \sum_l a_l \left|\phi_l\right\rangle \left\langle\phi_l\right|^{\otimes l} \right\|_1 \leq \frac{2^{2p+1}\sqrt{k-1}^3}{\sqrt{3}N} + \frac{2^{2p+1}k}{N}. \tag{6.50}$$

The notation here is adapted directly from Krumnow et al. [58] and warrants some explanation. As was emphasized in section 2.6, fermionic many-body Hilbert spaces do not have the tensor product structure and thus do not allow for simple tensor-product states[2]. However, as the states $|\phi_l\rangle$ respect the fermionic parity superselection rule, the density matrix $|\phi_l\rangle\langle\phi_l|$ consists of products of even number of mode operators. Products of even number of fermionic mode operators commute with each other, so a product of the density matrices $|\phi_l\rangle\langle\phi_l|$ for different sites will act exactly like a tensor product state. If we perform the Jordan-Wigner

[2]Instead, we can use *Slater determinants* [144], which have similar properties to product states.

transformation [145], we get indeed a tensor product state (of qudits). The operator $\mathrm{Tr}_{N\setminus k}\,\rho$ corresponds to the state ρ reduced from N to k sites.

The Fermionic de Finetti theorem implies that expectation values of operators with the permutationally invariant purification ground state can be approximated with error of order $\mathcal{O}\left(\frac{1}{N}\right)$ by a fermionic product state

$$\sigma = \sum_l a_l \left|\phi_l\right\rangle \left\langle\phi_l\right|^{\otimes N}. \tag{6.51}$$

In particular, let $|\Psi\rangle$ be the true purification ground state of the permutationally-invariant Hubbard model. The expectation value of the chemical-potential term with the ground state $|\Psi\rangle$ is

$$\left\langle\Psi\right| \sum_{i=1}^{N} \frac{\mu}{N} \left(a_{i,1}^\dagger a_{i,1} + a_{i,2}^\dagger a_{i,2}\right) \left|\Psi\right\rangle = \mu \left\langle\Psi\right| \left(a_{1,1}^\dagger a_{1,1} + a_{1,2}^\dagger a_{1,2}\right) \left|\Psi\right\rangle. \tag{6.52}$$

Here we used the fact that $|\Psi\rangle$ is permutationally invariant to replace the sum with just one of its summands. The expectation value of the same term with the state σ from equation (6.51) is

$$\mathrm{Tr}\left\{\sigma \sum_{i=1}^{N} \frac{\mu}{N} \left(a_{i,1}^\dagger a_{i,1} + a_{i,2}^\dagger a_{i,2}\right)\right\} = \mu\,\mathrm{Tr}\left\{\sigma \left(a_{1,1}^\dagger a_{1,1} + a_{1,2}^\dagger a_{1,2}\right)\right\}. \tag{6.53}$$

We can again use the fact that σ is permutationally invariant to replace the sum with just one of its summands. The difference of the expectation values in (6.52) and (6.53) is equal to

$$\begin{aligned}
&\mu\,\mathrm{Tr}\left\{\left(\left|\Psi\right\rangle\left\langle\Psi\right| - \sigma\right)\left(a_{1,1}^\dagger a_{1,1} + a_{1,2}^\dagger a_{1,2}\right)\right\} && (6.54)\\
&= \mu\,\mathrm{Tr}\left\{\mathrm{Tr}_{N\setminus 1}\left\{\left|\Psi\right\rangle\left\langle\Psi\right| - \sigma\right\}\left(a_{1,1}^\dagger a_{1,1} + a_{1,2}^\dagger a_{1,2}\right)\right\} && (6.55)\\
&\leq \mu \left\|\mathrm{Tr}_{N\setminus 1}\left\{\left|\Psi\right\rangle\left\langle\Psi\right| - \sigma\right\}\right\|_1 \left\|a_{1,1}^\dagger a_{1,1} + a_{1,2}^\dagger a_{1,2}\right\|_\infty. && (6.56)
\end{aligned}$$

From theorem 6 we know that the trace norm scales as $\mathcal{O}\left(\frac{1}{N}\right)$ whereas the other terms do not scale with N. This means that for large N the difference of the expectation values will go to 0 (for a correctly chosen σ).

As we are eventually interested in minimizing the energy, we may simplify σ further. Instead of considering a generic mixture of product states (6.51), we may instead only consider pure product states of the form $|\phi\rangle\langle\phi|^{\otimes N}$.

The same conclusion can be made for the on-site interaction term and for the off-site interaction term. However, it does not work for the hopping term. In that case the analogous term in $\|\cdot\|_\infty$ would scale linearly with N, negating the $\mathcal{O}\left(\frac{1}{N}\right)$ scaling of the $\|\cdot\|_1$ term.

6.2.3 Hopping

So far we have shown that we can approximate the ground state expectation value of the interaction terms and the chemical-potential term by a mixture of product

states σ (6.51). The same cannot be said of the hopping term. For that we show a different result.

First, we start with subtracting

$$t\sum_{i=1}^{N}\left(a_{i,1}^{\dagger}a_{i,1}+a_{i,2}^{\dagger}a_{i,2}\right) \tag{6.57}$$

from the chemical-potential term and adding it to the hopping term. This changes the chemical-potential term to

$$H'_{\mu}=(\mu-t)\sum_{i=1}^{N}\left(a_{i,1}^{\dagger}a_{i,1}+a_{i,2}^{\dagger}a_{i,2}\right) \tag{6.58}$$

and the hopping term to

$$H'_{t}=t\sum_{i=1}^{N}\sum_{j=1}^{N}\left(a_{i,1}^{\dagger}a_{j,1}+a_{i,2}^{\dagger}a_{j,2}\right), \tag{6.59}$$

so that now the sum includes the terms when $i=j$. This allows us to re-write the hopping term as

$$H'_{t}=tN\left(A_{1}^{\dagger}A_{1}+A_{2}^{\dagger}A_{2}\right) \tag{6.60}$$

by using

$$A_{k}=\sum_{i=1}^{N}\frac{a_{i,k}}{\sqrt{N}}. \tag{6.61}$$

Note that the operators A_k satisfy the standard fermionic anti-commutation relations:

$$\{A_{k},A_{l}^{\dagger}\}=\delta_{kl}, \tag{6.62}$$

$$\{A_{k},A_{l}\}=0. \tag{6.63}$$

The operator $A_k^{\dagger}A_k$ (for $k\in\{1,2\}$) is a positive operator with unit norm. Therefore the lowest / highest expectation value it can yield is 0 and 1 respectively. We will show that we can slightly tweak σ to force the expectation values to be either 0 or 1 without affecting the expectation values of the other terms by more than $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$. The slight tweak is to add the operators A_k or $A_k^{\dagger}$. To show how this works, we first consider the following two density matrices:

$$\sigma'=\frac{A_{1}^{\dagger}\sigma A_{1}}{\mathrm{Tr}\left\{A_{1}^{\dagger}\sigma A_{1}\right\}}, \tag{6.64}$$

$$\sigma''=\frac{A_{1}\sigma A_{1}^{\dagger}}{\mathrm{Tr}\left\{A_{1}\sigma A_{1}^{\dagger}\right\}}. \tag{6.65}$$

These two matrices were designed to yield expectation values 0 and 1 with the operator $A_1^{\dagger}A_1$. We can check that it is indeed the case by calculating

$$\mathrm{Tr}\left\{\sigma' A_{1}^{\dagger}A_{1}\right\}=\frac{\mathrm{Tr}\left\{A_{1}^{\dagger}\sigma A_{1}A_{1}^{\dagger}A_{1}\right\}}{\mathrm{Tr}\left\{A_{1}^{\dagger}\sigma A_{1}\right\}}=\frac{\mathrm{Tr}\left\{A_{1}^{\dagger}\sigma A_{1}\right\}}{\mathrm{Tr}\left\{A_{1}^{\dagger}\sigma A_{1}\right\}}=1, \tag{6.66}$$

$$\mathrm{Tr}\left\{\sigma'' A_{1}^{\dagger}A_{1}\right\}=\frac{\mathrm{Tr}\left\{A_{1}\sigma A_{1}^{\dagger}A_{1}^{\dagger}A_{1}\right\}}{\mathrm{Tr}\left\{A_{1}^{\dagger}\sigma A_{1}\right\}}=\frac{0}{\mathrm{Tr}\left\{A_{1}^{\dagger}\sigma A_{1}\right\}}=0. \tag{6.67}$$

Now we check the expectation values of the other terms of the Hamiltonian with these density matrices. For the following, define $h_I = a^\dagger_{1,1} a_{1,1} a^\dagger_{1,2} a_{1,2}$ and

$$\tilde{A}_k = A_k - \frac{a_{1,k}}{\sqrt{N}} = \sum_{i=2}^{N} \frac{a_{i,k}}{\sqrt{N}}. \tag{6.68}$$

The expectation value of the on-site interaction term is

$$\operatorname{Tr}\Big\{\sigma' \frac{H_U}{N}\Big\} = \frac{\operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1 H_U\Big\}}{N \operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1\Big\}} \tag{6.69}$$

$$= U \frac{\operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1 a^\dagger_{1,1} a_{1,1} a^\dagger_{1,2} a_{1,2}\Big\}}{\operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1\Big\}} \tag{6.70}$$

$$= \frac{U \operatorname{Tr}\Big\{\tilde{A}^\dagger_1 \sigma \tilde{A}_1 h_I\Big\}}{\operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1\Big\}} + \frac{\operatorname{Tr}\Big\{a^\dagger_{1,1} \sigma A_1 h_I + A^\dagger_1 \sigma a_{1,1} h_I\Big\}}{\sqrt{N} \operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1\Big\}} - \frac{\operatorname{Tr}\Big\{a^\dagger_{1,1} \sigma a_{1,1} h_I\Big\}}{N \operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1\Big\}}. \tag{6.71}$$

The second fraction on the line (6.71) is $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$ and the last fraction is $\mathcal{O}\left(\frac{1}{N}\right)$. To simplify the first fraction, we use the property that σ is a product state, so an expectation value of a product of non-overlapping observables factorizes

$$\operatorname{Tr}\Big\{\tilde{A}^\dagger_1 \sigma \tilde{A}_1 h_I\Big\} = \operatorname{Tr}\Big\{\tilde{A}^\dagger_1 \sigma \tilde{A}_1\Big\} \operatorname{Tr}\{\sigma h_I\}. \tag{6.72}$$

Finally, $\operatorname{Tr}\Big\{\tilde{A}^\dagger_1 \sigma \tilde{A}_1\Big\}$ is $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$-close to $\operatorname{Tr}\Big\{A^\dagger_1 \sigma A_1\Big\}$, which makes the final expectation value $U \operatorname{Tr}\{\sigma h_I\} = \operatorname{Tr}\Big\{\sigma \frac{H_U}{N}\Big\} + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$. Naturally, the same calculation can be done for σ'' and for the expectation values of H'_μ/N and H_V/N^2.

The only possible problem with σ' and σ'' is that they might not be defined if $\operatorname{Tr}\Big\{A^\dagger_k \sigma A_k\Big\}$ or $\operatorname{Tr}\Big\{A_k \sigma A^\dagger_k\Big\}$ is equal to 0. However, that means that either $\operatorname{Tr}\Big\{\sigma A^\dagger_k A_k\Big\} = 0$ or $\operatorname{Tr}\Big\{\sigma A_k A^\dagger_k\Big\} = 0$, which implies that σ already minimizes / maximizes the expectation value of $A^\dagger_k A_k$.

The density matrices σ' and σ'' only minimize / maximize $A^\dagger_1 A_1$. If we want to minimize / maximize $A^\dagger_2 A_2$ at the same time, we need to consider a matrix like $A_1 A_2 \sigma A^\dagger_2 A^\dagger_1$. In fact, we can add any number of the operators A_k, $A^\dagger_k$ (each at most once because of the anti-commutation relations) on top of σ without influencing the expectation values of H'_μ, H_U or H_V by more than $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$. In the final state, we will call σ the "even" part of the state and the added operators A_k, $A^\dagger_k$ the "odd" part of the state.

In previous research on the permutationally-invariant Hubbard model van Dongen and Vollhardt [129] reached the same conclusion as we did here. However, their approach was different. They showed that in the thermodynamic limit the kinetic part of the Hamiltonian commutes with the rest, so it can be minimized independently.

6.2.4 Restricting to Symmetric Purifications

Having described how the unitaries of the symmetries act on the Hilbert space of the purifications, we now search for those purifications which have all those

symmetries (including the phase). To simplify notation, in this phase we will relabel the creation / annihilation operators of the ancillary fermions. In chapter 5 and in subsection 6.2.1, these were labelled with the letter b to emphasize that they live in a subsystem B, e.g. $b_{i,2}$ is the ancillary fermion of species 2 on site i. However, here it will be more convenient to label them as the 3rd and 4th fermionic species, using the annihilation / creation operators $a_{i,3}, a_{i,3}^\dagger, a_{i,4}, a_{i,4}^\dagger$.

Because of the observations made in the subsections 6.2.2 and 6.2.3, we may take the following as the ansatz of our permutationally-invariant family of states:

$$\prod_{x=1}^{4} \left(A_x^\dagger\right)^{n_x} \prod_{x=1}^{4} (A_x)^{m_x} \prod_{i=1}^{N} \left[\prod_{\vec{x}\text{ even}} \left(\mathbb{1}_i + \frac{\alpha_{\vec{x}}}{\sqrt{N}} \prod_{j=1}^{|\vec{x}|} a_{i,x_j}^\dagger \right) \right] |\Omega\rangle \,. \tag{6.73}$$

The bracket $\prod_{i=1}^{N} [...]$ creates the product-state part of the state, also known as the "even" part. It only contains products of even number of creation operators to conform with the fermionic super-selection rule and the permutation invariance. Because of the anti-commutation relations of the operators A_k, the coefficients n_x and m_x can only attain values 0 or 1. This ansatz corresponds to the family of *fermionic symmetric basic states* which was introduced by Kraus et al. [143].

To simplify the description, we expand the product over $\prod_{\vec{x}\text{ even}}$ and relabel the coefficients:

$$\prod_{x=1}^{4} \left(A_x^\dagger\right)^{n_x} \prod_{x=1}^{4} (A_x)^{m_x} \prod_{i=1}^{N} \Big[\alpha_0 \mathbb{1}_i + \alpha_{12} a_{i,1}^\dagger a_{i,2}^\dagger + \alpha_{13} a_{i,1}^\dagger a_{i,3}^\dagger + \alpha_{14} a_{i,1}^\dagger a_{i,4}^\dagger \tag{6.74}$$

$$+\alpha_{23} a_{i,2}^\dagger a_{i,3}^\dagger + \alpha_{24} a_{i,2}^\dagger a_{i,4}^\dagger + \alpha_{34} a_{i,3}^\dagger a_{i,4}^\dagger + \alpha_{1234} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \Big] |\Omega\rangle \,. \tag{6.75}$$

All the coefficients α as well as n_x, m_x are still free variables. In this subsection, we examine which values of them are allowed in order to keep the state symmetric under the $\mathbf{U(1)}$ and $\mathbf{SU(2)}$ symmetries[3]. We do this by applying the symmetry transformations as defined in equations (6.34)–(6.37) and (6.46)–(6.49).

Let us start with $\mathbf{SU(2)}$. By choosing $\gamma = 1, \beta = 0$ in equations (6.46)–(6.49), we get a transformation that transforms $a_{i,1}^\dagger a_{i,3}^\dagger$ into $a_{i,2}^\dagger a_{i,4}^\dagger$. That implies that $\alpha_{13} = \alpha_{24}$. Setting $\beta = \mathrm{i}, \gamma = 0$ transforms $a_{i,1}^\dagger a_{i,4}^\dagger$ into $-a_{i,1}^\dagger a_{i,4}^\dagger$ and $a_{i,2}^\dagger a_{i,3}^\dagger$ into $-a_{i,2}^\dagger a_{i,3}^\dagger$, forcing $\alpha_{14} = \alpha_{23} = 0$. By similar arguments, the $\mathbf{U(1)}$ symmetry requires that $\alpha_{12} = \alpha_{34} = 0$. Consequently, the symmetry restrictions therefore reduce the ground-state search to states with the following "even" part:

$$|\Psi\rangle = \prod_{i=1}^{N} \left[\alpha_0 \mathbb{1}_i + \alpha_{13} \left(\frac{a_{i,1}^\dagger a_{i,3}^\dagger + a_{i,2}^\dagger a_{i,4}^\dagger}{\sqrt{2}} \right) + \alpha_{1234} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \right] |\Omega\rangle \,. \tag{6.76}$$

For consistency, we will enforce $|\alpha_0|^2 + |\alpha_{13}|^2 + |\alpha_{1234}|^2 = 1$, so that the "even" part of the state is normalized.

[3]The state is already permutationally symmetric by construction.

6.3 Ground State Phase Diagram

Using the variational family described in equation (6.76), we will now find the ground state depending on the parameters t, U, μ, V of the Hamiltonian. First, we calculate the energy of the generic state in equation (6.76). As was mentioned in subsection 6.2.4, the "odd" part of the state does not influence the expectation value of the H'_μ, H_U, H_V terms of the Hamiltonian. Let $|\Psi\rangle$ be a state of the form (6.76). We have

$$\langle\Psi| \frac{H_U}{N} |\Psi\rangle = U\alpha^*_{1234}\alpha_{1234}, \tag{6.77}$$

$$\langle\Psi| \frac{H'_\mu}{N} |\Psi\rangle = (\mu - t)\left(2\alpha^*_{1234}\alpha_{1234} + \alpha^*_{13}\alpha_{13}\right), \tag{6.78}$$

$$\langle\Psi| \frac{H_V}{N^2} |\Psi\rangle = V\left(2\alpha^*_{1234}\alpha_{1234} + \alpha^*_{13}\alpha_{13}\right)^2. \tag{6.79}$$

The expectation value of the hopping term H'_t can be manipulated purely by changing the "odd" part of the wavefunction. Specifically $A_1A_2A_3A_4$ forces the hopping term expectation value to be 0 whereas $A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger$ makes it t — the choice depends only on the sign of t. Both $A_1A_2A_3A_4$ and $A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger$ have all of the symmetries of the Hamiltonian.

We can now minimize the energy. Based on the resulting $\alpha_{1234}, \alpha_{13}, \alpha_0$ we construct the phase diagram. It is important to note that regardless of these coefficients, each phase is further split into two based on the "odd" part of the wavefunction. We may call these two situations *fermion hopping* and *hole hopping*, corresponding to $A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger$ and $A_1A_2A_3A_4$ respectively. Note that for $t \neq 0$, fermion/hole hopping will make a relevant change in the energy, even though the operators $A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger$ and $A_1A_2A_3A_4$ only add / remove 4 particles to / from the system. However, the hopping will *not* be visible in the thermodynamic limit. Adding fermion / hole hopping to a state cannot be detected by any local observable (in the thermodynamic limit) and thus it will make no difference for the ground state. Note that in the language of C*-algebras, there is no definition for the Hamiltonian or the ground state energy in the thermodynamic limit — only the ground state itself. See appendix B for a more thorough explanation of the C*-algebra approach to the thermodynamic limit.

As the calculation of the phase diagram is somewhat tedious, we leave it over to appendix A and only present the results here. The ground state phase space is parametrized by the 4 paramters μ, V, t, U, but the phase diagram is fully characterized by its cut in the $t - U$ plane for positive and negative V separately.

The phase diagram for positive V is depicted in figure 6.2 and the phase diagram for negative V is depicted in figure 6.3. The phases are separated by a number of hyperplanes whose equations are provided below in equations (6.80)–(6.89).

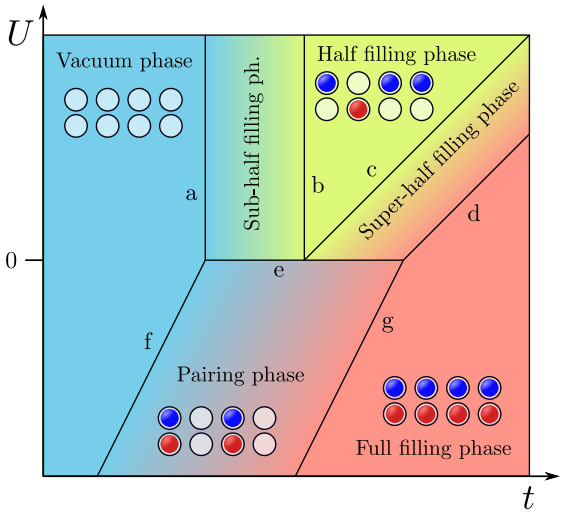


Figure 6.2: The ground state phase diagram for the permutationally-invariant Hubbard model in the $t - U$ plane for $V > 0$. The lines a, b, c, d, e, f, g are described in equations (6.80)–(6.86).

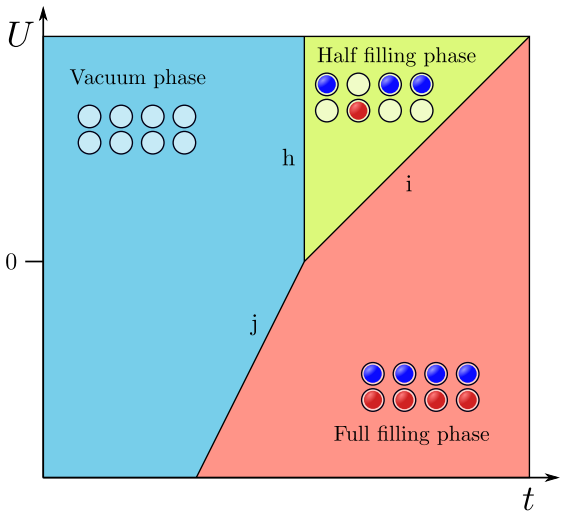


Figure 6.3: The ground state phase diagram for the permutationally-invariant Hubbard model in the $t - U$ plane for $V \leq 0$. The lines h, j, i are described in equations (6.87)–(6.89).

$$\text{a}: \ t = \mu \tag{6.80}$$
$$\text{b}: \ t = \mu + 2V \tag{6.81}$$
$$\text{c}: \ U = t - \mu - 2V \tag{6.82}$$
$$\text{d}: \ U = t - \mu - 4V \tag{6.83}$$
$$\text{e}: \ U = 0 \tag{6.84}$$
$$\text{f}: \ U = 2t - 2\mu \tag{6.85}$$
$$\text{g}: \ U = 2t - 2\mu - 8V \tag{6.86}$$
$$\text{h}: \ t = \mu + 2V \tag{6.87}$$
$$\text{i}: \ U = t - \mu - 2V \tag{6.88}$$
$$\text{j}: \ U = 2t - 2\mu - 4V \tag{6.89}$$

6.4 Interpretation of Results

In the phase diagrams 6.2 and 6.3 we can see a lot of interesting phases. In this section we go over parts of the phase diagram and discuss the physics of the various phases and the symmetries they spontaneously break. For a deeper look into how these results were obtained, see appendix A.

6.4.1 Full Filling and Vacuum

These two phases are relatively simple from the physical point of view. They correspond to the only non-zero coefficient being either α_0 or α_{1234} in equation (6.76). This makes their wavefunctions (without hopping)

$$|\Phi_{\text{vac}}\rangle = |\Omega\rangle \,, \tag{6.90}$$
$$|\Phi_{\text{full}}\rangle = \prod_{i=1}^{N} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \, |\Omega\rangle \,. \tag{6.91}$$

This means that the system is either completely full or completely empty of fermionic particles. As such, these states are product states without any entanglement or symmetry breaking. Depending on the sign of t, the vacuum ground state may have fermion hopping and the full filling state may have hole hopping:

$$|\Phi'_{\text{vac}}\rangle = A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger \, |\Omega\rangle \,, \tag{6.92}$$
$$|\Phi'_{\text{full}}\rangle = A_1 A_2 A_3 A_4 \prod_{i=1}^{N} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \, |\Omega\rangle \,. \tag{6.93}$$

6.4.2 Half Filling

The half filling phase corresponds to the situation of each site being occupied by just one particle. It corresponds to α_{13} being the only non-zero coefficient in

equation (6.76). That makes the "even" part of the wavefunction

$$|\Phi_{\text{half}}\rangle = \prod_{i=1}^{N} \left(\frac{a_{i,1}^\dagger a_{i,3}^\dagger + a_{i,2}^\dagger a_{i,4}^\dagger}{\sqrt{2}} \right) |\Omega\rangle \tag{6.94}$$

The ground state wavefunction itself will have $A_1 A_2 A_3 A_4$ or $A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger$ in front (depending on the sign of t). We can see that $|\Phi_{\text{half}}\rangle$ is entangled between the physical and ancillary subspaces, so the reduced state onto the physical subsystem is necessarily mixed. Specifically, it looks like

$$\rho_{\text{half}} = \prod_i \frac{1}{2} \left(a_{i,1}^\dagger a_{i,1} a_{i,2} a_{i,2}^\dagger + a_{i,1} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,2} \right). \tag{6.95}$$

This mixed state may be interpreted as a mixture of many different ground state (there is large degeneracy). Each of the ground states has 1 particle per site, conforming to the **U(1)** particle number symmetry. However they differ in which species are which of the particles. For example a state that is fully filled with fermions of species 1,

$$|\Phi_1\rangle = \prod_{i=1}^{N} a_{i,1}^\dagger |\Omega\rangle , \tag{6.96}$$

belongs to this phase. While this state has the permutation symmetry (up to a phase), it breaks the **SU(2)** symmetry by favouring one fermionic species over the other. Furthermore, as N grows this state becomes a rare abnormality. For large N, most of the ground states will have roughly half of the sites occupied by fermions of species 1 and the other half by fermions of species 2, which breaks the permutation symmetry.

6.4.3 Sub- and Super-Half Filling

The sub- and super-half filling phases correspond to both α_{13} and either α_0 or α_{1234} non-zero. The magnitude of α_{13} decreases as we move away from half-filling towards vacuum / full filling in the phase diagram. The "even" parts of the wavefunctions look like

$$|\Phi_{\text{sub}}\rangle = \prod_{i=1}^{N} \left[\alpha_0 \mathbb{1}_i + \alpha_{13} \left(\frac{a_{i,1}^\dagger a_{i,3}^\dagger + a_{i,2}^\dagger a_{i,4}^\dagger}{\sqrt{2}} \right) \right] |\Omega\rangle , \tag{6.97}$$

$$|\Phi_{\text{super}}\rangle = \prod_{i=1}^{N} \left[\alpha_{13} \left(\frac{a_{i,1}^\dagger a_{i,3}^\dagger + a_{i,2}^\dagger a_{i,4}^\dagger}{\sqrt{2}} \right) + \alpha_{1234} \left(a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \right) \right] |\Omega\rangle . \tag{6.98}$$

Again, on top of the wavefunctions in equations (6.97) and (6.98), these states will contain hole hopping or fermion hopping.

These phases correspond to ground states where the proportion of half-occupied sites is strictly between 0 and 1 while the rest is empty / fully occupied. As is shown in the appendix section A.3, the states in this phase break all of the **U(1)**, **SU(2)** and permutation symmetries. The reduced states on the physical subsystem look like

$$\rho_{\text{sub}} = \prod_{i=1}^{N} \left[|\alpha_0|^2 a_{i,1} a_{i,1}^\dagger a_{i,2} a_{i,2}^\dagger + \frac{|\alpha_{13}|^2}{\sqrt{2}} \left(a_{i,1}^\dagger a_{i,1} a_{i,2} a_{i,2}^\dagger + a_{i,1} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,2} \right) \right], \quad (6.99)$$

$$\rho_{\text{super}} = \prod_{i=1}^{N} \left[|\alpha_{1234}|^2 a_{i,1}^\dagger a_{i,1} a_{i,2}^\dagger a_{i,2} + \frac{|\alpha_{13}|^2}{\sqrt{2}} \left(a_{i,1}^\dagger a_{i,1} a_{i,2} a_{i,2}^\dagger + a_{i,1} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,2} \right) \right]. \quad (6.100)$$

These states may again be interpreted as mixtures of many different ground states. Most of these will have a proportion of $|\alpha_{13}|^2$ sites occupied by one fermion and the rest either empty of full (doubly occupied). Note that as $|\alpha_{13}|$ decreases, the entanglement weakens (and the ground state degeneracy becomes smaller).

6.4.4 Pairing

The pairing phase corresponds to states with both α_0 and α_{1234} non-zero. That means that the "even" part of the wavefunction of the ground state in this phase is

$$\prod_{i=1}^{N} \left[\alpha_0 \mathbb{1}_i + \alpha_{1234} \left(a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \right) \right] |\Omega\rangle . \quad (6.101)$$

The relative magnitude continuously shifts from α_0 to α_{1234} as we move from vacuum to the fully filled phase in the phase diagram. Physically, those states correspond to the situation in which no site is half-occupied while some are empty and some are fully occupied. This phase only occurs for negative interaction energy U, which encourages the fermions to pair together.

Again, there is entanglement between the ancillary and physical subspace and again the strength of the entanglement changes throughout the phase. As is shown in appendix section A.3, there is spontaneous breaking of $\mathbf{U(1)}$ and permutation symmetries. However, now the $\mathbf{SU(2)}$ symmetry is preserved, which comes from the fact that the fermions stick together in pairs. Among the many ground states in this phase, two are particularly interesting:

$$|\Psi_1\rangle = \prod_{i=1}^{|\alpha_{1234}|^2 N} \left(a_{i,1}^\dagger a_{i,2}^\dagger \right) |\Omega\rangle , \quad (6.102)$$

$$|\Psi_2\rangle = \prod_{i=1}^{N} \left(\alpha_0 + \alpha_{1234} a_{i,1}^\dagger a_{i,2}^\dagger \right) |\Omega\rangle . \quad (6.103)$$

The state $|\Psi_1\rangle$ preserves the $\mathbf{U(1)}$ symmetry while spontaneously breaking the permutation symmetry. On the other hand, the state $|\Psi_2\rangle$ preserves the permutation symmetry at the cost of breaking the $\mathbf{U(1)}$ symmetry. The state $|\Psi_1\rangle$ corresponds to $|\alpha_{1234}|^2 N$ sites being fully occupied with the rest empty. The state $|\Psi_2\rangle$ corresponds to each site being in a superposition of being fully occupied and empty. The state $|\Psi_2\rangle$ contains pairing as defined in [146], and it gives the phase its name.

Chapter 7

Hubbard Model in Infinite Dimensions

This chapter summarizes our investigation of the Hubbard model in infinite dimensions. The standard method to solve this problem is the *dynamical mean field theory* (DMFT) [147]. The research in this chapter was not motivated by the goal to improve upon this method, but by the idea to approach the problem from the direction of quantum information theory instead. The hope was that applying quantum-information-theoretic methods to this problem would give us more insight into the structure of the ground state, its entanglement and similar properties. This research was done in parallel with the research into models with permutation symmetry, however it was not published, as it did not lead to any particularly novel results.

First in section 7.1 we introduce the Hubbard Hamiltonian in infinite dimensions. Then in section 7.2 we briefly describe the idea behind DMFT. In section 7.3 we discuss the problem without on-site interactions and provide an analytical solution. Section 7.4 showcases the use of a variational family called *fermionic symmetric basic states* to provide a lower bound on the ground state energy. Then in section 7.5 we attempt to approach the ground state energy from the top with a few basic variational families of states generalized for infinite dimensions. Finally, section 7.6 provides a comparison of our approaches with DMFT.

7.1 Hamiltonian

The Hubbard Hamiltonian on a cubic lattice in d dimensions consists of the familiar hopping term

$$t \sum_{<i,j>} \left(a_{i,1}^\dagger a_{j,1} + a_{i,2}^\dagger a_{j,2} \right), \tag{7.1}$$

the interaction term

$$U \sum_i a_{i,1}^\dagger a_{i,1} a_{i,2}^\dagger a_{i,2}, \tag{7.2}$$

and the chemical-potential term

$$\mu \sum_i \left(a_{i,1}^\dagger a_{i,1} + a_{i,2}^\dagger a_{i,2} \right). \tag{7.3}$$

The sum over i runs over all sites of the entire system whereas the sum over $< i, j >$ runs over all ordered pairs of nearest neighbours. For simplicity of notation we still label the sites by a single letter i, but now these indices can be understood as a vector of d integers enumerating the coordinates of the sites. For future reference, we label $\mathcal{A}$ the set of all such vectors.

We want to investigate this Hamiltonian in the limit of an infinitely large cubic lattice in infinite dimensions. To get there, we will start with a finite cubic lattice of edge length N in d dimensions with periodic boundary conditions. Therefore, the system contains a total of N^d sites. We are interested in the limits $N \to +\infty$ and $d \to +\infty$. As one might expect, the order in which we take the limits is relevant. We will first take $d \to +\infty$ followed by $N \to +\infty$ to simplify the calculations.

As in the permutationally-invariant case, we need to pay attention to the scaling of the various terms in order to make them all relevant within the limits. Overall the hopping term will contain a sum of $2dN^d$ terms whereas the interaction term and the chemical-potential term will both contain just a sum of N^d terms. We want the energy to be extensive in the number of sites, so the scaling with N^d is desired. Similarly to the permutationally-invariant case though, the proper scaling of the hopping term now is $\frac{1}{\sqrt{2d}}$ [137, 147].

7.2 Dynamical Mean Field Theory

The standard method for solving the Hubbard model and related models in a large number of dimensions is dynamical mean field theory (DMFT). The results of DMFT become exact as the number of dimensions approaches infinity, but the theory is used for finite-dimensional systems as well [50, 148–151].

To explain properly the theory behind DMFT is out of scope of this thesis, so we just briefly summarize the approach here. For a more detailed look into DMFT, we recommend [147].

To derive the DMFT equations, we look at the Green's functions of the system, in particular at its self-energy $\Sigma_{ij\sigma}$ — the difference between the inverted free Green's function $G^0_{ij\sigma}$ and the invereted interacting Green's function $G_{ij\sigma}$. This definition of the self-energy $\Sigma_{ij\sigma}$

$$\Sigma_{ij\sigma} = \left(G^0\right)^{-1}_{ij\sigma} - (G)^{-1}_{ij\sigma} \tag{7.4}$$

is also called the *Dyson equation*. Here i, j refer to two site indices and σ indicates the species of the fermions. By examining the Feynman diagrams that make up the self-energy, it becomes apparent that a lot of them vanish in the $d \to +\infty$ limit. In fact, only the terms of the self-energy which are diagonal in the site indices i, j remain (this is where the *mean field* name comes from), rendering the self-energy independent of momentum. This allows us to obtain a compact relation between the (local) self energy and the (local) Green's function:

$$G_{ii\sigma}(\mathrm{i}\omega) = \int \mathrm{d}\varepsilon \frac{\rho(\varepsilon)}{\mathrm{i}\omega + \mu - \Sigma_{ii\sigma}(\mathrm{i}\omega) - \varepsilon}. \tag{7.5}$$

Here $\mathrm{i}\omega$ is the imaginary frequency that the Green's functions as well as the self-energy depend on and $\rho(\varepsilon)$ is the free density of states at energy ε. This equation together with (7.4) and the definition of Green's function

$$G_{ii\sigma}(\mathrm{i}\omega) = \langle a_{i\sigma}(\mathrm{i}\omega) a^\dagger_{i\sigma}(\mathrm{i}\omega)\rangle \tag{7.6}$$

form a trio of equations to solve self-consistently to find the (free) Green's function and the self-energy. The task is simplified by noting that the equations look identical to the equations obtained when solving the Anderson impurity model [152].

While the results of dynamical mean field theory gives good results in many areas [50, 148–151], it is very reliant on numerical calculations due to the complexity of the aforementioned equations. In the rest of this chapter, we try to approach the problem from a more information-theoretic perspective, trying out different ansätze to minimize the energy in hopes of getting more insight into the structure of the ground state.

For our research we collaborated with Andreas Gleis, a PhD student in the von Delft group at LMU. He provided us DMFT numerical results. His results are summarized in table 7.1 in section 7.6.

7.3 No Interactions

Setting $U = 0$ is a significant simplification of the Hubbard model. Removing the interaction effectively makes the fermions free [130]. It makes the Hamiltonian look like

$$H = \frac{t}{\sqrt{2d}} \sum_{<i,j>} \left(a^\dagger_{i,1} a_{j,1} + a^\dagger_{i,2} a_{j,2} \right) + \mu \sum_i \left(a^\dagger_{i,1} a_{i,1} + a^\dagger_{i,2} a_{i,2} \right). \tag{7.7}$$

Now we can apply the Fourier transform to momentum space (for $\sigma \in \{1, 2\}$),

$$a_{i,\sigma} = \sum_j b_{j,\sigma} \frac{\mathrm{e}^{2\pi \mathrm{i} \frac{j \cdot i}{N}}}{\sqrt{N^d}}. \tag{7.8}$$

Here $b_{j,\sigma}$ is the annihilation operator corresponding to momentum labelled by $j \in \mathcal{A}$ and $j \cdot i$ is a scalar product of i and j, which are both elements of $\mathcal{A}$ — real d-dimensional vectors of integers with values between 1 and N (included)[1].

The Fourier transform keeps the chemical-potential term invariant, since

$$\sum_i a^\dagger_{i,\sigma} a_{i,\sigma} \to \sum_i \sum_j b^\dagger_{j,\sigma} \frac{\mathrm{e}^{-2\pi \mathrm{i} \frac{j \cdot i}{N}}}{\sqrt{N^d}} \sum_k b_{k,\sigma} \frac{\mathrm{e}^{2\pi \mathrm{i} \frac{k \cdot i}{N}}}{\sqrt{N^d}} \tag{7.9}$$

$$= \sum_{i,j,k} b^\dagger_{j,\sigma} b_{k,\sigma} \frac{\mathrm{e}^{2\pi \mathrm{i} \frac{(k-j) \cdot i}{N}}}{N^d} \tag{7.10}$$

$$= \sum_{j,k} b^\dagger_{j,\sigma} b_{k,\sigma} \delta_{j,k} = \sum_j b^\dagger_{j,\sigma} b_{j,\sigma}. \tag{7.11}$$

[1] To clarify notation, in this chapter we do not use ancillary Hilbert spaces and purifications. Therefore we can use $b_{i,\sigma}$ and $b^\dagger_{i,\sigma}$ for the annihilation and creation operators in momentum space, even though these symbols are reserved for ancillary fermions in chapter 5.

Before we look at how the hopping term transforms, we slightly adapt our notation. Let e_x be a unit vector pointing in the direction of the x-th dimension (i.e. the vector has all components equal to 0 except for the x-th component which is equal to 1). With this, the sum over all ordered pairs of nearest neighbours can be rephrased as

$$\sum_{<i,j>} a^\dagger_{i,\sigma} a_{j,\sigma} = \sum_i \sum_{x=1}^{d} \left(a^\dagger_{i,\sigma} a_{i+e_x,\sigma} + a^\dagger_{i,\sigma} a_{i-e_x,\sigma} \right). \tag{7.12}$$

This simplifies the Fourier transform

$$\begin{aligned}
\sum_{<i,j>} a^\dagger_{i,\sigma} a_{j,\sigma} &\to \sum_{i,j,k} \sum_{x=1}^{d} b^\dagger_{j,\sigma} b_{k,\sigma} \left(\frac{\mathrm{e}^{2\pi \mathrm{i} \frac{i\cdot(k-j)+e_x\cdot k}{N}}}{N^d} + \frac{\mathrm{e}^{2\pi \mathrm{i} \frac{i\cdot(k-j)-e_x\cdot k}{N}}}{N^d} \right) && (7.13)\\
&= \sum_{j,k} \sum_{x=1}^{d} b^\dagger_{j,\sigma} b_{k,\sigma} \left(\delta_{k,j} \mathrm{e}^{2\pi \mathrm{i} \frac{e_x\cdot k}{N}} + \delta_{k,j} \mathrm{e}^{-2\pi \mathrm{i} \frac{e_x\cdot k}{N}} \right) && (7.14)\\
&= \sum_{k} \sum_{x=1}^{d} b^\dagger_{k,\sigma} b_{k,\sigma} \left(\mathrm{e}^{2\pi \mathrm{i} \frac{e_x\cdot k}{N}} + \mathrm{e}^{-2\pi \mathrm{i} \frac{e_x\cdot k}{N}} \right) && (7.15)\\
&= \sum_{k} b^\dagger_{k,\sigma} b_{k,\sigma} \sum_{x=1}^{d} 2\cos\left(2\pi \frac{k_x}{N} \right). && (7.16)
\end{aligned}$$

Here k_x is the x-th component of the vector k.

Putting these two together, the Hamiltonian reduces to

$$H = \sum_{k,\sigma} b^\dagger_{k,\sigma} b_{k,\sigma} \left[\mu + \frac{\sqrt{2}}{\sqrt{d}} t \sum_{x=1}^{d} \cos\left(2\pi \frac{k_x}{N} \right) \right] \equiv \sum_{k,\sigma} E(k) b^\dagger_{k,\sigma} b_{k,\sigma}. \tag{7.17}$$

Here we introduced $E(k)$ as the energy of the k-th momentum mode. This Hamiltonian is diagonal in the momentum labels k, making it very easy to find the ground state. The ground state is obtained by simply filling the *Fermi sea*, i.e. by starting with the vacuum state $|\Omega\rangle$ and adding all of the momentum modes which have negative energy, i.e.

$$|\Psi_0\rangle = \prod_{k\in\mathcal{K}} \prod_{\sigma} b^\dagger_{k,\sigma} |\Omega\rangle . \tag{7.18}$$

These states may be described as *momentum Fock states*. Here the set $\mathcal{K}$ is defined as the set of all possible k which fulfill

$$E(k) = \sum_{x=1}^{d} \frac{\sqrt{2}}{\sqrt{d}} t \cos\left(2\pi \frac{k_x}{N} \right) + \mu \leq 0. \tag{7.19}$$

This is a non-linear condition binding the various components of k together. That makes it difficult to describe the set $\mathcal{K}$ in a more elegant way. However, that is not needed. The ground state energy E_0 is equal to

$$E_0 = \sum_{k\in\mathcal{K}} \left[\sum_{x=1}^{d} \frac{\sqrt{2}}{\sqrt{2d}} t \cos\left(2\pi \frac{k_x}{N} \right) + \mu \right] = \sum_{k\in\mathcal{K}} E(k). \tag{7.20}$$

Fortunately, we are able to evaluate the formula (7.20) (at least in the $d \to +\infty$ limit) without having to describe the set $\mathcal{K}$ for a finite d any better than in equation (7.19). In order to do that, we first take a step back. For all the possible $k \in \mathcal{A}$, the values of k_x are distributed uniformly from 1 to N for all x. For a given x, the distribution of $\cos\left(2\pi\frac{k_x}{N}\right)$ is not uniform and it averages out to 0 with a variance of 1/2.

Now consider the sum $\sum_{x=1}^{d} \frac{\sqrt{2}}{\sqrt{d}} \cos\left(2\pi\frac{k_x}{N}\right)$. By the central limit theorem, as $d \to +\infty$, the distribution of this variable will approach the Gaussian distribution with mean 0 and variance 1. Multiplying it by t changes its variance to t^2 while adding μ changes the mean.

This means that in $d \to +\infty$, the energies of the modes $E(k)$ are distributed according to Gaussian distribution with mean μ and variance t^2. Equation (7.20) is then simply a sum of all of those energies which are negative. It can therefore be rephrased (in the $d \to +\infty$ limit) as

$$E_0 = \int_{-\infty}^{0} \mathrm{d}E(k) \left[E(k) \, \frac{1}{t\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{E(k)-\mu}{t}\right)^2} \right] \tag{7.21}$$

$$= -t\mathrm{e}^{-\frac{\mu^2}{2t^2}} + \frac{1}{2}\left[\mathrm{erf}\left(\frac{-\mu}{t\sqrt{2}}\right) + 1 \right]. \tag{7.22}$$

Interestingly, this does not depend on N. Therefore taking the $d \to +\infty$ limit first trivializes the $N \to +\infty$ limit.

7.4 FSBS as Lower Bound

One way to lower-bound the ground state energy is by using *fermionic symmetric basic states* (FSBS) introduced by Kraus et al. [143]. This family of states was designed as an ansatz for solving permutationally-invariant problems and we used it for the Hubbard model with permutation symmetry in chapter 6. In appendix B of [143], Kraus et al. demonstrated how FSBS may be used to produce a lower bound on the energy of infinite-dimensional cubic models. We follow that approach in this section.

Before we start, we need to define a few new terms. We will split the Hamiltonian H into a sum of terms h_{ij} which act on two neighbouring sites i and j, such that

$$H = \sum_{(i,j)} h_{ij}. \tag{7.23}$$

Here the sum is no longer over *ordered* pairs of sites, but *unordered* pairs of sites so the sum does not contain both h_{ij} and h_{ji}. This two-body Hamiltonian h_{ij} looks like

$$h_{ij} = \frac{t}{\sqrt{2d}} \left(a_{i,1}^\dagger a_{j,1} + a_{i,2}^\dagger a_{j,2} + \mathrm{h.c.} \right) \tag{7.24}$$

$$+ \frac{\mu}{2d} \left(a_{i,1}^\dagger a_{i,1} + a_{i,2}^\dagger a_{i,2} + a_{j,1}^\dagger a_{j,1} + a_{j,2}^\dagger a_{j,2} \right) \tag{7.25}$$

$$+ \frac{U}{2d} \left(a_{i,1}^\dagger a_{i,1} a_{i,2}^\dagger a_{i,2} + a_{j,1}^\dagger a_{j,1} a_{j,2}^\dagger a_{j,2} \right). \tag{7.26}$$

The single-site operators get a factor of $\frac{1}{2d}$ because they appear in $2d$ nearest-neighbour pairs.

Let $\mathcal{N}_i$ be the set of all sites which are the nearest neighbours of site i. With this, the sum in equation (7.23) can be rewritten as

$$H = \sum_i \sum_{j \in \mathcal{N}_i} \frac{h_{ij}}{2} \equiv \sum_i h_i, \tag{7.27}$$

where we promptly defined

$$h_i = \sum_{j \in \mathcal{N}_i} \frac{h_{ij}}{2}. \tag{7.28}$$

The factor $\frac{1}{2}$ in equation (7.27) exists precisely to cancel double counting.

Let $\mathcal{S}_H$ be the set of all states with the same symmetries as the Hamiltonian. The ground state energy E_0 can be expressed as

$$E_0 = \min_{\rho \in \mathcal{S}_H} \mathrm{Tr}\{\rho H\} = N \min_{\rho \in \mathcal{S}_H} \mathrm{Tr}\{\rho h_i\} = N \min_{\sigma_i \in \mathcal{S}'_H} \mathrm{Tr}\{\sigma_i h_i\}. \tag{7.29}$$

Here we defined σ_i to be the reduced density matrix of ρ onto the sites $i \cup \mathcal{N}_i$. Consequently, $\mathcal{S}'_H$ is the set of density matrices which were obtained from matrices in $\mathcal{S}_H$ by reducing them to sites $i \cup \mathcal{N}_i$. This set $\mathcal{S}'_H$ can be relatively difficult to describe. So we relax the restriction at the cost of transforming the equality into an inequality,

$$E_0 \geq N \min_{\sigma_i} \mathrm{Tr}\{\sigma_i h_i\}. \tag{7.30}$$

Now σ_i is allowed to be any density matrix on sites $i \cup \mathcal{N}_i$. When minimizing the energy for Hamiltonian h_i, we can take advantage of its symmetries. As can be seen from the defining equation (7.28), the Hamiltonian h_i is invariant with respect to any permutation of the sites in $\mathcal{N}_i$. We can therefore assume the same about σ_i. A symmetric density matrix has a symmetric purification, so we can write

$$E_0 \geq N \min_{|\Psi\rangle \in \mathcal{H}'} \langle \Psi | \, h_i \, | \Psi \rangle \,. \tag{7.31}$$

Here $\mathcal{H}'$ is the Hilbert space of all pure states on sites $i \cup \mathcal{N}_i$ with 4 fermionic species per site (two of which are physical and two of which are ancillary), which are invariant with respect to any permutation of the sites in $\mathcal{N}_i$. This set of states is easier to characterize and we will use FSBS as the basis of the $\mathcal{N}_i$ subsystem. For the purpose of this section we define

$$A_k = \sum_{j \in \mathcal{N}_i} \frac{a_{j,k}}{\sqrt{2d}}. \tag{7.32}$$

With this the Hamiltonian h_i can be rewritten as

$$h_i = \frac{t}{2} \left(a^\dagger_{i,1} A_1 + a^\dagger_{i,2} A_2 + \mathrm{h.c.} \right) \tag{7.33}$$

$$+ \frac{\mu}{2} \left(a^\dagger_{i,1} a_{i,1} + a^\dagger_{i,2} a_{i,2} \right) + \frac{\mu}{4d} \sum_{j \in \mathcal{N}_i} \left(a^\dagger_{j,1} a_{j,1} + a^\dagger_{j,2} a_{j,2} \right) \tag{7.34}$$

$$+ \frac{U}{2} \left(a^\dagger_{i,1} a_{i,1} a^\dagger_{i,2} a_{i,2} \right) + \frac{U}{4d} \sum_{j \in \mathcal{N}_i} \left(a^\dagger_{j,1} a_{j,1} a^\dagger_{j,2} a_{j,2} \right). \tag{7.35}$$

We want to minimize the energy of h_i with respect to states in $\mathcal{H}'$. Those states will be a combination of FSBS on sites $\mathcal{N}_i$ with anything on the site i. That is still somewhat complicated to solve by hand, but it is really simple to search numerically. The results are contained in table 7.1 in section 7.6.

7.5 Ground State Ansätze

In the following few sections, we go over a few possible ansätze for the ground state of the Hubbard Hamiltonian in infinite dimensions, namely momentum Fock states and Gaussian states. From DMFT we know accurately the ground state energy, so we will investigate how close to it do we get within our ground state ansätze.

As the DMFT results were calculated for a fixed $t = -1$ and $\mu = 0$ (with a few different values of U) we will use the same configuration of the parameters U, t and μ when we apply our ground state ansätze.

7.5.1 Momentum Fock States

In this subsection we attempt to minimize the energy of the infinite-dimensional Hubbard model within the space of momentum Fock states. These states are the exact ground states in the case of $U = 0$, as was demonstrated in section 7.3. It is natural to ask how close will those states get to the ground state energy for $U \neq 0$.

We start by transforming the Hubbard Hamiltonian (with $\mu = 0$) into momentum space, using the Fourier transform described in (7.8),

$$H = \frac{t}{\sqrt{2d}} \sum_{k,\sigma} \sum_{x=1}^{d} 2\cos\left(2\pi \frac{k_x}{N}\right) b^\dagger_{k,\sigma} b_{k,\sigma} + \frac{U}{N^2} \sum_{\substack{k,l,m,n \\ \text{such that} \\ k-l+m-n=0}} b^\dagger_{k,1} b_{l,1} b^\dagger_{m,2} b_{n,2}, \tag{7.36}$$

where the sum over σ runs over $\sigma \in \{1, 2\}$.

Generally speaking, the momentum Fock states have the following form:

$$|\Psi\rangle = \prod_{k \in \mathcal{K}_1} b^\dagger_{k,1} \prod_{l \in \mathcal{K}_2} b^\dagger_{l,2} |\Omega\rangle \tag{7.37}$$

Here $\mathcal{K}_1 \subset \mathcal{A}$ and $\mathcal{K}_2 \subset \mathcal{A}$ are two sets of momentum labels.

Looking at the Hamiltonian (7.36), we can expect that the energy will contain terms like $\langle\Psi| b^\dagger_{k,1} b_{l,1} b^\dagger_{m,2} b_{n,2} |\Psi\rangle$. However, keeping in mind the ansatz (7.37), those terms will be equal to 0 unless $k = l$ and $m = n$. This allows us to simplify the Hamiltonian

$$H = \frac{t}{\sqrt{2d}} \sum_{k,\sigma} \sum_{x=1}^{d} 2\cos\left(2\pi \frac{k_x}{N}\right) b^\dagger_{k,\sigma} b_{k,\sigma} + \frac{U}{N^d} \sum_{k,m} b^\dagger_{k,1} b_{k,1} b^\dagger_{m,2} b_{m,2}. \tag{7.38}$$

As in section 7.3, we define the hopping energy corresponding to a momentum label k as

$$E(k) = t \sum_{i=1}^{d} \frac{\sqrt{2}\cos\left(2\pi \frac{k_i}{a}\right)}{\sqrt{d}}. \tag{7.39}$$

In section 7.3 we showed that as $d \to +\infty$, the distribution of $E(k)$ approaches the normal distribution, by the central limit theorem. Plugging that into the Hamiltonian gives us:

$$H = \sum_{k,\sigma} E(k) b^\dagger_{k,\sigma} b_{k,\sigma} + \frac{U}{N^d} \sum_l b^\dagger_{l,1} b_{l,1} \sum_m b^\dagger_{m,2} b_{m,2} \tag{7.40}$$

As we have shown in section 7.3, in the non-interacting case ($U = 0$), the ground state is easy to find — we just fill in the Fermi sea. However, with the interaction, we should be more careful about which fermionic modes we add to minimize the energy. We can see that the interaction term in equation (7.40) does not depend on momentum — it depends only on the number of particles of both of the fermionic species. When filling the fermionic modes with particles, their momentum distribution is not relevant for the interaction term. We simply pick some hopping energy $E(k)$ threshold and fill in all particles with energies lower than that.

In the following we assume $t > 0$. For $t < 0$, the calculations would be nearly the same, the only difference would be a few inverted inequality signs and some cumulative distributions functions replaced the complementary cumulative distribution functions. We define E_1 and E_2 as "energy thresholds" defining which fermionic modes of each species we add. This allows us to characterize the sets

$$\mathcal{K}_1 = \{k : E(k) \leq E_1\} \tag{7.41}$$

and

$$\mathcal{K}_2 = \{k : E(k) \leq E_2\}. \tag{7.42}$$

With such defined sets $\mathcal{K}_1$ and $\mathcal{K}_2$, the energy of the state $|\Psi\rangle$ from equation (7.37) can be re-expressed as

$$\langle\Psi| H |\Psi\rangle = \sum_{k \in \mathcal{K}_1} E(k) + \sum_{l \in \mathcal{K}_2} E(l) + \frac{U}{N^d} \sum_{k \in \mathcal{K}_1} \sum_{l \in \mathcal{K}_2} 1. \tag{7.43}$$

This is relatively easy to analyze. $\sum_{k \in \mathcal{K}_1} \sum_{l \in \mathcal{K}_2} 1$ sums up to $|\mathcal{K}_1||\mathcal{K}_2|$, i.e. the product of the cardinalities of the two sets. Given that $E(k)$ is distributed like a normal distribution, the cardinalities of $\mathcal{K}_1$ and $\mathcal{K}_2$ are simply

$$|\mathcal{K}_1| = N^d \mathrm{cdf}(E_1) \qquad \text{and} \qquad |\mathcal{K}_2| = N^d \mathrm{cdf}(E_2). \tag{7.44}$$

Here $\mathrm{cdf}(\cdot)$ is the cumulative distribution function of the normal distribution, i.e. the proportion of momentum labels k which are in the corresponding set $\mathcal{K}_1$ or $\mathcal{K}_2$. N^d is the total number of momentum labels k, so multiplied by the proportion it gives the number of momentum labels in the sets $\mathcal{K}_1$ and $\mathcal{K}_2$. The term

$$\sum_{k \in \mathcal{K}_1} E(k) \tag{7.45}$$

is the sum of the value $E(k)$ for all elements in the set $\mathcal{K}_1$. An alternative way to get this sum is to take the number of elements in the set $\mathcal{K}_1$ and multiply it by

the average value of $E(k)$ across this set (i.e. its expected value, in the language of probability distributions). Therefore

$$\sum_{k\in\mathcal{K}_1} E(k) = \langle E(k)\rangle_{\mathcal{K}_1} N^d \operatorname{cdf}(E_1) = t\frac{-\operatorname{pdf}(E_1)}{\operatorname{cdf}(E_1)} N^d \operatorname{cdf}(E_1) = -tN^d \operatorname{pdf}(E_1), \tag{7.46}$$

where $\operatorname{pdf}(\cdot)$ is the probability density function of the normal distribution and the formula for $\langle E_k\rangle_{\mathcal{K}_A}$ is a standard formula for *truncated normal distribution* [153].

Taken all together, the energy of a momentum Fock state per site is

$$\frac{\langle\Psi| H |\Psi\rangle}{N^d} = -t\left[\operatorname{pdf}(E_1) + \operatorname{pdf}(E_1)\right] + U\operatorname{cdf}(E_1)\operatorname{cdf}(E_1). \tag{7.47}$$

To find the lowest-energy state, we just need to minimize this with respect to E_1 and E_1. For the results, see section 7.6.

7.5.2 Gaussian States

One idea on how to improve the upper bound is to use Gaussian states instead of Fock states (Fock states being a special case of Gaussian states). Here we will use Gaussian states with the same symmetries as the Hamiltonian. Since this research was done before we found about the connections between purifications and spontaneous symmetry breaking described in chapter 5, we did not consider ground state purifications here. As we find out later in the section, Gaussian states with the same symmetries as the Hamiltonian are not a good variational family for this problem, because the imaginary time evolution (restricted by the symmetries) does not evolve anywhere.

As was described in section 3.1, one way to define Gaussian states is by their covariance matrix, Γ. To define Γ, we first need to switch from creation / annihilation operators to Majorana fermions. As we have two different species of physical fermions, we will have four different species of Majorana fermions

$$a^\dagger_{i,1} = \frac{q_i - \mathrm{i}p_i}{\sqrt{2}}, \tag{7.48}$$

$$a_{i,1} = \frac{q_i + \mathrm{i}p_i}{\sqrt{2}}, \tag{7.49}$$

$$a^\dagger_{i,2} = \frac{Q_i - \mathrm{i}P_i}{\sqrt{2}}, \tag{7.50}$$

$$a_{i,2} = \frac{Q_i + \mathrm{i}P_i}{\sqrt{2}}. \tag{7.51}$$

The covariance matrix is defined using the Majorana fermions

$$\Gamma_{xy} = -2i\langle\psi| xy |\psi\rangle + \mathrm{i}\delta_{x,y}. \tag{7.52}$$

Here x, y stands for any of $\{p_i, q_j, P_k, Q_l\}$. These label all of the Majorana fermionic indices. The letter p, q, P, Q labels the species of the Majorana fermion and the sub-index labels the lattice site on which this Majorana fermion lives.

The Hubbard Hamiltonian has several symmetries which translate to the symmetries of the covariance matrix as

- Translational symmetry, i.e. $\Gamma_{p_i q_j} = \Gamma_{p_{i+k} q_{j+k}}$.
- Mirror symmetry, i.e. $\Gamma_{p_i q_j} = \Gamma_{p_{-i} q_{-j}}$.
- Spin flip symmetry (special case of **SU(2)**), i.e. $\Gamma_{p_i q_j} = \Gamma_{P_i Q_j}$.

This severely restricts the degrees of freedom of the covariance matrix, for example

$$\Gamma_{p_i q_j} = \Gamma_{p_j q_i} = \Gamma_{P_j Q_i}, \tag{7.53}$$
$$\Gamma_{p_i p_j} = \Gamma_{p_j p_i} = -\Gamma_{p_i p_j} = 0, \tag{7.54}$$
$$\Gamma_{p_i P_j} = \Gamma_{P_i p_j} = -\Gamma_{p_j P_i} = -\Gamma_{p_i P_j} = 0. \tag{7.55}$$

In equation (7.54) we first used a combination of rotation / translation symmetry to switch i and j (which is always possible) and then the anti-symmetry of Γ. In equation (7.55) we used the same together with the spin flip symmetry.

Equation (3.10) gives the formula for the imaginary time evolution of Γ:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\Gamma = -4\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma} + \Gamma\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\Gamma\right), \tag{7.56}$$

where E is the energy of the state, i.e. $\langle\psi| H |\psi\rangle$. $\frac{\mathrm{d}E}{\mathrm{d}\Gamma}$ is a matrix where each component corresponds to the derivative of the energy E with respect to the corresponding component of Γ. To calculate that, we first need to obtain the formula for the energy in terms of the components of Γ.

The Hubbard Hamiltonian in position space looks like

$$H = \frac{t}{\sqrt{2d}} \sum_{<i,j>} \left(a^\dagger_{i,1} a_{j,1} + a^\dagger_{i,2} a_{j,2}\right) + U \sum_i a^\dagger_{i,1} a_{i,1} a^\dagger_{i,2} a_{i,2}. \tag{7.57}$$

From the definition of the Majorana fermions in equations (7.48)–(7.51) we get

$$a^\dagger_{i,1} a_{j,1} = \frac{(q_i - \mathrm{i}p_i)(q_j + \mathrm{i}p_j)}{2} = \frac{q_i q_j + \mathrm{i}q_i p_j - \mathrm{i}p_i q_j + p_i p_j}{2}, \tag{7.58}$$
$$a^\dagger_{i,2} a_{j,2} = \frac{(Q_i - \mathrm{i}P_i)(Q_j + \mathrm{i}P_j)}{2} = \frac{Q_i Q_j + \mathrm{i}Q_i P_j - \mathrm{i}P_i Q_j + P_i P_j}{2}. \tag{7.59}$$

We use this to express the Hamiltonian in terms of the Majorana fermions,

$$H = \frac{t}{\sqrt{2d}} \sum_{<i,j>} (\mathrm{i}q_i p_j + \mathrm{i}Q_i P_j) + U \sum_i \left(\frac{1}{2} + \mathrm{i}q_i p_i\right)\left(\frac{1}{2} + \mathrm{i}Q_i P_i\right). \tag{7.60}$$

The expected value of (7.60) gives us the definition of various elements of the covariance matrix Γ (together with using Wick's theorem to split the 4-point expectation value into products of Γ's):

$$E = \frac{U}{4} \sum_i (1 - \Gamma_{q_i p_i} - \Gamma_{Q_i P_i} + \Gamma_{q_i p_i}\Gamma_{Q_i P_i} + \Gamma_{q_i P_i}\Gamma_{p_i Q_i}) \\ - \frac{t}{2\sqrt{2d}} \sum_{<i,j>} (\Gamma_{q_i p_j} + \Gamma_{Q_i P_j}). \tag{7.61}$$

This allows us to get the components of $\frac{\mathrm{d}E}{\mathrm{d}\Gamma}$:

$$\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_iq_i} = \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{P_iQ_i} = \frac{U}{2}(1+\Gamma_{p_iq_i}), \tag{7.62}$$

$$\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_iQ_i} = \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{P_iq_i} = -\frac{U}{2}\Gamma_{p_iQ_i}, \tag{7.63}$$

$$\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_iq_j} = \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{P_iQ_j} = -\frac{t}{\sqrt{2d}}. \tag{7.64}$$

The matrix $\frac{\mathrm{d}E}{\mathrm{d}\Gamma}$ inherits the symmetries from Γ, i.e.

- Translation symmetry: $\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_iq_j} = \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_{i+k}q_{j+k}}$.
- Mirror symmetry: $\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_iq_j} = \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_{-i}q_{-j}}$.
- Spin flip symmetry: $\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{p_iq_j} = \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{P_iQ_j}$.

If we show that Γ and $\frac{\mathrm{d}E}{\mathrm{d}\Gamma}$ commute, that implies that the right-hand side of (7.56) is identically 0, implying that the imaginary time evolution does not evolve the state at all. That is exactly what we are going to do here in order to show that symmetric Gaussian states are not a good ansatz. For $\alpha, \beta \in \{p, q, P, Q\}$ we look at

$$\left(\Gamma\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\alpha_i\beta_j} = \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_i\gamma_k}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\beta_j}. \tag{7.65}$$

We define an operation tilde $\tilde{}$ that changes between the species of the Majorana fermions as $\tilde{p} = P, \tilde{P} = p, \tilde{q} = Q, \tilde{Q} = q$. From (7.65) we can already see that either $\alpha = \beta$, $\alpha = \tilde{\beta}$ or the expression is identically zero (because then either $\Gamma_{\alpha_i\gamma_k}$ or $\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\beta_j}$ are zero). First assume $\alpha = \beta$. We perform a long chain of steps based

on the symmetries of Γ and $\frac{\mathrm{d}E}{\mathrm{d}\Gamma}$:

$$\left(\Gamma\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\alpha_i\alpha_j} = \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_i\gamma_k}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\alpha_j} \tag{7.66}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_0\gamma_{k-i}}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\alpha_j} \tag{7.67}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_0\gamma_k}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_{k+i}\alpha_j} \tag{7.68}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_0\gamma_k}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\alpha_{j-i}} \tag{7.69}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_0\gamma_{-k}}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_{-k}\alpha_{i-j}} \tag{7.70}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_0\gamma_k}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\alpha_{i-j}} \tag{7.71}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_0\gamma_k}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_{k+j}\alpha_i} \tag{7.72}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_0\gamma_{k-j}}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\alpha_i} \tag{7.73}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \Gamma_{\alpha_j\gamma_k}\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\gamma_k\alpha_i} \tag{7.74}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\alpha_i\gamma_k}\Gamma_{\gamma_k\alpha_j} \tag{7.75}$$

$$= \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\Gamma\right)_{\alpha_i\alpha_j}. \tag{7.76}$$

For the case of $\alpha = \tilde{\beta}$ most of the steps are the same, but in the end there's a

few extra steps,

$$\left(\Gamma\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\alpha_i\tilde{\alpha}_j} = ... \tag{7.77}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\tilde{\alpha}_i\gamma_k}\Gamma_{\gamma_k\alpha_j} \tag{7.78}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\alpha_i\tilde{\gamma}_k}\Gamma_{\gamma_k\alpha_j} \tag{7.79}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\alpha_i\gamma_k}\Gamma_{\tilde{\gamma}_k\alpha_j} \tag{7.80}$$

$$= \sum_{\gamma\in\{p,q,P,Q\}}\sum_k \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\right)_{\alpha_i\gamma_k}\Gamma_{\gamma_k\tilde{\alpha}_j} \tag{7.81}$$

$$= \left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\Gamma\right)_{\alpha_i\tilde{\alpha}_j}. \tag{7.82}$$

This means that the imaginary time derivative of the covariance matrix vanishes

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\Gamma = -4\left(\frac{\mathrm{d}E}{\mathrm{d}\Gamma} + \Gamma\frac{\mathrm{d}E}{\mathrm{d}\Gamma}\Gamma\right) = 0 \tag{7.83}$$

This implies that within the manifold of symmetric Gaussian states, the imaginary time evolution does not work at all. While the reason for this is not entirely understood, we believe that the symmetries constrain the class of Gaussian states too strongly, making it impossible to evolve the system while preserving all of the symmetries and the Gaussianity at the same time.

7.6 Comparison

In this section we compare the numerical results obtained from the methods described in sections 7.2, 7.4 and 7.5.1. The results are presented in table 7.1 and plotted in figure 7.1. For all of the methods, we compare the energy per site, i.e. for the FSBS lower bound, it is the expectation value of h_i (7.33) from equation (7.31) and for the momentum Fock states it is the expectation value of $\frac{\langle\Psi|H|\Psi\rangle}{N^d}$ from equation (7.47). For all methods, the chosen parameters were $t = -1, \mu = 0$ and U increasing from 0 to 3 in increments of 0.5. As the parameteres of the Hamiltonian are dimensionless, so is the energy.

The DMFT results were provided by Andreas Gleis. They contain not only the energy per site E, but also the kinetic contribution to the energy E_t, the interaction contribution E_U and the expected number of fermions per site n. Comparing the DMFT results with the FSBS lower bound shows that the lower bound is not very tight. On the other hand, the momentum Fock state ansatz provides good agreement with the DMFT numerics, at least for low values of U. Note that the energy value for $U = 0$ can be found analytically to be $-\sqrt{2/\pi}$.

	DMFT				FSBS	Mom. Fock
U	E_t	E_U	E	n	E	E
0	−0.7979	0	−0.7979	1	−1	−0.797885
0.5	−0.7759	0.0768	−0.6991	0.8376	−0.944425	−0.693609
1	−0.7350	0.0974	−0.6377	0.7348	−0.899802	−0.618636
1.5	−0.6949	0.0970	−0.5979	0.6703	−0.862956	−0.561657
2	−0.6610	0.0901	−0.5709	0.6286	−0.831819	−0.516552
2.5	−0.6337	0.0819	−0.5517	0.6006	−0.805006	−0.479738
3	−0.6118	0.0743	−0.5375	0.5809	−0.781560	−0.448977

Table 7.1: The results comparing the energy of infinite-dimensional Hubbard model obtained by the various methods.

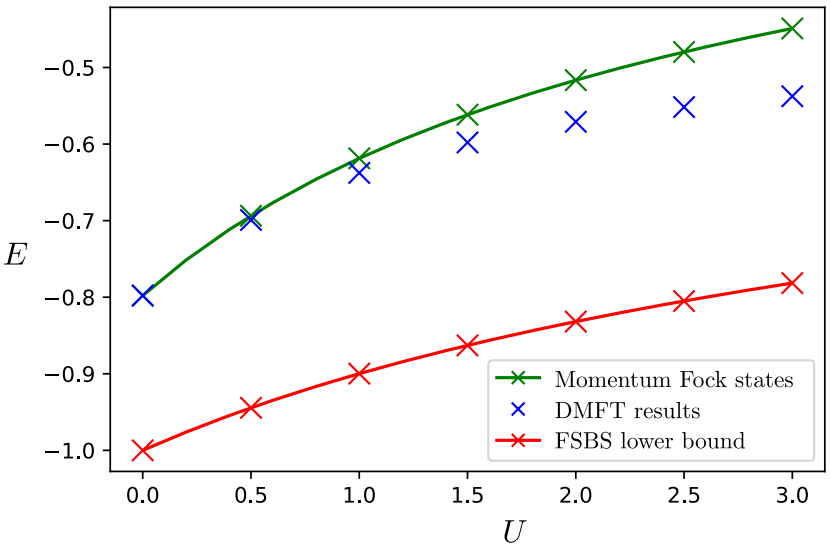


Figure 7.1: A graph depicting the values from table 7.1.

By further analysis of the momentum Fock space ansatz, we found that at $U/t = 3.3252...$ a phase transition occurs. For $U/t < 3.3252...$, the energy thresholds E_1 and E_2 are the same. For $U/t > 3.3252...$ one of them becomes larger while the other one quickly becomes small (spontaneous symmetry breaking). This is simple to interpret physically. As the interaction becomes strong, at one point it becomes energetically beneficial to favour one species over the other one. Unfortunately, further analysis of this behaviour is beyond the scope of this thesis.

Chapter 8

Conclusion

This thesis started by asking the question of how can methods of quantum information theory help us solve difficult problems in quantum many-body physics. The bulk of this thesis challenged two seemingly-unrelated problems: approximability of mixed states by tensor networks and detection of spontaneous symmetry breaking in highly symmetric systems. In solving both of these problems, we used similar approaches — we purified our states and we looked at the entanglement of the purification.

We have found that the scaling of the entanglement of purification may be used as a sufficient condition for a mixed state to have an efficient MPO approximation. This result closed a 15-year long gap since an analogous condition was discovered for pure states. Aside from the theoretical significance of this result, questions remain about its practical usability. The entanglement of purification is not a very popular measure of correlations and it remains open whether similar condition based on mutual information exists. Furthermore, the scaling required from the entanglement of purification is stricter than is necessary in the pure-state case, suggesting potential for improvement. On the other hand, the scaling was obtained by concatenating many single-cut approximations, suggesting that maybe it comes from the fundamental differences between pure states and operators.

We described a new method to approach the problem ground states of symmetric systems. Our approach has two primary uses. First, it allows to restrict the ground state search to purifications with all of the desired symmetries of the Hamiltonian without the possibility of missing any ground states. Second, it gives us a way to detect spontaneous symmetry breaking from the entanglement properties of the purification ground state and even detect which sub-symmetry was broken in case of multiple symmetries. The method cannot be used on its own, so its utility is tied to how well it can synergize with other tools for solving the ground-state problem. We showcased the use of our method by applying it to various systems and demonstrated that it shines when used for permutationally-invariant systems, for which it can be combined with the quantum de Finetti theorem. But there also seems to be potential for combining our method with tensor networks — research into which remains in the future. However, even if the practical use of our method turns out to be limited, the conceptual connection between entanglement, purifications and spontaneous symmetry breaking is of independent interest.

We applied our method to solve the ground state phase diagram of the permutationally-invariant Hubbard model. Even though this is mostly a toy model with very limited applications beyond theoretical studies, it does show some interesting physics. We observe a half-filling phase spontaneously breaking the **SU(2)** symmetry of the model as well as a pairing phase spontaneously breaking the **U(1)** particle-number symmetry. While this model appears completely solved now, it is always possible to construct and solve other, more complicated permutationally-invariant models, e.g. by adding more exotic interaction terms to the Hamiltonian.

Our results have highlighted the deep connections between the fields of quantum information theory and quantum many-body physics. Concepts like state purification, majorization and Rényi entropies find clear applications in relevant models describing real physical systems. Open questions remain, but unanswered questions today are new directions of research tomorrow. Our contribution expands the boundaries of scientific knowledge, but there is still a lot of research to be done on Quantum Information Methods in Many-Body Physics.

Appendix A

Hubbard Model Phase Diagram

This appendix expands the calculations regarding the permutationally-invariant Hubbard model of chapter 6. In here we will go over all the areas of the phase diagram of the model. For each area, we will find the lowest-energy state and then determine the phase from the wavefunction. Here we minimize only the "even" part of the state (6.76)

$$|\Psi\rangle = \prod_{i=1}^{N} \left[\alpha_0 \mathbb{1}_i + \alpha_{13} \left(\frac{a_{i,1}^\dagger a_{i,3}^\dagger + a_{i,2}^\dagger a_{i,4}^\dagger}{\sqrt{2}}\right) + \alpha_{1234} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger\right] |\Omega\rangle \,. \tag{A.1}$$

as the "odd" part may be minimized independently

Recall that the expected values of some of the terms of the Hubbard Hamiltonian with the state ansatz are

$$\langle\Psi| \frac{H_U}{N} |\Psi\rangle = U\alpha_{1234}^* \alpha_{1234}, \tag{A.2}$$

$$\langle\Psi| \frac{H'_\mu}{N} |\Psi\rangle = (\mu - t)\left(2\alpha_{1234}^* \alpha_{1234} + \alpha_{13}^* \alpha_{13}\right), \tag{A.3}$$

$$\langle\Psi| \frac{H_V}{N^2} |\Psi\rangle = V\left(2\alpha_{1234}^* \alpha_{1234} + \alpha_{13}^* \alpha_{13}\right)^2. \tag{A.4}$$

To simplify notation, we label $D = \alpha_{1234}^* \alpha_{1234} \geq 0$ and $S = \alpha_{13}^* \alpha_{13} \geq 0$ (for doubly and single-occupied sites, respectively). We also split the phase space into two halves (based on the sign of V) and solve the phase diagram on each individually. For some of the calculations below we use

$$D + S \leq 1, \tag{A.5}$$

which is implied by the normalization of the even part of the wavefunction $|\alpha_0|^2 + |\alpha_{13}|^2 + |\alpha_{1234}|^2 = 1$.

A.1 Repulsive Off-Site Interaction

We begin by considering $V > 0$.

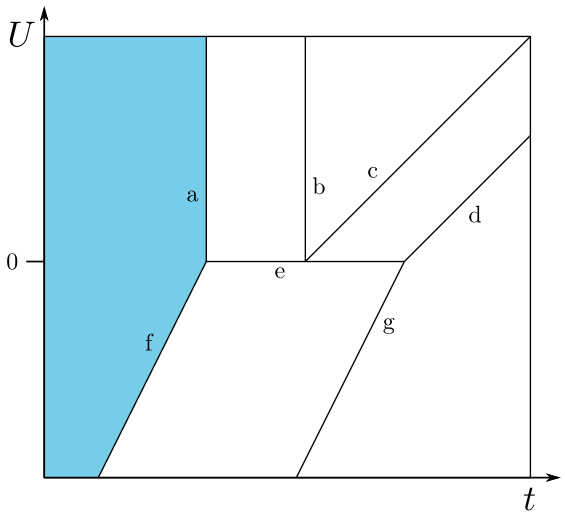


Figure A.1: The area of the phase diagram investigated in subsection A.1.1

A.1.1 Vacuum

We start by looking at the area bounded by the following two inequalities:

- $\mu > t$,
- $U > 2t - 2\mu$.

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\begin{aligned}
&\langle\Psi| H_U/N + H'_\mu/N + H_V/N^2 |\Psi\rangle && \text{(A.6)}\\
&\quad = UD + (\mu - t)(2D + S) + V(2D + S)^2 && \text{(A.7)}\\
&\quad \geq UD + (\mu - t)(2D + S) && \text{(A.8)}\\
&\quad \geq UD + 2(\mu - t)D && \text{(A.9)}\\
&\quad = [U + 2(\mu - t)]\, D && \text{(A.10)}\\
&\quad \geq 0. && \text{(A.11)}
\end{aligned}$$

This lower bound can be saturated by setting $D = S = 0$, which is the case for the state $|\Psi\rangle = |\Omega\rangle$, i.e. the vacuum state.

A.1.2 Full Filling

Now we look at the area bounded by the following two inequalities:

- $t - \mu - 4V > U$,
- $2t - 2\mu - 8V > U$.

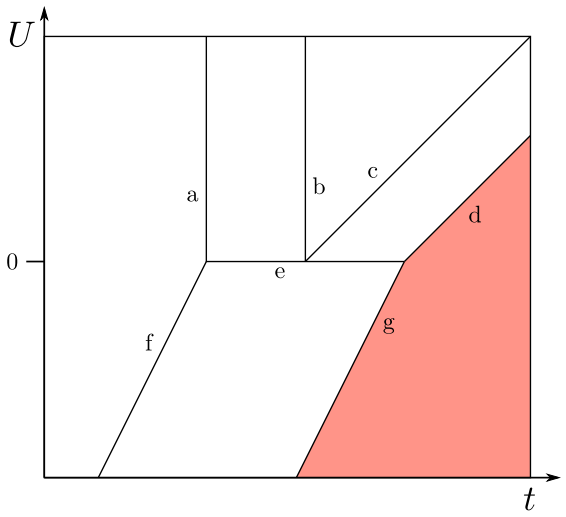


Figure A.2: The area of the phase diagram investigated in subsection A.1.2

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\langle\Psi| H_U/N + H'_\mu/N + H_V/N^2 |\Psi\rangle \tag{A.12}$$
$$= UD + (\mu - t)(2D + S) + V(2D + S)^2 \tag{A.13}$$
$$\geq UD + (\mu - t)(2D + S) + V(2D + S)^2 + S(U + \mu + 4V - t) \tag{A.14}$$
$$\geq D(2t - U - 2\mu - 8V) + (2D + S)(U + 2\mu + 4V - 2t) + V(2D + S)^2 \tag{A.15}$$
$$\geq (D + S)(U + 2\mu + 8V - 2t) + V(2D + S)^2 - 4V(2D + S) \tag{A.16}$$
$$\geq U + 2\mu + 8V - 2t + V(2D + S)^2 - 4V(2D + S) \tag{A.17}$$
$$\geq U + 2\mu + 4V - 2t. \tag{A.18}$$

This lower bound can be saturated by setting $D = 1$, which is the case for the state

$$|\Psi\rangle = \prod_{i=1}^{N} a^\dagger_{i,1} a^\dagger_{i,2} a^\dagger_{i,3} a^\dagger_{i,4} |\Omega\rangle , \tag{A.19}$$

i.e. the full filling state.

A.1.3 Half Filling

Now we look at the area bounded by the following two inequalities:

- $t > \mu + 2V$,
- $U > t - \mu - 2V$.

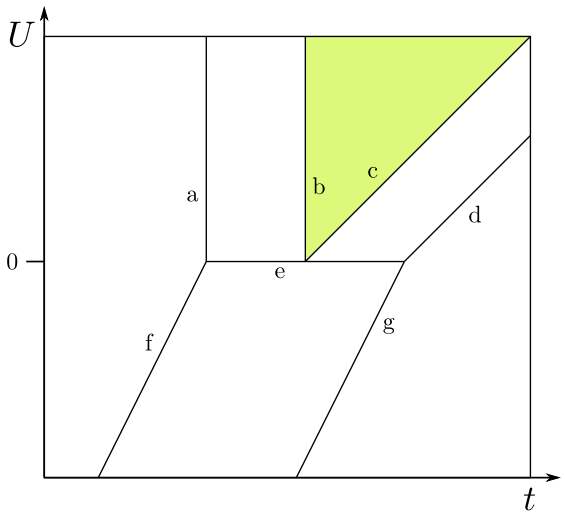


Figure A.3: The area of the phase diagram investigated in subsection A.1.3

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\langle\Psi|H_U/N + H'_\mu/N + H_V/N^2\,|\Psi\rangle \tag{A.20}$$
$$= UD + (\mu - t)\,(2D + S) + V\,(2D + S)^2 \tag{A.21}$$
$$\geq (t - \mu - 2V)D + (\mu - t)\,(2D + S) + V\,(2D + S)^2 \tag{A.22}$$
$$\geq (t - \mu - 2V)(2D + S - 1) + (\mu - t)\,(2D + S) + V\,(2D + S)^2 \tag{A.23}$$
$$\geq 2V + \mu - t - 2V(2D + S) + V\,(2D + S)^2 \tag{A.24}$$
$$\geq V + \mu - t + V\,(2D + S - 1)^2 \tag{A.25}$$
$$\geq V + \mu - t. \tag{A.26}$$

This lower bound can be saturated by setting $S = 1$, which is the case for the state $|\Psi\rangle = \prod_{i=1}^{N}\left(\frac{a^\dagger_{i,1}a^\dagger_{i,3}+a^\dagger_{i,2}a^\dagger_{i,4}}{\sqrt{2}}\right)|\Omega\rangle$, i.e. the half filling state.

A.1.4 Pairing

Now we look at the area bounded by the following three inequalities:

- $2t - 2\mu > U$,
- $U > 2t - 2\mu - 8V$,
- $0 > U$.

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\langle\Psi|\,H_U/N + H'_\mu/N + H_V/N^2\,|\Psi\rangle \tag{A.27}$$
$$= UD + (\mu - t)\,(2D + S) + V\,(2D + S)^2 \tag{A.28}$$
$$\geq \frac{U}{2}(2D + S) + (\mu - t)\,(2D + S) + V\,(2D + S)^2 \tag{A.29}$$
$$\geq -\frac{[U + 2(\mu - t)]^2}{16V}. \tag{A.30}$$

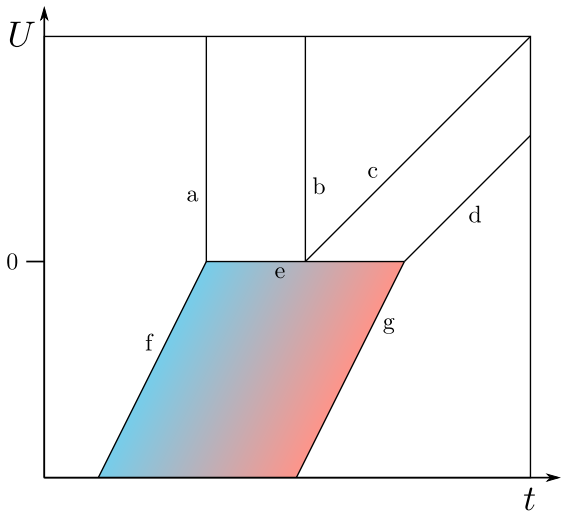


Figure A.4: The area of the phase diagram investigated in subsection A.1.4

Here in last inequality we replace the quadratic expression for $(2D + S)$ with its minimum achieved for $(2D + S) = \frac{2t-2\mu-U}{4V}$. In fact, the lower bound is achieved for $S = 0$ and $D = \frac{2t-2\mu-U}{8V}$. That corresponds to the state

$$|\Psi\rangle = \prod_{i=1}^{N} \left[\alpha_0 \mathbb{1}_i + \alpha_{1234} \left(a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger\right)\right] |\Omega\rangle \tag{A.31}$$

with $\alpha_{1234} = \sqrt{\frac{2t-2\mu-U}{8V}}$ and $\alpha_0 = \sqrt{1 - \alpha_{1234}^2}$. We call this phase *pairing*.

A.1.5 Sub-Half Filling

Now we look at the area delimited by the following three inequalities:

- $U > 0$,
- $\mu + 2V > t$,
- $t > \mu$.

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\begin{aligned}
&\langle\Psi| H_U/N + H'_\mu/N + H_V/N^2 |\Psi\rangle && \text{(A.32)}\\
&\quad = UD + (\mu - t)(2D + S) + V(2D + S)^2 && \text{(A.33)}\\
&\quad \geq (\mu - t)(2D + S) + V(2D + S)^2 && \text{(A.34)}\\
&\quad \geq -\frac{(\mu - t)^2}{4V}. && \text{(A.35)}
\end{aligned}$$

Similarly to before, in the last step we replace the quadratic expression for $(2D+S)$ with its minimum. The minimum is achieved for $(2D + S) = \frac{t-\mu}{2V}$, specifically for $D = 0$ and $S = \frac{t-\mu}{2V}$, corresponding to the state

$$|\Psi\rangle = \prod_{i=1}^{N} \left[\alpha_0 \mathbb{1}_i + \alpha_{13} \left(\frac{a_{i,1}^\dagger a_{i,3}^\dagger + a_{i,2}^\dagger a_{i,4}^\dagger}{\sqrt{2}}\right)\right] |\Omega\rangle \tag{A.36}$$

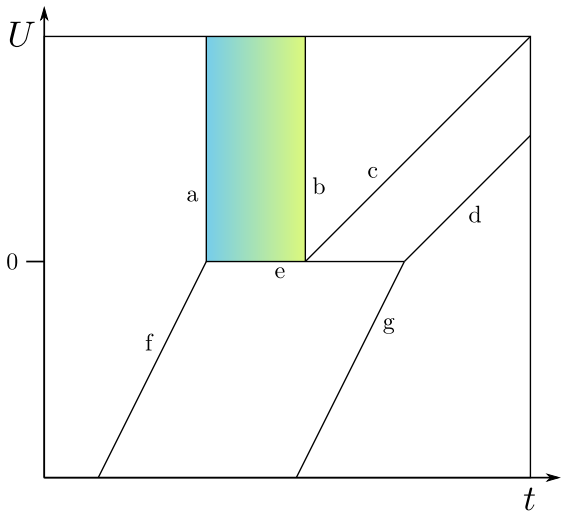


Figure A.5: The area of the phase diagram investigated in subsection A.1.5

with $\alpha_{13} = \sqrt{\frac{t-\mu}{2V}}$ and $\alpha_0 = \sqrt{1-\alpha_{13}^2}$. We call this phase sub-half filling.

A.1.6 Super-Half Filling

Now we look at the area delimited by the following three inequalities:

- $U > 0$,
- $t - \mu - 2V > U$,
- $t - \mu - 4V > U$.

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\begin{aligned}
&\langle\Psi| H_U/N + H'_\mu/N + H_V/N^2 |\Psi\rangle && \text{(A.37)}\\
&\quad = UD + (\mu - t)(2D + S) + V(2D + S)^2 && \text{(A.38)}\\
&\quad \geq U(2D + S - 1) + (\mu - t)(2D + S) + V(2D + S)^2 && \text{(A.39)}\\
&\quad \geq -U - \frac{(\mu - t + U)^2}{4V}. && \text{(A.40)}
\end{aligned}$$

Again, the last step consists of replacing the quadratic expression of $(2D + S)$ with its minimum. The minimum here is achieved for $S = \frac{U+\mu-t+8V}{4V}$ and $D = \frac{t-\mu-U-4V}{4V}$, corresponding to the state

$$\prod_{i=1}^{N} \left[\alpha_{13} \left(\frac{a_{i,1}^\dagger a_{i,3}^\dagger + a_{i,2}^\dagger a_{i,4}^\dagger}{\sqrt{2}}\right) + \alpha_{1234} \left(a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger\right)\right] |\Omega\rangle \qquad \text{(A.41)}$$

with $\alpha_{13} = \sqrt{\frac{U+\mu-t+8V}{4V}}$ and $\alpha_{1234} = \sqrt{\frac{t-\mu-U-4V}{4V}}$. We call this phase super-half filling.

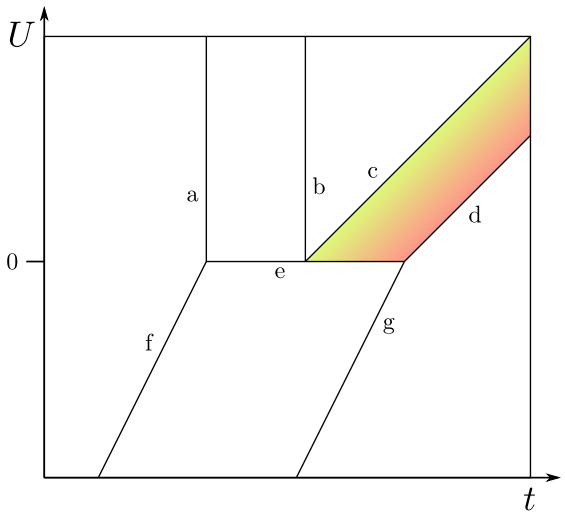


Figure A.6: The area of the phase diagram investigated in subsection A.1.6.

A.2 Attractive Off-Site Interaction

For non-positive V, the phase diagram looks similar to $V > 0$, but some phases are missing.

A.2.1 Vacuum

We start by looking at the area bounded by the following two inequalities:

- $\mu + 2V > t$,
- $U > 2t - 2\mu - 4V$.

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\begin{aligned}
&\langle\Psi| H_U/N + H'_\mu/N + H_V/N^2 |\Psi\rangle && \text{(A.42)}\\
&\quad = UD + (\mu - t)(2D + S) + V(2D + S)^2 && \text{(A.43)}\\
&\quad \geq UD + (\mu - t + 2V)(2D + S) && \text{(A.44)}\\
&\quad \geq UD + 2D(\mu - t + 2V) && \text{(A.45)}\\
&\quad = [U + 2(\mu - t + 2V)]\, D && \text{(A.46)}\\
&\quad \geq 0. && \text{(A.47)}
\end{aligned}$$

This expected value can be achieved by setting $D = S = 0$, which is the case for the state $|\Psi\rangle = |\Omega\rangle$, i.e. the vacuum state.

A.2.2 Half Filling

Now we look at the area bounded by the following two inequalities:

- $t > \mu + 2V$,
- $U > t - \mu - 2V$.

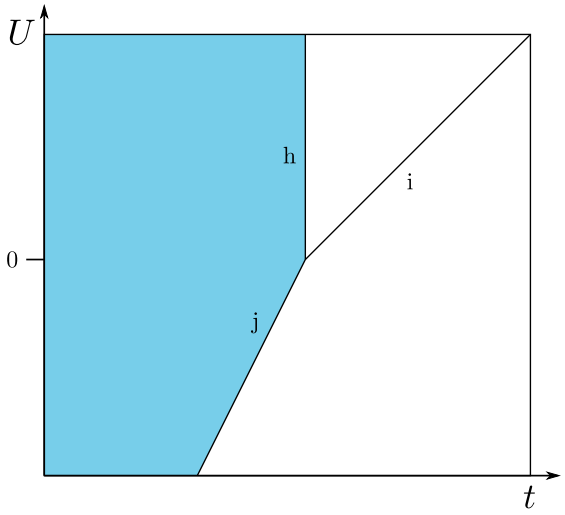


Figure A.7: The area of the phase diagram investigated in subsection A.2.1.

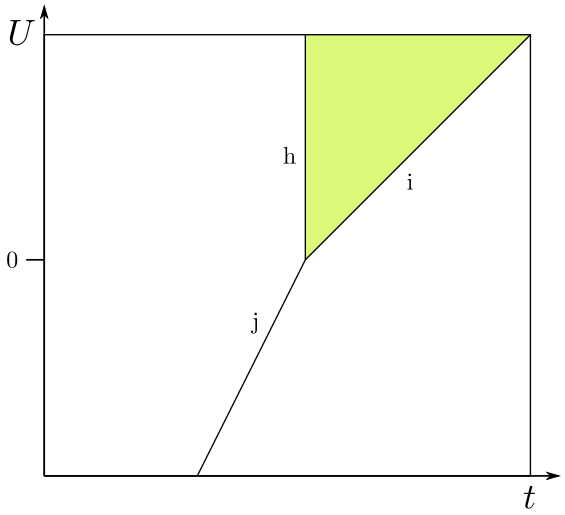


Figure A.8: The area of the phase diagram investigated in subsection A.2.2.

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\begin{aligned}
&\langle\Psi|H_U/N + H'_\mu/N + H_V/N^2\,|\Psi\rangle && \text{(A.48)}\\
&\quad = UD + (\mu - t)\,(2D + S) + V\,(2D + S)^2 && \text{(A.49)}\\
&\quad \geq (t - \mu - 2V)D + (\mu - t)\,(2D + S) + V\,(2D + S)^2 && \text{(A.50)}\\
&\quad \geq (t - \mu - 2V)(2D + S - 1) + (\mu - t)\,(2D + S) + V\,(2D + S)^2 && \text{(A.51)}\\
&\quad \geq 2V + \mu - t - 2V(2D + S) + V\,(2D + S)^2 && \text{(A.52)}\\
&\quad = V + \mu - t + V\,(2D + S - 1)^2 && \text{(A.53)}\\
&\quad \geq V + \mu - t. && \text{(A.54)}
\end{aligned}$$

This minimum is achieved for $S = 1$, which corresponds to the half filling state again (A.19).

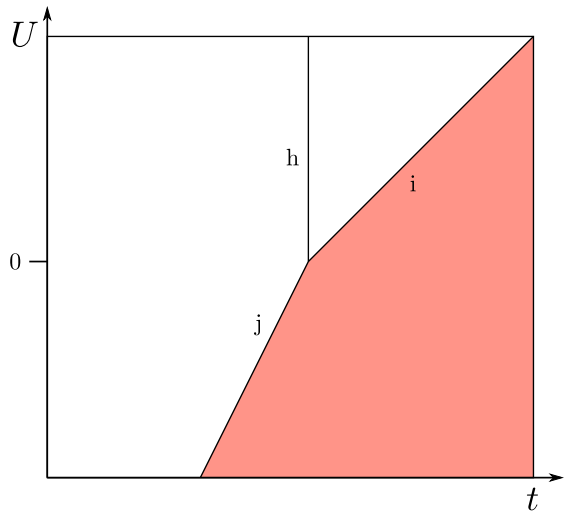


Figure A.9: The area of the phase diagram investigated in subsection A.2.3.

A.2.3 Full filling

Now we look at the area bounded by the following two inequalities:

- $t - \mu - 2V > U$,
- $2t - 2\mu - 4V > U$.

In this area, the expected value of $H_U/N + H'_\mu/N + H_V/N^2$ with the state $|\Psi\rangle$ is

$$\langle\Psi| H_U/N + H'_\mu/N + H_V/N^2 |\Psi\rangle \tag{A.55}$$
$$= UD + (\mu - t)(2D + S) + V(2D + S)^2 \tag{A.56}$$
$$\geq UD + (\mu - t)(2D + S) + V(2D + S)^2 + S(U + \mu + 2V - t) \tag{A.57}$$
$$\geq D(2t - U - 2\mu - 4V) + (U + 2\mu + 2V - 2t)(2D + S) + V(2D + S)^2 \tag{A.58}$$
$$\geq (D + S)(U - 2t + 2\mu + 4V) + -2V(2D + S) + V(2D + S)^2 \tag{A.59}$$
$$\geq U - 2t + 2\mu + 4V + -2V(2D + S) + V(2D + S)^2 \tag{A.60}$$
$$\geq U - 2t + 2\mu + 4V. \tag{A.61}$$

This minimum is achieved for $D = 1$, which corresponds to the full filling state.

A.3 Symmetry Breaking

In this section, we look at the various phases of the phase diagram and investigate which (if any) symmetries of the Hamiltonian they spontaneously break.

First we note that that hopping part of the state wavefunction ($A_1A_2A_3A_4$ for hole hopping and $A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger$ for fermion hopping) has all of the symmetries of the Hamiltonian, so no symmetry breaking occurs there.

Of the above ground states, some are product states between the physical and ancillary space. Those correspond to a ground state that is not degenerate and that respects all the symmetries of the Hamiltonian. Those are the phases of vacuum and full filling.

- Vacuum

$$|\Phi\rangle = |\Omega\rangle \qquad \text{or} \qquad |\Phi\rangle = A_1^\dagger A_2^\dagger A_3^\dagger A_4^\dagger |\Omega\rangle . \tag{A.62}$$

- Full filling

$$|\Phi\rangle = \prod_i a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger |\Omega\rangle \quad \text{or} \quad |\Phi\rangle = A_1 A_2 A_3 A_4 \prod_i a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger |\Omega\rangle . \tag{A.63}$$

The rest of the phases contain some entanglement between the physical and ancillary space and therefore correspond to ground state degeneracy and symmetry breaking. As was described in subsection 5.3.1 we can examine which symmetries are broken by applying the symmetry action to only half of the purified ground state. Here we use the unitaries from subsection 6.1.3 to represent the symmetries of the Hamiltonian:

$$U_{(lj)} = \prod_{k=1}^{2} \left(a_{l,k} a_{l,k}^\dagger a_{j,k} a_{j,k}^\dagger + a_{l,k}^\dagger a_{l,k} a_{j,k}^\dagger a_{j,k} + a_{l,k}^\dagger a_{j,k} + a_{j,k}^\dagger a_{l,k} \right), \tag{A.64}$$

$$U_{\mathbf{U(1)}}(\alpha) = \prod_{i=1}^{N} \prod_{k=1}^{2} \left(a_{i,k} a_{i,k}^\dagger + \mathrm{e}^{-\mathrm{i}\alpha} a_{i,k}^\dagger a_{i,k} \right), \tag{A.65}$$

$$U_{\mathbf{SU(2)}}(\beta,\gamma) = \prod_{i=1}^{N} \Big(a_{i,1} a_{i,1}^\dagger a_{i,2} a_{i,2}^\dagger + a_{i,1}^\dagger a_{i,1} a_{i,2}^\dagger a_{i,2} + \beta a_{i,1}^\dagger a_{i,1} a_{i,2} a_{i,2}^\dagger \tag{A.66}$$

$$+ \beta^* a_{i,1} a_{i,1}^\dagger a_{i,2}^\dagger a_{i,2} - \gamma^* a_{i,2}^\dagger a_{i,1} + \gamma a_{i,1}^\dagger a_{i,2} \Big). \tag{A.67}$$

Let us investigate the remaining phases one-by-one and see which of them break which symmetries.

Pairing

- Permutation symmetry. To simplify notation, we label π the permutation that swaps sites l and j:

$$U_{(lj)} |\Phi\rangle = U_{(lj)} \prod_{i=1}^{N} \left(\sin(\theta)\mathbb{1} + \cos(\theta) a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \right) |\Omega\rangle \tag{A.68}$$

$$= \prod_{i=1}^{N} \left(\sin(\theta)\mathbb{1} + \cos(\theta) a_{\pi(i),1}^\dagger a_{\pi(i),2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \right) |\Omega\rangle \not\propto |\Phi\rangle . \tag{A.69}$$

- **U(1)** symmetry:

$$U_{\mathbf{U(1)}}(\alpha) |\Phi\rangle = U_{\mathbf{U(1)}}(\alpha) \prod_{i=1}^{N} \left(\sin(\theta)\mathbb{1} + \cos(\theta) a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \right) |\Omega\rangle \tag{A.70}$$

$$= \prod_{i=1}^{N} \left(\sin(\theta)\mathbb{1} + \mathrm{e}^{2\mathrm{i}\alpha} \cos(\theta) a_{i,1}^\dagger a_{i,2}^\dagger a_{i,3}^\dagger a_{i,4}^\dagger \right) |\Omega\rangle \not\propto |\Phi\rangle \tag{A.71}$$

- **SU(2)** symmetry:

$$U_{\mathbf{SU(2)}}(\beta,\gamma)\,|\Phi\rangle = U_{\mathbf{U(1)}}(\theta)\prod_{i=1}^{N}\left(\sin(\theta)\mathbb{1}+\cos(\theta)a_{i,1}^{\dagger}a_{i,2}^{\dagger}a_{i,3}^{\dagger}a_{i,4}^{\dagger}\right)|\Omega\rangle \quad \text{(A.72)}$$

$$= \prod_{i=1}^{N}\left(\sin(\theta)\mathbb{1}+\cos(\theta)a_{i,1}^{\dagger}a_{i,2}^{\dagger}a_{i,3}^{\dagger}a_{i,4}^{\dagger}\right)|\Omega\rangle = |\Phi\rangle\,. \quad \text{(A.73)}$$

As we can see, the **SU(2)** symmetry is preserved while the **U(1)** and the permutational symmetries are broken in the pairing phase.

Half Filling

- Permutation symmetry:

$$U_{(lj)}\,|\Phi\rangle = U_{(lj)}\prod_{i=1}^{N}\left[a_{i,1}^{\dagger}a_{i,3}^{\dagger}+a_{i,2}^{\dagger}a_{i,4}^{\dagger}\right]|\Omega\rangle \quad \text{(A.74)}$$

$$= \prod_{i=1}^{N}\left[a_{\pi(i),1}^{\dagger}a_{i,3}^{\dagger}+a_{\pi(i),2}^{\dagger}a_{i,4}^{\dagger}\right]|\Omega\rangle \not\propto |\Phi\rangle\,. \quad \text{(A.75)}$$

- **U(1)** symmetry:

$$U_{\mathbf{U(1)}}(\alpha)\,|\Phi\rangle = U_{\mathbf{U(1)}}(\alpha)\prod_{i}\left[a_{i,1}^{\dagger}a_{i,3}^{\dagger}+a_{i,2}^{\dagger}a_{i,4}^{\dagger}\right]|\Omega\rangle \quad \text{(A.76)}$$

$$= \prod_{i}\left[\mathrm{e}^{\mathrm{i}\alpha}a_{i,1}^{\dagger}a_{i,3}^{\dagger}+\mathrm{e}^{\mathrm{i}\alpha}a_{i,2}^{\dagger}a_{i,4}^{\dagger}\right]|\Omega\rangle \quad \text{(A.77)}$$

$$= \prod_{i}\mathrm{e}^{\mathrm{i}\alpha}\left[a_{i,1}^{\dagger}a_{i,3}^{\dagger}+a_{i,2}^{\dagger}a_{i,4}^{\dagger}\right]|\Omega\rangle \propto |\Phi\rangle\,. \quad \text{(A.78)}$$

  We can see that the half filling phase only has this symmetry up to a global complex phase. However, as was shown in section 5.3.1, this is allowed in this case.

- **SU(2)** symmetry:

$$U_{\mathbf{SU(2)}}(\beta,\gamma)\,|\Phi\rangle = U_{\mathbf{SU(2)}}(\beta,\gamma)\prod_{i}\left[a_{i,1}^{\dagger}a_{i,3}^{\dagger}+a_{i,2}^{\dagger}a_{i,4}^{\dagger}\right]|\Omega\rangle \quad \text{(A.79)}$$

$$= \prod_{i}\left[\beta a_{i,1}^{\dagger}a_{i,3}^{\dagger}-\gamma^{*}a_{i,2}^{\dagger}a_{i,3}^{\dagger}+\beta^{*}a_{i,2}^{\dagger}a_{i,4}^{\dagger}+\gamma a_{i,1}^{\dagger}a_{i,4}^{\dagger}\right]|\Omega\rangle \not\propto |\Phi\rangle\,. \quad \text{(A.80)}$$

In this case we can see that the permutation symmetry and the **SU(2)** symmetries are broken while the **U(1)** is preserved.

Sub- and Super-half Filling

For the sub- and super-half filling phases, the symmetry breaking analysis may be done without tedious calculations. First, note that in these phases, the ground state is a superposition of vacuum / full filling with half filling. Vacuum and full filling

both have all symmetries while half filling does not have the permutation symmetry and the **SU(2)** symmetry. This implies that their superposition cannot have those two symmetries either. Furthermore, the action of the **U(1)** symmetry introduces a global complex phase in the half filling phase. When we take a superposition of this state with either vacuum / full filling, this global complex phase will become relative complex phase, which means that the symmetry will be broken.

Appendix B

Selected Topics from Functional Analysis

In this appendix we will go over some important results in functional analysis that play a role in our research. The provided citations refer to when these results were first discovered. Their formulation included here has been adapted from [154–157].

Theorem 7 (Hahn-Banach Theorem [158–160])**.** *Let M be a subspace of a vector space X, p a seminorm on X and f a linear functional acting on elements of M such that*

$$|f(x)| \leq p(x) \qquad \forall x \in M. \tag{B.1}$$

Then f extends to a linear functional F on X that satisfies

$$|F(x)| \leq p(x) \qquad \forall x \in X. \tag{B.2}$$

Throughout the thesis the Hahn-Banach theorem is used two times. It is used for the construction of an MPO approximation in subsection 4.3.1 and also when properly defining the thermodynamic limit in the section B.1.

Lemma 3 (Auerbach's Lemma [161])**.** *Let V be an n-dimensional normed vector space. Then there are unit vectors $x_1, ..., x_n \in V$ and unit functionals $x_1^*, ..., x_n^* \in V^*$ such that*

$$x_i^*(x_j) = \delta_{ij}. \tag{B.3}$$

The vectors $x_1, ..., x_n \in V$ are referred to as Auerbach's basis. The important part of the lemma is that the dual basis $x_1^*, ..., x_n^* \in V^*$ is also unit norm. In Hilbert spaces this can be easily achieved by taking any orthonormal basis with its dual, but in general normed vector spaces, the existence of Auerbach's basis is not so trivial.

Proof. (adapted from [156]) Let $\{v_i\}_{i=1}^n$ be a basis of the vector space V. For any (other) set of vectors $\{u_i\}_{i=1}^n$ in V define the function $\det(u_1, u_2, ..., u_n)$ as the determinant of the matrix whose jth column are the coefficients of u_j when expressed in the basis $\{v_i\}_{i=1}^n$. Let B_V be the closed unit ball in V and let $B_V^n = B_V \times B_V \times ... \times B_V$ be the set of all n-tuples of vectors, each of which is within the unit ball. The set B_V^n is compact, so the continuous function $\det(\cdot)$

attains its supremum there. Let $(x_1, x_2, ..., x_n) \in B_V^n$ be the point where $\det(\cdot)$ attains its supremum within B_V^n. Because of the multilinearity of determinant, the vectors $\{x_i\}_{i=1}^n$ have to be on the surface of the unit ball $\|x_i\| = 1$. Because the determinant is non-zero, the vectors $\{x_i\}_{i=1}^n$ have to be linearly independent and therefore they form another basis of V.

Define the functionals $f_i \in X^*$ as

$$f_i(x) = \frac{\det(x_1, ..., x_{i-1}, x, x_{i+1}, ..., x_n)}{\det(x_1, ..., x_n)} \quad \forall x \in V. \tag{B.4}$$

For $x \in B_V$, the following holds $|f_i(x)| \leq 1$ with equality attained for $x = x_i$, so the functionals f_i are unit norm. By construction, they fulfill $f_i(x_j) = \delta_{ij}$, so they are the dual basis to $\{x_i\}_{i=1}^n$. The set $\{x_i\}_{i=1}^n$ is therefore the Auerbach basis. □

Auerbach's Lemma is crucial in proving our theorem about approximability of mixed states by MPO.

Theorem 8 (Banach-Alaoglu Theorem [162])**.** *Let X be a normed vector space with a dual space X^*. Let B_{X^*} be the closed unit ball in X^* with respect to the operator norm. Then B_{X^*} is closed with respect to the weak-* topology.*

This theorem is used in this appendix in section B.1 to help properly define the thermodynamic limit in the language of C*-algebras.

B.1 C*-Algebras and Thermodynamic Limit

If one wants to take the thermodynamic limit of a Hamiltonian like (3.25), it is tempting to simply replace the upper bound of the summation to $+\infty$. However, there are many difficulties in actually taking the limit $N \to +\infty$. What object will this Hamiltonian be? How will it be normalized? To address all of that, we turn to C*-algebras. First we start with some basic definitions.

Definition 21. *Let $\mathcal{A}$ be a vector space equipped with a binary operation acting $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ (often called multiplication). If the operation is bilinear, $\mathcal{A}$ is an* algebra*. We denote the product of $x \in \mathcal{A}$ and $y \in \mathcal{A}$ as $xy \in \mathcal{A}$.*

If the multiplication operation is associative, then $\mathcal{A}$ is an associative algebra*.*

If $\mathcal{A}$ is an associative algebra equipped with a norm $\|\cdot\|$ satisfying $\|xy\| \leq \|x\|\|y\|$, then it is a normed algebra*.*

*If the underlying field of $\mathcal{A}$ is $\mathbb{C}$ and $\mathcal{A}$ is equipped with an involution * satisfying:*

$$(x^*)^* = x, \tag{B.5}$$
$$(x + y)^* = x^* + y^*, \tag{B.6}$$
$$(xy)^* = y^* x^*, \tag{B.7}$$
$$(\lambda x)^* = \lambda^* x^*. \tag{B.8}$$

for any $x, y \in \mathcal{A}$ and $\lambda \in \mathbb{C}$, then $\mathcal{A}$ is a *-algebra*.*

*If $\mathcal{A}$ is a complete normed *-algebra satisfying $\|x^*x\| = \|x\|^2$ for any $x \in \mathcal{A}$, then it is a* C*-algebra*.*

The definition above fits some classes of quantum operators, equipped with the operator norm $\|\cdot\|_\infty$ and with the Hermitian conjugate $\dagger$ as the involution. For example quantum operators acting on N sites form a C*-algebra. If the operators are elements of a C*-algebra, then what are the quantum states?

Definition 22. *A functional $\rho \in \mathcal{A}^*$ is called* positive *if $\rho(x^*x) \geq 0$ for all $x \in \mathcal{A}$.*
A positive functional ρ is called a state *if it has unit norm:*

$$\|\rho\| = \sup_{x\in\mathcal{A}} \frac{|\rho(x)|}{\|x\|} = 1. \tag{B.9}$$

If a state cannot be expressed as a non-trivial convex combination of two other states, it is called pure.

This algebraic definition of state coincides perfectly with our quantum-mechanical understanding of what a state is. States (as members of $\mathcal{A}^*$) acting on operators (as members of $\mathcal{A}$) can be interpreted as taking the expectation value of that state with the corresponding operator. In this interpretation a state ρ is not defined by a wavefunction vector or a density matrix, but by its expectation values.

In the thermodynamic limit, we want the number of sites to be infinite (but still countable). We can use the set of natural numbers $\mathbb{N}$ to label the sites. Let $\mathcal{F}(\mathbb{N})$ be the set of all finite subsets of $\mathbb{N}$. For any $\Lambda \in \mathcal{F}(\mathbb{N})$ we define $\mathcal{A}_\Lambda$ to be the C*-algebra of all observables acting on the sites contained in Λ. Note that for any $\Lambda' \supset \Lambda$ an observable $x \in \mathcal{A}_\Lambda$ can be identified with another observable $x_\Lambda \otimes \mathbb{1}_{\Lambda'\setminus\Lambda} \in \mathcal{A}_{\Lambda'}$.

With this we can define the algebra of all local observables,

$$\mathcal{A}_{\text{loc}} = \bigcup_{\Lambda\in\mathcal{F}(\mathbb{N})} \mathcal{A}_\Lambda. \tag{B.10}$$

In order for $\mathcal{A}_{\text{loc}}$ to be a well-defined algebra, it needs to be equipped with the required algebra operations.

Definition 23. *For any $x \in \mathcal{A}_\Lambda$ and $y \in \mathcal{A}_\Gamma$, we define the addition and multiplication in $\mathcal{A}_{loc}$ as*

$$x + y \equiv \left(x_\Lambda \otimes \mathbb{1}_{\Gamma\setminus\Lambda}\right) + \left(y_\Gamma \otimes \mathbb{1}_{\Lambda\setminus\Gamma}\right) \in \mathcal{A}_{\Lambda\cup\Gamma} \subset \mathcal{A}_{loc}, \tag{B.11}$$

$$xy \equiv \left(x_\Lambda \otimes \mathbb{1}_{\Gamma\setminus\Lambda}\right)\left(y_\Gamma \otimes \mathbb{1}_{\Lambda\setminus\Gamma}\right) \in \mathcal{A}_{\Lambda\cup\Gamma} \subset \mathcal{A}_{loc}. \tag{B.12}$$

The algebra of local operators can be equipped with the operator norm $\|\cdot\|_\infty$ and the * involution, making it a normed *-algebra. However, it is not complete and therefore it is not a C*-algebra. To make it a C*-algebra, we consider its closure $\overline{\mathcal{A}_{\text{loc}}}$ called the *algebra of quasi-local operators*. Hereinafter the term quasi-local is used to include local.

The corresponding space of functionals $\left(\overline{\mathcal{A}_{\text{loc}}}\right)^*$ corresponds to the space quantum states in the thermodynamic limit. Note that the states are again defined by their expectation values (on all quasi-local observables). In particular, if two states give the same expectation values, they are equal.

Example 9 (Equality in the thermodynamic limit)
Consider the following two (families of) states — one mixed and one pure:

$$|\phi\rangle = \frac{1}{\sqrt{2}} |0\rangle^{\otimes N} + \frac{1}{\sqrt{2}} |1\rangle^{\otimes N}, \tag{B.13}$$

$$\rho = \frac{1}{2} |0\rangle \langle 0|^{\otimes N} + \frac{1}{2} |1\rangle \langle 1|^{\otimes N}. \tag{B.14}$$

As we take the thermodynamic limit $N \to +\infty$, these states will give the same expectation values for any local observable. Therefore those two families of states correspond to the *same state* in the thermodynamic limit (which is mixed).

Many physical Hamiltonians are a sum of local terms, e.g.

$$H_N = \sum_{i,j=1}^{N} h_{ij}. \tag{B.15}$$

As N goes to infinity, we can interpret this as a sequence in $\overline{\mathcal{A}_{\text{loc}}}$. However, this sequence does not converge, which prevents us from properly defining the Hamiltonian in the thermodynamic limit. The object that we would get just by replacing the N in (B.15) with $+\infty$ is not well defined.

However, each of the Hamiltonians $H_N \in \mathcal{A}_{\{1,\dots,N\}}$ has a ground state $\phi_N \in \mathcal{A}^*_{\{1,\dots,N\}}$. Since $\mathcal{A}_{\{1,\dots,N\}}$ is a subspace of $\overline{\mathcal{A}_{\text{loc}}}$, we can apply the Hahn-Banach theorem. In this case the theorem states that we can extend the states ϕ_N into $\left(\overline{\mathcal{A}_{\text{loc}}}\right)^*$. We will call the extensions $\Phi_N \in \left(\overline{\mathcal{A}_{\text{loc}}}\right)^*$. These extensions Φ_N form a well-defined sequence in $\Phi_N \in \left(\overline{\mathcal{A}_{\text{loc}}}\right)^*$. Thanks to the Banach-Alaoglu Theorem, the set of states is weakly-* compact, meaning that every sequence of states has at least one accumulation point (if not a limit). We interpret these accumulation points as the ground state(s) in the thermodynamic limit. This way we define the *ground state in the thermodynamic limit* without having a proper definition of the Hamiltonian itself in the thermodynamic limit.

Appendix C

Rényi Mutual Information

This section contains some original results on the Rényi mutual information. These results were obtained as part of the research into approximability of mixed states by tensor networks (chapter 4). However, these results were never published. They do not fit into chapter 4 because eventually we chose the Rényi entanglement of purification as our criterion of approximability, not the Rényi mutual information.

In section C.1 we discuss the definition of the Rényi mutual information, settling on its version with the sandwiched Rényi relative entropy. Then in section C.2 we calculate the Rényi mutual information for a generic bipartite pure state. In section C.3 we use the result from section C.2 to prove the inequality between the Rényi mutual information and the Rényi entanglement of purification, analogous to equation (2.41). In section C.4 we calculate to which does the Rényi mutual information reduces on classical systems and in section C.5 we prove the Rényi mutual information area law for classical matrix-product density operators.

C.1 Definition

The Rényi mutual information is a generalizatation of the mutual information (2.30) to the Rényi setting. We require this quantity to have some of the following properties:

- When applied to pure states, it should reduce to twice the Rényi entropy of the reduced state, just like mutual information reduces to von Neumann entropy on pure states.
- It should be non-negative and it should be non-increasing under quantum channels applied to both systems (data processing inequality).
- It should be defined for $\alpha \in (0,1) \cup (1,+\infty)$ and the limit $\alpha \to 1$ should give the (von Neumann) mutual information.

The naïve way to do this generalization would be just to turn the von Neumann entropies in the definition of mutual information (2.30) into Rényi entropies, i.e.

$$I_\alpha(A:B)_\psi \stackrel{?}{=} S_\alpha(\rho_A) + S_\alpha(\rho_B) - S_\alpha(\rho_{AB}). \tag{C.1}$$

However, this generalization loses several crucial properties of the mutual information. It may sometimes be negative [163] and it is not non-increasing under quantum channels anymore.

To amend that, we instead generalize a different definition of the Rényi mutual information. Recall that there are four other alternative definitions of mutual information (2.31)–(2.34). It is standard to use the definition (2.32) and replace the relative entropy with the Rényi relative entropy [164]:

$$I_\alpha(A:B)_\psi = \min_{\sigma_B} D_\alpha(\rho_{AB}||\rho_A \otimes \sigma_B). \tag{C.2}$$

Unfortunately for us, there are several definitions of the Rényi relative entropy. The first one was introduced by Petz [165]:

$$D_\alpha(\rho||\sigma) = \frac{1}{\alpha-1}\log \mathrm{Tr}\Big\{\rho^\alpha \sigma^{1-\alpha}\Big\}. \tag{C.3}$$

While this can be defined for any $\alpha \in [0,1)\cup(1,\infty)$, it obeys the data processing inequality only for $\alpha \in [0,1)\cup(1,2]$ [165]. A different Rényi relative entropy generalization is called the sandwiched Rényi relative entropy:

$$\tilde{D}_\alpha(\rho||\sigma) = \frac{1}{\alpha-1}\log \mathrm{Tr}\Big\{\Big(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\Big)^\alpha\Big\}. \tag{C.4}$$

This quantity obeys the data processing inequality for $\alpha \in [1/2,1)\cup(1,\infty)$ [166]. These two quantities were later even further generalized into a two parameter family of relative entropies [167]

$$D_{\alpha,z}(\rho||\sigma) = \frac{1}{\alpha-1}\log \mathrm{Tr}\Big\{\Big(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\Big)^z\Big\}. \tag{C.5}$$

In the rest of this chapter, we define our Rényi mutual information with the sandwiched Rényi relative entropy. For more discussion about the various definitions of the Rényi mutual information, see Scalet et al. [102].

C.2 Rényi Mutual Information for Pure States

As was shown in example 5, for pure states, the mutual information reduces to twice the von Neumann entropy of the reduced density matrices. In this section we investigate whether the same holds for the Rényi mutual information. Start with a pure state $\rho_{AB} = |\psi\rangle\langle\psi|_{AB}$ and its reduced density matrices ρ_A, ρ_B.

As a bipartite pure state, $|\psi\rangle$ has a Schmidt decomposition

$$|\psi\rangle_{AB} = \sum_i \sqrt{\lambda_i}\,|a_i\rangle\,|b_i\rangle\,, \tag{C.6}$$

where $\{|a_i\rangle\}_i$ and $\{|b_i\rangle\}_i$ are orthonormal bases on the two parts of the system. As density matrices, ρ_{AB} and ρ_A may be written as:

$$\rho_{AB} = |\psi\rangle\langle\psi| = \sum_{i,j}\sqrt{\lambda_i\lambda_j}\,|a_i\rangle\,,|b_i\rangle\langle a_j|\langle b_j| \tag{C.7}$$

$$\rho_A = \mathrm{Tr}_B\,|\psi\rangle\langle\psi| = \sum_i \lambda_i\,|a_i\rangle\langle a_i|\,. \tag{C.8}$$

We slowly start plugging in the formula for the Rényi mutual information (C.2) using the sandwiched Rényi relative entropy (C.4), starting with

$$\left(\rho_A^{\frac{1-\alpha}{2\alpha}} \otimes \mathbb{1}_B\right) \rho_{AB} \left(\rho_A^{\frac{1-\alpha}{2\alpha}} \otimes \mathbb{1}_B\right) \tag{C.9}$$

$$= \left(\sum_l \lambda_l^{\frac{1-\alpha}{2\alpha}} |a_l\rangle \langle a_l|\right) \sum_{i,j} \sqrt{\lambda_i \lambda_j} |a_i\rangle |b_i\rangle \langle a_j| \langle b_j| \left(\sum_k \lambda_k^{\frac{1-\alpha}{2\alpha}} |a_k\rangle \langle a_k|\right) \tag{C.10}$$

$$= \sum_{l,i,j,k} \lambda_l^{\frac{1-\alpha}{2\alpha}} \sqrt{\lambda_i \lambda_j} \lambda_k^{\frac{1-\alpha}{2\alpha}} |a_l\rangle |b_i\rangle \langle b_j| \langle a_k| \delta_{jk} \delta_{li} \tag{C.11}$$

$$= \sum_{i,j} (\lambda_i \lambda_j)^{\frac{1}{2\alpha}} |a_i\rangle |b_i\rangle \langle a_j| \langle b_j| \equiv \rho'_{AB}. \tag{C.12}$$

Here we defined ρ'_{AB} to simplify notation in the future. Then we can write

$$I_\alpha(A:B)_\psi = \min_{\sigma_B} \left\{ \frac{1}{\alpha - 1} \log \operatorname{Tr} \left[\left(\mathbb{1} \otimes \sigma_B^{\frac{1-\alpha}{2\alpha}}\right) \rho'_{AB} \left(\mathbb{1} \otimes \sigma_B^{\frac{1-\alpha}{2\alpha}}\right) \right]^\alpha \right\}. \tag{C.13}$$

Now we're facing the optimization over all possible σ_B matrices. To simplify the task, we use lemma 4.

Lemma 4. *The optimal matrix σ_B is diagonal in the $\{|b_i\rangle\}_i$ basis.*

Proof. First note that the state ρ'_{AB} is an unnormalized pure state

$$\rho'_{AB} = |\zeta\rangle \langle \zeta| \qquad \text{for} \qquad |\zeta\rangle = \sum_i \lambda_i^{\frac{1}{2\alpha}} |a_i\rangle |b_i\rangle . \tag{C.14}$$

Assume a general density matrix $\sigma_B = \sum_k \mu_k |c_k\rangle \langle c_k|$. For now we fix the coefficients μ_k and just optimize over the basis $\{|c_k\rangle\}_k$. After applying the matrix $\mathbb{1} \otimes \sigma_B^{\frac{1-\alpha}{2\alpha}}$ to the state $|\zeta\rangle$ we get

$$\sigma_B^{\frac{1-\alpha}{2\alpha}} |\zeta\rangle = \sum_{i,k} \lambda_i^{\frac{1}{2\alpha}} \mu_k^{\frac{1-\alpha}{2\alpha}} |a_i\rangle |c_k\rangle \langle c_k|b_i\rangle \equiv |\zeta'\rangle . \tag{C.15}$$

The normalization of the vector $|\zeta'\rangle$ is

$$\langle \zeta'|\zeta'\rangle = \sum_{i,k,j,l} \lambda_i^{\frac{1}{2\alpha}} \mu_k^{\frac{1-\alpha}{2\alpha}} \lambda_j^{\frac{1}{2\alpha}} \mu_l^{\frac{1-\alpha}{2\alpha}} \langle c_k|b_i\rangle \langle b_j|c_l\rangle \langle a_j|a_i\rangle \langle c_l|c_k\rangle \tag{C.16}$$

$$= \sum_{ik} \lambda_i^{\frac{1}{\alpha}} \mu_k^{\frac{1-\alpha}{\alpha}} |\langle c_k|b_i\rangle|^2. \tag{C.17}$$

So overall we have

$$\operatorname{Tr} \left[\left(\mathbb{1} \otimes \sigma_B^{\frac{1-\alpha}{2\alpha}}\right) \rho'_{AB} \left(\mathbb{1} \otimes \sigma_B^{\frac{1-\alpha}{2\alpha}}\right) \right]^\alpha = \operatorname{Tr} \left(|\zeta'\rangle \langle \zeta'|\right)^\alpha \tag{C.18}$$

$$= \operatorname{Tr} \left(\frac{|\zeta'\rangle \langle \zeta'|}{\langle \zeta'|\zeta'\rangle} \right)^\alpha \langle \zeta'|\zeta'\rangle^\alpha \tag{C.19}$$

$$= \langle \zeta'|\zeta'\rangle^\alpha . \tag{C.20}$$

We want to choose such σ_B (specifically, the basis $\{|b_i\rangle\}_i$) so that this value is minimized. Define $M_{ki} = |\langle c_k|b_i\rangle|^2$. This is a doubly-stochastic matrix, i.e. a matrix of non-negative entries such that each row and column sums up to 1. Doubly-stochastic matrices are a convex hull of permutation matrices. Note that

$$\langle\zeta'|\zeta'\rangle = \sum_{ik} \lambda_i^{\frac{1}{\alpha}} \mu_k^{\frac{1-\alpha}{\alpha}} M_{ki} \tag{C.21}$$

is a linear function of the entries of M_{ki}. The extreme of a linear function on a convex set is attained in one of the points of its convex hull. Those are the permutation matrices, which correspond to the case where the basis $\{|c_k\rangle\}_k$ is equal to $\{|b_i\rangle\}_i$ up to permutation of elements. So we see that for any fixed set of eigenvalues μ_k, the optimal matrix σ_B is be diagonal in the $\{|c_i\rangle\}_i$ basis. □

From lemma 4 we know that we should consider $\sigma_B = \sum_k^n \mu_k |b_k\rangle\langle b_k|$ (with the optimal values of μ_k unknown so far). From the proof of lemma 4 we know that we can write

$$\left(\mathbb{1}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\sigma_{AB}\left(\mathbb{1}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right) = |\zeta'\rangle\langle\zeta'| \tag{C.22}$$

for

$$|\zeta'\rangle = \sum_i^n \mu_i^{\frac{1-\alpha}{2\alpha}} \lambda_i^{\frac{1}{2\alpha}} |a_i\rangle|b_i\rangle \tag{C.23}$$

with the norm

$$\langle\zeta'|\zeta'\rangle = \sum_i^n \mu_i^{\frac{1-\alpha}{\alpha}} \lambda_i^{\frac{1}{\alpha}}. \tag{C.24}$$

Therefore we have

$$\mathrm{Tr}\left[\left(\mathbb{1}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\sigma_{AB}\left(\mathbb{1}\otimes\sigma_B^{\frac{1-\alpha}{2\alpha}}\right)\right]^\alpha = \mathrm{Tr}\left(|\zeta'\rangle\langle\zeta'|\right)^\alpha = \langle\zeta'|\zeta'\rangle^\alpha, \tag{C.25}$$

which makes the minimization over σ_B look like

$$I_\alpha(A:B)_\psi = \min_{\{\mu_i\}_i} \frac{1}{\alpha-1}\log\left[\sum_i^n \mu_i^{\frac{1-\alpha}{\alpha}}\lambda_i^{\frac{1}{\alpha}}\right]^\alpha \tag{C.26}$$

$$= \min_{\{\mu_i\}_i} \frac{\alpha}{\alpha-1}\log\sum_i^n \mu_i^{\frac{1-\alpha}{\alpha}}\lambda_i^{\frac{1}{\alpha}}. \tag{C.27}$$

Now we need to extremize $\sum_i^n \mu_i^{\frac{1-\alpha}{\alpha}}\lambda_i^{\frac{1}{\alpha}}$ (minimize for $\alpha > 1$ and maximize for $\alpha < 1$). This is a problem for the Lagrange coefficients method with the constraint $\sum_l \mu_l - 1 = 0$. We define

$$f(\mu_1, ..., \mu_n) = \sum_i^n \mu_i^{\frac{1-\alpha}{\alpha}}\lambda_i^{\frac{1}{\alpha}}, \tag{C.28}$$

$$g(\mu_1, ..., \mu_n) = \sum_i^n \mu_i - 1, \tag{C.29}$$

so that the Lagrangian is

$$\mathcal{L} = f(a_1, ..., a_n) - \nu g(a_1, ..., a_n) \tag{C.30}$$

and we get

$$0 = \nabla_{\mu_1,...,\mu_n,\nu}\mathcal{L} = \begin{pmatrix} \left(\frac{1-\alpha}{\alpha}\right)\mu_1^{\frac{1-2\alpha}{\alpha}}\lambda_1^{\frac{1}{\alpha}} - \nu \\ \vdots \\ \left(\frac{1-\alpha}{\alpha}\right)\mu_n^{\frac{1-2\alpha}{\alpha}}\lambda_n^{\frac{1}{\alpha}} - \nu \\ -g(\mu_1,...,\mu_n) \end{pmatrix}. \tag{C.31}$$

For $\alpha = 1/2$, this reduces to the following set of equations:

$$0 = \begin{pmatrix} \lambda_1^2 - \nu \\ \vdots \\ \lambda_n^2 - \nu \\ -g(\mu_1,...,\mu_n) \end{pmatrix}. \tag{C.32}$$

This is impossible to satisfy unless all of the λ_i's are identical, which is a trivial case. Hence for $\alpha = 1/2$, the Lagrange multipliers method does not find any stationary point.

For $\alpha \neq 1/2$ we express μ_i as

$$\mu_i = \left(\frac{\nu\alpha}{1-\alpha}\right)^{\frac{\alpha}{1-2\alpha}} \lambda_i^{\frac{1}{2\alpha-1}} \quad \forall i \in \{1,...,n\}. \tag{C.33}$$

To enforce $g(\mu_1,...,\mu_n) = 0$, we choose the Lagrange multiplier ν such that

$$\mu_i = \frac{\lambda_i^{\frac{1}{2\alpha-1}}}{\sum_l^n \lambda_l^{\frac{1}{2\alpha-1}}} \quad \forall i \in \{1,...,n\}. \tag{C.34}$$

This is one stationary point of the Lagrangian and therefore a possible solution.

The function $f(\mu_1,...,\mu_n)$ is a sum of individual functions which we can call $f_i(\mu_1,...,\mu_n) = \mu_i^{\frac{1-\alpha}{\alpha}}\lambda_i^{\frac{1}{\alpha}}$. For $\alpha \in (0,1/2) \cup (1,+\infty)$, these individual functions are convex and so is their sum. For $\alpha \in (1/2,1)$ they are concave and so is their sum. This implies that for $\alpha \in (0,1/2) \cup (1,+\infty)$, the stationary point is the global minimum of the function f, while for $\alpha \in (1/2,1)$, it is the maximum.

The formula for the Rényi mutual information looks like (C.27)

$$I_\alpha(A:B)_\psi = \min_{\{\mu_i\}_i} \frac{\alpha}{\alpha-1} \log \sum_i^n a_i^{\frac{1-\alpha}{\alpha}}\lambda_i^{\frac{1}{\alpha}}. \tag{C.35}$$

Logarithm is a monotonic function, which will have no influence on the position of the minimum or maximum. However, the term $\frac{\alpha}{\alpha-1}$ flips the sign of the entire expression for $\alpha \in (0,1/2) \cup (1/2,1)$, thus converting the global maximum for $\alpha \in (1/2,1)$ into a global minimum and vice versa for $\alpha \in (0,1/2)$.

Therefore for $\alpha \in (1/2,1) \cup (1,+\infty)$ we have found a density matrix σ_B which minimizes the expression for mutual information. The mutual information will

then be

$$I_\alpha(A:B)_\psi = \frac{\alpha}{\alpha-1}\log\frac{\sum_i \lambda_i^{\frac{1}{2\alpha-1}}}{\left(\sum_k \lambda_k^{\frac{1}{2\alpha-1}}\right)^{\frac{1-\alpha}{\alpha}}} \tag{C.36}$$

$$= \frac{\alpha}{\alpha-1}\left[\log\sum_i \lambda_i^{\frac{1}{2\alpha-1}} - \log\left(\sum_k \lambda_k^{\frac{1}{2\alpha-1}}\right)^{\frac{1-\alpha}{\alpha}}\right] \tag{C.37}$$

$$= \frac{\alpha}{\alpha-1}\log\sum_i \lambda_i^{\frac{1}{2\alpha-1}} + \log\sum_k \lambda_k^{\frac{1}{2\alpha-1}} \tag{C.38}$$

$$= \frac{2\alpha-1}{\alpha-1}\log\sum_i \lambda_i^{\frac{1}{2\alpha-1}} \tag{C.39}$$

$$= 2\cdot\frac{1}{1-\frac{1}{2\alpha-1}}\log\sum_i \lambda_i^{\frac{1}{2\alpha-1}} \tag{C.40}$$

$$= 2S_{\frac{1}{2\alpha-1}}(A). \tag{C.41}$$

This means that the Rényi mutual information turns into the Rényi entropy for pure states, but with different α. The mapping $\alpha \to \frac{1}{2\alpha-1}$ maps the interval $(1/2, +\infty)$ into $(0, +\infty)$ with 1 being the fixed point. This is another indication that for our Rényi mutual information, only $\alpha > 1/2$ is relevant.

C.3 Entanglement of Purification and Rényi Mutual Information

As was mentioned in section 2.3.1, the regular entanglement of purification is bounded from the bottom by half the mutual information,

$$E_p(\rho_{AB}) \geq \frac{1}{2}I(A:B)_\rho. \tag{C.42}$$

Here we prove that the same holds for the Rényi mutual information and the Rényi entanglement of purification (defined in equation (2.51)). Assume that $|\Xi_\alpha\rangle_{AA'BB'}$ is the purification of ρ_{AB} that achieves the minimum Rényi entropy in the definition of the Rényi entanglement of purification. We have

$$2E_p^\alpha(\rho_{AB}) = 2S_\alpha(\mathrm{Tr}_{AA'}|\Xi_\alpha\rangle\langle\Xi_\alpha|) = I_{\frac{1+\alpha}{2\alpha}}(AA':BB')_\Xi \geq I_{\frac{1+\alpha}{2\alpha}}(A:B)_\rho. \tag{C.43}$$

The second equality comes from the equation (C.36). The inequality follows from the fact that the mutual information (for $\alpha \geq 1/2$, which is guaranteed here) is contractive under local quantum channels (including the partial trace).

C.4 Rényi Mutual Information for Classical Systems

In this section we examine how the Rényi mutual information behaves when the system in question is classical i.e. with density matrix diagonal in the computational basis.

We start by changing our perspective from diagonal density matrices to classical probability distributions (and correspondingly relabelling the eigenvalues from λ to p):

$$\rho_{AB} = \sum_{i,j} p_{ij} \left|ij\right\rangle \left\langle ij\right|, \tag{C.44}$$

$$\rho_A = \mathrm{Tr}_B\, \rho_{AB} = \sum_{i,j} p_{ij} \left|i\right\rangle \left\langle i\right| \equiv \sum_i p_i \left|i\right\rangle \left\langle i\right|. \tag{C.45}$$

Importantly for mutual information calculations, now the matrix ρ_{AB} commutes with $\rho_A \otimes \mathbb{1}$.

Lemma 5. *The optimal matrix σ_B is diagonal in the $\{|j\rangle\}_j$ basis.*

Proof. First assume a general density matrix σ_B

$$\sigma_B = \sum_j q_j \left|e_j\right\rangle \left\langle e_j\right|. \tag{C.46}$$

with some eigenvalues q_j and eigenvectors $|e_j\rangle$. The Rényi mutual information has the form

$$I_\alpha(A:B)_\rho = \min_{\sigma_B} \tilde{D}_\alpha(\rho_{AB}||\rho_A \otimes \sigma_B). \tag{C.47}$$

For $\alpha \in [1/2, 1) \cup (1, +\infty)$ the Rényi relative entropy is contractive under completely positive trace-preserving linear maps applied to both of its arguments [166]. We will take advantage of that with the map

$$\mathcal{N}(\cdot) = \mathrm{id}_A \otimes \left(\sum_i \left|i\right\rangle \left\langle i\right| \cdot \left|i\right\rangle \left\langle i\right| \right)_B, \tag{C.48}$$

which is a decohering map on the B subsystem (acting as identity on the A subsystem). We then have

$$\tilde{D}_\alpha(\rho_{AB}||\rho_A \otimes \sigma_B) \geq \tilde{D}_\alpha(\mathcal{N}(\rho_{AB})||\mathcal{N}(\rho_A \otimes \sigma_B)) \tag{C.49}$$

$$= \tilde{D}_\alpha(\rho_{AB}||\rho_A \otimes \sum_i \left|i\right\rangle \left\langle i\right| \sigma_B \left|i\right\rangle \left\langle i\right|). \tag{C.50}$$

We see that if we replace the density matrix σ_B simply by its diagonal (which is what we are doing by applying the map $\mathcal{N}$, the relative entropy does not increase. Therefore if we want to minimize $\tilde{D}_\alpha(\rho_{AB}||\rho_A \otimes \sigma_B)$ over all density matrices σ_B, we may choose to only consider matrices diagonal in the computational basis. □

Thanks to lemma 5, the matrices ρ_{AB} and $\rho_A \otimes \sigma_B$ are diagonal in the same basis, therefore they commute. That allows us to significantly simplify the formula for the mutual information:

$$I_\alpha(A:B)_\rho = \min_{\{q_k\}_k} \frac{1}{\alpha-1} \log \mathrm{Tr}\left\{\left[(\rho_A \otimes \sigma_B)^{\frac{1-\alpha}{2\alpha}} \rho_{AB} (\rho_A \otimes \sigma_B)^{\frac{1-\alpha}{2\alpha}}\right]^\alpha\right\} \tag{C.51}$$

$$= \min_{\{q_k\}_k} \frac{1}{\alpha-1} \log \mathrm{Tr}\left\{\left[(\rho_A \otimes \sigma_B)^{\frac{1-\alpha}{\alpha}} \rho_{AB}\right]^\alpha\right\} \tag{C.52}$$

$$= \min_{\{q_k\}_k} \frac{1}{\alpha-1} \log \mathrm{Tr}\left\{(\rho_A \otimes \sigma_B)^{1-\alpha} \rho_{AB}^\alpha\right\} \tag{C.53}$$

$$= \min_{\{q_k\}_k} \frac{1}{\alpha-1} \log \sum_{i,j} (p_i q_j)^{1-\alpha} p_{ij}^\alpha. \tag{C.54}$$

Now our goal is to find the optimal set $\{q_j\}_j$. Our goal is to extremize the function

$$f(\{q_k\}_k) = \sum_{i,j} (p_i q_j)^{1-\alpha} p_{ij}^\alpha. \tag{C.55}$$

As before, we want to minimize it for $\alpha > 1$ and maximize it for $\alpha < 1$ because of the $\frac{1}{\alpha-1}$ factor. We have the constraint that $g(\{q_k\}_k) = \sum_j q_j - 1 = 0$.

Again, we use the Lagrange multipliers method. The Lagrangian is

$$\mathcal{L} = f(\{q_k\}_k) - \nu g(\{q_k\}_k) = \sum_{i,j} (p_i q_j)^{1-\alpha} p_{ij}^\alpha - \nu \left(\sum_k q_k - 1\right) \tag{C.56}$$

and its gradient is

$$0 = \nabla_{q_1,\ldots,q_n,\nu} \mathcal{L} = \begin{pmatrix} (1-\alpha) \sum_i p_i^{1-\alpha} q_1^{-\alpha} p_{i1}^\alpha - \nu \\ \vdots \\ (1-\alpha) \sum_i p_i^{1-\alpha} q_n^{-\alpha} p_{in}^\alpha - \nu \\ \sum_k q_k - 1 \end{pmatrix}. \tag{C.57}$$

The first n rows in (C.57) can be rearranged into

$$q_j = \left[\frac{(1-\alpha) \sum_i p_i^{1-\alpha} p_{ij}^\alpha}{\nu}\right]^{\frac{1}{\alpha}}. \tag{C.58}$$

The constraint requires that the q_j's sum to 1. The Lagrange coefficient ν must therefore be set to satisfy this constraint:

$$\nu = (1-\alpha) \left[\sum_j \left(\sum_i p_i^{1-\alpha} p_{ij}^\alpha\right)^{\frac{1}{\alpha}}\right]^\alpha, \tag{C.59}$$

which gives

$$q_j = \frac{\left[\sum_i p_i^{1-\alpha} p_{ij}^\alpha\right]^{\frac{1}{\alpha}}}{\sum_k \left[\sum_l p_l^{1-\alpha} p_{lk}^\alpha\right]^{\frac{1}{\alpha}}}. \tag{C.60}$$

As a solution of the Lagrange multipliers method, this point is a stationary point of the mutual information. The objective function

$$f(\{q_k\}_k) = \sum_{i,j} (p_i q_j)^{1-\alpha} p_{ij}^\alpha \tag{C.61}$$

is convex for $\alpha \in (1, \infty)$ and concave for $\alpha \in (0, 1)$, so for $\alpha \in (1, \infty)$ the stationary point is a minimum while for $\alpha \in (0, 1)$ it is a maximum. Since the form of the mutual information is

$$I_\alpha(A:B)_\rho = \min_{\{q_k\}_k} \frac{1}{\alpha - 1} \log f(\{q_k\}_k), \tag{C.62}$$

we know that the stationary point is the minimizing point for all α.

Now substituting $Q_j = \sum_i p_{ij}^\alpha p_i^{1-\alpha}$, we can express the mutual information as

$$\begin{aligned}
I_\alpha(A:B)_\rho &= \frac{1}{\alpha - 1} \log \sum_j Q_j q_j^{1-\alpha} && \text{(C.63)}\\
&= \frac{1}{\alpha - 1} \log \sum_j Q_j \left(\frac{Q_j^{1/\alpha}}{\sum_k Q_k^{1/\alpha}} \right)^{1-\alpha} && \text{(C.64)}\\
&= \frac{1}{\alpha - 1} \log \sum_j \frac{Q_j^{1/\alpha}}{\left(\sum_k Q_k^{1/\alpha}\right)^{1-\alpha}} && \text{(C.65)}\\
&= \frac{1}{\alpha - 1} \log \sum_j \frac{Q_j^{1/\alpha}}{\left(\sum_k Q_k^{1/\alpha}\right)^{1-\alpha}} && \text{(C.66)}\\
&= \frac{\alpha}{\alpha - 1} \log \left(\sum_k Q_k^{1/\alpha} \right) && \text{(C.67)}\\
&= H_{1/\alpha}(\{Q_k\}_k), && \text{(C.68)}
\end{aligned}$$

where H_α is the classical Rényi entropy. Note that this is a slight abuse of notation because the Q_k's do not sum up to 1, so they are not a valid probability distribution.

C.5 Area Law for Classical Matrix Product Density Operators

A classical system $\rho_{AB} = \sum_{i,j} p_{ij} |ij\rangle \langle ij|$ is a 2-site matrix product density operator (MPDO)[1] with bond dimension D if and only if the matrix p_{ij} has rank D. This can be seen by using the fact that a matrix with rank D can be written as a sum of D matrices of rank 1, which is equal to the MPDO representation. If we prove the area law for 2-site MPDOs, it will automatically hold for multi-site MPDOs as well, since we can always interpret all the sites on one side of the cut as just a single site (of correspondingly higher physical dimension).

Consider that the matrix p_{ij} has rank D. We want to bound the mutual information between the two parts of the system. We will do that by bounding the max-mutual information

$$I_{\max}(A:B) = \min_{\{q_j\}_j} \log \max_{ij} p_{ij} q_j^{-1} p_i^{-1}. \tag{C.69}$$

[1] See definition 19 for the definition of a matrix product (density) operator

Before we can do that, we need to calculate the optimal q_j. We get that from the limit $\alpha \to +\infty$ applied to formula (C.60), which yields

$$q_j = \frac{\max_i p_{ij}/p_i}{N}, \qquad \text{where} \qquad N = \sum_j \max_i p_{ij}/p_i. \tag{C.70}$$

The formula for the max-mutual information (C.69) is effectively looking for the largest entry of the matrix

$$p_{ij} q_j^{-1} p_i^{-1} = N \frac{p_{ij}/p_i}{\max_i p_{ij}/p_i} \leq N \tag{C.71}$$

with the indices i and j denoting the rows and columns of the matrix.

Thus our goal is to bound N. First assume that the matrix p_{ij} has rank 1. Therefore $p_{ij} = a_i b_j$ for some non-negative sets of values $\{a_i\}_i$, $\{b_j\}_j$ (which we can understand as a row and column vector) and consequently $p_i = a_i \sum_j b_j$.

$$N = \sum_j \max_i \frac{p_{ij}}{p_i} = \sum_j \max_i \frac{a_i b_j}{a_i \sum_k b_k} = \sum_j \frac{b_j}{\sum_k b_k} = 1. \tag{C.72}$$

Now assume that the matrix p_{ij} has rank D. We can use the following lemma:

Lemma 6. *Any matrix of rank D with non-negative elements can be written as a sum of D rank 1 matrices:*

$$p_{ij} = \sum_{k=1}^{D} a_i^k b_j^k, \tag{C.73}$$

where $a_i^k b_j^k$ is the ij-th element of the k-th rank 1 matrix and all the numbers $\{a_i^k\}_{i,k}$ and $\{b_j^k\}_{j,k}$ are non-negative.

Proof. Say that the matrix p_{ij} has n columns. That means that its rows as vectors live in a vector space of n-dimensional real vectors. Moreover, they are all elements of the convex cone (which we call C from now on) of vectors with non-negative elements. The convex cone C is the conical hull of the standard basis elements $(1, 0, 0, ...), (0, 1, 0, ...), ...$

Consider the convex hull of the standard basis elements. This is an $(n-1)$-dimensional simplex (hereafter referred to as S) with the standard basis elements as vertices. Any non-zero vector in C lies on S after normalization in the 1-norm (for vectors).

The fact that the matrix p_{ij} has rank D means that its rows span a D-dimensional subspace (hereafter called Z) of the n-dimensional vector space. Since the non-zero rows of p_{ij} lie in C, we know that their normalized versions lie on S, which implies that Z intersects S. On the other hand we know that the origin lies on Z, but not on the affine hull of S.

The affine hull of S is an $(n-1)$-dimensional hyperplane obtained by taking all of the affine combinations of the elements of S. Z is a D-dimensional plane that contains some elements not contained in the affine hull of S, but it does contain at least D linearly-independent elements which are in the affine hull of

S. Therefore the intersection of these two planes must be a plane with dimension $D-1$. Hereafter we call this plane P.

Consider the intersection of P and S. This is an intersection of a $(D-1)$-dimensional plane with an $(n-1)$-dimensional simplex. This set can be at most $(D-1)$ dimensional simplex (hereafter called Q). Therefore there exist (at most) D vectors such that Q is the convex hull of them. Call these vectors $\{b^k\}_{k=1}^D$. Note that

$$\{b^k\}_{k=1}^D \subset Q \subset S \tag{C.74}$$

so they have non-negative components.

Recall the definition of Q:

$$Q = P \cap S = [Z \cap (\text{affine hull of } S)] \cap S = Z \cap S. \tag{C.75}$$

This implies that the normalized versions of the non-zero rows of p_{ij} lie in Q (as they lie in both S and Z). That means that the normalized versions of the non-zero rows of p_{ij} can be expressed as a convex combination of $\{b^k\}_{k=1}^D$. That means that their un-normalized versions can be expressed as a conical combination of $\{b^k\}_{k=1}^D$. Let's call the coefficients for the i-th row $\{a_i^k\}_{k=1}^D$. Then we can write

$$p_{ij} = \sum_{k=1}^{D} a_i^k b_j^k \tag{C.76}$$

for $\{a_i^k\}_{i,k}$ and $\{b_j^k\}_{j,k}$ non-negative.

□

Now we can write

$$N = \sum_j \max_i \frac{p_{ij}}{p_i} \tag{C.77}$$

$$= \sum_j \max_i \frac{\sum_{k=1}^D a_i^k b_j^k}{\sum_{m=1}^D a_i^m \sum_l b_l^m} \tag{C.78}$$

$$\leq \sum_j \sum_{k=1}^{D} \max_i \frac{a_i^k b_j^k}{\sum_{m=1}^D a_i^m \sum_l b_l^m} \tag{C.79}$$

$$\leq \sum_j \sum_{k=1}^{D} \max_i \frac{a_i^k b_j^k}{a_i^k \sum_l b_l^k} \tag{C.80}$$

$$= \sum_{k=1}^{D} \sum_j \frac{b_j^k}{\sum_l b_l^k} \tag{C.81}$$

$$= \sum_{k=1}^{D} 1 = D. \tag{C.82}$$

Here the first inequality comes from switching the sum over k with the maximization over i. The second inequality comes from omitting all but the k-th element of the sum over m in the denominator. This can be done because all the terms there are non-negative.

Thus the Rényi mutual information for a classical MPDO with bond dimension D is bounded by $\log(D)$.

Bibliography

[1] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[2] B. Schumacher. Quantum coding. *Phys. Rev. A*, 51:2738–2747, Apr 1995.

[3] Seth Lloyd. Capacity of the noisy quantum channel. *Physical Review A*, 55(3):1613–1622, Mar 1997.

[4] C. Adami and N. J. Cerf. von Neumann capacity of noisy quantum channels. *Physical Review A*, 56(5):3470–3483, Nov 1997.

[5] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[6] C. H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.

[7] J. Eisert and M. B. Plenio. A comparison of entanglement measures. *Journal of Modern Optics*, 46(1):145–154, 1999.

[8] M. B. Plenio and S. S. Virmani. An introduction to entanglement theory. In *Quantum information and coherence*, pages 173–209. Springer, 2014.

[9] A. Einstein, M. Born, H. Born, and I. Born. *The Born-Einstein Letters: Correspondence Between Albert Einstein and Max and Hedwig Born from 1916-1955, with Commentaries by Max Born.* Macmillan, 1971.

[10] D. Deutsch and R. Jozsa. Rapid solution of problems by quantum computation. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 439(1907):553–558, 1992.

[11] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, Oct 1997.

[12] T. Ono, R. Okamoto, and S. Takeuchi. An entanglement-enhanced microscope. *Nature Communications*, 4(1), Sep 2013.

[13] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger. Experimental quantum teleportation. *Nature*, 390(6660):575–579, Dec 1997.

[14] R. Valivarthi, S. I. Davis, C. Peña, S. Xie, N. Lauk, L. Narváez, J. P. Allmaras, A. D. Beyer, Y. Gim, M. Hussein, G. Iskander, H. L. Kim, B. Korzh, A. Mueller, M. Rominsky, M. Shaw, D. Tang, E. E. Wollman, C. Simon, P. Spentzouris, D. Oblak, N. Sinclair, and M. Spiropulu. Teleportation systems toward a quantum internet. *PRX Quantum*, 1:020317, Dec 2020.

[15] S. Annesley. *A Supplement to the Morning-exercise at Cripple-gate: Or, Several More Cases of Conscience Practically Resolved by Sundry Ministers*. Early English Books Online / EEBO. Thomas Cockerill, 1674.

[16] M. B. Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024–P08024, Aug 2007.

[17] S. R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[18] R. Orús and G. Vidal. Infinite time-evolving block decimation algorithm beyond unitary evolution. *Phys. Rev. B*, 78:155117, Oct 2008.

[19] S. R. White. Minimally entangled typical quantum states at finite temperature. *Phys. Rev. Lett.*, 102:190601, May 2009.

[20] F. Verstraete, J. J. Garcí a-Ripoll, and J. I. Cirac. Matrix product density operators: Simulation of finite-temperature and dissipative systems. *Physical Review Letters*, 93(20), Nov 2004.

[21] E. M. Stoudenmire and S. R. White. Minimally entangled typical thermal state algorithms. *New Journal of Physics*, 12(5):055026, May 2010.

[22] W. Li, S.-J. Ran, S.-S. Gong, Y. Zhao, B. Xi, F. Ye, and G. Su. Linearized tensor renormalization group algorithm for the calculation of thermodynamic properties of quantum lattice models. *Phys. Rev. Lett.*, 106:127202, Mar 2011.

[23] M. Binder and T. Barthel. Minimally entangled typical thermal states versus matrix product purifications for the simulation of equilibrium states and time evolution. *Phys. Rev. B*, 92:125119, Sep 2015.

[24] P. Czarnik and J. Dziarmaga. Variational approach to projected entangled pair states at finite temperature. *Phys. Rev. B*, 92:035152, Jul 2015.

[25] B.-B. Chen, L. Chen, Z. Chen, W. Li, and A. Weichselbaum. Exponential thermal tensor network approach for quantum lattice models. *Phys. Rev. X*, 8:031082, Sep 2018.

[26] A. Kshetrimayum, M. Rizzi, J. Eisert, and R. Orús. Tensor network annealing algorithm for two-dimensional thermal states. *Phys. Rev. Lett.*, 122:070502, Feb 2019.

[27] A. M. Alhambra and J. I. Cirac. Locally accurate tensor networks for thermal states and time evolution. *PRX Quantum*, 2(4), Nov 2021.

[28] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal. Time-dependent density-matrix renormalization-group using adaptive effective hilbert spaces. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(04):P04005, Apr 2004.

[29] A. E. Feiguin and S. R. White. Time-step targeting methods for real-time dynamics using the density matrix renormalization group. *Phys. Rev. B*, 72:020404, Jul 2005.

[30] J. J. García-Ripoll. Time evolution of matrix product states. *New Journal of Physics*, 8(12):305–305, Dec 2006.

[31] G. Vidal. Classical simulation of infinite-size quantum lattice systems in one spatial dimension. *Phys. Rev. Lett.*, 98:070201, Feb 2007.

[32] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete. Time-dependent variational principle for quantum lattices. *Phys. Rev. Lett.*, 107:070601, Aug 2011.

[33] M. P. Zaletel, Roger S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann. Time-evolving a matrix product state with long-ranged interactions. *Phys. Rev. B*, 91:165112, Apr 2015.

[34] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete. Unifying time evolution and optimization with matrix product states. *Phys. Rev. B*, 94:165116, Oct 2016.

[35] B. Vanhecke, L. Vanderstraeten, and F. Verstraete. Symmetric cluster expansions with tensor networks. *Phys. Rev. A*, 103:L020402, Feb 2021.

[36] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig. Time-evolution methods for matrix-product states. *Annals of Physics*, 411:167998, 2019.

[37] L. Hulthén. *Über das Austauschproblem eines Kristalles: Akad. avh. Sthlms högskola.* Arkiv för matematik, astronomi och fysik. Almqvist & Wiksell, 1938.

[38] R. Orbach. Linear antiferromagnetic chain with anisotropic coupling. *Phys. Rev.*, 112:309–316, Oct 1958.

[39] J. des Cloizeaux and J. J. Pearson. Spin-wave spectrum of the antiferromagnetic linear chain. *Phys. Rev.*, 128:2131–2135, Dec 1962.

[40] E. H. Lieb and W. Liniger. Exact analysis of an interacting bose gas. i. the general solution and the ground state. *Phys. Rev.*, 130:1605–1616, May 1963.

[41] R. B. Griffiths. Magnetization curve at zero temperature for the antiferromagnetic heisenberg linear chain. *Phys. Rev.*, 133:A768–A775, Feb 1964.

[42] E. H. Lieb and F. Y. Wu. Absence of mott transition in an exact solution of the short-range, one-band model in one dimension. *Phys. Rev. Lett.*, 20:1445–1448, Jun 1968.

[43] M. Ogata, M. U. Luchini, S. Sorella, and F. F. Assaad. Phase diagram of the one-dimensional t-j model. *Phys. Rev. Lett.*, 66:2388–2391, May 1991.

[44] H. A. Bethe. Zur theorie der metalle; 1, eigenwerte und eigenfunktionen der linearen atomkette. *Z. Phys.*, 71:205–226, 1931.

[45] A. Georges and G. Kotliar. Hubbard model in infinite dimensions. *Phys. Rev. B*, 45:6479–6483, Mar 1992.

[46] Y. Asano. *Mean-Field Theory of Superconductivity*, pages 21–37. Springer Singapore, Singapore, 2021.

[47] P. Ring. Relativistic mean field theory in finite nuclei. *Progress in Particle and Nuclear Physics*, 37:193–263, 1996.

[48] J. W. Negele. The mean-field theory of nuclear structure and dynamics. *Rev. Mod. Phys.*, 54:913–1015, Oct 1982.

[49] M. Bender, P.-H. Heenen, and P.-G. Reinhard. Self-consistent mean-field models for nuclear structure. *Rev. Mod. Phys.*, 75:121–180, Jan 2003.

[50] A. Rüegg, H.-H. Hung, E. Gull, and G. A. Fiete. Comparative DMFT study of the eg-orbital Hubbard model in thin films. *Physical Review B*, 89(8), Feb 2014.

[51] J.-M. Lasry and P.-L. Lions. Self-consistent mean-field models for nuclear structure. *Japanese Journal of Mathematics*, 2:229–260, Mar 2007.

[52] J.-Y. Le Boudec, D. McDonald, and J. Mundinger. A generic mean field convergence result for systems of interacting objects. In *Fourth International Conference on the Quantitative Evaluation of Systems (QEST 2007)*, pages 3–18, 2007.

[53] T. Parr, N. Sajid, and K. J. Friston. Modules or mean-fields? *Entropy*, 22(5), 2020.

[54] M. Christandl, R. König, G. Mitchison, and R. Renner. One-and-a-half quantum de finetti theorems. *Communications in Mathematical Physics*, 273(2):473–498, Mar 2007.

[55] R. Koenig and G. Mitchison. A most compendious and facile quantum de finetti theorem. *Journal of Mathematical Physics*, 50(1), Jan 2009.

[56] R. Renner. *Security of Quantum Key Distribution*. PhD thesis, Swiss Federal Institute of Technology Zurich, 2005.

[57] R. Renner. Symmetry of large physical systems implies independence of subsystems. *Nature Physics*, 3(9):645–649, Jul 2007.

[58] C. Krumnow, Z. Zimborás, and J. Eisert. A fermionic de Finetti theorem. *Journal of Mathematical Physics*, 58(12):122204, Dec 2017.

[59] B. de Finetti. La prévision : ses lois logiques, ses sources subjectives. *Annales de l'institut Henri Poincaré*, 7(1):1–68, 1937.

[60] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri. Detecting multipartite entanglement. *Physical Review A*, 71(3), Mar 2005.

[61] C. de Groot. *Topological Phases, Symmetries and Open Systems*. PhD thesis, Technical University Munich, 2022.

[62] J. Guth Jarkovský, A. Molnár, N. Schuch, and J. I. Cirac. Efficient description of many-body systems with matrix product density operators. *PRX Quantum*, 1:010304, Sep 2020.

[63] J. Guth Jarkovský, L. Piroli, and J. I. Cirac. Detecting spontaneous symmetry breaking with purifications. 2022. Manuscript in preparation.

[64] M. M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

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

[66] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000.

[67] A. Rényi. On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 547–561. Berkeley, California, USA, 1961.

[68] R. V. L. Hartley. Transmission of information 1. *Bell System technical journal*, 7(3):535–563, 1928.

[69] E. H. Lieb and M. B. Ruskai. Proof of the strong subadditivity of quantum-mechanical entropy. *Journal of Mathematical Physics*, 14(12):63–66, Dec 1973.

[70] B. M. Terhal, M. Horodecki, D. W. Leung, and D. P. DiVincenzo. The entanglement of purification. *Journal of Mathematical Physics*, 43(9):4286–4298, Sep 2002.

[71] P. Nguyen, T. Devakul, M. G. Halbasch, M. P. Zaletel, and Brian Swingle. Entanglement of purification: from spin chains to holography. *Journal of High Energy Physics*, 2018(1), Jan 2018.

[72] P. Liu, Y. Ling, C. Niu, and J.-P. Wu. Entanglement of purification in holographic systems. *Journal of High Energy Physics*, 2019(9), Sep 2019.

[73] K. Umemoto and Y. Zhou. Entanglement of purification for multipartite states and its holographic dual. *Journal of High Energy Physics*, 2018(10), Oct 2018.

[74] P. Caputa, M. Miyaji, T. Takayanagi, and K. Umemoto. Holographic entanglement of purification from conformal field theories. *Phys. Rev. Lett.*, 122:111601, Mar 2019.

[75] R.-Q. Yang, C.-Y. Zhang, and W.-M. Li. Holographic entanglement of purification for thermofield double states and thermal quench. *Journal of High Energy Physics*, 2019(1), Jan 2019.

[76] K. B. Velni, M. R. M. Mozaffar, and M. H. Vahidinia. Some aspects of entanglement wedge cross-section. *Journal of High Energy Physics*, 2019(5), May 2019.

[77] C. Akers and P. Rath. Entanglement wedge cross sections require tripartite entanglement. *Journal of High Energy Physics*, 2020(4), Apr 2020.

[78] J. Kudler-Flam and S. Ryu. Entanglement negativity and minimal entanglement wedge cross sections in holographic theories. *Phys. Rev. D*, 99:106014, May 2019.

[79] K. Umemoto and T. Takayanagi. Entanglement of purification through holographic duality. *Nature Physics*, 14(6):573–577, Mar 2018.

[80] J. Maldacena. The large N limit of superconformal field theories and supergravity. *International Journal of Theoretical Physics*, 38(4):1113–1133, 1999.

[81] S. Ryu and T. Takayanagi. Holographic derivation of entanglement entropy from the anti–de Sitter space/conformal field theory correspondence. *Physical Review Letters*, 96(18), May 2006.

[82] S. W. Hawking. Information loss in black holes. *Physical Review D*, 72(8), Oct 2005.

[83] S. Sachdev. Strange metals and the AdS/CFT correspondence. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(11):P11022, Nov 2010.

[84] M. A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83:436–439, Jul 1999.

[85] Y. He, A. B. Hamza, and H. Krim. A generalized divergence measure for robust image registration. *IEEE Transactions on Signal Processing*, 51(5):1211–1220, 2003.

[86] L. J. Rogers. A generalized divergence measure for robust image registration. *Messenger of Mathematics*, XVII(10):145–150, 1888.

[87] P. Coleman. *Introduction to Many-Body Physics*. Cambridge University Press, 2015.

[88] G. C. Wick. The evaluation of the collision matrix. *Phys. Rev.*, 80:268–272, Oct 1950.

[89] L. Hackl, T. Guaita, T. Shi, J. Haegeman, E. Demler, and J. I. Cirac. Geometry of variational methods: dynamics of closed quantum systems. *SciPost Phys.*, 9:48, 2020.

[90] M. C. Gutzwiller. Effect of correlation on the ferromagnetism of transition metals. *Phys. Rev. Lett.*, 10:159–162, Mar 1963.

[91] T. Guaita, L. Hackl, T. Shi, E. Demler, and J. I. Cirac. Generalization of group-theoretic coherent states for variational calculations. *Physical Review Research*, 3(2), May 2021.

[92] T. Guaita. *Quantum Variational Methods for Gaussian States and Beyond*. PhD thesis, Technical University Munich, 2022.

[93] R. Orús. Tensor networks for complex quantum systems. *Nature Reviews Physics*, 1(9):538–550, Aug 2019.

[94] U. Schollwöck. The density-matrix renormalization group in the age of matrix product states. *Annals of Physics*, 326(1):96–192, Jan 2011.

[95] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Reviews of Modern Physics*, 93(4), Dec 2021.

[96] R. Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, Oct 2014.

[97] M. Fannes, B. Nachtergaele, and R. Werner. Finitely correlated states on quantum spin chains. *Communications in Mathematical Physics*, 144, 03 1992.

[98] D. Perez-Garcia, F. Verstraete, M. Wolf, and J. Cirac. Matrix product state representations. *Quantum Information & Computation*, 7:401–430, Jul 2007.

[99] K. Noh, L. Jiang, and B. Fefferman. Efficient classical simulation of noisy random quantum circuits in one dimension. *Quantum*, 4:318, Sep 2020.

[100] D. N. Page. Average entropy of a subsystem. *Physical Review Letters*, 71(9):1291–1294, Aug 1993.

[101] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac. Area laws in quantum systems: Mutual information and correlations. *Phys. Rev. Lett.*, 100:070502, Feb 2008.

[102] S. O. Scalet, A. M. Alhambra, G. Styliaris, and J. I. Cirac. Computable Rényi mutual information: Area laws and correlations. *Quantum*, 5:541, Sep 2021.

[103] G. De las Cuevas, N. Schuch, D. Pé rez-García, and J I. Cirac. Purifications of multipartite states: limitations and constructive methods. *New Journal of Physics*, 15(12):123021, Dec 2013.

[104] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac. Computational complexity of projected entangled pair states. *Physical Review Letters*, 98(14), Apr 2007.

[105] L. G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8(2):189–201, 1979.

[106] G. Scarpa, A. Molná r, Y. Ge, J. J. García-Ripoll, N. Schuch, D. Pérez-García, and S. Iblisdir. Projected entangled pair states: Fundamental analytical and numerical limitations. *Physical Review Letters*, 125(21), Nov 2020.

[107] J.-Y. Chen, L. Vanderstraeten, S. Capponi, and D. Poilblanc. Non-Abelian chiral spin liquid in a quantum antiferromagnet revealed by an iPEPS study. *Phys. Rev. B*, 98:184409, Nov 2018.

[108] X. Liang, S.-J. Dong, and L. He. Hybrid convolutional neural network and projected entangled pair states wave functions for quantum many-particle states. *Phys. Rev. B*, 103:035138, Jan 2021.

[109] J.-G. Liu, Y.-H. Zhang, Y. Wan, and L. Wang. Variational quantum eigensolver with fewer qubits. *Phys. Rev. Research*, 1:023025, Sep 2019.

[110] W.-Y. Liu, S.-S. Gong, Y.-B. Li, D. Poilblanc, W.-Q. Chen, and Z.-C. Gu. Gapless quantum spin liquid and global phase diagram of the spin-1/2 j1-j2 square antiferromagnetic heisenberg model. *Science Bulletin*, 67(10):1034–1041, 2022.

[111] S. Lu, M. Kanász-Nagy, I. Kukuljan, and J. I. Cirac. Tensor networks and efficient descriptions of classical data, 2021.

[112] F. Verstraete and J. I. Cirac. Matrix product states represent ground states faithfully. *Physical Review B*, 73(9), Mar 2006.

[113] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac. Entropy scaling and simulability by matrix product states. *Physical Review Letters*, 100(3), Jan 2008.

[114] J. Sieber. *Recovery Maps for Approximate Markov Chains —– a Matrix Product State Perspective*. Bachelor's thesis, Technical University of Munich, 2017.

[115] O. Fawzi and R. Renner. Quantum conditional mutual information and approximate Markov chains. *Communications in Mathematical Physics*, 340(2):575–611, Sep 2015.

[116] T. Kuwahara, K. Kato, and F. G. S. L. Brandão. Clustering of conditional mutual information for quantum Gibbs states above a threshold temperature. *Physical Review Letters*, 124(22), Jun 2020.

[117] M. B. Hastings and T. Koma. Spectral gap and exponential decay of correlations. *Communications in Mathematical Physics*, 265(3):781–804, Apr 2006.

[118] I. Arad, A. Kitaev, Z. Landau, and U. Vazirani. An area law and sub-exponential algorithm for 1D systems, 2013.

[119] F. G. S. L. Brandão and M. Horodecki. Exponential decay of correlations implies area law. *Communications in Mathematical Physics*, 333(2):761–798, Jan 2015.

[120] A. Beekman, L. Rademaker, and J. van Wezel. An introduction to spontaneous symmetry breaking. *SciPost Physics Lecture Notes*, Dec 2019.

[121] W. Ludwig and C. Falter. *Symmetries in Physics: Group Theory Applied to Physical Problems*. Springer Series in Solid-State Sciences. Springer Berlin Heidelberg, 2012.

[122] R. Li and B. Bowerman. *Symmetry Breaking in Biology: A Subject Collection from Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Laboratory Series. Cold Spring Harbor Laboratory Press, 2010.

[123] J. Fraser. Spontaneous symmetry breaking in finite systems. *Philosophy of Science*, 83, Oct 2015.

[124] C. K. Majumdar and D. K. Ghosh. On next-nearest-neighbor interaction in linear chain. i. *Journal of Mathematical Physics*, 10(8):1388–1398, 1969.

[125] C. K. Majumdar. Antiferromagnetic model with known ground state. *Journal of Physics C: Solid State Physics*, 3(4):911–915, Apr 1970.

[126] J. Hubbard and B. H. Flowers. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 276(1365):238–257, 1963.

[127] H. A. Gersch and G. C. Knollman. Quantum cell model for bosons. *Phys. Rev.*, 129:959–967, Jan 1963.

[128] M. Takahashi. Half-filled Hubbard model at low temperature. *Journal of Physics C: Solid State Physics*, 10(8):1289–7301, Apr 1977.

[129] P. van Dongen and D. Vollhardt. Exact solution and thermodynamics of the Hubbard model with infinite-range hopping. *Phys. Rev. B*, 40:7252–7255, Oct 1989.

[130] R. T. Scalettar. An introduction to the hubbard hamiltonian. In *Quantum Materials: Experiments and Theory Modeling and Simulation, Vol. 6*. Verlag des Forschungszentrum Jülich, 2016.

[131] M. Raczkowski, R. Peters, T. T. Phù ng, N. Takemori, F. F. Assaad, A. Honecker, and J. Vahedi. Hubbard model on the honeycomb lattice: From static and dynamical mean-field theories to lattice quantum monte carlo simulations. *Physical Review B*, 101(12), Mar 2020.

[132] A. Mielke. Exact ground states for the Hubbard model on the Kagome lattice. *Journal of Physics A: Mathematical and General*, 25(16):4335–4345, Aug 1992.

[133] K. Ferhat and A. Ralko. Phase diagram of the 1/3-filled extended Hubbard model on the Kagome lattice. *Physical Review B*, 89(15), Apr 2014.

[134] A. Szasz and J. Motruk. Phase diagram of the anisotropic triangular lattice Hubbard model. *Physical Review B*, 103(23), Jun 2021.

[135] M. Gaudin. Un systeme a une dimension de fermions en interaction. *Physics Letters A*, 24(1):55–56, 1967.

[136] C. N. Yang. Some exact results for the many-body problem in one dimension with repulsive delta-function interaction. *Phys. Rev. Lett.*, 19:1312–1315, Dec 1967.

[137] W. Metzner and D. Vollhardt. Correlated lattice fermions in $d = \infty$ dimensions. *Phys. Rev. Lett.*, 62:324–327, Jan 1989.

[138] R. Németh. Hubbard model with infinite-range hopping. *Physica A: Statistical Mechanics and its Applications*, 163(2):672–682, 1990.

[139] F. S. Nogueira and E. V. Anda. Study of a toy model and its relation to the Hubbard model with infinite range hopping. *International Journal of Modern Physics B*, 10(27):3705—-3715, Dec 1995.

[140] G. Fano and F. Ortolani. Hubbard model with unconstrained hopping on a finite number of sites. *Phys. Rev. B*, 49:7205–7209, Mar 1994.

[141] M. Salerno. Exact analytical solutions for the Hubbard model with unconstrained hopping. *Physica Scripta*, 54(1), July 1996.

[142] P. Pieri. An exact result on ferromagnetism in the Hubbard model with an infinite-range hopping. *Physics Letters A*, 235(2):183–185, Oct 1997.

[143] C. V. Kraus, M. Lewenstein, and J. I. Cirac. Ground states of fermionic lattice Hamiltonians with permutation symmetry. *Physical Review A*, 88(2), Aug 2013.

[144] J. C. Slater. The theory of complex spectra. *Phys. Rev.*, 34:1293–1322, Nov 1929.

[145] P. Jordan and E. Wigner. Über das paulische Äquivalenzverbot. *Zeitschrift für Physik*, 47, 1928.

[146] C. V. Kraus, M. M. Wolf, J. I. Cirac, and G. Giedke. Pairing in fermionic systems: A quantum-information perspective. *Physical Review A*, 79(1), Jan 2009.

[147] D. Vollhardt, K. Byczuk, and M. Kollar. Dynamical mean-field theory. In *Springer Series in Solid-State Sciences*, pages 203–236. Springer Berlin Heidelberg, Aug 2011.

[148] J. Zang, J. Wang, J. Cano, A. Georges, and A. J. Millis. Dynamical mean-field theory of Moiré bilayer transition metal dichalcogenides: Phase diagram, resistivity, and quantum criticality. *Phys. Rev. X*, 12:021064, Jun 2022.

[149] A. Paul and T. Birol. Applications of DFT + DMFT in materials science. *Annual Review of Materials Research*, 49(1):31–52, Jul 2019.

[150] U. Schneider, L. Hackermüller, S. Will andHh. Best, I. Bloch, T. A. Costi, R. W. Helmes, D. Rasch, and A. Rosch. Metallic and insulating phases of repulsively interacting fermions in a 3D optical lattice. *Science*, 322(5907):1520–1525, Dec 2008.

[151] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner. Nonequilibrium dynamical mean-field theory and its applications. *Rev. Mod. Phys.*, 86:779–837, Jun 2014.

[152] P. W. Anderson. Localized magnetic states in metals. *Phys. Rev.*, 124:41–53, Oct 1961.

[153] N.L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Volume 1.* Wiley Series in Probability and Statistics. Wiley, 1994.

[154] J. Diestel, H. Jarchow, and A. Tonge. *Absolutely Summing Operators.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1995.

[155] W. Rudin. *Functional Analysis.* International Series in Pure & Applied Mathematics. McGraw-Hill Science/Engineering/Math, 2 edition, January 1991.

[156] A. Mandal. *Markushevich Bases and Auerbach Bases in Banach Spaces.* University of windsor, A Major Research Paper, 2018.

[157] P. Hajek, V.M. Santalucia, J. Vanderwerff, and V. Zizler. *Biorthogonal Systems in Banach Spaces*. CMS Books in Mathematics. Springer, 2008.

[158] H. Hahn. Über lineare Gleichungssysteme in linearen Räumen. *Journal für die reine und angewandte Mathematik*, 157:214–229, 1927.

[159] S. Banach. Sur les fonctionnelles linéaires. *Studia Mathematica*, 1(1):211–216, 1929.

[160] S. Banach. Sur les fonctionnelles linéaires ii. *Studia Mathematica*, 1(1):223–239, 1929.

[161] H. Auerbach. *O polu krzywych wypukłych o średnicach sprzężonych*. PhD thesis, University of Lviv, 1929.

[162] L. Alaoglu. Weak topologies of normed linear spaces. *The Annals of Mathematics*, 41(1):252–267, Jan 1940.

[163] M. Kormos and Z. Zimborás. Temperature driven quenches in the ising model: appearance of negative rényi mutual information. *Journal of Physics A: Mathematical and Theoretical*, 50(26):264005, Jun 2017.

[164] M. Berta, K. P. Seshadreesan, and M. M. Wilde. Rényi generalizations of the conditional quantum mutual information. *Journal of Mathematical Physics*, 56(2):022205, Feb 2015.

[165] D. Petz. Quasi-entropies for finite quantum systems. *Reports on Mathematical Physics*, 23(1):57–65, February 1986.

[166] F. Leditzky, C. Rouzé, and N. Datta. Data processing for the sandwiched Rényi divergence: a condition for equality. *Letters in Mathematical Physics*, 107(1):61–80, Nov 2016.

[167] K. Audenaert and N. Datta. α-z-Rényi relative entropies. *Journal of Mathematical Physics*, 56:022202, 02 2015.