Detecting and Understanding Efficient Structures in Finite Fermionic Systems

Christian Detlef Krumnow

Berlin, 2017

Dissertation zur Erlangung des Grades eines Doktors der Naturwissenschaften,
eingereicht am Fachbereich Physik der Freien Universität Berlin

Freie Universität Berlin
Dahlem Center for Complex Quantum Systems
14195 Berlin, Germany

Erstgutachter:	Prof. Dr. Jens Eisert, Freie Universität Berlin
Zweitgutachter:	Dr. Christoph Karrasch, Freie Universität Berlin

Tag der Disputation: 23.05.2018

Contents

1. Introduction **8**

2. Technical Introduction **12**
2.1. Basic Notation and Notions 12
2.1.1. Finite Quantum Systems of Distinguishable Particles 13
2.1.2. Bookkeeping and Order: Partitions and Permutations 14
2.2. Finite Fermionic Systems 15
2.2.1. Jordan-Wigner Transformation 18
2.2.2. Mode Transformations 19
2.2.3. Superselection Rule 22
2.3. Reduced Density Matrices 23
2.3.1. Reduction of Modes 23
2.3.2. Reduction of Particles 24
2.4. Free Fermions and Gaussian States 26
2.5. Interacting Fermions 29
2.5.1. Hartree-Fock Approximation 31
2.5.2. Complexity Theory and Interacting Fermions 32
2.6. Kinematic Constraints of Finite Fermionic Systems 34
2.6.1. One-Body Density Matrix Constraints 36
2.6.2. Two-Body Density Matrix Constraints 38
2.6.3. Local Constraints from Hermiticity 38
2.7. Summary 41

3. Capturing Correlations with Tensor Network States **42**
3.1. Tensor Network States 44
3.1.1. Tensor Network Decompositions 44
3.1.2. Correlations in Tensor Network States 47
3.1.3. Matrix Product States: Normal Forms and Operations 50
3.1.4. Symmetric Tensor Network States 54
3.2. Density Matrix Renormalization Group and Related Algorithms 57
3.2.1. Ground State Search 58
3.2.2. Convergence of DMRG 62
3.2.3. Excited States 64
3.3. Mode Transformation in Tensor Network States on Fermionic Systems . . 69
3.3.1. Cost Function 72
3.3.2. Optimization Set 72
3.3.3. Algorithm and Examples 74

3.3.4. The Bird's-Eye View . . . 82
3.4. Summary . . . 83

4. Towards an Understanding of Mean Field Approaches 85
4.1. Quantum de Finetti Theorems . . . 86
4.2. A Fermionic Mode de Finetti Theorem . . . 89
4.2.1. Suppression of Oddness and a Fermionic de Finetti Theorem . . . 92
4.2.2. Proof of Lem. 4 and Thm. 5 . . . 93
4.3. Structure of the Mode Product States and Applications . . . 98
4.3.1. Few Local Modes and Mean Field Approximations . . . 98
4.3.2. Extension of Hudson's Theorem . . . 99
4.4. Summary . . . 102

5. Gaussification: Relaxation under Free Hamiltonians 104
5.1. Equilibration and Thermalization of Closed Quantum Systems . . . 105
5.2. Relaxation of Free Fermionic Systems . . . 109
5.2.1. Conditions and Intuition for Gaussification . . . 110
5.2.2. Technical Comments on Gaussification . . . 115
5.2.3. Physical Implications and Relation to Thermalization . . . 115
5.3. Summary . . . 119

6. Outlook and Open Research Question 121

A. Wick's Theorem and Definitions of Gaussian States 138
A.1. Wick's Theorem . . . 138
A.2. Definitions of Gaussian States . . . 140

B. Details on the Two-Site DMRG 142
B.1. Presummed Operators and Costs of the Two Site DMRG . . . 142
B.2. Mode Transformation of Presummed Operators and Costs . . . 144

C. Comment on the Developed Code Structure for Tensor Network States 146

D. Optimizing with Unitary Constraints 148
D.1. Sets with Unitary Constraints . . . 148
D.2. Optimization over Grassmannians . . . 149

E. Proof of the Extension of Huddson's Central Limit Theorem 151

F. Proof of Gaussification 153
F.1. Reformulation of the Problem . . . 153
F.2. Restriction to the Lieb-Robinson Cone . . . 154
F.3. Factorization by the Exponentially Suppressed Correlations . . . 155
F.4. Suppression by Delocalizing Transport . . . 158
F.5. Putting Everything Together . . . 161

G. Delocalizing Transport in Free Models **162**

H. Backmatter **165**
H.1. Publications of the Author of the Thesis 165
H.2. Supervision of Bachelor and Master Students 166
H.3. Acknowledgment 167
H.4. Abstract 168
H.5. Zusammenfassung 169
H.6. Eigenständigkeitserklärung 170

1. Introduction

The concept of particles being truly indistinguishable from each other is one of the hallmarks and cornerstones of early quantum mechanics separating it from classical physics. The realization that additional symmetry constraints need to be imposed when describing multiple identical particles led to a deeper understanding and intuitive explanation of properties of matter. Especially the antisymmetry of fermionic systems has profound consequences. The Pauli exclusion principle allows us to (at least roughly) understand the formation of chemical bonds, the emergence of magnetism in solids or the stability and solidness of matter as such. Understanding fermionic systems and their proper description impacts our comprehension of physics on all scales; from the structure of molecules or the precision of a single electron transistor to insights about the stability of neutron stars. Furthermore we discovered and gained a deeper understanding into numerous ordinary and exotic phases of matter by the theoretical, numerical and experimental study of various models and systems in the past century. Hence, since the first days of quantum mechanics, it has been an important task of modern theoretical physics to comprehend and describe fermionic systems.

The importance and omnipresence of fermionic systems is, however, complemented by an intrinsic hardness of their description. This hardness is implicitly illustrated by the fact that over the past decades a zoo of numerical methods and schemes have been developed which are able to successfully simulate static and dynamic properties of fermionic systems in different limiting cases only and severe approximations have to be imposed for describing fermionic models analytically. Today we know that this seemingly hardness is not an illusion or lack of imagination as it can be confirmed and quantified using concepts of computational complexity theory. Even a quantum computer can not approximate the ground state of all interacting fermionic systems efficiently and classical computers are not able to describe all fermionic systems using the Hartree-Fock methods, i.e., within one of their roughest approximations [1–3]. Multiple insights and tools allow us to overcome these limitations. Restricting energy scales involved and realizing that many effects and properties of physical systems are only sensitive to certain degrees of freedom allows to formulate effective theories such as the Fermi liquid theory or to use approximations as the Hartree-Fock approximation in order to faithfully capture many aspects of a model [4, 5]. The possibility of such effective descriptions reveals that many systems possess and underlying (often hidden) additional structure. Such structures can for instance be given by symmetries or a locality of the Hamiltonian describing the system and impact the correlations present in natural states, e.g., thermal or ground states. Furthermore they emerge dynamically. The expected thermalization of generic quantum systems allows us often to restrict ourselves to a thermalized equilibrium description of a quantum system neglecting many or even any non-thermal parts of the initial state

and by this vastly simplifying its characterization.

In addition to these conceptual insight, powerful numerical tools have been developed for approximating the ground state of a fermionic system or capturing its time evolution. Due to the overall hardness of the tasks there is no ideal universal strategy to solve these problems but different methods which need different underlying structures in the simulated system operate in various niches. Methods such as density functional theory (DFT), the configuration interaction (CI) or coupled cluster method [6, 7] allow to tackle large but only moderately correlated systems. The correlation structure resolvable by DFT for instance depends strongly on the approximation to the universal functional used. Coupled cluster and CI on the other hand yield systematically extendable methods which in principle interpolate between the mean field Hartree-Fock approximation and exact solution of the system. Tensor network state (TNS) based methods are on the other hand more flexible and able to capture strongly correlated settings and states at the expense of higher computational costs. They are most efficient if the correlations of the state to be approximated are restricted by area laws [8] while in finite systems a systematic increase of the number of variational parameters interpolates between a mean field approach and the exact solution. Next to providing a numerical tool, the investigation of TNS motivated and led to different fundamental insights for instance about the correlation structure of ground states [9] or can even provide the exact ground state [10]. Their hybrid status of numerical and analytical tool allow to employ TNS to a wide range of physical applications beyond the direct simulation of different quantum systems such as the detection and classification of symmetry protected topological order [11, 12] or the investigation of the AdS/CFT correspondence (see for instance [13]).

The additional structures exploited, either implicitly or explicitly, by the various methods discriminate the respective system from the hard instances of, for instance, the interacting ground state problem and allow to either successfully describe its essential features in analytic terms or to simulate it numerically. Identifying and understanding these structures in more detail opens a window into interacting fermionic systems and their associated physics and allows us to formulate more powerful numerical schemes circumventing their generic hardness. It is the aim of this thesis to contribute further pieces to both of these endeavors and to extend our knowledge about and to improve our tools for simulating fermionic systems. The thesis contains three parts.

In the first part in Ch. 2 we provide an overall introduction into finite fermionic systems. We collect results from different communities underlining the general discussion above. We will carefully discuss the efficient solution of non-interacting systems using mode transformations as well as the complexity of interacting systems and general kinematic constraints such as the Pauli principle. Furthermore, we introduce the notation and concepts needed in the following two parts which should be general and flexible enough in order to numerically simulate realistic interacting systems and formulate mathematical assumptions and results in a rigorous as well as physically intuitive fashion at the same time. We find the second quantized formulation of finite fermionic systems to be ideal here and we introduce and discuss different notions of reduced states for them. We expect the reader to have a good foundation in second quantization and

linear algebra as we only remind on different concepts and link them among each other but will not carefully introduce them from first scratch. Next to presenting a comprehensive collection of, to us, important aspects of fermionic systems we also provide a few small novel insights into the structure of fermionic reduced states and the relation of antisymmetry and the Pauli principle.

In Ch. 3 we focus on the detection of suitable efficient structures in fermionic systems using tensor network methods. We start by reviewing TNS from a practical point of view as a numerical tool. We discuss in detail how a density matrix renormalization group algorithm (DMRG) can be implemented which allows for the simulation of general interacting fermionic systems. We cover the basic routines needed and review how to account for symmetries of the system. Furthermore, we discuss in detail how to extend heuristic error measures to excited states in general and highlight the structure of local minima of local TNS update schemes such as the DMRG. At the heart of this part however lies the realization that in fermionic systems TNS capture correlations partially in the wrong picture. In tensor network states the amount of correlation between different fermionic modes is limited such that even a single Slater determinant, a fermionic particle product state, might not be approximable by TNS if the wrong single particle basis is chosen. We present a scheme of how to combine TNS and mode transformations which allows to take into account high entanglement effects which go beyond strict area laws. We explain how to extend an existing DMRG implementation and by this develop ground state approximation and real time evolution schemes for non-local interacting fermionic systems which optimize the single particle basis to the correlation structure of the state. Their functioning is illustrated on physical examples and we show that if they are present as an underlying structure we are able to identify more optimal single particle basis which allow for a more efficient description of the system.

In the last part we want to understand the emergence of structures in fermionic systems on more rigorous grounds. It is divided into two subparts which study on the one hand structures induced by symmetries that allow for faithful mean field approximations and on the other hand structures that emerge dynamically in the spirit of thermalization. It is frequently observed that seemingly complex systems can be well approximated using mean field methods such as the Hartree-Fock approximation. However, we typically do not understand why such approximations work and can not predict the error obtained from imposing it on a given specific system. In Ch. 4 we discuss different quantum de Finetti theorems which are a tool to certify mean field approximations in specific settings and well established for distinguishable and bosonic particles. Concretely, they derive an i.i.d. product structure from a permutation invariance of the state of interest. We show that in fermionic systems the intrinsic antisymmetry allows to derive similar results even under the assumption of a relaxed definition of permutation invariance. We discuss how the obtained mode product states are related to Gaussian states and link the obtained theorem to mean field approximations and exemplify how it yields natural extensions to existing theorems about the structure of fermionic states. In the second subpart in Ch. 5 we investigate the non-equilibrium dynamics of fermionic systems. Despite significant progress over the past decades, important questions are still open

concerning the mechanism behind equilibration and thermalization of closed quantum systems. It is for instance unclear under which conditions and after which time a generic closed quantum system will thermalize and only conjectured how the system loses its memory on a majority of its initial conditions. In order to progress in these challenging problems, we consider the dynamics of free fermionic systems and find a to free systems particular relaxation process. Initial states with short range correlations evolved in a free system with sufficient transport will become locally Gaussian after a short time. Any non-Gaussian initial correlation is quickly smeared over the complete system and by this essentially lost for practical observables such that the non-equilibrium dynamics can be captured within a purely Gaussian settings. We carefully discuss the requirements and intuition behind the result. The Gaussification of free systems is reminiscent to the convergence to a generalized Gibbs ensemble and we explain how in special cases it links to equilibration and thermalization. It furthermore explains and predicts the occurrence of Gaussian states in natural settings.

We conclude our discussion with an outlook of future research direction extending the ideas of this thesis in Ch. 6. Furthermore we present details to the proofs and extended comments on various topics in the appendix.

2. Technical Introduction

In this first chapter we want to settle the basic notation and introduce the concepts used in the remainder of this thesis. The used formalism should hereby be general enough for describing realistic fermionic systems appearing from first principles in the context of quantum chemistry and, at the same time, allow for deeper theoretical insights valid in condensed matter physics systems. Concerning the technical aspects we hence aim to take the middle ground in the sense that our formulations should not be too technical but simultaneously also not based on uncontrolled approximations. Throughout this thesis we want to state all assumptions needed for the derived results as clearly as possible and still be able to connect them to intuitive physical properties.

We find that a formulation in terms of finite fermionic systems is best suited for our needs. It appears naturally in the context of quantum chemistry where the number of relevant degrees of freedom can often be assumed to be finite and allows to consider lattice systems of condensed matter physics by taking the thermodynamic limit if necessary. Furthermore, it allows us to complement analytic insights easily with numerical calculations.

In the following, we will put a special emphasis on reviewing the computational complexity of obtaining static and dynamic properties in fermionic systems. Whereas free systems are found to be efficiently solvable on the level of individual particles, interacting fermionic systems pose in general computationally hard problems. Furthermore, we will carefully discuss different aspects of the fermionic state space. Next to technical details such as the justification of the superselection rule or the reduction to subregions of a system or few particles we review several kinematic constraints which generalize and complement the well known Pauli exclusion principle. In order to keep the presentation compact, we assume the reader to have a good foundation in second quantization and linear algebra and repeat concepts only briefly in order to highlight important conceptual links or introduce the notation. For more details and background we recommend [14, 15].

2.1. Basic Notation and Notions

In this section we want to establish some of the basic notation and concepts used in the following of the thesis. We start by introducing our notation first for finite quantum systems of distinguishable particles and simultaneously review a few basic concepts and ideas which will reappear for fermionic systems later on. After that, we bundle a few further miscellaneous technical concepts, notations and conventions used at different places in this thesis concerning especially partitions of sets and the arrangement of sums

and products.

2.1.1. Finite Quantum Systems of Distinguishable Particles

All results contained in this thesis are obtained for finite-dimensional quantum systems. Therefore let us introduce the most basic notation and ideas used throughout this thesis in the following. In the setting of V distinguishable particles, each to be described by a finite-dimensional Hilbert space $\mathbb{C}^d$, the full Hilbert space of the system is given by $\mathcal{H} = \otimes_{k=1}^{V} \mathbb{C}^d$. We use the standard Dirac notation and denote state vectors and their dual by $|\psi\rangle \in \mathcal{H}$ and $\langle\psi|$ respectively, which makes it natural to denote the inner product of $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ by $\langle\psi|\phi\rangle$. If not stated otherwise or clear from the notation, state vectors will always be assumed to be normalized with respect to the norm induced from the inner product of $\mathcal{H}$. For any Hilbert space $\mathcal{H}$ we denote by $\mathcal{D}(\mathcal{H})$ the set of density matrices, so non-negative, Hermitian operators with trace 1, and by $\mathcal{B}(\mathcal{H})$ the set of linear operators on $\mathcal{H}$, bounded in operator norm, i.e.,

$$\forall A \in \mathcal{B}(\mathcal{H}) : \|A\| = \sup_{|\psi\rangle \in \mathcal{H}: \||\psi\rangle\|=1} \|A|\psi\rangle\| < \infty. \tag{2.1}$$

An observable of the system corresponds to a Hermitian operator, so $A \in \mathcal{B}(\mathcal{H})$ with $A^\dagger = A$, to which we associate the in principle via an experiment accessible expectation value $\mathrm{tr}(\rho A)$ with respect to the state $\rho \in \mathcal{D}(\mathcal{H})$. As usual for a general operator $A : (\mathbb{C}^d)^{\times V} \to (\mathbb{C}^d)^{\times V}$ we denote by $\mathrm{supp}(A) \subset [V]$ the support of A, i.e., the set of copies of $\mathbb{C}^d$ on which A does not act as the identity. Furthermore, $\mathrm{spec}(A)$ denotes the sequence of eigenvalues of A.

Given two states $\rho_1, \rho_2 \in \mathcal{D}(\mathcal{H})$ we want to measure their distance in the state space using how well we can distinguish between both using physical operations, i.e., probing expectation values of observables. When we want to quantify the difference of two states we will therefore be interested in the one-norm of the difference $\rho_1 - \rho_2$ which is connected to expectation values via

$$\|\rho_1 - \rho_2\|_1 = \sup_{\substack{A \in \mathcal{B}(\mathcal{H}): \\ A^\dagger = A, \|A\|=1}} |\,\mathrm{tr}\left((\rho_1 - \rho_2)A\right)|. \tag{2.2}$$

In addition we introduce the notion of locality. An operator A is called k-local, if it acts non-trivially on some set of at most k constituents of the systems only and as the identity for the remaining part of the system, i.e., $|\,\mathrm{supp}(A)| \leq k$. In many cases in the following the specific value of k is not important and we will often only speak of local operators. By this we mean that the operator under consideration is k-local for some small k which is independent of the system size. Given a subregion S of the full lattice $[V]$ and a state $\rho \in \mathcal{D}(\mathcal{H})$ we denote the reduced state on the subsystem associated to the region S as $\mathrm{tr}_{S^c}(\rho)$. Here and throughout the thesis we use for a given $N \in \mathbb{N}$ the notation $[N] = \{n \in \mathbb{N} | n \leq N\}$. The reduced state is then defined as the state in $\mathcal{D}(\mathcal{H}_S)$ which has the same expectation values as ρ for any local observable A supported on S only, so as the state which looks like ρ on S under any consideration. Denoting by

$\tau_S : \mathcal{B}(\mathcal{H}_S) \hookrightarrow \mathcal{B}(\mathcal{H})$ the embedding of operators acting on $\mathcal{H}_S$ into the ones acting on $\mathcal{H}$ by extending their actions trivially with the identity we the implicitly define $\mathrm{tr}_{S^c}(\rho)$ via the relation

$$\forall A \in \mathcal{B}(\mathcal{H}_S) \quad : \quad \mathrm{tr}(\mathrm{tr}_{S^c}(\rho)A) = \mathrm{tr}(\rho\tau_S(A)). \tag{2.3}$$

An explicit expression for $\mathrm{tr}_{S^c}(\rho)$ is easily constructed using the partial trace over S^c, hence the notation.

One very remarkable difference between quantum mechanics and classical physics is the appearance of new kinds of correlation phenomena arising from entanglement. These were early noted to give rise to seemingly non-local effects and constitute a resource used for quantum computation in modern conceptions [16]. For a pure state $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathcal{H}$ in a finite system, the amount of entanglement established between a subsystem S and the remaining systems can be measured considering entropies of the reduced state $\mathrm{tr}_{S^c}\rho$. Given a density matrix ρ the Rényi entropies are defined for $\alpha > 0$ by

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \ln \mathrm{tr}\, \rho^\alpha \tag{2.4}$$

where the well known von Neumann entropy S_{vN} is connected to the Rényi entropies by the limit

$$\lim_{\alpha\to 1} S_\alpha(\rho) = -\mathrm{tr}(\rho \ln \rho) = S_{\mathrm{vN}}(\rho). \tag{2.5}$$

The von Neumann entropy often plays an exceptional role among different correlation measures due to its connection to the thermodynamic entropy and the quantification of information and uncertainty on a conceptional side as well as special unique technical properties such as strong subadditivity [16].

In many cases it is natural to assign a geometric structure to the system which leads us to quantum lattice systems. In this case, the system consists of a finite patch of an n-dimensional lattice $\mathbb{Z}^n$. To each of the V sites we then associate a local Hilbert space $\mathbb{C}^d$ describing a particle on that site giving rise to the same global Hilbert space as above. In addition to the global Hilbert space the lattice structure gives rise to a natural definition of a geometric distance of different constituents of the system. Especially, we can then introduce geometric locality, e.g., spatially local Hamiltonians which can be written as $H = \sum_S h_S$ with h_S acting non-trivially only on region S which is assumed to have some finite diameter bounded by a constant. If not stated differently, in the following with local Hamiltonian we will refer to a geometrically local one while we will distinguish geometric and plain locality for general operators and observables if not clear from the context.

Before we turn to fermionic systems and discuss in detail in how far the concepts introduced here carry over or need to be adapted when additional antisymmetry constraints are imposed on the state space we want to introduce some further notation.

2.1.2. Bookkeeping and Order: Partitions and Permutations

The idea of grouping elements in patches using partitions in order to account for different contribution within a compact notation will repeatedly appear in different contexts in

the following chapters. We therefore want to introduce a bit of notation here which at the same time introduces some of the general mindset.

For $N \in \mathbb{N}$ we denote by S_N the symmetric group collecting all possible permutations of N elements. As usual $\mathrm{sign}(\pi)$ is the sign of a permutation $\pi \in S_N$, which is 1 if π involves an even number of transpositions and -1 otherwise. For a set J we denote the set of all partitions of J by

$$\mathcal{P}(J) = \{P : P \,\text{partition of}\, J\}. \tag{2.6}$$

Hereby $P = \{p^{(1)}, \ldots, p^{(|P|)}\}$ with $p^{(j)} \subset J$ for $j \in [|P|]$ is a partition of J iff $\cup_j p^{(j)} = J$ and all $p^{(j)}$ are disjoint among each other. We will use partitions in order to bookkeep fermionic operators and therefore want to specify how products and sequences arising are arranged in order to keep track of potential signs from the anti-commuting character of fermions. If J is ordered, we assume all parts $p \in P$ for any $P \in \mathcal{P}(J)$ to be ordered increasingly. Products, sums and sequences taken over the elements of a part of a partition are then assumed to be taken in that order. In addition we assume that products, sums and sequences taken over the partition, i.e., $\sum_{p\in P}$ for a $P \in \mathcal{P}(J)$, are taken in the order that the minimal elements of each of the p forms an increasing sequence. Furthermore, for a partition $P \in \mathcal{P}(J)$ we then define $\mathrm{sign(P)} \in \{\pm 1\}$ as the sign of the permutation that orders the sequence of elements of the different parts of the partition, i.e., the sequence $(j)_{j\in p, p\in P} = (p^{(1)}_1, \ldots, p^{(1)}_{|p^{(1)}|}, \ldots, p^{(|P|)}_{|p^{(|P|)}|})$. Next to that we define for $m \in [|J|]$ the following subsets of $\mathcal{P}(J)$

$$\mathcal{P}^e(J) = \{P \in \mathcal{P}(J) : \forall p \in P : |p| \,\text{is even}\}, \tag{2.7}$$

$$\mathcal{P}_m(J) = \{P \in \mathcal{P}(J) : \forall p \in P : |p| = m\}, \tag{2.8}$$

such that that for instance $\mathcal{P}_2(J)$ will be the set containing all possible ways of dividing J into pairs and for a pairing $P \in \mathcal{P}_m(J)$ we denote by $\mathrm{sign(P)}$ the sign of the permutation which brings the ordered elements of J in an order such that paired elements are placed next to each other (without changing the relative order inside each pair).

More specific notation which extends these definitions but is only sporadically used will be introduced when needed.

2.2. Finite Fermionic Systems

The quantum states of N spin-1/2 fermions can be captured by a many-body wave function $\Psi \in \bigwedge^N L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ which assigns to the N tuples of positions and spin labels $(\mathbf{x}_a, \sigma_a)$ for $a \in [N]$ a complex amplitude. Here $\bigwedge$ denotes the exterior power which ensures, that Ψ is antisymmetric with respect to the exchange of particles, i.e.,

$$\Psi(\ldots, (\mathbf{x}_a, \sigma_a), \ldots, (\mathbf{x}_b, \sigma_b), \ldots) = -\Psi(\ldots, (\mathbf{x}_b, \sigma_b), \ldots, (\mathbf{x}_a, \sigma_a), \ldots), \tag{2.9}$$

for any $a, b \in [N]$ where all other $(\mathbf{x}_c, \sigma_c)$ for $c \neq a, b$ are kept fixed in their position. In many applications it is useful or even necessary to change from the wave function

formulation of a quantum system to its second quantized formulation which we will use throughout this thesis.

In order to reduce a fermionic quantum system to a finite fermionic system with M modes, a set of orthonormal functions $\phi_j \in L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ for $j \in [M]$ is chosen[1] which serve as single-particle states and span the single-particle Hilbert space $\mathcal{H}_1 = \mathrm{span}(\{\phi_j | j \in [M]\})$. Based on these single-particle states one defines the Slater determinants

$$\Psi_{j_1,\ldots,j_N}((\mathbf{x}_1,\sigma_1),\ldots,(\mathbf{x}_N,\sigma_N)) = \frac{1}{\sqrt{N!}} \det\Big[(\phi_{j_a}(\mathbf{x}_b,\sigma_b))_{a,b\in[N]}\Big], \tag{2.10}$$

with $j_r \in M$ for $r \in [N]$. Note that $\Psi_{j_1,\ldots,j_N}((\mathbf{x}_1,\sigma_1),\ldots,(\mathbf{x}_N,\sigma_N)) = 0$ if any single-particle mode $j \in [M]$ is contained more the once in $\{j_1,\ldots,j_N\}$ and the Slater determinants are antisymmetric under the exchange of particles by the antisymmetry of the determinant. The Slater determinants span a finite-dimensional space $\mathcal{H}_N \subset \bigwedge^N L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$. The problem at hand, for instance describing the time evolution of a fermionic system or identifying the eigenstates of a Hamiltonian, is then solved in this subspace only, by considering a projected Schrödinger equation. This whole procedure is essentially a variant of the Galerkin method where test functions are used in order to discretize a continuous operator problem [18].

The quality and physical relevance of results obtained in this subspace, depend naturally strongly on its choice, i.e., the choice of the single-particle orbitals. This choice is usually based on available analytic solutions of related model-systems, or itself part of a variational principle. In the context of condensed matter systems, the localized Wannier functions and approximations to them obtained from the non-interacting problem are often chosen which lead to the tight-binding approximation where in the context of quantum chemistry single-particle orbitals based on the analytic solution of the hydrogen atom are usually used [4, 6]. Depending on the application the relevant orbitals are then selected based on a physical and chemical intuition or convention. Especially for accurate ab-initio simulations of molecular systems, being able to choose an as small as possible set of orbitals is closely tied to understanding the physical correlation properties of the systems, e.g. the bonding properties of valence electrons.

For a fixed set of single-particle orbitals ϕ_j, we can define the Hilbert space of N fermions for any $N \in [M]$ using the construction above. The collection of all these spaces gives rise to the fermionic Fock space of the chosen M modes $\mathcal{F}_M = \bigoplus_{N=0}^{M} \mathcal{H}_N$, where $\mathcal{H}_0 \simeq \mathbb{C}$. The Fock space has the structure of the exterior algebra of the finite-dimensional single-particle Hilbert space $\mathcal{H}_1$, i.e., $\mathcal{F}_M = \bigoplus_{N=0}^{M} \bigwedge^N \mathcal{H}_1$ with $\bigwedge^0 \mathcal{H}_1 = \mathcal{H}_0 \simeq \mathbb{C}$. Instead of using the above notation for Slater determinants it is convenient to employ the occupation number representation, which introduces a short-hand notation

[1] In practice, further regularity conditions need to be imposed on the functions ϕ depending on the analytic properties of the Hamiltonian (see [17, Ch. 1] for an introduction). As we will not be so much concerned with explicitly discretizing a given problem we skip these details here.[2]

[2] In this thesis we will repeatedly make use of footnotes in order to clarify certain aspects or provide more details which would obstruct the flow of the text to much. Usually you can skip a footnote if you are fine with the marked text.

for the in Eq. (2.10) defined Slater determinants according to

$$\Psi_{j_1,\ldots,j_N} = |i_1,\ldots,i_M\rangle, \tag{2.11}$$

with $i_j = 1$ if $j \in \{j_1,\ldots,j_N\}$ and $i_j = 0$ else and we assumed $j_1 < j_2 < \ldots < j_N$. If the mode indices are not ordered, a global sign may appear according to rules of the determinant in Eq. (2.10). Furthermore, we can define a set of creation and annihilation operators $\{f_j^\dagger\}_{j\in[M]}$ and $\{f_j\}_{j\in[M]}$ which connect different particle number sectors of the Fock space. Denoting the vacuum state by $|0,\ldots,0\rangle = |0\rangle \in \bigwedge^0 \mathcal{H}_1$, the action of the creation and annihilation operators is defined by their property to build up Slater determinants from the vacuum via

$$\Psi_{j_1,\ldots,j_N} = f_{j_1}^\dagger \ldots f_{j_N}^\dagger |0\rangle, \tag{2.12}$$

for $j_r \in [M]$ for $r \in [N]$. The antisymmetry of the many-body wave function imposes the canonical anti-commutation relations (CAR) on the creation and annihilation operators

$$\{f_j^\dagger, f_k\} = \delta_{j,k}, \qquad \{f_j, f_k\} = 0, \qquad \forall j,k \in [M]. \tag{2.13}$$

All bounded operators acting on the chosen subspace of $\bigoplus_{N=0}^M \bigwedge^N L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ are contained in the algebra spanned by $\{f_j^\dagger\}_{j\in[M]}$ and $\{f_j\}_{j\in[M]}$. In practice, using the notation of second quantization eliminates all references to the chosen single-particle orbitals from the state vectors $|\psi\rangle \in \mathcal{F}_M$ as simply the existence of an abstract set of M creation and annihilation operators acting on a Fock space of dimension 2^M is postulated. The concretely chosen single-particle basis influences then only the shape of physically relevant operators, i.e., the couplings present in an Hamiltonian or other observables.

In addition to allowing for an efficient bookkeeping of many-particle states with varying particle number, second quantization allows also to build model systems easily which are not directly derived from an explicit first quantized Hamiltonian and choice of single-particle orbitals but are meant to capture essential physical aspects of more realistic systems. In this thesis we will encounter both, systems which are derived from a first-principle formulation as explained above and models which are used to explore essential physical aspects of different quantum effects. In both cases we deal with finite fermionic systems with M modes to which we associate as explained above a set of creation and annihilation operators $\{f_j^\dagger\}_{j\in[M]}$ and $\{f_j\}_{j\in[M]}$ fulfilling the CAR in (2.13). It is often convenient to define a set of $2M$ Majorana operators via

$$m_{2j-1} = f_j^\dagger + f_j, \qquad m_{2j} = i\Big(f_j^\dagger - f_j\Big), \qquad \forall j \in [M], \tag{2.14}$$

which satisfy the anti-commutation relation

$$\{m_j, m_k\} = 2\delta_{j,k}, \qquad \forall j,k \in [M]. \tag{2.15}$$

We will refer to both, Eq. (2.13) and (2.15) as canonical anti-commutation relations as they are equivalent. The Majorana operators defined above form an orthonormal basis

of the operator algebra of fermionic operators with respect to the Hilbert-Schmidt scalar product. A general operator $A \in \mathcal{B}(\mathcal{F}_M)$ acting on the fermionic Fock space $\mathcal{F}_M$ can therefore be expanded uniquely as

$$A = \sum_{r=0}^{2M} \sum_{1 \leq j_1 < \ldots < j_r \leq 2M} a_{j_1,\ldots,j_r} m_{j_1} \ldots m_{j_r}, \tag{2.16}$$

where any physical operator involves only terms with r even as we will argue below. In the case of A conserving the particle number, i.e., $[A, \hat{N}] = 0$ with $\hat{N} = \sum_j f_j^\dagger f_j$ it is convenient to expand A in terms of normal ordered polynomials of the creation and annihilation operators

$$A = \sum_{r=0}^{M} \sum_{\substack{1 \leq j_1 < \ldots < j_r \leq M \\ 1 \leq k_1 < \ldots < k_r \leq M}} a_{j_1,\ldots,j_r,k_1,\ldots,k_r} f_{j_1}^\dagger \ldots f_{j_r}^\dagger f_{k_r} \ldots f_{k_1}. \tag{2.17}$$

In the context of fermionic systems the notion of support of an operator changes. Here we say that an operator A has a support $S \subset [M]$ if an expansion of the operator in terms of Majorana or creation and annihilation operators involves operators of the modes S only. All notions of locality discussed in Sec. 2.1.1 carry over accordingly as we will discuss in detail below.

In many practical applications, a concrete representation of a fermionic system, so of its states and operators, in terms of complex numbers is needed which can then be used to represent the system on a computer. One convenient and frequently used representation of fermions is provided by the Jordan-Wigner transformation.

2.2.1. Jordan-Wigner Transformation

The Fock space of M fermions is isomorphic to the Hilbert space of M qubits $\mathbb{C}^{2^M}$ as the Slater determinants written in occupation number representation can be directly identified with the standard basis vectors of $\mathbb{C}^{2^M}$. We can therefore find a representation of the fermionic creation and annihilation operators in terms of the Pauli matrices

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.18}$$

acting locally on the qubits.

This is achieved by the Jordan-Wigner transformation [19] which promises the following: Given a set of M fermionic creation and annihilation operators $\{f_j^\dagger\}_{j \in [M]}$ and $\{f_j\}_{j \in [M]}$ then

$$f_j^\dagger = \frac{1}{2} \left(\prod_{k=1}^{j-1} Z_k \right) (X_j - iY_j), \quad f_j = \frac{1}{2} \left(\prod_{k=1}^{j-1} Z_k \right) (X_j + iY_j) \tag{2.19}$$

provides a matrix representation of these operators which is acting on $\mathbb{C}^{2^M}$. Note that the term transformation is a bit misleading as the Jordan-Wigner transformation rather corresponds to the choice of a representation. Here, $X_j = \mathbb{1}_2^{\otimes(j-1)} \otimes X \otimes \mathbb{1}_2^{\otimes(M-j-1)}$ with $\mathbb{1}_2$ being the 2×2 identity matrix denotes the Pauli-X matrix acting on the j-th qubit.

In several applications there exists a natural blocking of the modes into V disjoint blocks of equal size p. The most common case would be the treatment of a set of spatial orbitals which have an additional spin-1/2 degree of freedom, i.e., $p = 2$. We are then interested in a representation of the fermionic operators acting on $\mathbb{C}^{(2^p)^V}$, i.e., a system of V sites with local dimension 2^p. Choosing a labeling of the pV modes such that the consecutive modes $p(j-1)+1, \ldots, pj$ for $j \in [V]$ are blocked together, the corresponding creation and annihilation operators are then represented by

$$f_{j,\sigma}^\dagger = \frac{1}{2}\left[\prod_{k=1}^{j-1}(Z^{\otimes p})_k\right]\left[Z^{\otimes(\sigma-1)} \otimes (X - iY) \otimes \mathbb{1}_2^{\otimes(p-\sigma)}\right]_j, \tag{2.20}$$

$$f_{j,\sigma} = \frac{1}{2}\left[\prod_{k=1}^{j-1}(Z^{\otimes p})_k\right]\left[Z^{\otimes(\sigma-1)} \otimes (X + iY) \otimes \mathbb{1}_2^{\otimes(p-\sigma)}\right]_j. \tag{2.21}$$

In the case of spatial orbitals with a spin-1/2 degree of freedom we therefore obtain for the creation operator of a spin-up and spin-down particle ($\sigma = 1, 2$) in the j-th orbital

$$f_{j,1}^\dagger = \frac{1}{2}\left[\prod_{k=1}^{j-1}(Z^{\otimes 2})_k\right][(X - iY) \otimes \mathbb{1}_2]_j, \tag{2.22}$$

$$f_{j,2}^\dagger = \frac{1}{2}\left[\prod_{k=1}^{j-1}(Z^{\otimes 2})_k\right][Z \otimes (X - iY)]_j. \tag{2.23}$$

2.2.2. Mode Transformations

The coefficients of a specific state vector in or operator acting on the Fock space depend on the choice of basis of the single-particle Hilbert space $\mathcal{H}_1$. Transforming the modes of the system by choosing a new basis of $\mathcal{H}_1$ will lead to a new set of fermionic creation and annihilation operators. Stated more explicitly, given a (special) unitary matrix $U \in SU(M)$ we define a new set of creation and annihilation operators, which for the time being we label by $\{f_j^{(U)\dagger}\}_{j\in[M]}$ and $\{f_j^{(U)}\}_{j\in[M]}$, in terms of the original ones denoted as $\{f_j^\dagger\}_{j\in[M]}$ and $\{f_j\}_{j\in[M]}$ via

$$f_j^{(U)} = \sum_{k=1}^{M} U_{j,k}^\dagger f_k, \quad f_j^{(U)\dagger} = \sum_{k=1}^{M} U_{k,j} f_k^\dagger. \tag{2.24}$$

It is easy to check that due to the unitarity of U the new operators indeed fulfill the CAR. On a more abstract level, there exists a wider class of linear transformations which

preserve the fermionic algebra than basis rotations in the single-particle Hilbert space. Given an arbitrary (special) orthogonal matrix $O \in SO(2M)$ the operators

$$m_j^{(O)} = \sum_{k=1}^{2M} O_{j,k} m_k \tag{2.25}$$

form again a set of Majorana operators fulfilling the respective anti-commutation relations in Eq. (2.15). Being more general, the transformation of the Majorana operators in Eq. (2.25) contains of course the basis change in the single-particle Hilbert space represented by Eq. (2.24). To be precise, choosing in (2.25) $O = e^{h_R \otimes \mathbb{1}_2 + i h_I \otimes Y}$ for $h_R, h_I \in \mathbb{R}^{M \times M}$ antisymmetric and symmetric respectively with $\mathrm{tr}(h_I) = 0$ and Y denoting the Pauli-Y matrix yields the transformation in Eq. (2.24) for $U = e^{-h_R + i h_I}$.

In the following if no confusion can arise, we will often make no distinction between the different operators $m^{(O)}$ and m as they span isomorphic algebras. A mode transformation as above induces a linear transformation of the Fock space which viewed as an active rotation in a fixed frame changes the coefficients of physical observables and state vectors. For later use in this thesis, where we will repeatedly encounter mode transformation in different readings, we want introduce a few more formal aspects here.

The following lemma captures the transformation induced by a mode transformation on the full fermionic Fock space.

Lemma 1 (Fock space representation of mode transformations)**.** *Given a fermionic system with M modes, Majorana operators m_j for $j \in [2M]$ and corresponding creation and annihilation operators ${f_k}^\dagger$ and f_k for $k \in [M]$ as well as $O \in SO(2M)$ with skew-symmetric real logarithm $\ln(O)$ then*

$$G(O) = \exp\left(\frac{1}{4} \sum_{j,k} m_j m_k \ln(O)_{j,k} \right) \tag{2.26}$$

corresponds to the transformation of the Fock space induced by the mode transformation O. For $O = e^{h_R \otimes \mathbb{1}_2 + h_I \otimes iY}$ with $h_R, h_I \in \mathbb{R}^{M \times M}$ antisymmetric and symmetric respectively and $\mathrm{tr}(h_I) = 0$, the transformation reduces to

$$G(U) = G(O) = \exp\left(\sum_{j,k} f_j^\dagger f_k \ln(U^\dagger)_{j,k} \right), \tag{2.27}$$

with $U = e^{h_R - i h_I}$.

Proof. Under the mode transformation O a product of Majorana operators transforms as

$$m_{j_1} \dots m_{j_r} = \sum_{k_1, \dots, k_r = 1}^{2M} O^T_{j_1,k_1} \dots O^T_{j_r,k_r} m^{(O)}_{k_1} \dots m^{(O)}_{k_r}. \tag{2.28}$$

Defining $h = \ln(O)$ and using that $h^T = -h$ we obtain for the nested commutator

$$\left[\sum_{j,k=1}^{2M} \frac{1}{4} m_j m_k h_{j,k}, m_l\right]_n = \sum_j m_j \big(h^n\big)_{j,l}. \tag{2.29}$$

By Hadamard's lemma we then conclude

$$G(O)\, m_l\, G(O^T) = \sum_{n=0}^{\infty} \sum_{j=1}^{2M} \frac{1}{n!} m_j \big(h^n\big)_{j,l} = \sum_{j=1}^{2M} m_j O_{j,l} = \sum_{j=1}^{2M} O^T_{l,j} m_j, \tag{2.30}$$

meaning that products of Majorana operators and by linearity a general operator acting A on $\mathcal{F}_M$ transform as $A \to G(O) A G(O^T)$.

The vector $\mathbf{m} = (m_j, \ldots, m_{2M})_{j\in[2M]}$ of Majorana operators and the vector $\mathbf{c} = (f_1^\dagger, f_1, \ldots, f_M^\dagger, f_M)$ of creation and annihilation operators are related by

$$\mathbf{m} = \bigoplus_{i=1}^{M} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \mathbf{c}. \tag{2.31}$$

Inserting this relation in the definition of $G(O)$ for $O = e^{h_R \otimes \mathbb{1}_2 + i h_I \otimes Y}$ yields the above result. □

Given a basis transformation of the single-particle Hilbert space $U \in SU(M)$, a single Slater determinant transforms under this basis change as

$$f_{j_1}^\dagger \ldots f_{j_r}^\dagger |0\rangle \to G(U) f_{j_1}^\dagger \ldots f_{j_r}^\dagger |0\rangle = \sum_{k_1,\ldots,k_r=1}^{M} U^\dagger_{k_1,j_1} \ldots U^\dagger_{k_r,j_r} f_{k_1}^\dagger \ldots f_{k_r}^\dagger |0\rangle. \tag{2.32}$$

For practical applications, e.g., within numerical schemes, it will however be useful to have an explicit and more compact matrix representation of $G(U)$ acting on $\mathcal{F}_M \simeq \mathbb{C}^{2^M}$. We obtain it by realizing that according to Eq. (2.32), $G(U)$ acts on $\mathcal{F}_M = \bigoplus_{N=0}^{M} \bigwedge^N \mathcal{H}_1$ as the natural endomorphism induced by the basis change $U : \mathcal{H}_1 \to \mathcal{H}_1$.[3] Therefore $G(U)$ acts on $\mathbb{C}^{2^M}$ as

$$G(U) = \bigoplus_{N=0}^{M} \bigwedge^N U^\dagger, \tag{2.33}$$

where $\bigwedge^N A \in \mathbb{C}^{\binom{M}{N}\times\binom{M}{N}}$ denotes the N-th compound matrix of $A \in \mathbb{C}^{M\times M}$ and its entries are given by determinants of the $N \times N$ submatrices of A:

$$\left(\bigwedge^N A\right)_{I,J} = \det((A_{i,j})_{i\in I, j\in J}), \tag{2.34}$$

[3] Given an endomorphism $f : \mathcal{H} \to \mathcal{H}$ on a vector or Hilbert space $\mathcal{H}$ the natural induced endomorphism $\bigwedge^N f : \bigwedge^N \mathcal{H} \to \bigwedge^N \mathcal{H}$ acts as $(\bigwedge^N f)(e_{j_1} \wedge \ldots \wedge e_{j_N}) = f(e_{j_1}) \wedge \ldots \wedge f(e_{j_N})$.

for $I, J \subset [M]$ with $|I| = |J| = N$. Equation (2.33) hence tells us how a mode transformation acts on state vectors after a Jordan-Wigner transformation.

Similarly, we will need the explicit transformation of the coefficients of an operator induced by a mode transformation. Above we argued that an operator A acting on $\mathcal{F}_M$ will transform under the mode transformation as

$$A \to G(U)AG(U^\dagger). \tag{2.35}$$

Expanding the operator in terms of the creation and annihilation operators

$$A = \sum_{r=1}^{M}\sum_{s=1}^{M}\sum_{1\leq j_1<\ldots<j_r\leq M}\sum_{1\leq k_1<\ldots<k_s\leq M} a^{r,s}_{(j_1,\ldots,j_r),(k_1,\ldots,k_s)} f^\dagger_{j_1}\ldots f^\dagger_{j_r} f_{k_1}\ldots f_{k_s}. \tag{2.36}$$

it is easy to see from Eq. (2.24) that the coefficients transform under the mode transformation as

$$a^{(U),r,s} = (U^\dagger)^{\otimes r} a^{r,s} U^{\otimes s}. \tag{2.37}$$

2.2.3. Superselection Rule

In fermionic systems various kinematic constraints arise. One very important constraint is given by the so-called parity superselection rule. Splitting the fermionic Fock space in two sectors $\mathcal{F}_M = \mathcal{F}_M^{\text{even}} \oplus \mathcal{F}_M^{\text{odd}}$ with $\mathcal{F}_M^{\text{even}} = \bigoplus_n \mathcal{H}_{2n}$ and $\mathcal{F}_M^{\text{odd}} = \bigoplus_n \mathcal{H}_{2n+1}$ the parity superselection rule states, that there exists no physical scheme to distinguish the states $(|\psi_1\rangle + |\psi_2\rangle)/\sqrt{2}$ and $(|\psi_1\rangle - |\psi_2\rangle)\sqrt{2}$ with $|\psi_1\rangle \in \mathcal{F}_M^{\text{even}}$ and $|\psi_2\rangle \in \mathcal{F}_M^{\text{odd}}$. It was first introduced by Wick, Wightman and Wigner [20] in the context of relativistic field theories where it is necessary in order to allow for an unambiguous implementation of Lorentz transformations. It was later noted that in the context of non-relativistic finite systems, the presence of the parity superselection rule is necessary in order to preserve the non-signaling assumption (see for instance [21, 22]). In finite systems the non-signaling assumption amounts to that two simultaneous disjoint physical operations should commute with each other. Let us illustrate the need of such a technical assumption in more physical terms. As argued in [22], without a parity superselection rule, the Hermitian and unitary operators $f_1^\dagger + f_1$ and $f_2^\dagger + f_2$ are admissible operations which clearly do not commute. As a consequence the application of $f_1^\dagger + f_1$ to the state $(f_2^\dagger|0\rangle + |0\rangle)/\sqrt{2}$ changes the expectation value of $f_2^\dagger + f_2$ from +1 to -1 and by this would change the local measurement statistics of an observer of the second mode instantaneously. In order to rule out such effects we will employ the parity superselection rule throughout this thesis. This can be done consistently by assuming that all physical operators such as observables consist of terms with an even number of Majorana operators only. A state vector $(|\psi_1\rangle + e^{i\phi}|\psi_2\rangle)\sqrt{2}$ with $\phi \in \mathbb{R}$, $|\psi_1\rangle \in \mathcal{F}_M^{\text{even}}$ and $|\psi_2\rangle \in \mathcal{F}_M^{\text{odd}}$ will then yield the same expectation values as the classical mixture $\rho = |\psi_1\rangle\langle\psi_1|/2 + |\psi_2\rangle\langle\psi_2|/2$. State vectors are therefore assumed to be either element of $\mathcal{F}_M^{\text{even}}$ or $\mathcal{F}_M^{\text{odd}}$ whereas we allow for classical mixtures of states from both sectors. Furthermore, any valid state ρ will then

be an even operator, i.e., consist of terms with an even number of Majorana operators only.

Note that the above argument implicitly assumes all observers to be equally powerful in order to motivate the superselection rule. We assume the scenario in which a single observer is able to perform any operation while all other parties are restricted to physical ones (which would also lead to no signaling) to be unphysical.

2.3. Reduced Density Matrices

In fermionic systems two different kinds of reduced states appear. The first is related to the notion of tracing out a certain number of modes of the systems and reducing a state ρ which is defined on M modes to a state which is supported on the modes $S \subset [M]$ only. We are therefore essentially interested in relating $\mathcal{D}(\mathcal{F}_M)$ and $\mathcal{D}(\mathcal{F}_{|S|})$. The restriction to a local subregion of the total system allows us to meaningfully analyze systems with a variable systems size and to capture their properties within the thermodynamic limit. Further, it induces the notion of correlation between different modes in fermionic systems; a topic which we will encounter throughout this thesis. The second kind of reduced density matrices is connected to tracing out particles, which is straightforward within the first quantized picture and in the setting of finite fermionic systems amounts to restricting the action of a state to the Hilbert space of the remaining number of particles $\mathcal{H}_N$. Reductions to few-particle settings are conceptually very natural for fermionic systems, as typical interactions such as the Coulomb repulsion may not feature a spatial locality but act on two particles only. In addition as we will see below, in the special class of non-interacting systems, a description of the single particle sector is often sufficient for capturing the dynamical evolution and for the description of Gaussian states.

2.3.1. Reduction of Modes

The reduction of modes is an useful tool to talk meaningfully about a subsystem of a larger fermionic system and allows us to address questions involving locality, e.g., the local distinguishability of two states, in these systems. Given a fermionic system with M modes, the reduction of a state to a subsystem $S \subset [M]$ is defined the easiest using an operational definition analogue to the one presented in Sec. 2.1.1. Given $S = \{s_1, \dots s_{|S|}\} \subset [M]$, we define the natural embedding $\tau_S : \mathcal{B}(\mathcal{F}_{|S|}) \hookrightarrow \mathcal{B}(\mathcal{F}_M)$ with $\tau_S : m_{j_1} \dots m_{j_r} \mapsto m_{s_{j_1}} \dots m_{s_{j_r}}$ for any $j_1, \dots, j_r \in [|S|]$. We then define for a given fermionic state $\rho \in \mathcal{D}(\mathcal{F}_M)$ the reduced state $\mathrm{tr}_{[M]\setminus S}\,\rho \in \mathcal{D}(\mathcal{F}_{|S|})$ implicitly by

$$\mathrm{tr}_S[(\mathrm{tr}_{[M]\setminus S}\,\rho)\,A] = \mathrm{tr}[\rho\,\tau_S(A)] \quad \forall A \in \mathcal{B}(\mathcal{F}_{|S|}). \tag{2.38}$$

Put differently, the reduced state is defined to reproduce all expectation values of ρ for operators with support $\mathrm{supp}(A) \subset S$, i.e., those which only involve the modes collected in S. If clear from the context, we omit the reference to the full set of modes and simply write S^c instead of $[M]\setminus S$.

An explicit description of the reduced state can be constructed in different ways. In general, using the characteristic function of a fermionic state within the Grassmann calculus [23] we obtain the reduced state by integrating out the degrees freedom corresponding to the chosen set of modes. Alternatively, we can relabel the modes of the system such that $S = \{1, \ldots, |S|\}$. In this case the reduced state can be calculated directly as the reduced state of the Jordan-Wigner transformed state using the ordinary partial trace.

From the implicit definition above one can show that the resulting object, $\mathrm{tr}_{S^c}\,\rho$ is indeed a valid state, so $\mathrm{tr}_{S^c}\,\rho \in \mathcal{D}(\mathcal{F}_{|S|})$.[4]

Having established how to reduce a state to a certain set of modes, let us also introduce the notation for multiple copies of a state. Assume we have V many copies of a system of p fermionic modes which we each denote by $S_k = \{kp, kp+1, \ldots, (k+1)p-1\} \subset [Vp]$. Given then a state $\rho \in \mathcal{D}(\mathcal{F}_p)$ on p fermionic modes, we denote by $\rho^{\otimes V} \in \mathcal{D}(\mathcal{F}_{Vp})$ the state on Vp modes which is completely uncorrelated across different S_k and locally on each S_k agrees with ρ, i.e., $\rho = \mathrm{tr}_{S_k{}^c}\,\rho^{\otimes V}$.[5]

2.3.2. Reduction of Particles

Just as the state reduced to a subset of modes is able to account for the action of local operators supported on these modes, the state reduced to k particles captures the effect of operators that act on k of fewer particles only. This notation is particularly useful for the description of Gaussian states, which arise essentially from a single-particle picture we will discuss in more detail below. It will also lead us to rather profound and important questions concerning the kinematic structure of fermionic systems.

Assume a fermionic state $\rho \in \mathcal{D}(\mathcal{H}_N)$ with a well defined particle number N. A k-particle operator with $k < N$ will not probe the full state but is only sensitive to certain averaged information present. We define the k-body covariance matrix $\gamma^{(k)}[\rho] \in \mathbb{C}^{M^k \times M^k}$ of ρ by the collection of all k-th moment of the state by

$$\gamma^{(k)}[\rho]_{J,L} = \mathrm{tr}\left(\prod_{j\in J} f_j^\dagger \prod_{l\in L^{-1}} f_l\,\rho\right) \tag{2.39}$$

for $J, L \in [M]^{\times k}$ being sequences of length k, L^{-1} denotes the reversed sequence to L (i.e., $(1,3,2)^{-1} = (2,3,1)$) and the products are performed in order of the corresponding sequence. Given the k-body covariance matrix we define the k-particle reduced state

[4] The normalization $\mathrm{tr}(\mathrm{tr}_{S^c}\,\rho) = 1$ follows from choosing $A = \mathbb{1}$ and the relation to original expectation values of ρ ensures $\mathrm{tr}_{S^c}(\rho)$ to be Hermitian and positive as $\tau_S(A)$ is Hermitian/positive for Hermitian/positive A respectively.

[5] Tensor products have to be used with some care in the context of fermionic states due to their intrinsic antisymmetry. Here we use them in the sense of the above implicit definition as combining states in an uncorrelated fashion. Furthermore, due to the chosen ordering of the fermionic modes and the superselection rule the $\otimes$ operation defined that way corresponds to the proper tensor product within the Jordan-Wigner representation.

$\rho^{(k)}$ as

$$\rho^{(k)} = \sum_{\substack{J,L\in[M]^{\times k}\\ J,L \text{ increasing}}} \frac{\gamma^{(k)}[\rho]_{J,L}}{\binom{N}{k}} \prod_{l\in L} f_l^\dagger |0\rangle\langle 0| \prod_{j\in J^{-1}} f_j, \tag{2.40}$$

where we denote a sequence J to be increasing if the elements are ordered increasingly. One can verify that indeed $\rho^{(k)} \in \mathcal{D}(\mathcal{H}_k)$, i.e., $\operatorname{tr}\rho^{(k)} = 1$, $\rho^{(k)\dagger} = \rho^{(k)}$, $\rho^{(k)} \geq 0$ and $\operatorname{tr}(\hat{N}^l\rho) = k^l$ for all $l \in \mathbb{N}$ and $\hat{N}$ denoting the total particle number operator.[6] In addition $\rho^{(k)}$ is related to ρ by having, up to a normalization factor, the same k-th moments, i.e.,

$$\gamma^{(k)}[\rho^{(k)}] = \frac{1}{\binom{N}{k}}\gamma^{(k)}[\rho]. \tag{2.43}$$

It is clear from its definition in Eq. (2.40) and (2.39) that $\rho^{(k)}$ meets the criterion of being able to reproduce all expectation values of ρ with k-particle operators. However, the notion of a particle-reduced density matrix can be made more concrete. In the first quantized picture we can relate $\rho^{(k)}$ to the integration over the degrees of freedom of excessive particles. Assume a pure state defined by its wave function $\Psi(x_1, \ldots, x_N)$, where x_i collects here for simplicity of notation all degrees of freedom of the j-th particle. We can then define the k-particle reduced state by its density operator as

$$\rho^{(k)}(x_1, \ldots, x_k, x'_1, \ldots, x'_k) = \int \Psi(x_1, \ldots, x_k, x'_{k+1}, \ldots, x'_N)\overline{\Psi}(x'_1, \ldots, x'_k, x'_{k+1}, \ldots, x'_N)dx'_{k+1}dx'_N, \tag{2.44}$$

where by the (anti-)symmetry of Ψ it is irrelevant which $N-k$ particles are integrated over. The action of any k-particle operator on Ψ can then be obtained from $\rho^{(k)}$ alone.

Assume now that Ψ is a state of a fermionic system with M modes $\phi_j(x)$ for $j \in [M]$, i.e., it can be expanded into the corresponding Slater determinants according to

$$\Psi(x_1, \ldots, x_N) = \sum_{\substack{J\in[M]^{\times N}\\ J \text{ increasing}}} \alpha_J \Psi_J(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \sum_{\substack{J\in[M]^{\times k}\\ J \text{ injective}}} \tilde{\alpha}_J \prod_{l=1}^{N} \phi_{j_l}(x_l), \tag{2.45}$$

where we call a sequence injective if no element appears twice and $\tilde{\alpha}_J = \alpha_{J^\uparrow}\operatorname{sign}(\pi_J)$ with $J^\uparrow$ denoting the sorted sequence with the elements of J and π_J is the sorting

[6] Verifying the positivity is the only non-trivial part. For any $|\psi\rangle \in \mathcal{H}_k$ we define the operator A_ψ by

$$|\psi\rangle = \sum_{\substack{J\in[M]^{\times k}\\ J \text{ increasing}}} \alpha_J \prod_{j\in J} f_j^\dagger |0\rangle = A_\psi|0\rangle. \tag{2.41}$$

We then obtain by the positivity of ρ

$$\langle\psi|\rho^{(k)}|\psi\rangle = \sum_{\substack{J,K\in[M]^{\times k}\\ J,L \text{ increasing}}} \overline{\alpha}_L \gamma^{(k)}[\rho]_{J,L}\alpha_J = \operatorname{tr}\left(A_\psi A_\psi^\dagger \rho\right) \geq 0. \tag{2.42}$$

permutation. The reduced state can then be written as

$$\rho^{(k)}(x_1,\ldots,x_k,x'_1,\ldots,x'_k) = \sum_{\substack{A,B\in[M]^{\times k}\\ A,B \text{ increasing}}} \rho^{(k)}_{A,B}\Psi_A(x_1,\ldots,x_k)\overline{\Psi}_B(x'_1,\ldots,x'_k), \quad (2.46)$$

with

$$\rho^{(k)}_{A,B} = \frac{1}{\binom{N}{k}} \sum_{\substack{J,K\in[M]^{\times N}\\ J,K \text{ increasing}\\ A\subset J, J\backslash A = K\backslash B}} \alpha_J\overline{\alpha}_K \mathrm{sign}(\pi_{\mathrm{A\,(J\backslash A)}})\mathrm{sign}(\pi_{\mathrm{B\,(K\backslash B)}}), \quad (2.47)$$

where $A \subset J$ and $J\backslash A$ have the obvious meaning of denoting that A is a subsequence of J and the sequence of elements of J not in A and $A\,J$ denotes the joint sequence $(a_1,\ldots,a_{|A|},j_1,\ldots,j_{|J|})$. Starting from

$$|\psi\rangle = \sum_{\substack{J\in[M]^{\times N}\\ J \text{ increasing}}} \alpha_J \prod_{j\in J} f_j^\dagger|0\rangle \quad (2.48)$$

we find by direct calculation of the entries of the covariance matrix that indeed

$$|\psi\rangle\langle\psi|^{(k)} = \sum_{\substack{A,B\in[M]^{\times k}\\ A,B \text{ increasing}}} \rho^{(k)}_{A,B} \prod_{a\in A} f_a^\dagger|0\rangle\langle 0| \prod_{b\in B^{-1}} f_b \quad (2.49)$$

with the same coefficients as above. The same holds true for mixed states by linearity and illustrates the interpretation of forming the reduced k-body states by integrating out a portion of the particles.

2.4. Free Fermions and Gaussian States

Non-interacting fermionic system turn out to provide an essential tool for understanding basic effects in fermionic systems appearing in nature. It is quite remarkable how many different important physical insights can be gained from investigating non-interacting fermions (see for instance numerous applications in [4]) despite the interactions present in every realistic natural setting. An independent particle assumption is for instance employed when we explain conduction properties of solids by their band structure. The low energy theory of excitations in a Fermi liquid can often be well approximated to be non-interacting; a general theme appearing at many places in condensed matter physics where the low energy sector of interacting theories can be captured by emergent collective excitations described by weakly or non-interacting quasiparticles. In addition one of the historic hallmarks of quantum mechanics, the description of the excitation spectrum of hydrogen, relies on the solution of a single (and therefore non-interacting) particle problem. The physics of multiple non-interacting fermions is then dominated by the antisymmetric character of the fermionic state space manifested in the well known Pauli

principle. We will give a more detailed account for these effects later in this chapter in Sec. 2.6.

Furthermore, free systems provide an important starting point for many more elaborate analytic and numerical calculations. Starting from the description of free systems, perturbation theory allows to include weak interaction effects. Mean field theory approaches to interacting systems ask for identifying the best non-interacting description to the system at hand as we will discuss in Sec. 2.5 and suitable single-particle orbitals used for the procedure of second quantization are often solution of non-interacting problems. Furthermore, deeper insights into the structure of interacting systems allow to approximate the ground state of an interacting system by iteratively identifying the ground state of a family of free proxy problems using the density functional theory (DFT) (see for instance [7, Ch. 1] and [24] for an introduction).

Solving even a non-interacting particle problem analytically in the continuous case is a demanding task and can only be done in a few selected instances. In finite systems on the other hand the structure of the problem allows for an efficient numerical simulation of single particle effects in these systems. From this, many properties can be calculated with an effort that scales polynomially in the system size M despite the exponential size of the Fock space as many aspects can be faithfully mapped to the single-particle Hilbert space of the problem as we will explain below. The combination of being easy to solve as well as able to capture essential effects of realistic systems render free systems to be a perfect tool for conceptual theoretic investigation of fermionic systems; a tool we will rely on later in Ch. 5.

The most general free Hamiltonians are given by

$$H = \sum_{j,k=1}^{2M} \frac{i}{4} h^{(2)}_{j,k} m_j m_k, \tag{2.50}$$

where without loss of generality we can always choose $h^{(2)} \in \mathbb{R}^{2M\times 2M}$ antisymmetric. As $h^{(2)}$ is antisymmetric there exists a mode transformation $O \in SO(2M)$ which block diagonalizes $h^{(2)}$ such that

$$O h^{(2)} O^T = \Lambda \otimes iY \tag{2.51}$$

with $\Lambda \in \mathbb{R}^{M\times M}$ diagonal. Denoting the fermionic creation and annihilation operators corresponding to the $m^{(O)}_j$ by $f^{(O)\dagger}_k$ and $f^{(O)}_k$ with $j \in [2M]$ and $k \in [M]$ we obtain that the transformed Hamiltonian

$$H = \sum_{j=1}^{M} \frac{i}{2} \Lambda_{j,j} m^{(O)}_{2j-1} m^{(O)}_{2j} = \sum_{j=1}^{M} \Lambda_{j,j} f^{(O)\dagger}_j f^{(O)}_j - \frac{\operatorname{tr}\Lambda}{2} \mathbb{1} \tag{2.52}$$

is diagonal and all eigenstates are given by Slater determinants formed with respect to the operators $f^{(O)\dagger}_j$ and $f^{(O)}_j$.[7] In addition by Lem. 1, the time evolution operator

$$e^{iHt} = G(e^{-h^{(2)}\, t}) \tag{2.53}$$

[7] Note that a general orthogonal rotation of the Majorana operators can change the vacuum state of the system as particles might be identified as holes and vice versa. The Slater determinants are

acts as a mode transformation on the Fock space. These two observations allow us to identify the eigenenergies and eigenstates of H as well as performing the time evolution induced by H solely by considering the mode space, i.e., single particle Hilbert space, instead of the full Fock space.

We have seen that the eigenstates of a non-interacting system are of particular simple structure as they take the form of a single Slater determinant if written in the correct basis. As these Slater determinants might be formulated with respect to an unphysical vacuum state it seems natural to ask, what parts of this structure are independent of the single-particle basis. This question can be answered easily by realizing that Slater determinants are important representatives of the larger class of fermionic Gaussian states. Next to Slater determinants, all thermal mixtures of the eigenstates of a free system are also Gaussian such that Gaussian states play an important role in the description of the equilibrium properties of free systems at any temperature. Furthermore, a deeper investigation of the structure of Gaussian states reveals (unsurprisingly regarding their context) that they are fully described by their single particle restrictions leading to a description of Gaussian states in terms of $\mathcal{O}(M^2)$ parameters.

Definition 1. *A state $\rho \in \mathcal{D}(\mathcal{F}_M)$ on the Fock space of M modes is a Gaussian state if there exists a mode transformation $O \in SO(2M)$ such that the state takes the form*

$$G(O)\,\rho\,G(O^T) = \prod_{j=1}^{M}\left(\frac{1}{2}\mathbb{1} + \frac{i}{2}\xi_j m_{2j-1}^{(O)} m_{2j}^{(O)}\right), \tag{2.54}$$

with $\xi_j \in [-1,1]$ for $j \in [M]$.

The thermal state $e^{-\beta H}/\mathcal{Z}$ with $\mathcal{Z} = \mathrm{tr}(e^{-\beta H})$ of a free Hamiltonian is a Gaussian state as we have

$$\frac{1}{\mathcal{Z}} G(O)\exp\left(\frac{i}{4}\sum_{j,k} h_{j,k}^{(2)} m_j m_k\right) G(O^T) = \prod_{j=1}^{M}\left(\frac{1}{2}\mathbb{1} + \frac{i}{2}\tanh\left(\Lambda_{j,j}/2\right) m_{2j-1}^{(O)} m_{2j}^{(O)}\right) \tag{2.55}$$

for $O \in SO(2M)$ block-diagonalizing $h^{(2)}$ and $Oh^{(2)}O^T = \Lambda \otimes iY$ with Λ diagonal. In addition Slater determinants and by this all eigenstates of a free Hamiltonian are pure Gaussian states and obviously every state that is related to a Gaussian state by a mode transformation will be again Gaussian by definition.

Compared to a general state whose number of parameters grows exponentially in the system size, a Gaussian state has a very specific structure. According to the above definition, a Gaussian state is defined by specifying M many ξ_j and a mode transformation $O \in SO(2M)$. Expectation values with respect to ρ are then related to these parameters

then constructed based on the vacuum of the creation and annihilation operators $f_j^{(O)\dagger}$ and $f_j^{(O)}$ associated to $m_j^{(O)}$ via Eq. (2.14) which is defined as the state $|\psi\rangle \in \mathcal{F}_M$ with $f_j^{(O)}|\psi\rangle = 0$ for all $j \in [M]$.

via Wick's theorem. Wick's theorem states that, as shown for instance in App. A, given a Gaussian state ρ and any operators $c_1 \dots c_{2r}$ which are linearly related to the Majorana operators $m_j^{(O)}$ from the definition of the Gaussian state then the expectation value of $c_1 \dots c_{2r}$ can be calculated from the second moments of ρ alone via

$$\mathrm{tr}(c_1 \dots c_{2r}\rho) = \sum_{P\in\mathcal{P}_m([2r])} \mathrm{sign(P)} \prod_{(p_1,p_2)\in P} \mathrm{tr}(c_{p_1}c_{p_2}\rho), \tag{2.56}$$

using the notation of Sec. 2.1.2.

Exploiting Wick's theorem a Gaussian state ρ is specified alone by its second moments in a given basis which are often collected in the covariance matrix

$$\gamma(\rho)_{j,k} = \frac{i}{2}\,\mathrm{tr}(\rho\,[m_j, m_k]). \tag{2.57}$$

For a state ρ with fixed particle number, the covariance matrix $\gamma(\rho)$ is uniquely defined by the single-body correlation matrix $\gamma^{(1)}[\rho]$ introduced above in Sec. 2.3.2 due to the linear relation of creation and annihilation operators and Majorana modes. Gaussian states with a fixed particle number are therefore uniquely specified by their correlation matrix $\gamma^{(1)}[\rho]$ and Wick's theorem. From the definition of a general Gaussian state one quickly derives that its covariance matrix fulfills the conditions $\gamma(\rho) \in \mathbb{R}^{2M\times 2M}$, $\gamma(\rho)^T = -\gamma(\rho)$ and $-\gamma(\rho)^2 \leq \mathbb{1}$. In addition for every such matrix $\gamma \in \mathbb{R}^{2M\times 2M}$ fulfilling these two conditions, there is a Gaussian state ρ with $\gamma(\rho) = \gamma$.

There are three further common definitions of Gaussian states. As a state is fully specified by all its expectation values a Gaussian state can be defined to be the state with a certain covariance matrix γ which fulfills Wick's theorem. As we define reduced states to subregions of the full systems in Sec. 2.3.1 by their expectation values, it can be directly seen that the reduction of a Gaussian state has to be Gaussian as well, as it necessarily fulfills Wick's theorem. Furthermore, a Gaussian state with a covariance matrix γ is the state which maximizes the von Neumann entropy given the second moments specified by γ. Introducing the calculus of anti-commuting variables, so-called Grassmann numbers, Gaussian states can also be defined as states with a Gaussian Grassmann representation [25]. All these definitions are equivalent (see App. A for details) and highlight different important aspects of Gaussian states.

2.5. Interacting Fermions

Non-interacting systems are, however, able to capture only certain effects of fermionic systems. Predictions made from non-interacting investigations usually have to be measured and tested against experimental investigations or predictions from interacting descriptions in selected cases. This can be illustrated on a currently very active field of research. The question in how far results on the localization of particles and a resulting breakdown of transport from disorder effects observed for non-interacting particles, so-called Anderson localization, carry over to interacting particles has gained a lot of

attention from an analytical, numerical and experimental perspective (see [26] and references therein). In other words, how stable are the predictions from free fermions under the inclusion of interaction effects. It is unclear if generic disorder alone prevents transport in realistic interacting materials, i.e., if a stable so-called many-body localized phase exists [27, 28].

Beyond that, including interactions properly into a model of the system is often crucial in order to obtain a correct description of physical effects. This starts at the correct description of quantum dots, effectively zero-dimensional objects whose physics can be dominated by interaction effects, goes over to the formation of a Mott-insulator in the course of an interaction driven quantum phase transition from a metallic into an insulating phase to the correct description of one-dimensional fermions appearing in Carbon nanotubes; in order to list only a few applications. Taking interactions into account properly in ab-initio simulations is also crucial for obtaining quantitatively correct results for binding energies of molecules or explaining their bonding structures. In the following, we therefore want to complement Sec. 2.4 about non-interacting fermions with an introduction into interaction fermions here.

We will discuss the particle number conserving case only. The general Hamiltonian of interacting fermions with two-particle interaction reads in its second quantized form

$$H = \sum_{j,k=1}^{M} t_{i,j} f_i^\dagger f_j + \sum_{i,j,k,l=1}^{M} v_{i,j,k,l} f_i^\dagger f_j^\dagger f_l f_k. \tag{2.58}$$

Note that, as H contains single and two-body terms only, we are able to write the expectation value of any state $\rho \in \mathcal{D}(\mathcal{F}_M)$ in terms of the single and two-body covariance matrices as

$$\mathrm{tr}(H\rho) = \mathrm{tr}(t\gamma^{(1)}[\rho]^T) + \mathrm{tr}(\mathrm{mat}(v)\gamma^{(2)}[\rho]^T), \tag{2.59}$$

where $\mathrm{mat}(v) \in \mathbb{C}^{M^2 \times M^2}$ denotes the rearrangement of the tensor v into a matrix with entries $\mathrm{mat}(v)_{(i,j),(k,l)} = v_{i,j,k,l}$. For a given state ρ with reduced 2-body density matrix $\rho^{(2)}$, we are able to calculate the energy expectation value of ρ for a general interacting Hamiltonian using Eq. (2.59) with knowing the $\mathcal{O}(M^4)$ entries of $\rho^{(2)}$ only – circumventing the exponential size of the full Fock space. However, we will see below, that this unfortunately does not yield a general efficient optimization method for the ground state problem of interacting fermions.

Typical models of interest for us in the following will be for instance second quantized versions of the electronic structure Hamiltonian for N electrons within the Born Oppenheimer approximation,

$$H = -\sum_{j=1}^{N} \frac{1}{2m_j}\Delta_i + V(\mathbf{x}_j) + \sum_{\substack{j,k=1 \\ j\neq k}}^{N} \frac{1}{2}\frac{1}{|\mathbf{x}_j - \mathbf{x}_k|}, \tag{2.60}$$

where $\mathbf{x}_j$ denotes the position of particle j and V is the external potential created from the nuclei. Using an appropriate single-particle basis a second quantized Hamiltonian as

in Eq. (2.58) is then derived as explained at the start of this chapter. These models allow us to perform ab-initio calculations for the electronic structure of molecular systems in order to identify, for instance, bonding and spectroscopic properties of molecules or solids. In addition we consider toy models of condensed matter systems. Most prominently different version of the Fermi-Hubbard model will appear, which in its simplest version for spin-polarized fermions in one spatial dimension reads

$$H = \sum_{j=1}^{M} \Big[t\big(f_j^\dagger f_{j+1} + f_{j+1}^\dagger f_j\big) + U n_j n_{j+1} \Big] \tag{2.61}$$

with $n_j = f_j^\dagger f_j$ denoting the occupation number operator of mode j and we employ either periodic ($f_{M+1} = f_1$) or open ($f_{M+1} = 0$) boundary conditions. The Hamiltonian in (2.61) provides one of the simplest toy models of interacting fermions, the first term capturing the kinetic energy of the particles while the second part implements a local two-body interaction, modeling for instance a Coulomb repulsion. However, the methods and results present in this thesis will not be restricted to specific Hamiltonians but be generally applicable. Specific systems provide us then with illustrative examples and motivate of course our investigations.

As in the case of non-interacting fermions we are interested in static properties, i.e., ground, low-lying excited or thermal states, as well as dynamic properties of interacting fermions. However, different from the free setting, obtaining these properties for interacting fermions either analytically or numerically is a very demanding and complex task and over the past decades the development of numerical and analytical approximation tools and ansatzes has been an important topic in theoretical physics. The question on exactly how demanding it is to solve these problems can be nicely captured using the tools of computational complexity theory as reviewed below. Furthermore, in the next chapter we will briefly review different numerical approaches, and go into quite some detail for one specific family of methods, employed for simulating interacting fermionic systems. Before we do that however, we first discuss one of the most important and at the same time simplest approximations imposed on interacting fermionic systems; the particle mean field or Hartree-Fock approximation as it will reappear multiple times in the course of this thesis.

2.5.1. Hartree-Fock Approximation

Mean field approximations in general ask for the best approximation of the state of a system by a product state. In fermionic systems the role of this product state is most commonly taken by an individual Slater determinant. The Hartree-Fock approximation in finite fermionic systems seeks then the best single-particle basis such that a property of choice, for instance the ground state energy, is best approximated using a fixed predefined Slater determinant.

Consider the case of a fixed particle number N. In case of approximating the ground state of the Hamiltonian given in Eq. (2.58), we choose without loss of generality $|HF\rangle = f_1^\dagger \dots f_N^\dagger |0\rangle$ in some initially chosen basis. We reach any Slater determinant with N

particles by varying the single-particle basis using mode transformations. The Hartree-Fock energy is hence defined as

$$E_{HF} = \min_{U \in U(M)} \Big[\mathrm{tr}\Big(U^\dagger t U \gamma^{(1)} [|HF\rangle\langle HF|]^T \Big) + \mathrm{tr}\Big((U^\dagger \otimes U^\dagger) \mathrm{mat}(v) (U \otimes U) \gamma^{(2)} [|HF\rangle\langle HF|]^T \Big) \Big]. \tag{2.62}$$

Inserting the concrete form of $|HF\rangle$ allows to simplify the expression above a bit further and there are several techniques of solving the above optimization problem numerically using self-consistency schemes (see e.g. [6, p. 146]) or direct optimization algorithms on Grassmann manifolds[8] [29].

Next to using it only for capturing static properties, the time evolution of a quantum system can also be performed within a mean field approximation. In the context of the Hartree-Fock approximation, the single-particle basis will vary in time mimicking the real time evolution of the interacting system by the one generated from a time-dependent free Hamiltonian. We will not be much concerned with the details of solving Hartree-Fock equations or schemes and therefore will not discuss them in further depth here.

In several instances does the Hartree-Fock approximation provide a surprisingly good account for essential features of the system. Dissociation curves, for instance, calculated based on the Hartree-Fock approximation can yield pretty accurate predictions of equilibrium configurations when compared with experimental data [6, Sec. 3.7]). However, there exists only little knowledge about when and why a Hartree-Fock approximation will work. We are unable to decide beforehand if the ground state energy of a given Hamiltonian can be well approximated within a Hartree-Fock approach or not. This lack of a general certificate of the Hartree-Fock approximation has a deeper reason as we will argue in the upcoming section. For a specific class of systems, however, we can obtain a bound to the error made by employing the Hatree Fock approximation in these systems as discussed in Ch. 4 by exploiting fermionic versions of de Finetti's theorem. There we find that systems in which all parts of the system are coupled equally to each other, a mean field description allows to capture local expectation values and that the Hartree-Fock approximation provides an accurate description of ground states energy densities.

2.5.2. Complexity Theory and Interacting Fermions

One might ask, how difficult are the problems regarding interacting fermions, e.g., calculating the electronic structure of molecules etc. Next to referring to a large zoo of methods solving certain instances of this problem illustrating that so far we were not able to find an unified efficient solution one can invoke results on the computational

[8] As any internal rotation inside the fully occupied or the unoccupied orbitals yields no change of the cost function, the optimization can be restricted to $Gr(N, M)$ (see also App. D for a short recap on Grassmann manifolds).

complexity of fermionic systems obtained in the past decade. These results tell us how efficient different computational models can solve a task, or more precisely, how difficult and hard it is to solve it. Most prominently, the classes NP and QMA contain all problems for which a suggested solution can be checked in a time scaling polynomially with the input size of the problem to be valid or not using a Turing machine or a quantum computer respectively – while it is highly expected that there are cases where these solution cannot be found in a polynomial time. For a complexity class C, a problem is C-hard if any problem in C could be solved with polynomial effort once an efficient solution for the initial problem is found and is denoted as C-complete if it is in addition a member of C. For a further introduction into the basic notation of complexity theory see for instance [3]. However, before we review the known results, let us note that these results have to be considered with some caution. They concern the worst case complexity of the problem at hand meaning that there can still be a method that solves the majority of instances but fails at tailored examples. Complexity results as presented here are therefore suited to formulate certain no-go arguments (such as: We will not be able to find a simple parametrization of all two-body density matrices arising from pure N-particle states on M modes for growing N and M).

We focus our discussion here solely on fermionic systems and only remark that the results below rely on complexity results obtained for spin systems. Concerning the approximation of ground states of interacting fermionic systems two fundamental results are known.

- Given as inputs a particle number N, a number of orbitals $M = \text{poly}(N)$, a Hamiltonian H as in (2.58) and energy E, then testing if the Hartree-Fock energy of H is below E up to a threshold ϵ scaling as $\epsilon = 1/\text{poly}(N)$ is NP-complete [1]. This result persists even if H is restricted to be translation invariant [30].

- Given as inputs a particle number N, a number of orbitals $M = \text{poly}(N)$, a Hamiltonian H as in (2.58) and energy E, then testing if the ground state energy of H is below E up to a threshold ϵ scaling as $\epsilon = 1/\text{poly}(N)$ is QMA-complete [2, 3].

Based on these results we can infer a few fundamental limitations to our ability of describing fermionic systems.

- Given as inputs a particle number N, a number of orbitals $M = \text{poly}(N)$, a single-particle density and a Hamiltonian H then evaluating the universal exchange correlation functional for the given single-particle density up to an error $\epsilon = 1/\text{poly}(N)$ is QMA-complete [1].

- Given as inputs a particle number N and a number of orbitals $M = \text{poly}(N)$, deciding if a matrix $\gamma \in \mathbb{C}^{M^2 \times M^2}$ is the two-body correlation matrix of any state $|\psi\rangle \in \mathcal{H}_N$, i.e., $\gamma^{(2)}[|\psi\rangle\langle\psi|] = \gamma$, up to an error $\epsilon = 1/\text{poly}(N)$ is QMA-hard [2].

These two insights limit the applicability of two different approaches of simplifying the computation of ground states of interacting fermionic systems. According to the first,

the universal functional of DFT can not be constructed and every DFT calculation needs to invoke approximations thereof. The second results shows us that despite being able to reduce the effort for evaluating the energy expectation value to scaling as $\mathcal{O}(M^4)$ with Eq. (2.59), we are not able to tell if the input $\gamma^{(2)}[\rho]$ is valid, i.e., belongs to a pure state with N particles on M modes. Note that from the hardness of the determination of the ground state energy follows an at least equal hardness of the time evolution as otherwise the energy could be inferred using phase estimation schemes [16, 31].

Instead of being demoralizing, these result should be seen as rough guides when dealing with the simulation of (fermionic) quantum systems and draw certain boundaries. It is quite notable and surprising that (classical) computers can not solve the ground state energy problem even within the crudest approximation, the Hartree-Fock approximation. The existence of numerous classical methods and successful applications of those however shows that the above results are indeed worst cases of a multifaceted problem. Furthermore, recently quantum algorithms were designed that aim for solving the ground state problem on a quantum computer – complementing the classical approaches with the potential of outperforming them in large classes of problems [32, 33] while staying of course unable to solve the full ground state energy problem in full generality. It is this certain boundary which motivates us to seek to understand additional structure of physical systems which allows for an often observed efficient description. In the last section of this introduction to finite fermionic system we want to account for a very generic structure present these systems which are captured by kinematic constraints. The obtained results are of a very fundamental nature and can provide a physical intuition for several effects but at the same time do not directly imply an efficient description due to their generality. The upcoming chapters focus then on more specific approaches to identify and understand the correlation structure of more restrictive settings and either provide concrete algorithms or schemes to simulate interacting fermionic systems more efficiently or aim at understanding rigorously the emergence of important structures.

2.6. Kinematic Constraints of Finite Fermionic Systems

The complexity and hardness of the description of fermionic systems makes it necessary to understand their structure in a more detailed manner. In the last part of this technical introduction we want to dwell a bit more on the underlying structure of the fermionic state space and its implication on physical quantities. The qualifying feature of a state to describe fermions is its intrinsic antisymmetry. A direct consequence of that is the famous Pauli exclusion principle which states that two fermions cannot occupy the same state at the same time. This simple principle directly links to fundamental aspects of physics on all scales. It directly implies the Aufbau principle accounting for the structure of the periodic table and explaining chemical and physical properties of the elements, it explains thermal and electric properties of solids by the formation of Fermi seas as well as the solidness of the matter surrounding us.

In finite fermionic systems the Pauli principle implies that for any fermionic state $\rho \in \mathcal{D}(\mathcal{F}_M)$ we can upper-bound the occupation of any mode i by 1, i.e., $\mathrm{tr}(f_i^\dagger f_i \rho) \leq 1$

$\forall i \in [M]$, irrespective of the chosen single-particle basis. Put differently the eigenvalues of the one-body covariance matrix are bounded by 1 from above, i.e., $\gamma^{(1)}[\rho] \leq \mathbb{1}$. We want to refer to such a constraint to be kinematic, as it is not specific to any dynamical or static properties of a specific system involving a specific Hamiltonian but originates from the state space structure itself. The particular constraint $\gamma^{(1)}[\rho] \leq \mathbb{1}$ will be denoted as Pauli constraint in the following.

It is often remarked that in fermionic systems the Pauli principle has to be replaced by the more general concept of having an antisymmetric wave function and the Pauli principle is then viewed among other results to be a consequence of this symmetry-restriction. This is in fact not true as one can argue easily at least for finite fermionic systems. Here, the Pauli principle directly implies the antisymmetric structure of the state space and both concepts are equivalent as can be seen as follows.[9]

Proof. Choose a set of single-particle orbitals $\phi_i = |i\rangle$ with $i \in [M]$ and consider the full N-particle Hilbert space of distinguishable particles $\mathcal{H}_1^{\otimes N} = \mathrm{span}(\{|i_1\rangle \otimes \cdots \otimes |i_N\rangle | i_1, \ldots, i_N \in [M]\})$. The Pauli exclusion principle states that only states are admissible in which no $|i\rangle$ appears more than once (otherwise two particles would occupy the same state). We therefore define two states $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}_1^{\otimes N}$ to be equivalent, denoted by $|\psi_1\rangle \sim |\psi_2\rangle$, if $|\psi_1\rangle = |\psi_2\rangle + |\varphi\rangle$ with $|\varphi\rangle \in I_N$ and $I_N = \mathrm{span}\{|j_1\rangle \otimes \cdots \otimes |j_k\rangle \otimes |j_k\rangle \otimes \cdots \otimes |j_{N-1}\rangle | k \in [N-1], |j_1\rangle, \ldots, |j_{N-1}\rangle \in \mathcal{H}_1\}$[10]. Note the subtlety that in the definition of I_N we allow for any states $|j_l\rangle$ in $\mathcal{H}_1$ which displays the fact that the Pauli principle applies to any single-particle basis. In contrast, in the case of hard-core bosons the equivalence would be defined by using states from one specific orthonormal basis of $\mathcal{H}_1$ only, the one in which the hard-core constraint is defined. The states respecting the Pauli principle are then the equivalence classes with respect to the relation above, i.e., $\mathcal{H}_N = \mathcal{H}_1^{\otimes N}/I_N$. This, however, is a standard definition for the exterior power, such that $\mathcal{H}_N = \bigwedge^N \mathcal{H}_1$ (see for instance [15, Sec. 6.4]). Performing this construction for all N yields the known Fock space. □

Therefore, the Pauli principle directly yields the antisymmetric structure of the Fock space without postulating it.

It was noticed early that in the case of the one-body reduced density matrix of mixed states the Pauli constraints are not only necessary but also sufficient meaning that the necessary and sufficient condition on $\gamma \in \mathbb{C}^{M \times M}$ such that there exists a state of N fermions in M modes $\rho \in \mathcal{D}(\mathcal{H}_N)$ with $\gamma^{(1)}[\rho] = \gamma$ are [34]

$$\gamma = \gamma^\dagger, \qquad \mathrm{tr}(\gamma) = N, \qquad 0 \leq \gamma \leq \mathbb{1}. \tag{2.63}$$

In light of Eq. (2.59) one might therefore ask how these conditions change if we restrict ourselves to pure states or want to assess the conditions on higher order reduced density matrices. This leads to the so-called k-body N-representability problems: What are the

[9] Thanks to Zoltán Zimborás for conjecturing this thought.

[10] It is sufficient to choose only neighboring orbitals to agree as we can always swap $|x\rangle, |y\rangle$ in $|x\rangle \otimes |y\rangle$, up to further elements contained in I_N by subtracting $(|x\rangle - |y\rangle) \otimes (|x\rangle - |y\rangle) \in I_N$ from it and therefore can construct for instance $|x\rangle \otimes |y\rangle \otimes |x\rangle$ from elements in I_N.

reduced k-body covariance matrices of a pure/mixed state $\rho \in \mathcal{D}(\mathcal{H}_N)$ of N particles in M modes? The interest in this question was actually mostly motivated by Eq. (2.59) as a solution of the N-representability problem for $k = 2$ would allow to eliminate the wave function from the ground state problem of interacting fermions. However, we have also already argued in Sec. 2.5.2 that any parametrization of the set of all admissible $\gamma^{(2)}[\rho]$ will be not efficient. Constraints such as the Pauli principle promise however to be interesting from a physical perspective beyond the technical aspects of Eq. (2.59). The range of fundamental and important physical effects which can be intuitively understood based on the Pauli principle illustrates the physical significance of even rather simple kinematic constraints. We will therefore shortly discuss the recent solution of the one-body N-representability problem for pure states in the following. Furthermore, despite the hardness of the two-body N-representability problem, it is of course possible to find necessary constraints such that outer approximations on the set of admissible $\gamma^{(2)}[\rho]$ can be formed as we will review below. We conclude this section with a discussion of constraints emerging from locality. Let us emphasis however that in many cases the practical applicability of the results we review here is often still to be shown and topic of recent research [35–38]. The following paragraphs serve therefore mostly as a collection of different approaches and constraints and discuss the current research.

2.6.1. One-Body Density Matrix Constraints

It was early noted that if we restrict ourselves to pure states, the solution of the one-body N-representability problem differs from the one found for mixed states. In the case of $N = 3$ and $M = 6$ it was found that a matrix $\gamma \in \mathbb{C}^{6\times 6}$ is 3-representable if it fulfills the conditions in Eq. (2.63) and in addition the sorted eigenvalues $\lambda_1 \geq \cdots \geq \lambda_6$ of γ have to fulfill [39, 40]

$$\lambda_1 + \lambda_6 = 1, \qquad \lambda_2 + \lambda_5 = 1, \qquad \lambda_3 + \lambda_4 = 1, \qquad \lambda_1 + \lambda_2 + \lambda_4 \leq 2. \tag{2.64}$$

This was the only known set of constraints until the one-body N-representability problem of pure states was completely solved [41–43]. It was found that the eigenvalues of all admissible one-body reduced density matrices, viewed as an vector in $\mathbb{R}^M$, lie in a convex polytope and that this polytope is indeed smaller as the one defined by the constraints of the mixed state problem in Eq. (2.63). The resulting condition on the spectrum of a matrix are then necessary and sufficient for N-representability and generalize the result for $N = 3$, $M = 6$. They are most conveniently cast into a form of a set of linear constraints, often called generalized Pauli constraints

$$\sum_{i\in[M]} a_i^{(k)} \lambda_i \geq a_0^{(k)}, \tag{2.65}$$

where k labels different constraints, whose intersection forms a convex polytope. The corresponding coefficients $a_i^{(k)}$ can be found numerically for given N and M. The result is based on an abstract investigation of the Hilbert space representation of mode transformations in Eq. (2.27). The adjoint of this map is used to project orbits in the

(N, M)	(3,6)	(3,7)	(4,8)	(4,9)	(5,10)
# constraints	4	4	15	60	161

Table 2.1. – Number of constraints found which characterize the polytope of sorted spectra of one-body covariance matrices of states $|\psi\rangle \in \mathcal{H}_N = \bigwedge^N \mathbb{C}^M$ for different N and M [42].

dual algebra of $U(\dim(\mathcal{H}_N))$, i.e., Hermitian operators in $\mathcal{B}(\mathcal{H}_N)$ with a fixed spectrum, onto the dual algebra of the mode transformations, the single-particle operators. It can be shown that this map associates to any state $\rho \in \mathcal{D}(\mathcal{H}_N)$ its single-particle covariance matrix. Using a generalization of Konstant's theorem it is then possible to characterize the orbits of single-particle covariance matrices defined by their corresponding spectrum contained in the projection of the orbit of a general state $\rho \in \mathcal{D}(\mathcal{H}_N)$ with a fixed spectrum to the single particle level. This investigation leads to constraints as above on the spectrum of the single-particle operator in order to be an admissible covariance matrix of an actual pure state [41–43].

This nicely solves the one-body N-representability problem from a mathematical point of view. However there are a few drawbacks from a practical point of view. There is no closed expression known for the coefficients $a_i^{(k)}$ and they have to be determined for every N and M. Even worse, the number of constraints seems to increase drastically in N and M as displayed in Tab. 2.1 and they are therefore only known for small systems. The scaling of the number of constraints as a function of N and M is up to now unknown and an understanding of it would yield important insight into the possibility of designing algorithms around the generalized Pauli principles.

It is currently one emerging area of research to try to infer how important these new constraints are in physical systems. Due to their single-particle nature, the constraints apply directly to non-interacting systems. If we minimize a linear function in $\gamma^{(1)}[\rho]$, e.g. the energy functional, the optimum will lie on a facet of the polytope. Then, some of the generalized Pauli constraints are saturated and they determine the structure of the corresponding ground states. However, in non-interacting systems ground states are always Slater determinants and the corresponding spectra are already captured by the Pauli principle, rendering this example trivial.

Surprisingly, specific interacting systems can be found in which these single-particle kinematic constraints are at least almost saturated. It was for instance argued that the first excited state of beryllium saturates a non-trivial generalized Pauli constraint [35]. Evaluating the generalized Pauli constraints under the inclusion of spin it was further suggested that the reduced saturation magnetization in different ferromagents can be explained by saturated single-particle constraints. Furthermore, it was argued that for increasing temperature the state of the system moves along the facets and that changes the magnetic properties of a material are connected to the lifting or saturation of a generalized Pauli constraint [36].

Furthermore, harmonically interacting particles were found to yield ground states that are close to a facet of the polytope [37]. Here the question arises whether a saturated

constraint is trivial, in the sense that the old fashioned Pauli constraints already imply them, or only captured by the new inequalities [38]. It is an open research question, if algorithms can be formulated that exploits this structure, optimizing for instance over individual facets only and if larger classes of interacting model exist whose physics is considerably and non-trivially constrained by these single-particle effects.

2.6.2. Two-Body Density Matrix Constraints

Concerning the investigation of interacting fermionic systems we would of course be more interested in the solution of 2-body N-representability problem. Here we already know that the parametrization of the exact set will be infeasible to obtain practically. However, it is still possible to formulate necessary conditions which allow to give outer approximations on the set of reduced two-body density matrices. Ultimately these conditions give rise to a hierarchy of constraints which in total are also sufficient. The constraints are formulated in the form of positivity relations of matrices which depend on the two-particle reduced density matrix. The basic insight underlying these conditions is that for any state $\rho \in \mathcal{D}(\mathcal{H}_N)$ not only $\gamma^{(2)}[\rho]$ has to be positive, i.e., $\gamma^{(2)}[\rho] \geq 0$, but all expectation values of the form $\mathrm{tr}(\sum_i A_i A_i^\dagger \rho)$ will be. Using $A_1 = \sum_{j,k} b_{j,k} f_j f_k$ or $A_1 = \sum_{j,k} b_{j,k} f_j^\dagger f_k$ and realizing that the expressions are positive for all coefficients b one obtains the so-called Q and G conditions [44]

$$\big(\mathrm{tr}(f_i f_j f_l^\dagger f_k^\dagger \rho)\big)_{(i,j),(k,l)} \geq 0, \qquad \big(\mathrm{tr}(f_i^\dagger f_j f_l f_k^\dagger \rho)\big)_{(i,j),(k,l)} \geq 0, \tag{2.66}$$

which, by exploiting the CAR of the fermionic operators, give positivity conditions on a matrix which depends linearly on $\gamma^{(2)}[\rho]$. Inserting $A_1 = \sum_{i,j,k} b_{i,j,k} f_i^\dagger f_j^\dagger f_k^\dagger$ and $A_2 = \sum_{i,j,k} \overline{b}_{i,j,k} f_i f_j f_k$ yields in the same way the T1 condition [44, 45] where the seemingly depends on $\gamma^{(3)}[\rho]$ cancels upon exploiting the CAR. In the same way the T2 and a hierarchy of higher conditions can be derived [44, 46]. Going a level higher in the hierarchy will take into account more of the antisymmetry restrictions of the full state and tightens the approximation on the set of reachable two-body reduced density matrices by the expense of the introducing more constraints. These positivity constraints can then be accounted for by semidefinite programming techniques (see e.g [47, 48]) and allow in principle to construct lower bounds on ground state energies which complement upper bounds on the energy from the optimization over restricted sets of states which will be explored in more detail in the upcoming chapter.

2.6.3. Local Constraints from Hermiticity

The previous two sections dealt with constraints on the full spectrum of particle reduced density matrices of fermionic states based on the antisymmetry of the state. The question arises whether these constraints can be combined with a local structure which is present in many physical systems. One simple result allowing for deriving local constraints is the result by Schur and Horn [49, 50]. It states that for any Hermitian matrix $A \in \mathbb{C}^{n \times n}$

the diagonal is majorized by its spectrum, i.e., $(A_{i,i})_{i\in[n]} \prec \mathrm{spec}(A)$, where for $a, b \in \mathbb{R}^n$ we say that a majorizes b, $b \prec a$, if

$$\sum_{i=1}^{k} b^{\downarrow} \leq \sum_{i=1}^{k} a^{\downarrow} \quad \forall k < n \qquad \text{and} \qquad \sum_{i=1}^{n} b^{\downarrow} = \sum_{i=1}^{n} a^{\downarrow} \tag{2.67}$$

with $b^{\downarrow}$ and $a^{\downarrow}$ denoting the decreasingly order sequences with the elements of a and b. The result is based solely on the Hermiticity of the matrix in question, e.g. a reduced density matrix. It can for instance be used to directly bound in any single-particle basis the occupation numbers $\mathrm{tr}(f_i^\dagger f_i \rho)$ of a Gaussian state with known spectrum of the one-body reduced density matrix, e.g., a thermal state of a free model.

The question arises if expectation values which are more complex than single densities can be bounded. Can we for instance establish bounds for operators supported on larger regions and account for the overlap of different regions consistently? One set of such bounds is provided by the following theorem.

The following result was obtained in collaboration with Etienne Werly but was unfortunately already known to Thompson [51].

Theorem 2. *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian and denote by $\lambda = \mathrm{spec}(A)$ the spectrum of A. For any $I \subset [n]$ we denote by λ_I the subsequence $(\lambda_i)_{i\in I}$ and by μ_I the spectrum of the principle submatrix of A with indices defined by I, i.e., $\mu_I = \mathrm{spec}((A_{i,j})_{i\in I, j\in I})$. Then for any $k, l \in \mathbb{N}$ with $l \leq k \leq n$*

$$(e_l(\mu_I))_{I\subset[n]:|I|=k} \prec (e_l(\lambda_I))_{I\subset[n]:|I|=k} \tag{2.68}$$

where $e_l(\,\cdot\,)$ denote the elementary symmetric polynomials defined by

$$e_l(x_1, \ldots, x_k) = \sum_{\substack{J\subset[k],\\ |J|=l}} \prod_{j\in J} x_j. \tag{2.69}$$

Proof. The simplest version of the proof follows partially along the argumentation of [51]. Let $U \in U(n)$ diagonalize A. Using the functionality of the wedge product we obtain for any $I \subset [n]$ with $|I| = k$

$$\Big(\bigwedge^k(\alpha\mathbb{1} - A)\Big)_{I,I} = \sum_{\substack{J\subset[n],\\ |J|=k}} \Big(\bigwedge^k U\Big)_{I,J} \Big(\bigwedge^k(\alpha\mathbb{1} - \Lambda)\Big)_{J,J} \Big(\bigwedge^k U^\dagger\Big)_{J,I}, \tag{2.70}$$

where Λ denotes the diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. From the relation of the elementary symmetric polynomials and the coefficients of the characteristic polynomial we conclude

$$\Big(\bigwedge^k(\alpha\mathbb{1} - A)\Big)_{I,I} = \det(\alpha\mathbb{1} - A|_{I,I}) = \sum_{l=0}^{k} (-1)^l e_l(\mu_I)\alpha^{k-l} \tag{2.71}$$

and a direct expansions shows that

$$\left(\bigwedge^k(\alpha\mathbb{1}-\Lambda)\right)_{J,J}=\prod_{j\in J}(\alpha-\lambda_j)=\sum_{l=0}^{k}(-1)^l e_l(\lambda_J)\alpha^{k-l}. \tag{2.72}$$

Comparing the coefficients of α^{k-l} yields that

$$e_l(\mu_I)=\sum_{\substack{J\subset[n],\\|J|=k}} W_{I,J}e_l(\lambda_J). \tag{2.73}$$

From $W_{I,J}=(\bigwedge^k U)_{I,J}\bigwedge^k U^\dagger_{J,I}$ being doubly-stochastic follows the majorization relation as claimed. □

This result extends the insights about possible occupation numbers in fermionic systems. It can for instance be applied to the setting of Gaussian states where expectation values split according to Wick's theorem into sums of products of second moments. To illustrate this, assume a pure Gaussian state ρ with single-particle covariance matrix $\gamma^{(1)}[\rho]$ and fixed particle number. Consider the expectation value of $f_i^\dagger f_j^\dagger f_j f_i$ which factors according to Wick's theorem into

$$\operatorname{tr}(\rho f_i^\dagger f_j^\dagger f_j f_i)=\operatorname{tr}(\rho f_i^\dagger f_i)\operatorname{tr}(\rho f_j^\dagger f_j)-\operatorname{tr}(\rho f_i^\dagger f_j)\operatorname{tr}(\rho f_j^\dagger f_i)=\det(\gamma(i,j)) \tag{2.74}$$

with $\gamma(i,j)=(\gamma^{(1)}[\rho]_{a,b})_{a\in\{i,j\},b\in\{i,j\}}$ being the 2×2 submatrix of $\gamma^{(1)}[\rho]$ containing the elements of the crossings of the i-th and j-th rows and columns only. The determinant $\det(\gamma(i,j))$ equals, however, $e_2(\operatorname{spec}(\gamma(i,j)))$ such that Thm. 2 bounds the possible values of such two- (or even more) particle operators. With this we are for instance able to bound Gaussian expectation values of local Hamiltonians. In practice, however, these bounds are only moderately useful as usually the exact expectation value can be derived by exploiting the structure of the Gaussian state which allows to reduce these calculations to the single particle Hilbert space (compare Sec. 2.4). Furthermore the insight above does not allow to formulate a variational scheme as the constraints of Thm. 2 are only necessary and not sufficient. It is therefore subject of current research to on the one hand site explore the structure of such constraints in more detail in order to potentially gain sets of sufficient constraints and on the other hand to obtain more general physical constraints along the lines of the argumentation above in non-trivial cases, e.g., in non-Gaussian settings or for higher order correlation functions. Furthermore, note that the constraints of Thm. 2 result from Hermiticity alone. They can hence also be applied to the two-particle reduced density matrix and result in new sets of constraints for those. Exploring in how far these constraints combined with the conditions derived from the antisymmetry are able bound two particle properties of general states presents an interesting direction of future research.

2.7. Summary

In this chapter we discussed the basic formalism needed in order to meaningfully capture different concepts of finite fermionic systems and introduced the notation that will reappear in the different parts of this thesis. We will see in the upcoming chapters, that the introduced notation is versatile enough to bridge the gap between very applied settings where we want to approximate static or dynamic properties of specific systems from first principles and very theoretical questions around the non-equilibrium dynamics of closed quantum systems and the validity of mean field approximations. We emphasized in particular the state space structure and the relation of different formulations and representations by introducing the Jordan-Wigner representation and rotations of the single-particle basis which will be useful later in numerical as well as theoretical investigations.

We then continued to discuss more physical aspects of fermionic systems by distinguishing free models from interacting fermions. We briefly explained how to solve free fermionic systems by choosing the appropriate single-particle basis, leading us to the concept of Gaussian states, and highlighted in contrast the underlying features and complexity of interacting fermionic systems. A deeper look at the state space structure revealed that the antisymmetry constraints of fermionic systems lead to a rich structure of kinematic constraints which extend the Pauli principle. We especially discussed the known constraints of the one and two-body covariance matrix as well as locally emerging constraints.

3. Capturing Correlations with Tensor Network States

Many problems in physics and applied mathematics give rise to very high-dimensional spaces which need to be accounted for by numerical methods. One of the most prominent examples of such problems is the approximation of the ground state of a quantum many-body system in a Hilbert space with a dimension that grows exponentially with the system size – but we could also think in general about the solution of large systems of differential equations, etc. [52]. Tackling these tasks by direct brute force approaches typically yields rather limited methods such as exact diagonalization for the ground state problem and the story would end here, if we could not reasonably assume that additional structures are present. One very common structure encountered is that the problem at hand typically contains only a limited number of parameters, as it could otherwise not even be formulated. Hamiltonians considered in many-body quantum mechanics are typically formulated with a number of parameters that scales at maximum polynomially in the system size. It is therefore natural to hope, that we are able to describe their ground states also with few parameters only once an appropriate parametrization is found. We have however also already argued in Sec. 2.5.2 that such a parametrization will not suffice to capture all systems of interacting fermions but at best only a large but restricted class of models.

In the context of simulating interacting fermionic systems, different approaches have been developed in the past decades in order to restrict the number of parameters in a physically meaningful manner which gave rise to a multitude of schemes and algorithms [6]. The Hartree-Fock method restricts its ansatz class to a single Slater determinant and optimizes the single-particle orbitals from which the Slater determinant is constructed. By this, effectively the best non-interacting description for the considered model is found. Physical quantities calculated based on a Hartree-Fock approximation often capture realistic systems quite accurately. However, they are prone to violating global symmetry constraints (see [5] for an instructive discussion) and neglect any non-Gaussian correlations present in the initial state. Next to being able to reproduce certain aspects of a realistic system, the Hartree-Fock solution of a system often serves as a basis for more exact schemes. The configuration interaction (CI) and coupled cluster (CC) methods apply excitation operators with a small number of free parameters and specific structure to a reference Slater determinant (for instance a Hartree-Fock solution) and optimize the corresponding parameters. Here, depending on the computational resources available, a systematic increase of the number of parameters is possible which interpolates between rough approximations and the exact simulation of the system. Both methods perform particularly well in systems where a good initial guess is available. Strongly correlated

settings where several Slater determinants are of equal high importance for the final state are however often only poorly captured and less efficient multi-reference methods need to be employed.

DFT exploits the Hohenberg-Kohn theorem which establishes a one-to-one correspondence between an external potential given for instance by positively charged cores and the one-particle reduced density matrix of the ground state wave function of interacting electrons subjected to the said potential [7, Ch. 1]. The problem then reduces to identifying the proper one-particle reduced density matrix which contains a number of parameters that grows only polynomially in the system size by solving a non-interacting proxy problem. Formally, however DFT is only capable of obtaining the ground state of an interacting problem and can only heuristically access excited states. Moreover, as discussed in Sec. 2.5.2, the universal exchange functional, needed in order to formulate the non-interacting proxy problem, can not be evaluated efficiently such that any DFT calculation relies on approximations thereof. In weakly correlated settings, heuristic functionals are known which allow for a fast and reliable simulation of large systems. The obtained results however formally remain heuristic and not based on first principles. Furthermore, in settings with strong correlations DFT calculations usually fail to produce correct results such that despite its fundamental beauty, DFT involves uncontrolled approximations and is applicable to specific cases only.

The approaches which gain efficiency though a restriction to certain ansatz-classes are contrasted by the schemes which try to circumvent the problem of the exponential size of the Hilbert space by using statistical techniques. Traditional Monte Carlo schemes however suffer in general from the sign problem in fermionic systems and are less versatile than in systems of distinguishable particles. Recently, schemes have been developed to partially overcome the sign problem by sampling directly coefficients in the abstract Fock space such that Monte Carlo methods can be applied for describing the electronic structure of small molecules [53] even in strongly correlated settings.

As the last set of methods let us introduce the schemes which we will discuss in detail in the upcoming chapter. In recent years it was shown that tensor network schemes can successfully be applied to simulate interacting fermionic systems and compensate some of the shortcomings of other established methods. From a technical point of view, efficient tensor network descriptions of a high-dimensional problem can be found once the coefficient tensor, be it the one of a quantum state or the solution of a general differential equation, has a low rank structure. In more physical terms this translates to the requirement that the state approximated by a tensor network decomposition is supposed to be only moderately correlated on large length scales, where we will make this notion more precise in the following. The restrictions imposed by the use of tensor network states (TNS) hence have a clear physical intuition. In the context of non-locally interacting fermionic systems, tensor network states are found to naturally capture strongly correlated multi-reference settings. This advantage compensates the higher computational costs of using tensor network based schemes such that methods as the density matrix renormalization group (DMRG) algorithm joined the canon of established tools [54–57].

In this chapter we review the basic concepts of tensor network methods from a practical

perspective with a particular focus on the simulation of non-local quantum systems. In doing so we also present a few smaller technical insights concerning the convergence behavior of the DMRG algorithm and the estimation of errors when computing excited states. We further argue that the requirement of TNS fulfilling an area law for all correlations leads to a too restrictive variational set of states in the context of fermionic systems. In the extreme, a simple Hartree-Fock approximation can outperform a tensor network approach if the wrong initial single-particle basis is chosen. We lift this strong dependence of the obtained results on the initial basis by extending the DMRG algorithm with rotations of the singe-particle basis. Combining multiple approaches for this task we set up a scheme that performs close to a black box tool and identifies the best set of orbitals in order to represent the target state as a TNS almost independently from the initially chosen single-particle basis. The presentation aims to be as self contained as possible and explains based on pseudo-code how to implement a ready-to-use DMRG method. In App. C we comment on the code structure developed in the course of this work in more detail.

3.1. Tensor Network States

In the following we introduce and review the basic concepts and notation needed in order to implement a working version of a DMRG algorithm for long range Hamiltonians. For this we will mostly keep the perspective of applied numerics but will also comment on the physical restrictions imposed by the use of tensor network states in order to allow for a more intuitive understanding. We will discuss tensor network states for finite-dimensional distinguishable particles. This contains by the Jordan-Wigner transformation finite fermionic models as explained in Sec. 2.2.1. Assume therefore a collection of V finite-dimensional quantum systems with Hilbert spaces $\mathcal{H}^{[j]} \simeq \mathbb{C}^{d_j}$ for $j \in [V]$. For each $j \in [V]$ let $\{|i_j\rangle | i_j \in [d_j]\}$ denote an orthonormal basis of $\mathcal{H}^{[j]}$. A general state vector $|\psi\rangle \in \bigotimes_{j=1}^{V} \mathcal{H}^{[j]}$ is then defined by a coefficient tensor $c \in \mathbb{C}^{d_1} \otimes \cdots \otimes \mathbb{C}^{d_V}$ as

$$|\psi\rangle = \sum_{i_1=1}^{d_1} \cdots \sum_{i_V=1}^{d_V} c_{i_1,\ldots,i_V} |i_1\rangle \otimes \cdots \otimes |i_V\rangle. \tag{3.1}$$

The coefficient tensor being exponentially large in the system size V renders it unpractical to use c directly in numerical schemes. One way forward to overcome this issue systematically is to find low tensor-rank decompositions or approximations to the coefficient tensor. The coefficient tensor is then represented as a tensor network.

3.1.1. Tensor Network Decompositions

We choose to represent a tensor network decomposition of a tensor c of order[1] V by a weighted graph $(\mathcal{V}, \mathcal{E}, \mathcal{K})$ with vertices $\mathcal{V}$, weighted edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V} \times \mathbb{N}$ and half edges

[1] The order of a tensor denotes the number of vector spaces it maps multilinearly to a scalar, i.e., the dimension of the array needed in order to represent the tensor.

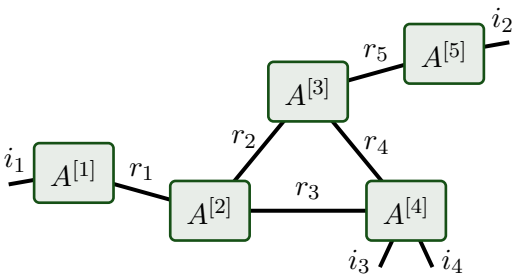

Figure 3.1. – Illustration of a tensor network decomposition of a tensor of order 4 by the corresponding graph. The full tensor c is then given by its components $c_{i_1,i_2,i_3,i_4} = \sum_{\alpha_1,\ldots,\alpha_5} A^{[1]}_{i_1,\alpha_1} A^{[2]}_{\alpha_1,\alpha_2,\alpha_3} A^{[3]}_{\alpha_2,\alpha_4,\alpha_5} A^{[4]}_{i_3,i_4,\alpha_3,\alpha_4} A^{[5]}_{i_4,\alpha_5}$ where the sum over α_k ranges from 1 to r_k. Note that in an abuse of notation, half edges are here labeled by the physical index i_j corresponding to the space $\mathcal{H}^{[j]}$ instead of the label j of the space. If they are not of concrete importance we will drop the labels on the half edges and edges in the following.

$\mathcal{K} \subset \mathcal{V} \times [V]$, where for every $j \in [V]$ there exists exactly one half edge $(v, j) \in \mathcal{K}$ [57, 58]. To each vertex v of the chosen decomposition we associate a component tensor $A^{[v]}$ of an order given by the coordination number of v. The dimension of each index of $A^{[v]}$ is hereby given by the weight of the corresponding edge or the dimension of the associated space $\mathcal{H}^{[j]}$ for a half edge. We will denote the indices $\{\alpha_j | j \in [|\mathcal{E}|]\}$ associated to edges to be virtual while the ones corresponding to half-edges as physical indices. We contract over all virtual indices such that two virtual indices joined by an edge are summed over while the open physical indices correspond to the indices of the full tensor c. Drawing the graph with all vertices and edges corresponds then essentially to the tensor network depicted in the Penrose graphical notation. This graphical notation provides a very compact and common tool which simplifies the notation as explained in Fig. 3.1. It is clear that given a tensor network decomposition $(\mathcal{V}, \mathcal{E}, \mathcal{K})$ of a high-dimensional tensor c, it might be possible to save computational resources by not specifying and modifying the large tensor c directly but only indirectly via the components $A^{[v]}$ of the tensor network. The efficiency advantage granted by the tensor network decomposition is then obviously controlled by the number of entries in the tensors $A^{[v]}$, or to be precise by the weights of the edges so the dimension $\{r_j | j \in [|\mathcal{E}|]\}$ of the virtual indices which are usually referred to as bond dimension or in the mathematical literature as rank. In the following we will either use the notation r_j for $j \in [|\mathcal{E}|]$ or r_e for $e \in \mathcal{E}$ in order to denote a specific bond dimension of a tensor network decomposition.

Clearly, there are a multitude of different decompositions of a given tensor as for instance purely virtual components, meaning tensors $A^{[v]}$ with only virtual indices, can be added in any number to the network. However there are only a few common decompositions used in the literature as they either allow for very compact decomposition or for the definition of a well defined tensor rank and other useful properties. The by far most common decomposition is the matrix product state (MPS) decomposition [8, 59–63], also known as tensor train decomposition in the mathematical community where it was

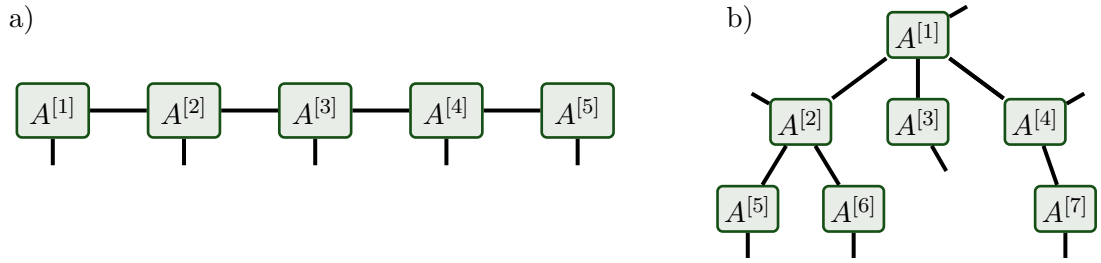

Figure 3.2. – We show the graphs of different tensor network decompositions. **a)** The open boundary MPS decomposition of a tensor of order 5 is shown. **b)** A general tree tensor network decomposition of an order 7 tensor is displayed with physical indices in each component $A^{[v]}$.

independently rediscovered [64]. Here the network is represented by a connected tree and the components $A^{[v]}$ are tensors with a maximal order 3 and exactly one physical index per component (compare Fig. 3.2). A second, more complex, decomposition with similar properties is given by the tree tensor network states (TreeTNS) decomposition. Here the graph of the decomposition is a general tree with again exactly one physical index per component (see again Fig. 3.2). Furthermore, there exists the hierarchical Tucker format which is mostly used in the applied mathematics community, where the graph of the decomposition is a tree with physical indices only at the leaves and purely virtual components in the bulk. The canonical decomposition of a tensor c defined by

$$c = \sum_{\alpha=1}^{r} A_\alpha^{[1]} \otimes A_\alpha^{[2]} \otimes \cdots \otimes A_\alpha^{[V]}, \tag{3.2}$$

for $A_\alpha^{[j]} \in \mathbb{C}^{d_j}$ can be viewed as a special case of the hierarchical Tucker format with one fixed core tensor joining all virtual legs. Note that every tensor can be decomposed in any of the formats if we do not restrict the allowed bond dimensions [52, 65].

On the formal side, many problems relevant for numerical applications involving tensors can be shown to be NP-hard [66], including for instance finding the best rank one approximation of a tensor of order > 2 or finding the minimal rank needed for a canonical decomposition of a tensor of order > 2. Let us however again emphasize that such hardness results, usually have to be considered with caution as discussed in Sec. 2.5.2.

Furthermore, from a numerical perspective it is of course interesting to have a closer look at the set of all possible tensors that can be reached by a fixed tensor network decomposition upon varying the entries of the components $A^{[v]}$ of the tensor network. For practical applications it is desirable that the corresponding sets of tensors form variational smooth manifolds which is indeed the case for most of the examples above [67, 68].[2] In the context of numerical variational schemes we also desire closedness

[2]To be precise, in order to obtain a smooth manifold we need an additional full rank assumption, meaning that we only consider tensors for which it would have been not possible to describe the same tensor with the same tensor network decomposition using lower weights on the edges. Considering the

properties of the corresponding sets. The canonical decomposition in Eq. (3.2), for instance, suffers from the so-called border rank problem. A sequence of tensors with a fixed bond dimension in the canonical decomposition can converge to a tensor with a larger bond dimension [52, Sec. 9.4]. This effect is usually based on cancellations and can result in numerical instabilities when we try to approximate tensors using tensor networks. Hence, in the context of TNS, the canonical decomposition is so far mostly of theoretical interest. In addition, the reachable set for a tensor network decomposition with cycles in the underlying graph is known to be not closed [52, 58] which can lead to similar instabilities.

Physicists are often less intrigued by those mathematical concerns and widely use different kinds of tensor network decompositions which contain cycles and find them to perform very well in practice. For two-dimensional lattice systems projected entangled pair states (PEPS) [59, 69] are frequently used, which are based on a decomposition according to a two-dimensional grid with a physical degree of freedom at each node and additional 4 virtual indices on tensors in the bulk. Furthermore, MPS with periodic boundary condition, [70, 71], so MPS where each tensor is of order 3, have been proven to be useful for one-dimensional systems with periodic boundary conditions and multiscale entanglement renormalization ansatz (MERA) states [72], which use a decomposition similar to the hierarchical Tucker format with additional nodes that form loops in the different levels of the tree, allow for the simulation of critical systems [73]. The decomposition chosen in order to represent quantum states with tensor networks depends therefore in practice usually on the geometry and symmetries of the underlying system. We will see in the next section that next to the obvious appeal of using a decomposition graph similar to the lattice of the system this choice ensures that the TNS is able to represent natural correlation structures of the system as the network inherits the physical locality structure.

3.1.2. Correlations in Tensor Network States

Given a tensor network decomposition $(\mathcal{V}, \mathcal{E}, \mathcal{K})$, we defined the corresponding set of tensor network states as the states $|\psi\rangle \in \bigotimes_{j=1}^{V} \mathcal{H}^{[j]}$ whose coefficient tensor c can be decomposed according to $(\mathcal{V}, \mathcal{E}, \mathcal{K})$. For explicit considerations we will restrict ourselves mostly to MPS. The set of MPS with bond dimensions $\{r_j | j \in [V]\}$ is then given by all states which are of the form

$$|\psi\rangle = \sum_{i_1,\ldots,i_V} \sum_{\alpha_1=1}^{r_1} \cdots \sum_{\alpha_{V-1}=1}^{r_{V-1}} A^{[1]}_{i_1,\alpha_1} A^{[2]}_{i_2,\alpha_1,\alpha_2} \cdots A^{[V]}_{i_V,\alpha_{V-1}} |i_1\rangle \otimes \cdots \otimes |i_V\rangle \tag{3.3}$$

with arbitrary components $A^{[j]} \in \mathbb{C}^{d_j \times r_{j-1} \times r_j}$, with formally $r_0 = r_V = 1$ and $\{|i_j\rangle | j \in [d_j]\}$ with $d_j = \dim(\mathcal{H}^{[j]})$ being an orthogonal basis of $\mathcal{H}^{[j]}$. Given a set of component

set of all possible tensors that can be reached with a fixed tensor network decomposition, including the ones that would have a simpler description, gives rise to algebraic varieties with those tensors that have a simpler description as singular points. Note that the set of all tensors with a fixed tensor network decomposition does not constitute a linear subspace or convex set as we will explicitly see in upcoming section for the MPS decomposition.

tensors $\{A^{[j]}|j \in [V]\}$ with appropriate dimensions we denote by $|(A^{[j]})_j\rangle$ the corresponding MPS which might be unnormalized.

For MPS we can easily see that the restriction of the maximal bond dimension to a fixed value leads directly to a bound on how correlated subsystems can be. Let $|\psi\rangle$ be an MPS of maximal bond dimension r. Given any $j \in [V-1]$, we split the systems into the sites $[j]$ and the remaining part $[j]^c$. From the definition of an MPS we directly obtain a decomposition of $|\psi\rangle$ into

$$|\psi\rangle = \sum_{\alpha=1}^{r_j} |A^{[1]}, \ldots, A^{[j]}\rangle_\alpha \otimes |A^{[j+1]}, \ldots, A^{[V]}\rangle_\alpha \tag{3.4}$$

with unnormalized

$$|A^{[1]}, \ldots, A^{[j]}\rangle_\alpha = \sum_{i_1,\ldots,i_j} \sum_{\alpha_1,\ldots,\alpha_{j-1}} A^{[1]}_{i_1,\alpha_1} \cdots A^{[j]}_{i_j,\alpha_{j-1},\alpha} |i_1\rangle \otimes \cdots \otimes |i_j\rangle, \tag{3.5}$$

$$|A^{[j+1]}, \ldots, A^{[V]}\rangle_\alpha = \sum_{i_{j+1},\ldots,i_V} \sum_{\alpha_{j+1},\ldots,\alpha_{V-1}} A^{[j+1]}_{i_{j+1},\alpha,\alpha_{j+1}} \cdots A^{[V]}_{i_V,\alpha_{V-1}} |i_{j+1}\rangle \otimes \cdots \otimes |i_V\rangle. \tag{3.6}$$

The Schmidt rank of the state for such a bipartition is therefore upper bounded by r, as $|(A^{[k]})_{k\in[j]}\rangle_\alpha$ span for $\alpha \in [r_j]$ a subspace of $\otimes_{k=1}^{j} \mathcal{H}^{[k]}$ of maximal dimension $r_j \leq r$. As a consequence the von Neumann entropy S_{vN} of the reduced state $\mathrm{tr}_{[j]} |\psi\rangle\langle\psi|$ has then by the Schur-concavity of S_{vN} the maximal value of $\ln(r)$ irrespective of the subsystem size. Therefore the physical consequence of enforcing a specific tensor network description on a state $|\psi\rangle$ is a limitation of the correlations present in $|\psi\rangle$. The same argument and result hold for TreeTNS with a maximal bond dimension, only that the bipartitions need to be chosen according to the arrangement of the sites inside the tree.

For more general bipartitions of the corresponding system and other tensor network states such as MERA states and PEPS, the Schmidt rank of chosen bipartitions are bounded by the product of the dimension of the cut edges in the graph. In the case of PEPS the Schmidt rank and by this for instance also the von Neumann entropy scale at maximum with the length of the cut through the network, but not with the volume of the patch enclosed by the cut. MPS and PEPS are therefore said to fulfill an area law for the contained correlations as the amount of correlation present in these states for any bipartition only scales with the size of the surface area and not with the volume of its parts. This is in stark contrast to generic quantum states. Drawing a quantum state at random from $\otimes_{j=1}^{V} \mathcal{H}^{[j]}$ one finds with high probability a state whose correlations do not fulfill an area law but a volume law [74, 75], meaning that the amount of correlations of a subsystem with its surrounding will scale with the total size of the subsystem so that every constituent is correlated with every part of the system in an unstructured way. Such states illustrate the limits and restrictions of tensor network methods as they are not efficiently representable and typically also only badly approximable by tensor network states as essentially the full exponential complexity of the Hilbert space is explored.

However, naturally occurring states can often be well approximated by tensor network states. On the one hand, this is seen simply from the success of different tensor network

methods in many instances [8, 59–63], on the other hand it is proven by rigorous results on the correlation structure of ground states of one-dimensional local Hamiltonians [9, 76]. Assuming a gap between the unique ground state and the first excited state it is possible to show that the Rényi entropies S_α defined in Eq. (2.4) fulfill an area law for $\alpha \leq \alpha_{\text{crit}} < 1$ [9]. On the other hand, if a state exhibits an area law for any Rényi entropy with $\alpha < 1$ then it can be efficiently approximated with an MPS with a bond dimension that scales only polynomially with the system size and algebraically in the desired error threshold [9, 77]. Local properties of such a state, so expectation values with local observables like the terms of a local Hamiltonian, can however already be correctly reproduced with a bond dimension independent of the system size which explains the enormous success of MPS as numerical tool for obtaining ground state energies of local one-dimensional systems. Recently, algorithms have been formulated which reliably identify MPS approximations to ground states of local one-dimensional gapped systems of distinguishable particles in a time that scales only polynomially in the system size [78, 79]. These algorithm complement more heuristic methods such as the more efficient but less controlled DMRG which we will discuss below in more detail.

In higher-dimensional systems neither of the two results holds in such generality. It is still at maximum conjectured that typical ground states of local systems fulfill an area law in two dimensions and even an area law of every Rényi entropy will not imply the approximability of the corresponding state by a PEPS [80], which would seem to be the natural candidate.

In the MERA, additional purely virtual components are added to the network which renormalize the entanglement on different length scales [72]. These additional degrees of freedom allow MERA states in one-dimensional systems to contain correlations that go beyond a strict area law and allow for correction to the area law that scale logarithmically with the subsystem size even with fixed bond dimensions. MERA is therefore well suited to treat one-dimensional critical systems where such violations of area laws are expected [73, 81]

The result above concerning the natural occurrence of MPS in physical systems needs locality as an underlying structure in the Hamiltonian. For many systems of interest, originating for instance from the investigation of solids, this is a sane assumption (at least within approximative descriptions). However, only little is known once we reach the realm of non-local systems, which are encountered for instance in the context of quantum chemistry. If we are interested in the ground state configuration of N electrons inside a molecule, one generically encounters non-local Hamiltonians without a strict underlying geometry, i.e., we can usually not speak naturally of neighboring sites or having a n-dimensional system. In these cases no general results on the correlation structure of the ground states of such systems are known. However, tensor network methods can still be applied to these systems as they provide a numerical tool, capable of representing strong correlations once the bond dimension is high enough and intrinsically allow for an error control by scaling the bond dimension appropriately. If applied to non-local systems, tensor network state methods do provide controlled schemes which allow for approximations of ground and excited states. However, the bond dimensions needed

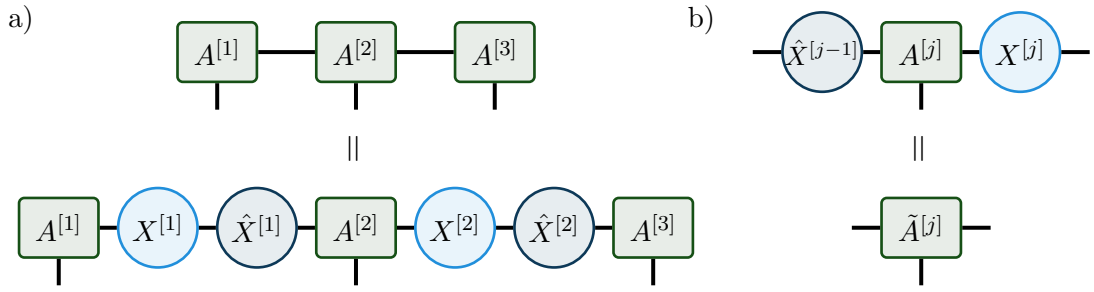

Figure 3.3. – Illustration of the gauge freedom of MPS. For a given MPS decomposition with components $\{A^{[j]}|j \in [V]\}$ we can insert right-invertible transformations $X^{[j]}$ and their right inverse $\hat{X}^{[j]}$ as shown in a) on 3 sites without altering the global tensor. For the new MPS components $\tilde{A}^{[j]}$ defined in b) we obtain then that $|(A^{[j]})_{j\in[V]}\rangle = |(\tilde{A}^{[j]})_{j\in[V]}\rangle$.

are often very large such that in these cases the used code is operated at the limits of the corresponding implementation as the correlation structure in realistic states often only poorly matches the one of MPS. The method developed in this chapter allows to partially overcome this limitation of tensor network methods and to efficiently handle highly entangled states if an additional structure is present. In order to discuss the resulting algorithm in detail we will first introduce further technical details about MPS and existing algorithms for approximating ground states with them. In the remainder of this section we further discuss the structure of MPS and introduce normal forms that are crucial for practical purposes and review the incorporation of global symmetries, such as having states with a fixed particle number, into tensor network states.

3.1.3. Matrix Product States: Normal Forms and Operations

Repeating the definition for convenience, given for a specific set of bond dimensions $\{r_j|j \in [V-1]\}$ valid component tensors $\{A^{[j]}|j \in [V]\}$ of an MPS, the MPS $|(A^{[j]})_{j\in[V]}\rangle \in \bigotimes_{j=1}^{V} \mathcal{H}^{[j]}$ is defined by

$$|(A^{[j]})_{j\in[V]}\rangle = \sum_{i_1=1}^{d_1} \cdots \sum_{i_V=1}^{d_V} \sum_{\alpha_1=1}^{r_1} \cdots \sum_{\alpha_{V-1}=1}^{r_{V-1}} A^{[1]}_{i_1,\alpha_1} A^{[2]}_{i_2,\alpha_1,\alpha_2} \cdots A^{[V]}_{i_V,\alpha_{V-1}} |i_1\rangle \otimes \cdots \otimes |i_V\rangle \quad (3.7)$$

with $\{|i_j\rangle|j \in [d_j]\}$ denoting an orthonormal basis of $\mathcal{H}^{[j]}$. The decomposition of a specific state $|\psi\rangle$ into an MPS is not unique meaning that given any MPS with components $A^{[j]}$ with $j \in [V]$, we can construct new component tensors $\tilde{A}^{[j]}$ with $|(A^{[j]})_{j\in[V]}\rangle = |(\tilde{A}^{[j]})_{j\in[V]}\rangle$. As depicted in Fig. 3.3 we can insert for any $j \in [V-1]$ between the two components any transformation $X^{[j]} \in \mathbb{C}^{r_j \times k}$ with a right-inverse $\hat{X}^{[j]}$, i.e., $X^{[j]}\hat{X}^{[j]} = \mathbb{1}_{r_j}$ without altering the global tensor resulting from contracting the MPS over all virtual indices. For general TNS this of course generalizes to the freedom that we can insert such transformations at any edge in the network. Staring from some initial

Figure 3.4. – Illustration of the condition on the component $A^{[j]}$ of an MPS for being left normalized (shown in a)) and right normalized (shown in b)).

component tensors $\{A^{[j]}|j \in [V]\}$, we can define a new set of tensors $\{\tilde{A}^{[j]}|j \in [V]\}$ by contracting the original components with transformations $X^{[j]}$ of our choice as illustrated in Fig. 5.1 while the MPS stays unchanged meaning that $|(A^{[j]})_{j\in[V]}\rangle = |(\tilde{A}^{[j]})_{j\in[V]}\rangle$. This non-uniqueness of a tensor network representation and the resulting gauge freedom of tensor network decompositions has several implications.

First of all note that the bond dimension can be altered. By choosing transformations $X \in \mathbb{C}^{r_j \times \tilde{r}_j}$ with $\tilde{r}_j \geq r_j$, the new components $\tilde{A}^{[j]}$ constitute an MPS representation of $|(A^{[j]})_{j\in[V]}\rangle$ with bond dimensions $\tilde{r}_j$. It is easy to see that there exists for every tensor c an MPS decomposition with minimal bond dimensions. If $c \in \bigotimes_{j=1}^{V} \mathbb{C}^{d_j}$ is the tensor to be decomposed, then for every MPS decomposition of c the bond dimensions r_j are restricted by the ranks of different matrices obtained from rearranging the elements of c according to

$$r_j \geq \operatorname{rank}\Big((c_{i_1,\ldots,i_V})_{(i_1,\ldots,i_j)\in[d_1]\times\cdots\times[d_j],(i_{j+1},\ldots,i_V)\in[d_{j+1}]\times\cdots\times[d_V]}\Big), \tag{3.8}$$

by the uniqueness of the matrix rank. In addition such a decomposition can always be constructed [65]. Note that this result is expected from the relation of correlation, i.e., entanglement, in an MPS and its bond dimensions. We can therefore define the set of all MPS that have a sequence of minimal bond dimensions $(r_j)_{j\in[V]}$ and denote it by $\mathcal{M}((r_j)_{j\in[V]})$.

However, even after fixing the bond dimensions, the MPS representation of a certain state $|\psi\rangle \in \mathcal{M}((r_j)_{j\in[V]})$ is still not unique. Given a valid MPS representation of $|\psi\rangle$ with the components $A^{[j]}$ for $j \in [V]$, the components

$$\tilde{A}^{[j]}_{i_j,\tilde{\alpha}_{j-1},\tilde{\alpha}_j} = \sum_{\alpha_{j-1}=1}^{r_{j-1}} \sum_{\alpha_j=1}^{r_j} [(X^{[j-1]})^{-1}]_{\tilde{\alpha}_{j-1},\alpha_{j-1}} A^{[j]}_{i_j,\alpha_{j-1},\alpha_j} X^{[j]}_{\alpha_j,\tilde{\alpha}_j} \tag{3.9}$$

for arbitrary $X^{[j]} \in GL(r_j, \mathbb{C})$ with $|X^{[0]}| = |X^{[V]}| = 1$, yield the same MPS, i.e., $|(A^{[j]})_{j\in[V]}\rangle = |(\tilde{A}^{[j]})_{j\in[V]}\rangle = |\psi\rangle$. In order to investigate this on the level of the component tensors let us define the spaces of components $\mathcal{A}^{[j]} \subset \mathbb{C}^{d_j \times r_{j-1} \times r_j}$ compatible with a given set of minimal bond dimensions via

$$\begin{aligned}\mathcal{A}^{[j]} = \{A \in \mathbb{C}^{d_j \times r_{j-1} \times r_j} | \operatorname{rank}((A_{i_j,\alpha_{j-1},\alpha_j})_{(\alpha_{j-1},i_j),(\alpha_j)}) = r_j \wedge \\ \operatorname{rank}((A_{i_j,\alpha_{j-1},\alpha_j})_{(\alpha_{j-1}),(i_j,\alpha_j)}) = r_{j-1}\},\end{aligned} \tag{3.10}$$

Algorithm 1 – Algorithm to left and right normalize an MPS until site j

procedure LEFT NORMALIZE($(A^{[k]})_{k\in[V]}, j$)
  for $k = 1$ to j **do**
    $U^{[k]}, \Sigma^{[k]}, V^{[k]} \leftarrow \mathrm{SVD}\Big((A^{[k]}_{i_k,\alpha_{k-1},\alpha_k})_{(i_k,\alpha_{k-1}),(\alpha_k)}\Big)$
    $A^{[k]} \leftarrow (U^{[k]}_{(i_k,\alpha_{k-1}),\alpha_k})_{i_k,\alpha_{k-1},\alpha_k}$
    if $k+1 \leq V$ **then**
      $A^{[k+1]} \leftarrow \left(\sum_{\alpha=1}^{r_k} \Sigma^{[k]}_{\alpha_k,\alpha_k} V^{[k]}_{\alpha_k,\alpha} A^{[k+1]}_{i_{k+1},\alpha,\alpha_{k+1}}\right)_{i_{k+1},\alpha_k,\alpha_{k+1}}$
    else
      $A^{[V]} \leftarrow V^{[V]} A^{[V]}$
procedure RIGHT NORMALIZE($(A^{[k]})_{k\in[V]}, j$)
  for $k = V$ to j **do**
    $U^{[k]}, \Sigma^{[k]}, V^{[k]} \leftarrow \mathrm{SVD}\Big((A^{[k]}_{i_k,\alpha_{k-1},\alpha_k})_{(\alpha_{k-1}),(i_k,\alpha_k)}\Big)$
    $A^{[k]} \leftarrow (V^{[k]}_{\alpha_{k-1},(i_k,\alpha_k)})_{i_k,\alpha_{k-1},\alpha_k}$
    if $k-1 \geq 1$ **then**
      $A^{[k-1]} \leftarrow \left(\sum_{\alpha=1}^{r_{k-1}} A^{[k-1]}_{i_{k-1},\alpha_{k-2},\alpha} U^{[k]}_{\alpha,\alpha_{k-1}} \Sigma^{[k]}_{\alpha_{k-1},\alpha_{k-1}}\right)_{i_{k-1},\alpha_{k-2},\alpha_{k-1}}$
    else
      $A^{[1]} \leftarrow A^{[1]} U^{[1]}$

such that the $\mathcal{A}^{[j]}$ for $j \in [V]$ contain all tensors which yield a full rank MPS with minimal bond dimension $\{r_j | j \in [V]\}$. We can now divide out the from the transformation in Eq. (3.9) induced action by of $\bigtimes_{j=1}^{V} GL(r_j, \mathbb{C})$ on the parameter manifold $\bigotimes_{j=1}^{V} \mathcal{A}^{[j]}$ of the component which gives $\mathcal{M}((r_j)_{j\in[V]}) = \bigotimes_{j=1}^{V} \mathbb{C}^{r_{j-1}\times d_j \times r_j} / \bigtimes_{j=1}^{V} GL(r_j, \mathbb{C})$ the structure of a smooth manifold [57, 68]. The closure $\overline{\mathcal{M}((r_j)_{j\in[V]})}$ corresponds to the set of all states which can be represented as an MPS with bond dimensions $(r_j)_{j\in[V]}$ and forms an algebraic variety (see for instance [57]). In practice we can lift the gauge freedom of MPS by choosing canonical forms [82] and use it in order to define normal forms. We will only need the latter for the following. A component $A^{[j]}$ is called left or right normalized if it fulfills the conditions depicted in Fig. 3.4. Using the gauge invariance of an MPS we can bring it into a mixed normalized form with respect to any $j \in [V]$, meaning that all components on sites $k < j$ are left normalized and all components on sites $k > j$ are right normalized. In practice, depending on the desired normalization, the individual components $A^{[j]}$ are iteratively rearranged into matrices, decomposed using the singular value decomposition (SVD)[3] and parts of the results of the SVD are absorbed into the component $A^{[j+1]}$ or $A^{[j-1]}$ as explained in Alg. 1. The normalized form has several immediate advantages and applications. If an MPS is given in a mixed normalized form with respect to site j, the calculation of expectation

[3]Every matrix $A \in \mathbb{C}^{n\times m}$ with rank r can be decomposed into U, Σ, V with $U\Sigma V = A$, $U \in \mathbb{C}^{n\times r}$,$\Sigma \in \mathbb{R}^{r\times r}$ and $V \in \mathbb{C}^{r\times m}$ with $U^\dagger U = \mathbb{1}_r$, $VV^\dagger = \mathbb{1}_r$ and Σ diagonal with $\Sigma_{j,j} > 0$ for $j \in [r]$.

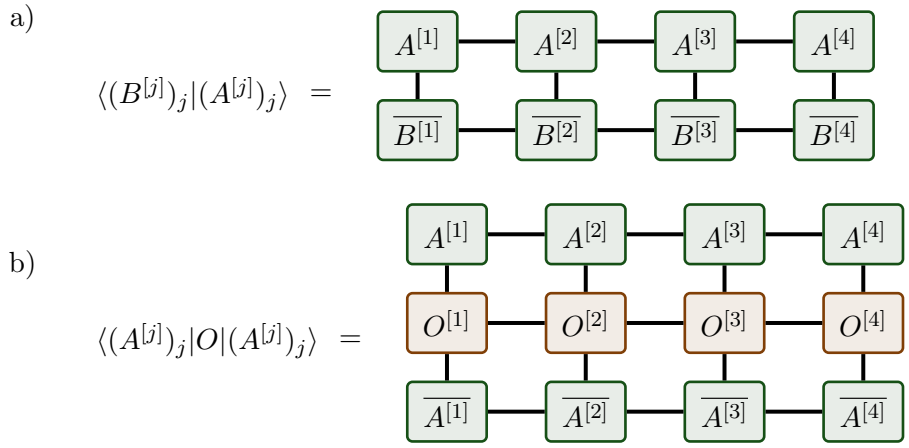

Figure 3.5. – Illustration of the tensor networks for calculating the scalar product of the two MPS $|(A^{[j]})_j\rangle$ and $|(B^{[j]})_j\rangle$ in a) and in b) the expectation value of the MPO decomposed operator O with respect to the MPS $|(A^{[j]})_j\rangle$. In b) a general MPO decomposition of an operator O is shown. The resulting components $O^{[j]}$ are tensors of order 3 or 4 after splitting the physical indices into two indices as explained in Sec. 3.1.3.

values of observables supported on j only simplifies drastically and can be done in a time independent of the system size as we can insert the relations in Fig. 3.4 starting from the left and right boundary. Secondly, if a state is completely right or left normalized the vectors $|(A^{[k]})_{k\in[j]}\rangle_\alpha$ or $|(A^{[k]})_{k\in[V]\setminus[j]}\rangle_\alpha$ defined in Eq. (3.5) and (3.6) are orthonormal for all $j \in [V]$. Starting from a right normalized MPS and applying the left normalization until site j, will not only result in a mixed normalized MPS with respect to site $j+1$ but also all singular value matrices $\Sigma^{[k]}$ obtained during the left normalization, correspond to the Schmidt values of the state with respect to the bipartition $[k]$ and $[k]^c$. Converting a given MPS with bond dimensions r_j and physical dimensions d_j into a normalized form and by this the calculation of the Schmidt coefficients can be done in a time scaling as $\mathcal{O}(Vdr^3)$ with $d = \max(\{d_j | j \in [V]\})$ and $r = \max(\{r_j | j \in [V-1]\})$.

Next to determining the Schmidt coefficients with respect to a left-right partition of the system, we are able to perform other operations efficiently within the MPS framework [8]. Consider two MPS of the same physical system with maximal physical dimensions $d = \max(\{d_j | j \in [V]\})$ and bond dimensions $r_j^{(1)}$ and $r_j^{(2)}$ with $j \in [V-1]$ and $r = \max(\{r_j^{(a)} | j \in [V-1], a \in \{1,2\}\})$. The two MPS can be added to a new MPS, where the bond dimensions r_j of the new MPS are upper bounded by $r_j \leq r_j^{(1)} + r_j^{(2)}$ and the time needed scales as $\mathcal{O}(Vdr^2)$. Note that as in general the bond dimension increases under addition, $\overline{\mathcal{M}((r_j)_{j\in[V]})}$ is neither a linear subspace of the full Hilbert space nor convex. Furthermore, we can evaluate the scalar product of the two MPS according to the scheme showed in Fig. 3.5. If performed in the correct order the contraction of the network in Fig. 3.5 a) can be achieved in a time scaling as $\mathcal{O}(Vdr^3)$.

The construction of MPS is quite general. It can also be applied to the Hilbert space of

operators acting on $\bigotimes_{j=1}^{V} \mathcal{H}^{[j]}$. In order to avoid confusion the resulting decomposition is denoted as matrix product operator (MPO) decomposition and the physical indices of dimension d_j^2 are split into two indices in order to highlight the character of an MPO of being a linear map between MPS (see also Fig. 3.5 b)). As an example, note that for $\mathcal{H}^{[j]} = \mathbb{C}^2$ for all $j \in [V]$ a product of Pauli operators such as $Z \otimes \cdots \otimes Z \otimes (X - iY) \otimes \mathbb{1}_2 \otimes \ldots$ can be written as an MPO with maximal bond dimension $r^{(O)} = 1$ due to the product structure. In addition every operator that acts only on a finite region $S \in [V]$ acquires an MPO decomposition with a maximal bond dimension $r^{(O)} \leq d^{|S|}$. Given an MPO of bond dimensions $r_j^{(O)}$ with $r^{(O)} = \max(\{r_j^{(O)} | j \in [V-1]\})$ we are able to calculate its expectation value with an MPS in a time scaling as $\mathcal{O}(V r^{(O)} [d^2 r^{(O)} r^2 + d r^3])$ by contracting the network displayed in Fig. 3.5 b).

3.1.4. Symmetric Tensor Network States

The incorporation of symmetries of a system such as a fixed particle number or spatial reflection symmetries into MPS or TNS in general is interesting and important from a conceptional and a practical point of view. On a conceptual level the investigation of how symmetry groups act on TNS allows for an detection and analysis of symmetry protected topological order theoretically [11] and also practically [12].

The implementation of symmetries into TNS and by this the possibility to restrict ourselves to a defined symmetry sector has on the other hand clear practical advantages for numerical investigations. First, the concrete question at hand, often includes the restriction to a symmetry sector, e.g. find the lowest eigenenergy of a state with N electrons. Secondly, the restriction to a specific symmetry sector allows to reduce the number of parameters as only a fraction of the full Hilbert space is concerned, which however will in general still be exponentially large in the system size.

It is possible to incorporate general symmetries that act locally on the individual constituents $\mathcal{H}^{[j]}$ of the full Hilbert space and split them into irreducible representations (irreps). To be precise we need that the unitary representation of the symmetry operation U_g on the full Hilbert space decomposes as $U_g = \bigotimes_{j=1}^{V} U_g^{[j]}$ for all $g \in G$ with G being the symmetry group under consideration which is assumed to be finite or compact in the following.[4] A state $|\psi\rangle \in \bigotimes_{j=1}^{V} \mathcal{H}^{[j]}$ is invariant under this symmetry if

$$\bigotimes_{j=1}^{V} U_g^{[j]} |\psi\rangle = e^{i\phi_g} |\psi\rangle \tag{3.11}$$

where the $e^{i\phi_g}$ forms a one-dimensional unitary representation of G [5]. We can ensure

[4] As an example consider the case of a fixed magnetization of the total m_z quantum number in a spin-1/2 system (which translates by the Jordan-Wigner transformation to a fixed total particle number in the fermionic system). The corresponding symmetry group would be $U(1)$ with the representation $\bigotimes_{j=1}^{V} e^{ig(Z+\mathbb{1}_2)/2}$ with Z being the Pauli-Z matrix and $g \in [0, 2\pi] \simeq U(1)$ that splits the local space $\mathbb{C}^2$ in the two irreducible components $\mathrm{span}(|m_z^{[j]} = -1/2\rangle)$ and $\mathrm{span}(|m_z^{[j]} = 1/2\rangle)$.

[5] Continuing with the example of a fixed magnetization, if $|\psi\rangle$ has the magnetization m then $e^{i\phi_g} = e^{ig(m+V)/2}$ with V denoting the system size which is an irrep of $U(1)$ labeled by $(m+V)/2 \in \mathbb{N}$

this property for a TNS if we demand it to be composed of symmetric components $A^{[j]}$ (see for instance [83, 84]). Assume a tensor network decomposition according the graph $(\mathcal{V}, \mathcal{E}, \mathcal{K})$. Assign then a direction to each edge of the graph and label all half edges as ingoing. We denote by $\mathrm{in}(v)$ and $\mathrm{out}(v)$ the set of in- and out-going edges and half edges to a vertex $v \in \mathcal{V}$ and denote by $\mathcal{H}^{[e]}$ the vector space associated to the edge or half edge $e \in \mathcal{E} \cup \mathcal{K}$ with $\dim(\mathcal{H}^{[e]}) = r_e$ for edges and $\dim(\mathcal{H}^{[e]}) = d_e$ for half edges. Furthermore, we need unitary representations of the group on the virtual spaces $\mathcal{H}^{[e]}$ with $e \in \mathcal{E}$ which is induced by the representation on the physical spaces.[6] From the above we then have an unitary representation of G associated to every edge and half edge which we denote by $U^{[e]}$ with $e \in \mathcal{E} \cup \mathcal{K}$. A component tensor $A^{[v]}$ for $v \in \mathcal{V}$ is then called symmetric if it is invariant under the joint action of the representation of the same element $g \in G$ on all spaces as depicted in Fig. 3.6 a). A TNS constructed from symmetric components will then be a symmetric state as argued in Fig. 3.6 b).

The condition Fig. 3.6 a) imposes a structure on the component $A^{[j]}$ by Schur's lemma [85, p. 57][7]. In order to see this note that due to the representations of G, the physical spaces $\mathbb{C}^{d_j}$ and virtual space $\mathbb{C}^{r_k}$ for $j \in [V]$ and $k \in [|\mathcal{E}|]$ decompose into multiples of irreducible subspaces $\mathcal{H}_\lambda$, where λ labels the according irrep. To be precise, every space $\mathcal{H}^{[e]}$ with $e \in \mathcal{E} \cup \mathcal{K}$ decomposes as

$$\mathcal{H}^{[e]} = \bigoplus_\lambda D_\lambda^{[e]} \otimes \mathcal{H}_\lambda \tag{3.12}$$

where degeneracy space $D_\lambda^{[e]}$ accounts for the fact that $\mathcal{H}_\lambda$ might appear multiple times in $\mathcal{H}^{[e]}$.[8] The representation of G on $\mathcal{H}^{[e]}$ splits then as

$$U_g^{[e]} = \bigoplus_\lambda \mathbb{1}_{\dim(D_\lambda^{[e]})} \otimes U_{\lambda,g}. \tag{3.13}$$

Collecting for a given vertex $v \in \mathcal{V}$ all spaces corresponding to ingoing and outgoing edges and half edges into two big tensor product spaces $\mathcal{H}_{\mathrm{in}}^{[v]} = \bigotimes_{e \in \mathrm{in}(v)} \mathcal{H}^{[e]}$ and $\mathcal{H}_{\mathrm{out}}^{[v]} = \bigotimes_{e \in \mathrm{out}(v)} \mathcal{H}^{[e]}$, the component $A^{[v]}$ can be viewed as a linear transformation from $\mathcal{H}_{\mathrm{in}}^{[v]}$ to $\mathcal{H}_{\mathrm{out}}^{[v]}$. The group G then acts by the product representation on $\mathcal{H}_{\mathrm{in}}^{[v]}$ and $\mathcal{H}_{\mathrm{out}}^{[v]}$ which again decompose into irreducible subspaces. Note that if either $\mathrm{in}(v)$ or $\mathrm{out}(v)$ are empty we set $\mathcal{H}_{\mathrm{in}}^{[v]} = \mathbb{C}$ or $\mathcal{H}_{\mathrm{out}}^{[v]} = \mathbb{C}$ accordingly with G acting trivially on it. If $A^{[v]}$ is a symmetric tensor, it commutes with the action of G and decomposes therefore by Schur's lemma into a structure factor determined by the irreducible subspaces and into a parameter

[6] The possible representations on the virtual spaces originate from the ones on the physical spaces. In the case of a fixed magnetization and an MPS, the first virtual space can contain the components $|m_z = -1/2\rangle$ and $|m_z = 1/2\rangle$. The second virtual space can, after the addition of a second spin, contain the contributions $|m_z = 0\rangle$ and $|m_z = \pm 1\rangle$ and so on.

[7] Schur's lemma states that given a group G that acts with the irreps U^1 and U^2 on the complex vector spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ and a linear transformation $A : \mathcal{H}_1 \to \mathcal{H}_2$ with $U_g^1 A = A U_g^2$ for all $g \in G$ then either $A = 0$ or $\dim(\mathcal{H}_1) = \dim(\mathcal{H}_2)$, U^1 and U^2 are equivalent and $A = c\mathbb{1}$ for some $c \in \mathbb{C}$.

[8] In the example of a fixed magnetization, the space of two qubits $\mathbb{C}^2 \otimes \mathbb{C}^2$ splits into $\mathrm{span}(|m_z = -1\rangle) \oplus (\mathbb{C}^2 \otimes \mathrm{span}(|m_z = 0\rangle)) \oplus \mathrm{span}(|m_z = 1\rangle)$.

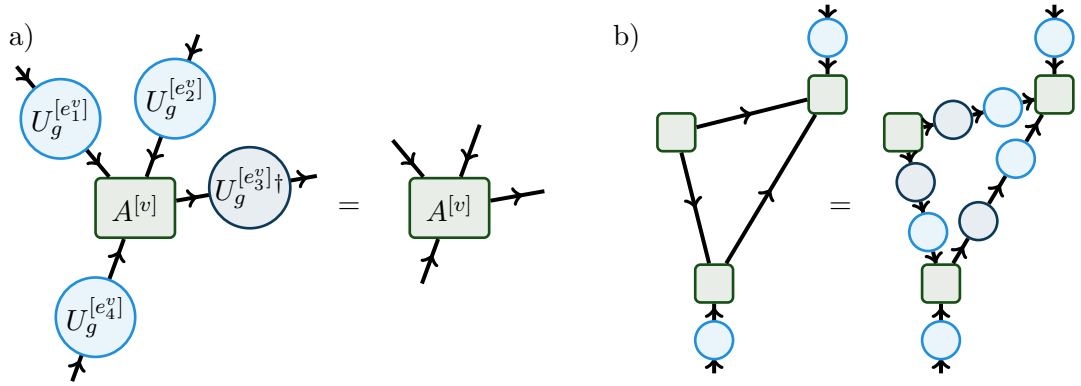

Figure 3.6. – a) Illustration of the conditions on the components $A^{[v]}$ of a tensor network for being a symmetric tensor. Given the component $A^{[v]}$ with ingoing edges or half edges e_1^v, e_2^v and e_4^v and an outgoing edge e_3^v, $A^{[v]}$ is symmetric if it is invariant under the joint multiplication with the unitary representations as shown above. Note that on ingoing edges the representation is chosen to act as $U^{[e_k^v]}$, while on outgoing edges with the adjoint $U^{[e_k^v]\dagger}$ (which are indicated in darker color). The condition generalizes to more or fewer ingoing and outgoing edges in the obvious way. b) Illustration of a tensor network with two physical indices and three components where we suppress all labels for better visibility. If we act with the symmetry operators $U_g^{[1]}$ and $U_g^{[2]}$ on the physical spaces, we can always insert the according representatives on the virtual spaces in the network due to the unitarity of the representations, where we use the same color coding as in a). If all components are now symmetric tensors as displayed in a), the resulting state will fulfill $U_g^{[1]} \otimes U_g^{[2]} |A^{[1]}, A^{[2]}, A^{[3]}\rangle = |A^{[1]}, A^{[2]}, A^{[3]}\rangle$, i.e., it is invariant with $\phi_g = 0$ for all $g \in G$ in Eq. (3.11). We can account for the additional phase factor in Eq. (3.11) by adding one additional index of dimension 1 to any of the components and let the symmetry group G act on it with the representation $e^{i\phi_g}$ – in an MPS decomposition we would typically add this index to the last site.

factor originating from the degeneracy spaces. $A^{[v]}$ shaped as the linear transformation from $\mathcal{H}_{\text{in}}^{[v]}$ to $\mathcal{H}_{\text{out}}^{[v]}$ as explained above then decomposes into a direct sum of independent contributions and we obtain the sparsity structure imposed on the level of the tensor $A^{[v]}$ by splitting the combined indices again. The structure appearing for general groups on the level of the tensor is then quite involved as general Clebsch-Gordan coefficients and multiplicities will appear in the product representation. It is worked out in full generality with details on the practical implementation in Ref. [84]. For Abelian groups on the other hand the resulting scheme is quite straightforward [86]. The reason for this is that all irreps are one-dimensional and tensor product of two irreps is again an irrep. We denote by $\lambda_1 \otimes \cdots \otimes \lambda_k$ the label of the unique irrep that results from building the tensor product of the irreps corresponding to all $\lambda_1, \ldots, \lambda_k$. All spaces decompose again as in Eq. (3.12) with $\dim(\mathcal{H}_\lambda) = 1$ and therefore $\sum_\lambda \dim(D_\lambda^{[e]}) = \dim(\mathcal{H}^{[e]})$. We then split the index corresponding to $\mathcal{H}^{[e]}$ into a tuple $(\lambda, \alpha_\lambda)$ where λ ranges over all irreps appearing in the decomposition of $\mathcal{H}^{[e]}$ and $\alpha_\lambda \in [\dim(D^{[e]})_\lambda]$. Given any $v \in \mathcal{V}$ and

abbreviating $|\text{in}(v)| = a$ and $|\text{out}(v)| = b$, the symmetric component $A^{[v]}$ then splits into

$$A^{[v]}_{(\lambda_1^i,\alpha_{\lambda_1^i}),\ldots,(\lambda_a^i,\alpha_{\lambda_a^i}),(\lambda_1^o,\alpha_{\lambda_1^o}),\ldots,(\lambda_b^o,\alpha_{\lambda_b^o})} = A^{[v,\lambda_1^i,\ldots,\lambda_a^i,\lambda_1^o,\ldots,\lambda_b^o]}_{\alpha_{\lambda_1^i},\ldots,\alpha_{\lambda_a^i},\alpha_{\lambda_1^o},\ldots,\alpha_{\lambda_b^i}} \delta_{\lambda_1^i\otimes\cdots\otimes\lambda_a^i,\lambda_1^o\otimes\cdots\otimes\lambda_b^o} \ , \quad (3.14)$$

where the entries of the tensors $A^{[v,\lambda_1^i,\ldots,\lambda_a^i,\lambda_1^o,\ldots,\lambda_b^o]}$ are up to the usual gauge freedom of tensor networks, independent of each other and we denote by λ^i and λ^o the irreps of the corresponding ingoing and outgoing spaces. The component $A^{[v]}$ splits therefore in a direct sum of smaller tensors $A^{[v,\lambda_1^i,\ldots,\lambda_a^i,\lambda_1^o,\ldots,\lambda_b^o]}$ which can be practically implemented by splitting all indices into tuples as in Eq. (3.14) and use a block-sparse tensor format for the direct implementation. All calculations can then be reduced to the smaller component tensor and all we need to specify of a given symmetry is how two representations are fused for edges that go into the same direction, i.e., we need to specify $\lambda_1 \otimes \lambda_2$, and for edges that go into opposite directions and how the local physical spaces decompose into different symmetry sectors.[9]

3.2. Density Matrix Renormalization Group and Related Algorithms

The DMRG algorithm was first developed independently of tensor network states. In its original formulation [87] the DMRG algorithm relied on an iterative update of the single or two-site reduced density matrix and truncation of the space it is supported on. The DMRG turned out to be able to approximate the ground state of local spin systems such as the Heisenberg model with an impressive accuracy. The connection to MPS was established in the following [88, 89] and lead to a deeper understanding of the method and also at least partially motivated the research on the correlation structure in one-dimensional systems summarized in Sec. 3.1.2. In its modern formulation [8], the DMRG algorithm is one of several tensor network algorithms used in computational physics. In essence, it constitutes a heuristic iterative eigensolver of a high-dimensional matrix acting on a tensor product space which approximates low-lying eigenvectors by an MPS. The second very popular approach of finding ground states in MPS manifolds, relies on the imaginary time evolution. This can be either done using the time evolving block decimation (TEBD) method by applying decompositions of the time evolution operator for small time steps [65] or by integrating the on the MPS manifold projected flow generated by the Schrödinger equation [90]. Recently, a connection between the imaginary time evolution based on TEBD and DMRG has been established [91]. This scheme allows us now to also perform a (real or imaginary) time evolution using a DMRG for all Hamiltonians the DMRG algorithm can be applied to (including non-local ones).

In this section we focus on the basic techniques needed in order to understand and implement a DMRG algorithm for long range Hamiltonians, with a specific focus on the

[9] Again for the fixed magnetization the physical space of a single qubit splits into span $(|m_z = -1/2\rangle)$ and span $(|m_z = 1/2\rangle)$ with labels $\lambda = 0, 1$ and the irrep labels of two edges going in the same direction fuse as $\lambda_1 + \lambda_2$ where for edges with opposite direction we have $\lambda_{\text{ingoing}} - \lambda_{\text{outgoing}}$.

Hamiltonian of interacting fermions in Eq.(2.58). The essential steps are described and the resulting computational costs and heuristic error measure discussed. In addition we present insights that have been gained concerning the convergence properties of DMRG. More specifically we will find that a badly initialized DMRG does not converge to the correct solution although we do not restrict the computational resources, and we comment on how to extend the error measures to individual excited states.

3.2.1. Ground State Search

Within the MPS picture [8], the s-site DMRG algorithm corresponds to an iterative update of s many component tensors on consecutive sites such that the energy expectation value of the resulting MPS is approximately minimal, given that the component tensors on all other sites are kept unchanged. Usually we set $s = 1, 2$.

3.2.1.1. Micro Iteration Step

Fix $s \in [V]$ and $k \in [V - s]$ and assume we are given an MPS with the components $A^{[j]}$ with $j \in [V]$ in a mixed normalized form with respect to site k. We introduce for $X \in \mathbb{C}^{d_k \cdots d_{k+s} \times r_{k-1} \times r_{k+s}}$ the abbreviation

$$|X\rangle_{k,s}^{(A^{[j]})_j} = |A^{[1]}, \ldots, A^{[k-1]}, X, A^{[k+s]}, \ldots, A^{[V]}\rangle. \tag{3.15}$$

We are then interested in the optimal tensor X_{opt} which minimizes the energy expectation value of $|X\rangle_{k,s}^{(A^{[j]})_j}$ while being normalized. Defining the Lagrangian

$$\mathcal{L}_{k,s}(X, \overline{X}) = \langle X|_{k,s}^{(A^{[j]})_j} H |X\rangle_{k,s}^{(A^{[j]})_j} + \lambda(\langle X|_{k,s}^{(A^{[j]})_j} |X\rangle_{k,s}^{(A^{[j]})_j} - 1) \tag{3.16}$$

the optimal tensor X_{opt} minimizes (3.16) under the given constraint. Arranging all entries of X in a vector and denoting it by $\mathbf{X}$, we can write (3.16) in the form $\mathbf{X}^\dagger H^{(k,s)} \mathbf{X} - \lambda(\mathbf{X}^\dagger \mathbf{X} - 1)$ where the matrix $H^{(j,k)}$ is defined in Fig. 3.7 a) and the scalar product in (3.16) simplifies due to the normalization as displayed in Fig. 3.7 b). Taking the derivative with respect to $\mathbf{X}^\dagger$ [92] we find that the extrema are given by the solutions of

$$H^{(k,s)} \mathbf{X} - \lambda \mathbf{X} = 0, \tag{3.17}$$

which is an ordinary eigenvalue problem of $H^{(k,s)}$. Solving for the normalized eigenvector $\mathbf{X}_{\min}$ corresponding to the smallest eigenvalue $\lambda_{\min}$, we find that $\mathbf{X}_{\text{opt}} = \mathbf{X}_{\min}$ with $\langle X_{\text{opt}}|_{k,s}^{(A^{[j]})_j} H |X_{\text{opt}}\rangle_{k,s}^{(A^{[j]})_j} = \lambda_{\min}$.

3.2.1.2. Single Site DMRG

For a given k in case of the single-site DMRG ($s = 1$) the resulting optimum $\mathbf{X}_{\text{opt}}$ can be reshaped into a tensor of order 3 (or 2 at the boundaries) again and we update the MPS by replacing $A^{[k]}$ by X. Bringing the resulting MPS in the mixed normalized form with respect to site $k + 1$ we identify and insert the optimal tensor for site $k + 1$ by the

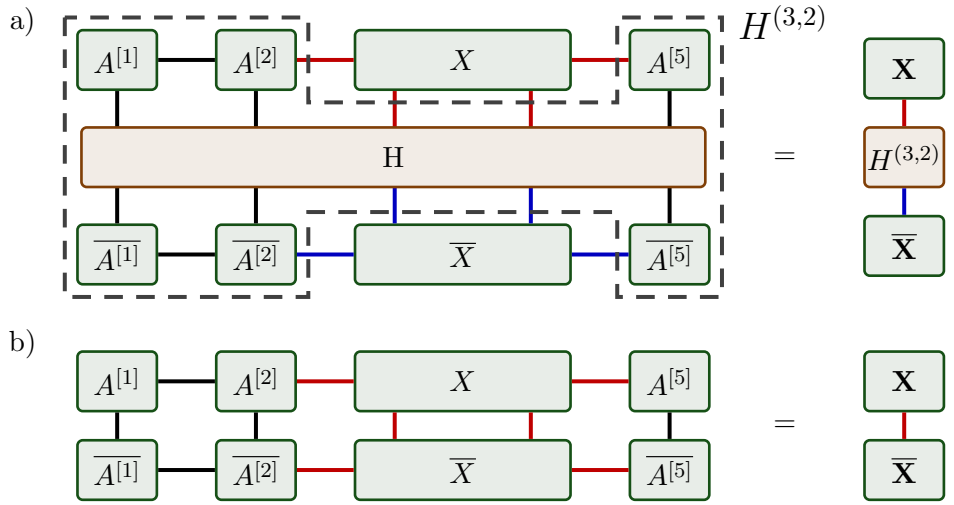

Figure 3.7. – a) Illustration of the energy expectation value $\langle X|_{k,s}^{(A^{[j]})_j} H |X\rangle_{k,s}^{(A^{[j]})_j}$ for $k = 3$ and $s = 2$. We define the matrix $H^{(3,2)}$ to be the part of the tensor network encircled by the dashed line, where all red and blue edges are combined to one index respectively. The energy expectation value can then be written as $\mathbf{X}^\dagger H^{(3,2)} \mathbf{X}$. b) Assuming that the MPS is given in a mixed normalized form with respect to site 3 (or 4), the norm of $|X\rangle_{k,s}^{(A^{[j]})_j}$ can be simplified as shown. Inserting the normalization conditions Fig. 3.4 on the left and right iteratively, yields that $\langle X|_{k,s}^{(A^{[j]})_j} |X\rangle_{k,s}^{(A^{[j]})_j} = \mathbf{X}^\dagger \mathbf{X}$ by combining all red edges to one index.

above scheme. This process is iterated over the full system. Once the right boundary of the system is reached, we continue by performing updates of the tensors on sites $V-1$, $V-2$, etc. We sweep over the system back and forth and perform local updates. We denote the cycle of starting from site 1 and reaching it again after $2(V-s)$ steps as a macro iteration. Using the single-site DMRG we will in each micro iteration step decrease the value of the energy expectation value while staying inside the manifold $\mathcal{M}((r_j)_{j\in[V]})$ where $(r_j)_{j\in[V]}$ are the bond dimensions of the initial MPS. The initial components can be chosen either at random, i.e. random tensors that might respect the symmetry constraints, or based on an initial intuition or precalculations.

3.2.1.3. Two Site DMRG: Projections and Discarded Weights

In case of the two-site DMRG, i.e., $s = 2$, an additional step is needed. The optimum $\mathbf{X}_{\text{opt}} \in \mathbb{C}^{d_k d_{k+1} r_{k-1} r_{k+1}}$ found along the lines explained in the previous section can not be directly inserted as an MPS component but needs to be decomposed first. Reshaping the tensor X_{opt} by combining the left virtual and physical index and the right virtual and physical index we obtain the matrix $\text{mat}(X_{\text{opt}})$ with entries $\text{mat}(X_{\text{opt}})_{(i_k,\alpha_{k-1}),(i_{k+1},\alpha_{k+1})} = \mathbf{X}_{\text{opt}(i_k,i_{k+1},\alpha_{k-1},\alpha_{k+1})}$. Note that the singular values of $\text{mat}(X_{\text{opt}})$ correspond to the Schmidt spectrum of $|X_{\text{opt}}\rangle_{k,2}^{(A^{[j]})_j}$ for the bipartition of the system in $[k]$ and $[k]^c$ as the MPS is in a mixed normalized form. However, typically

$\operatorname{rank}(\operatorname{mat}(X_{\text{opt}})) \geq r_k$ such that if we decompose $\operatorname{mat}(X_{\text{opt}})$ by an SVD and reshape the results into new components $A^{[k]}$ and $A^{[k+1]}$ we would locally increase the bond dimension. We can deal with this issue in three different ways. We keep the exact solution X_{opt} at the cost of an increasing bond dimension. We approximate the $\operatorname{mat}(X_{\text{opt}})$ up to a predefined error with a matrix of lower rank before the decomposition or we best-approximate the matrix $\operatorname{mat}(X_{\text{opt}})$ by a matrix of fixed predefined rank.

The first option is rather unpractical as $\operatorname{mat}(X_{\text{opt}})$ will be typically full rank due to small numerical errors and the bond dimension would grow exponentially during several macro iterations. The second and third option are similar in spirit and lead to the two different strategies which are actually used for a two-site DMRG. Due to the normalized form we obtain for an error induced on the states by changing the tensor X

$$\| |X_1\rangle_{k,2}^{(A^{[j]})_j} - |X_2\rangle_{k,2}^{(A^{[j]})_j} \|_2 = \|\operatorname{mat}(X_1) - \operatorname{mat}(X_2)\|_2. \tag{3.18}$$

It is therefore practical to either for a given error ϵ find a the lowest rank possible which allows for a low rank approximation of $\operatorname{mat}(X_{\text{opt}})$ with an error not larger than ϵ by solving

$$\text{find minimal } r \in \mathbb{N} : \min_{\substack{X \in W: \\ \operatorname{rank}(\operatorname{mat}(X)) \leq r}} \|\operatorname{mat}(X_{\text{opt}}) - \operatorname{mat}(X)\|_2 \leq \epsilon \tag{3.19}$$

and denote the minimizing X by $X_{\text{opt}}^{\text{approx}}$ or for a predefined maximal bond dimension r minimize the error directly according to

$$X_{\text{opt}}^{\text{approx}} = \operatorname*{arg\,min}_{\substack{X \in W: \\ \operatorname{rank}(\operatorname{mat}(X)) \leq r}} \|\operatorname{mat}(X_{\text{opt}}) - \operatorname{mat}(X)\|_2, \tag{3.20}$$

where $W = \mathbb{C}^{d_k \times d_{k+1} \times r_{k-1} \times r_{k+1}}$. Both solutions can be found easily using the SVD. Given a matrix $A \in \mathbb{C}^{n \times m}$ with SVD $U\Sigma(A)V = A$ and denote by $\Sigma_r(A)$ the matrix in which the smallest $\operatorname{rank}(A) - r$ singular values are set to zero then

$$U\Sigma_r(A)V = \operatorname*{arg\,min}_{\substack{X \in \mathbb{C}^{n \times m}: \\ \operatorname{rank}(X) = r}} \|A - X\|_2 \tag{3.21}$$

and the error is given by the discarded singular values according to $\|A - U\Sigma_r(A)V\|_2^2 = \sum_{l=r+1}^{\operatorname{rank}(A)} \sigma_l^{\downarrow}(A)^2$, where $\sigma^{\downarrow}(A)$ denotes the sequence of decreasingly ordered singular values of A [93]. Based on this result, we can solve the above minimization problems either by discarding singular values until the error reaches ϵ or by simply discarding the smallest $\operatorname{rank}(\operatorname{mat}(X_{\text{opt}})) - r$ singular values. The discarded singular values constitute the discarded weight

$$\epsilon_k = \sum_{l=r+1}^{\operatorname{rank}(\operatorname{mat}(X)_{\text{opt}})} \sigma_l^{\downarrow}(\operatorname{mat}(X)_{\text{opt}})^2, \tag{3.22}$$

which allows for a heuristic error estimation [94, 95] as it captures the truncation error induced by projecting the solution of the micro step $|X_{\text{opt}}\rangle_{k,2}^{(A^{[j]})_j}$ to an MPS manifold with lower bond dimensions.

Algorithm 2 – Basic components of a DMRG micro step. First, given the current MPS the locally optimal update is calculated. Secondly, a found update is projected to a lower rank depending on the direction of the sweep, where we use the convention +1 for a right sweep and −1 for a left sweep. U_r, Σ_r and V_r are the restrictions of U, Σ and V to r rows and/or columns.

procedure DMRG STEP($(A^{[j]})_{j\in[V]}, k$)
  bring $|(A^{[j]})_{j\in[V]}\rangle$ in mixed normalized form with respect to site k
  prepare $H^{(k,2)}$ (see App. B.1)
  $\mathbf{X}_{\text{opt}} \leftarrow$ lowest eigenvector($H^{(k,2)}$)
  return X_{opt}
procedure PROJECT(X, direction)
  $U, \Sigma, V \leftarrow \text{SVD}(\text{mat}(X))$
  fix r by restricting ϵ_k or set to $\min(r_{\max}, \text{rank}(\text{mat}(X)))$
  if direction $= -1$ **then**
    return $U_r\Sigma_r$, V_r
  else
    return U_r, $\Sigma_r V_r$

This notion of being a projection can be made more precise. We can view the rank r solution $\text{mat}(X_{\text{opt}}^{\text{approx}}) = U\Sigma_r(\text{mat}(X_{\text{opt}}))V$, where $U\Sigma(\text{mat}(\text{opt}))V = \text{mat}(\text{opt})$, as the result of a projection as

$$\text{mat}(X_{\text{opt}}^{\text{approx}}) = U_r {U_r}^\dagger \text{mat}(X_{\text{opt}}), \tag{3.23}$$

with U_r being the matrix containing the columns of U associated to the largest r singular values in $\Sigma(X_{\text{opt}})$. $U_r {U_r}^\dagger$ is then an orthogonal projection connecting $\text{mat}(X_{\text{opt}}^{\text{approx}})$ with $\text{mat}(X_{\text{opt}})$.

After this projection, we step to the next site $k \pm 1$ and continue with sweeps as in the case of the single-site DMRG – Alg. 2 summarizes the essential steps of the micro step of a two-site DMRG.

3.2.1.4. Application to Quantum Chemistry: Efficiency and Costs

The DMRG performs local updates of the components and its structure allows for a very efficient approximation of the ground states of local one-dimensional systems. For local Hamiltonians, the micro iteration step can be performed in a time that is independent of the system size if we use disk memory of the order of $\mathcal{O}(Vr^2)$ with $r = \max(\{r_j | j \in [V-1]\})$. We are therefore in the very comfortable situation of having an efficient method in order to approximate ground states of local Hamiltonians within an ansatz class which provably allows for efficient approximation of the desired states.

This picture changes for non-local Hamiltonians. As explained above we have no rigorous bound on the needed bond dimensions. Secondly, the costs of performing a DMRG calculation increases and will scale with higher powers of the system size.

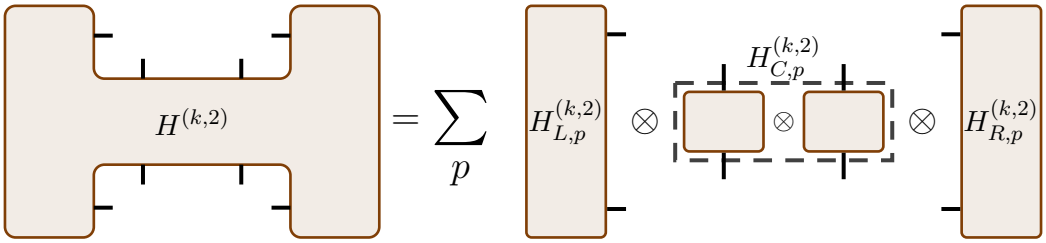

Figure 3.8. – Illustration of the decomposition needed of the matrix $H^{(k,2)}$ introduced in Fig. 3.7 in order to implement an practical two-site DMRG. The matrix $H^{(k,2)}$ is split into triples $(H^{(k,2)}_{L,p}, H^{(k,2)}_{C,p}, H^{(k,2)}_{R,p})$ which allow to bookkeep for instance all terms that act on the central and right part of the system always with the same operators and summarize their action on the left part as explained in App. B.1 for interacting electrons. The scheme presented in App. B.1 can be directly generalized to other Hamiltonians given by coefficient tensors and low polynomials of commuting or anticommuting on-site operators.

The most time consuming part in a micro step is the calculation of $\mathbf{X}_{\text{opt}}$. Using iterative eigenvalue solvers such as the Lanczos or the Davidson method we can calculate the eigenvector corresponding to the lowest eigenvalue of a matrix A solely by having an oracle that can calculate $A\mathbf{X}$ for any vector $\mathbf{X}$.[10] The number of times this oracle is called depends only weakly on the size of the matrix if a reasonable initial guess is provided.[11] In order keep the matrix vector product efficient we factor the partially contracted Hamiltonian $H^{(k,s)}$ as shown in Fig. 3.8 [96], as we can then calculate $H^{(k,s)}\mathbf{X}$ in a time $\mathcal{O}(Pd^2r^2(d+r))$ where P denotes the number of terms in the decomposition and r and d are the maximal bond and physical dimension occurring as usual. In App. B.1 we explain how to obtain a decomposition of $H^{(k,2)}$ of an interacting fermionic Hamiltonian as in Eq. (2.58) on M modes that contains $\mathcal{O}(M^2)$ terms where usually $M \propto V$. We obtain this decomposition in a time $\mathcal{O}(M^2(dr^3 + d^2r^2) + M^3r^2)$, where $r = \max(\{r_j | j \in [V-1]\})$ and $d = \max(\{d | j \in [V]\})$. The evaluation of the matrix vector product $H^{(k,2)}\mathbf{X}$ can then be performed in a time scaling as $\mathcal{O}(M^2(d^2r^3 + d^3r^2))$.

In the course of the research documented in this thesis a substantial code structure has been developed. We present the developed code and discuss included features in App. C. If not referenced otherwise, results presented in the following have been obtained with this code.

3.2.2. Convergence of DMRG

The result for local minima on low rank manifolds for the DMRG algorithm originates from several discussions with Max Pfeffer. The convergence behavior of the DMRG algorithm is mostly only understood from practical investigation of specific examples

[10] Note that storing the full matrix $H^{(k,s)}$ in memory or fully diagonalizing it is only possible for small bond dimensions which would strongly limit the applicability of DMRG.

[11] After the first sweep, the current component, or product of components for a two-site DMRG, usually provides a good initial guess.

(e.g. [94, 95]) and a rigorous convergence analysis of the scheme is still lacking. If applied to natural examples, the truncated weight of a converged DMRG calculation is found to be related to the relative energy error by a polynomial dependence [94, 95], which indicates a generic steady convergence upon increasing the available resources, i.e., bond dimension. The rigorous understanding and formal convergence behavior is on the other hand only poorly understood. Existing results are rather restrictive. The analysis of the easier problem of finding the best MPS approximation to a given state using a single-site DMRG like update scheme [97] provides a first step but needs rather restrictive assumptions and is only valid on fixed rank manifolds. Algorithms which yield a guaranteed best approximation of the ground state of local gapped one-dimensional systems [78, 79] are less efficient than the DMRG and their results do not easily carry over to more general settings such as non-local Hamiltonians. One of the main obstacles for a deeper theoretical investigation of the DMRG algorithm is the structure of the set of states. In the vicinity of rank deficient MPS, i.e., states which can be well approximated using an MPS with smaller bond dimension, the curvature of the MPS fixed rank manifold diverges. Furthermore, rank adaptive schemes such as the two-site DMRG do not operate on simple manifolds but algebraic varieties. In order to overcome these complications, recently a desingularization method for matrix varieties has been designed which operators on a larger flattened manifold structure [98]. In addition the geometric structure of tensor varieties has been more rigorously investigated leading to a description of the tangent cones [99]. Unifying these current trends in one combined picture has the potential of achieving a more rigorous and deeper understanding of DMRG and tensor network methods as a whole in the near future and is subject of current research.

Concerning the general convergence of the DMRG algorithm we here want to comment shortly on the best case scenario and want to investigate a rather simple question. Given unlimited resources, i.e., allowing for an arbitrary growth of the bond dimensions and runtime, does the two-site DMRG algorithm converge to the global optimum, given that the initial state is not perfectly orthogonal to the ground state? The answer to this question is, surprisingly, no. The local update structure of the DMRG induces local minima from which the algorithm can not escape. The simplest example for this can already be given on 3 sites with physical dimension $d = 2$.[12] Consider the non-orthogonal target and initial state

$$|\psi_\mathrm{t}\rangle = \frac{1}{\sqrt{5}}(|0\rangle \otimes |0\rangle \otimes |0\rangle + 2|1\rangle \otimes |1\rangle \otimes |1\rangle), \qquad |\psi_\mathrm{i}\rangle = |0\rangle \otimes |0\rangle \otimes |0\rangle, \tag{3.24}$$

which are for instance realized by a Hamiltonian $H = -|\psi_\mathrm{t}\rangle\langle\psi_\mathrm{t}|$ and a ground state search calculation initialized with $|\psi_\mathrm{i}\rangle$. Note that $|\psi_\mathrm{t}\rangle$ and $|\psi_\mathrm{i}\rangle$ can be represented by MPS with bond dimension 2 and 1 respectively. If we however now use a single- or two-site DMRG in order to approximate the ground state of H starting from $|\psi_\mathrm{i}\rangle$ one can easily verify that $|\psi_\mathrm{i}\rangle$ is a local minimum of the DMRG. The calculation will not move away from $|\psi_\mathrm{i}\rangle$ and the algorithm is stuck at $|\psi_\mathrm{i}\rangle$ for the Hamiltonian above. Even

[12] Thanks to Benjamin Kutschan for proposing the following refined example.

worse, $|\psi_\mathrm{i}\rangle$ is not even the best bond dimension 1 approximation to $|\psi_\mathrm{t}\rangle$ which is clearly given by $|1\rangle \otimes |1\rangle \otimes |1\rangle$. In order to leave the above local minimum, we would need to update all three sites simultaneously and allow for a growth of both bond dimensions, r_1 and r_2, in one single step. From this we conclude that due to the local update structure of the DMRG additional local minima on low rank manifolds are induced which are not even the optimum on the corresponding low rank manifold. Note that this is, as usual, different to the matrix case, i.e., a problem on 2 sites. If we find convergence of the two-site DMRG on a manifold of matrices with a rank lower than the maximal admissible bond dimension, then true optimality follows and the minimum is not induced by the optimization scheme itself [100]. This result is quite intuitive as in this case a two-site scheme implements a global optimization step [100].

The example above can obviously be generalized to more complex settings and corresponds to an early developed intuition that if certain information is missing in the initial or current state, then the DMRG can get stuck in local minima. Therefore noise is often added to the solution of the micro step in order to be able to escape such minima [96]. It is however unclear if this can be done consistently. The question to be answered would then be: Are local minima in the spirit of the one constructed above a null set on the corresponding fix rank manifolds and are these all of the local minima? These questions are currently open and need to be answered in view of obtaining a full fletched convergence analysis of any local update scheme for tensor network states such as the DMRG.

3.2.3. Excited States

Using DMRG it is possible to not only approximate the ground state of a system but also a few of the lowest excited states. The resulting family of states $|\psi_k\rangle$ are approximated by a family of MPS which differ in only one component tensor, e.g., $|\psi_k\rangle = |A^{[1,k]}, A^{[2]}, \dots, A^{[V]}\rangle$, where we typically combine the different $A^{[j,k]}$ into a tensor of order 4 (order 3 at the boundaries) and by this add to one component an additional index which labels the different states. The additional index is always placed on one of the components we update in a current micro iteration step and the additional index is shifted along the chain using SVDs as illustrated in Fig. 3.9 during a sweep. Assume that we want to calculate the m lowest energy states. The micro iteration step then changes from calculating only the lowest eigenvector to calculating the lowest m eigenvectors $\mathbf{X}^{(m)}$ of $H^{(k,s)}$. The resulting tensor is decomposed as shown in Fig. 3.9 and small or excessive singular values are truncated depending on the scheme used as described in Sec. 3.2.1.3.[13] The truncated weights obtained are however only a joint measure of accuracy for all states but contains no information about the projection error of each of the targeted states.

The following result on how to access the projection error of each individual excited states was obtained together with Max Pfeffer.

[13] Note that now also in the case of a single-site DMRG a truncation scheme needs to be employed as the additional index implies an increase of the bond dimensions during the sweeps otherwise.

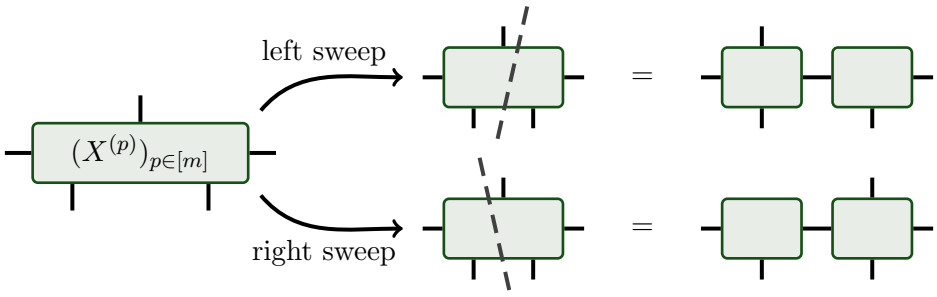

Figure 3.9. – During the calculation of m low energy states with a DMRG we obtain in a micro step the joint tensor $(X^{(p)})_{p\in[m]}$ of the lowest eigenvector of $H^{(k,s)}$. We consider the case of a two-site DMRG here. Depending on the direction of the sweep, the additional index labeling the different states is pushed into the left or right new MPS component by performing an SVD for a matrixfication with respect to the left and right of the dashed lines.

It is possible to access the individual projection error for each state. As explained in Sec. 3.2.1.3, if the projected solution $\mathrm{mat}(X_{\mathrm{opt}}^{\mathrm{approx}}) \in \mathbb{C}^{d_k r_{k-1} \times d_{k+1} r_{k+1}}$ has a rank r we find a projection $U_r U_r{}^\dagger$ with $U_r \in St(r, d_k r_{k-1})$, where $St(p,n)$ denotes the Stiefel manifolds (see also App. D), such that $\mathrm{mat}(X_{\mathrm{opt}}^{\mathrm{approx}}) = U_r U_r{}^\dagger \mathrm{mat}(X_{\mathrm{opt}})$. In addition we find that U_r is the minimizer of

$$\epsilon_k = \min_{U \in St(r, d_k r_{k-1})} \|\mathrm{mat}(X_{\mathrm{opt}}) - UU^\dagger \mathrm{mat}(X_{\mathrm{opt}})\|_2^2, \tag{3.25}$$

with ϵ_k being the truncated weight.

Denoting the matrixfications discussed in Fig. 3.9 of the joint tensor $(X^{(p)})_{p\in[m]}$ by $\mathrm{mat}((X^{(p)})_{p\in[m]})$ we obtain the global truncated weight ϵ_k as

$$\epsilon_k = \min_{U \in St(r,h)} \|\mathrm{mat}((X_{\mathrm{opt}}^{(p)})_{p\in[m]}) - UU^\dagger \mathrm{mat}((X_{\mathrm{opt}}^{(p)})_{p\in[m]})\|_2^2, \tag{3.26}$$

with $h = d_k r_{k-1}$ or $h = d_k r_{k-1} m$ depending on the direction of the sweep. By the structure of the cost function it is easy to show that

$$\epsilon_k = \min_{U \in St(r,h)} \sum_{p=1}^{m} \|\mathrm{mat}(X_{\mathrm{opt}}^{(p)}) - UU^\dagger \mathrm{mat}(X_{\mathrm{opt}}^{(p)})\|_2^2, \tag{3.27}$$

meaning that the global truncation error is the sum of all individual projection errors $\epsilon_k^{(p)}$. Denoting the minimizer of Eq. (3.27) by U_r we can directly calculate the individual truncation errors by

$$\epsilon_k^{(p)} = \|\mathrm{mat}(X_{\mathrm{opt}}^{(p)}) - U_r U_r{}^\dagger \mathrm{mat}(X_{\mathrm{opt}}^{(p)})\|_2^2. \tag{3.28}$$

It is easy to show along the lines of the argumentation to Eq. (3.18), that $\epsilon_k^{(p)}$ corresponds to the error of the individual excited states when approximating $\mathrm{mat}((X_{\mathrm{opt}}^{(p)})_{p\in[m]})$ with

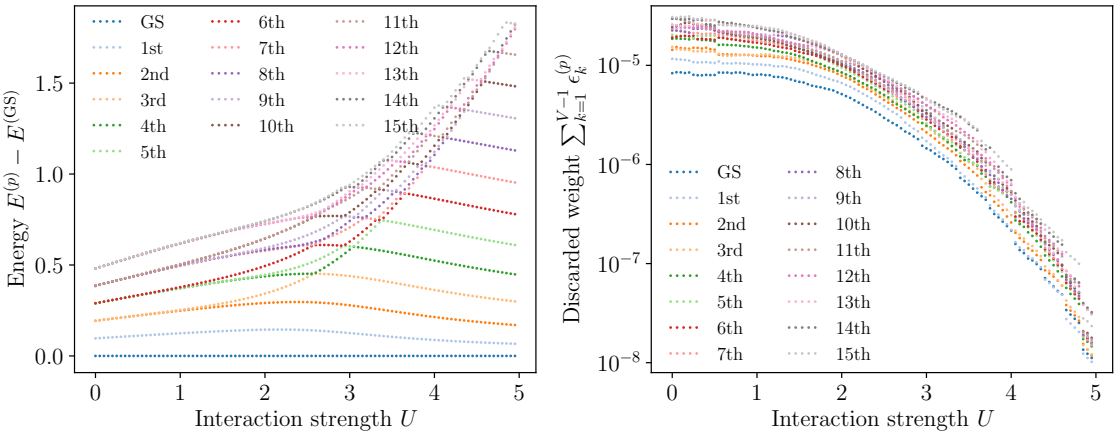

Figure 3.10. – Investigation of the 16 lowest lying eigenenergies of the one-dimensional spinless Fermi-Hubbard model defined in Eq. (2.61) with open boundary conditions on 64 sites for varying interaction strength U at half-filling. All eigenstates are approximated during one DMRG run with maximal bond dimension $r_{\max} = 256$. On the left we show the in sum minimal energy configuration reached during 10 sweeps for different interaction strengths where the hopping amplitude t was set to 1. We subtract the approximated ground state energy $E^{(GS)}$ at each U for a clearer presentation. On the right we show the projection error of each excited state as defined in Eq. (3.28) summed up over the full sweep in which the minimal energy is encountered.

the lower rank optimum $U_r U_r{}^\dagger \mathrm{mat}((X^{(p)}_{\mathrm{opt}})_{p\in[m]})$. In practice U_r is obtained from the SVD explained in Fig. 3.9 and truncation schemes explained in Sec. 3.2.1.3.

The knowledge of $\epsilon_k^{(p)}$ allows then for a heuristic convergence analysis of each excited state as presented in Fig. 3.10 and to address questions around the way individual excited states are approximated within a DMRG as shown in Fig. 3.11 and discussed below. In Fig. 3.10 we show that for a local Hamiltonian we are able to approximate low lying excited states accurately already at rather low bond dimensions. If we would naively split the maximal bond dimension of 256 between all 16 states shown in Fig. 3.10, we would obtain an effective bond dimension of 16 per state. However, due to the locality of the Hamiltonian, we can expect that the ground state can be well approximated by an MPS and that low lying excited states are close to elements of the tangent space of the ground state [68, 101]. Hence ground and excited states are expected to be closely related and are anticipated and found to be well approximable by a family of MPS of the form $|A^{[1,k]}, A^{[2]}, \ldots, A^{[V]}\rangle$ with common $A^{[2]}, \ldots, A^{[V]}$ due the shared features. In the case considered in Fig. 3.10 we find that the approximability of the low energy sector varies strongly with the interaction strength of the Fermi-Hubbard model. The projection errors encountered decrease by orders of magnitude with increasing U which is to be expected as the insulating phase should allow for an efficient real space representation of the eigenstates. Note for instance that once an eigenstate state of the low U limit leaves the range of the tracked 16 states with increasing U we see clear drops of the discarded weights as the new states are expected to be more localized in real space. Furthermore,

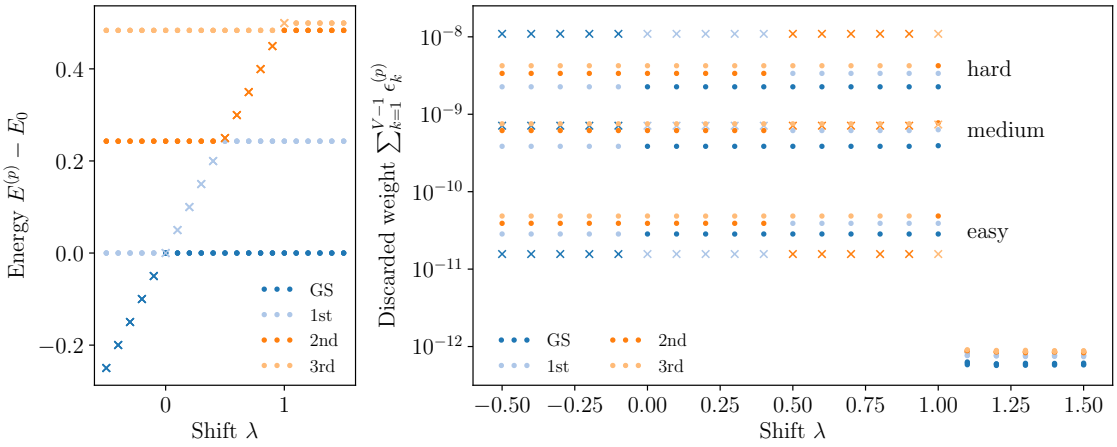

Figure 3.11. – Investigation of the 4 lowest lying eigenenergies of the one-dimensional spinless Fermi-Hubbard model with open boundary conditions on 32 sites with an additional state shifted into the spectrum as explained in Eq. (3.29). The interaction strength U and hopping amplitude t of the Fermi-Hubbard model are set to 1. We choose three different states $|\psi_s\rangle$ which are easily, moderately and with some difficulty approximable by an MPS as discussed in Fig. 3.12. The eigenenergy E of the shifted state $|\psi_s\rangle$ is chosen to be $E = \lambda(E_3 - E_0) + E_0$ with E_0 and E_3 being the ground state energy and energy of the 3rd excited state of the unperturbed Hamiltonian H. In the left panel we show the eigenenergies of the four lowest eigenstates at half filling in dependence of the shift λ. For clarity we mark the energy of the shifted state by a cross and the spectrum is the same for all three shifted states. In the right panel we show the projection error of each excited state (again summed over the full sweep in which the minimal energy is found) for different shifts λ. In order to allow for a better comparison we plot the results for all three different shifted states $|\psi_s\rangle$ in one panel. For $\lambda \leq 1$, i.e., $E \leq E_3$ we find that the three different states lead to different projection errors which accumulate around $5 \cdot 10^{-10}$ in the "easy" case, around 10^{-9} in the "medium" and around $5 \cdot 10^{-8}$ in the "hard" case. For $\lambda > 1$ all three settings lead to the same projection error. We indicate again the projection error associated to the shifted state by a cross.

we find that all states are almost equally well approximated and the projection errors are typically of the same order of magnitude. However, the lowest eigenstates are approximated the best and we find a slightly but systematically increasing error with increasing energy. This could have two causes. Firstly, the corresponding eigenstates could simply be less efficiently approximable by MPS and hence lead to larger projection errors in a fixed bond dimension calculation. Secondly, the DMRG could favor low lying states. Such a bias could for instance result from an internal hierarchy of orthogonality relations, e.g., the constraint that higher excited states need to be approximately orthonormal to the states with lower energy which is a limiting factor for naive schemes that determine the eigenvectors in separate calculations.

In order to address this question further we perform the test shown in Fig. 3.11. We alter the Hamiltonian of the system by replacing H with

$$H' = H + \left(E + \langle\psi_s|H|\psi_s\rangle\right)|\psi_s\rangle\langle\psi_s| - H|\psi_s\rangle\langle\psi_s| - |\psi_s\rangle\langle\psi_s|H \tag{3.29}$$

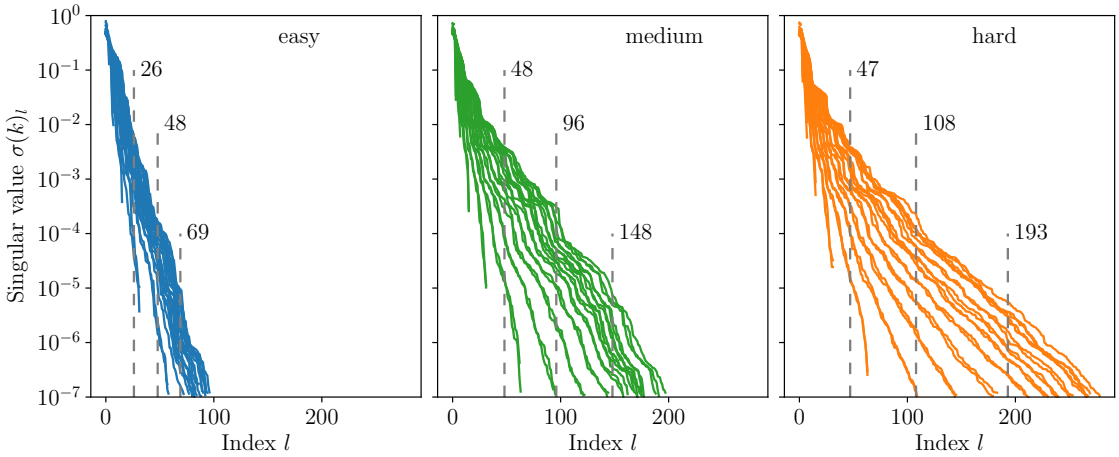

Figure 3.12. – Sorted Schmidt spectra $\sigma(k)$ of the three states used in Fig. 3.11 for each bipartition of the chain into $[k]$ and $[k]^c$ with $k = 1, \ldots, 31$. The dashed vertical lines at indices r indicate the number of Schmidt values needed in order to keep the weight of the discarded Schmidt values $\sum_k \sum_{l=r+1}^{\infty} \sigma(k)_l^2$ below 10^{-3}, 10^{-6} and 10^{-9} respectively (where the left line corresponds to the value r needed for the threshold 10^{-3}, the central line to 10^{-6} and the right one to 10^{-9} – the numbers indicate the corresponding index r). We find that the shifted state $|\psi_s\rangle$ in the "easy" case on the left is expected to be well approximable by an MPS as the Schmidt spectrum decays strongly. The "medium" and "hard" instance in the center and on the right are increasingly difficult to approximate as the Schmidt spectrum is still chosen to decay exponentially but falls of more slowly.

for a given state $|\psi_s\rangle$. It is easy to check that $|\psi_s\rangle$ is an eigenvector of H' with eigenvalue E. Furthermore, if we choose $|\psi_s\rangle$ orthonormal to the low lying eigenstates if H, we find that H' has the same low energy sector except that $|\psi_s\rangle$ is shifted into the spectrum of H' at a position controlled by the parameter E. In Fig. 3.11 we show the results for the eigenenergies and projection errors for three different states $|\psi_s\rangle$ which are easily, moderately or with some difficulty approximable by an MPS as discussed in Fig. 3.12.[14] From the projection errors shown in Fig. 3.11 we can conclude two general aspects of the DMRG.

Firstly, we find that if one target state is difficult to approximate, the errors of all approximated states increases. Furthermore does the error of an individual state depend on the difficulty to approximate it as for instance the projection error of $|\psi_s\rangle$ increases relative to the error of the other states with increasing difficulty.

Secondly, we see that the projection error of a state is independent of the position of the state in the spectrum. For an increasing shift λ in Fig. 3.11 we find that $|\psi_s\rangle$ takes the position of the ground state as well as first, second and third excited state. However, the projection errors are (up to the drop at $\lambda = 1$) independent of the shift λ and by

[14] The different states were constructed as the sum of multiple MPS with exponentially decaying prefactors. By increasing the bond dimension of the random summands, we obtain a slower decay of the Schmidt spectra of the corresponding resulting state as displayed in Fig. 3.12. Each resulting state has an overlap of at most 10^{-12} with a low lying eigenstate of H.

this the position of the shifted state $|\psi_\mathrm{s}\rangle$ within the spectrum.

From this we conclude, that the DMRG approximates multiple eigenstates of an Hamiltonian truly jointly within one run and does not possess an internal orthonormality hierarchy. Furthermore, it distributes its available resources over the different states in order to allow for an approximately equal approximation.

3.3. Mode Transformation in Tensor Network States on Fermionic Systems

The method presented in this section and results displayed in Fig. 3.16 originate from joint work with Libor Veis, Örs Legeza and Jens Eisert which is published in [102].

For non-local fermionic systems originating for instance from quantum chemistry, we have no promise that ground or low-lying excited states are well approximable by MPS. However, taking a step back, DMRG is simply an eigensolver which allows us to approximate the few lowest eigenvalues and eigenvectors and systematically increase the accuracy. It is therefore natural to test DMRG outside of its comfort zone – with remarkable success (see for instance the reviews [54–57]). In the context of electronic structure calculation from quantum chemistry DMRG turns out to be applicable to strongly correlated problems for which other traditional methods fail to provide exact results [54–57]. The same applies for systems in higher-dimensional lattices or problems originating from nuclear physics [103, 104]. On the other hand, the resources needed to run a DMRG calculation increase quickly with the system size, such that only small systems can be considered. Current state of the art implementations typically are used for systems of ≈ 100 spin orbitals and ≈ 50 electrons in them and are quickly outperformed by CC, DFT or other methods for weakly correlated systems.

If a fermionic problem lacks locality, new challenges arise in the context of TNS based methods. TNS and in particular MPS are formulated with respect to a fixed decomposition of the global Hilbert space into local spaces which translates by the Jordan-Wigner transformation to a specific single-particle basis of the fermions. However it is unclear in advance, which single-particle basis we should pick.[15] If a fixed single particle basis is chosen we cannot efficiently represent many rotated Slater determinants using MPS. In the extreme case, a poorly chosen initial single-particle basis could result in an approximation of the ground state that is less accurate than a simple Hartree-Fock approximation despite a large bond dimension. This issue, that seemingly simple states (individual Slater determinants), may not be captured efficiently results from the fact that TNS are designed to efficiently represent states with limited correlations between individual modes but may struggle with correlations which are simple within the par-

[15] That the choice of the single-particle basis is important becomes clear when we are asked to find the best MPS of bond dimension 1 with a fixed particle number that approximates the ground state. The resulting state will be a Slater determinant in the chosen basis and we know that the resulting minimal energy found depends strongly on the single-particle basis (compare to the Hartree-Fock methods discussed in Sec. 2.5.1 whose objective it is to identify the optimal single-particle basis for a Slater determinant approximation of the ground state).

ticle picture of fermionic systems. Here we show that we can overcome this limitation of TNS and lift the ambiguity of the initial single particle basis. By combining MPS with mode transformations we develop a tensor network based scheme that allows to go a step beyond the restrictive area laws of correlations in TNS and at the same time enables to represent a large portion of the correlations of fermionic systems in the correct way. In doing so we in addition are able to identify single-particle basis sets in which correlations are stronger localized which has the potential to give us further insights into the correlation structure of molecules and bondings in future work. When TNS and mode transformation are combined they complement and compensate weaknesses of each other. Mode transformations can generate high entanglement between different modes in a state (which is limited if only the TNS is considered) and the underlying TNS is able to capture very general non-Gaussian correlation which can not be captured by the Hartree-Fock method, i.e., a mode transformation alone.

Take a fermionic system with M modes and V sites, where we assume for the simplicity of notation that each site supports p modes, i.e., $M = Vp$. Given an interacting fermionic Hamiltonian H as in Eq. (2.58) and a sequence of maximal bond dimensions $(r_j)_{j\in[V]}$ we aim for solving

$$\left(|(A^{[j]}_{\min})_{j\in[V]}\rangle, U_{\min}\right) = \underset{\substack{|(A^{[j]})_j\rangle\in\mathcal{M}((r_j)_j)\\ U\in U(M)}}{\arg\min} \langle (A^{[j]})_{j\in[V]}|G(U)^\dagger H G(U)|(A^{[j]})_{j\in[V]}\rangle, \quad (3.30)$$

where the $G(U)$ is the transformation induced on the Fock space by the mode transformation U introduced in Sec. 2.2.2.

There are already different schemes in use which optimize the single-particle basis intertwined with a DMRG. The most common ansatz is to use a specific starting basis that based on some intuition should allow for an efficient MPS approximations of the ground state. Natural candidates which are in use are localized or split-localized orbitals, where orbitals that are occupied and unoccupied in the Hartree-Fock solution are jointly or separately localized, plain Hartree-Fock orbitals or different types of natural orbitals [105]. We then perform simple DMRG runs for one or multiples of these bases and continue to work with the best result. In a second approach one starts from a fixed basis and tries to optimize the ordering of the sites such that strongly correlated sites are placed next to each other [106, 107]. For this, one calculates for the final state $|\psi\rangle$ of a DMRG run the mutual information

$$I(|\psi\rangle)_{j,k} = S_{\mathrm{vN}}(\mathrm{tr}_{\{j,k\}^c}|\psi\rangle\langle\psi|) - S_{\mathrm{vN}}(\mathrm{tr}_{\{j\}^c}|\psi\rangle\langle\psi|) - S_{\mathrm{vN}}(\mathrm{tr}_{\{k\}^c}|\psi\rangle\langle\psi|) \quad (3.31)$$

between all sites, which measures how correlated two sites are by essentially testing how close their joint reduced state comes to a product state. Subsequently, an ordering is identified which minimizes cost functions such as $\sum_{j,k} I_{j,k}|j-k|^2$, i.e., minimizing the distance of strongly correlated sites, and an additional DMRG run is performed in the reordered basis. Both schemes need multiple DMRG runs in order to optimize a given basis without changing the computational costs of the DMRG, but can already yield significant improvements if applied to realistic systems [108, 109]. A third scheme

optimizes the basis slightly more intertwined with the DMRG but increases its computational costs. Here the one- and two-body correlation matrices $\gamma^{(1)}$ and $\gamma^{(2)}$ of the current state are formed such that the energy can be computed according to Eq. (2.59) which gives rise to the DMRG self-consistent field (SCF) method in more general settings [110]. Similar to the Hartree-Fock method, one solves the problem

$$U_{\min} = \underset{U \in U(M)}{\arg\min} \operatorname{tr}(U^\dagger t U \gamma^{(1)}) + \operatorname{tr}(U^\dagger \otimes U^\dagger \operatorname{mat}(v) U \otimes U \gamma^{(2)}), \tag{3.32}$$

using for instance a conjugate gradient scheme on the manifold $U(M)$ [111]. Updating the couplings of the Hamiltonian will globally adapt the basis to the partially converged state and can be done multiple times during a DMRG run. However computing the reduced density matrices, rotating the orbitals and updating the presummed operators induces a computational overhead which adds to the normal scaling of the DMRG a term of order $\mathcal{O}(M^5)$ [110]. All these approaches are prone to get stuck in local minima for the basis optimization – which is not very surprising as already finding the optimal Hartree-Fock basis is NP complete (compare Sec. 2.5.2). In addition their implementation is rather complementary to the DMRG and they either increase the computational costs or consider only very restrictive sets of unitaries, e.g. permutations or choosing the best basis from a small set of guesses. We present a method that optimizes within a DMRG run both, the MPS and the physical basis iteratively without increasing the computational costs. In addition the scheme can be combined with the ones above to a joint method that in practice gets less trapped in local minima.

In order to naturally combine mode transformations and DMRG, the problem of finding a global mode transformation needs to be split up into local sub-problems. We here focus on the two-site DMRG. Given a site $k \in [V-1]$ for which we are updating the components $A^{[k]}$ and $A^{[k+1]}$ in the current micro step of the DMRG, we will solve for $X \in \mathbb{C}^{d_k \times d_{k+1} \times r_{k-1} \times r_{k+1}}$ local problems of the form

$$U_{\text{opt}}^{(k,2)} = \underset{U \in U(2p)}{\arg\min} f^{(k)}(X(U)), \tag{3.33}$$

where $f^{(k)}$ is a cost function discussed below and

$$X(U)_{i_k, i_{k+1}, \alpha_{k-1}, \alpha_{k+1}} = \sum_{j_k, j_{k+1}=1}^{2^p} G(U)_{(i_k, i_{k+1}),(j_k, j_{k+1})} X_{j_k, j_{k+1}, \alpha_{k-1}, \alpha_{k+1}}. \tag{3.34}$$

Note that in Eq. (3.34) we assume G to be the transformation induced on the Fock space $\mathcal{F}_{2p}$ of $2p$ modes and $X(U)$ is constructed such that it is related to the transformation of the global space via

$$|X(U)\rangle_{k,2}^{(A^{[j]})_j} = G(\mathbb{1}_{pk} \otimes U \otimes \mathbb{1}_{M-(k+2)p}) |X\rangle_{k,2}^{(A^{[j]})_j}. \tag{3.35}$$

Solving such local problems repeatedly we build up a global unitary $U_{\text{tot}} = \left(U_{\text{opt}}^{(1,2)} \otimes \mathbb{1}_{M-2p}\right)\left(\mathbb{1}_{2p} \otimes U_{\text{opt}}^{(2,2)} \otimes \mathbb{1}_{M-4p}\right) \dots$ (compare Fig. 3.13 a)).

3.3.1. Cost Function

The micro step of the DMRG already yields the optimal update for the (joint) components which minimizes the energy – hence, the energy is not a convenient cost function for the mode transformations here. However, the solution $\mathbf{X}_{\text{opt}}$ of a micro step needs to be projected to a lower rank tensor as explained in Sec. 3.2.1.3. $\mathbf{X}_{\text{opt}}$ contains therefore residual information of the additional important directions in the global Hilbert space which are projected out by the restrictions of the ansatz. We therefore choose a cost function which tries to minimize the projection error and by this tries to use as much as possible of this traditionally discarded residuals. In total, the global mode transformation should then adapt the single-particle basis such that the target state can be approximated by an MPS more efficiently.

The most convincing results where obtained when choosing the cost function

$$f_1^{(k)}(X) = \|\Sigma(\text{mat}(X))\|_1, \tag{3.36}$$

where $\Sigma(A)$ denotes again the singular values of A. $f_1^{(k)}$ is related to the Rényi-1/2 entropy of $|X\rangle_{k,2}^{(A^{[j]})_j}$ for a bipartition of the system into $[k]$ and $[k]^c$ via

$$S_{1/2}\Big(\text{tr}_{[k]^c}\Big[|X\rangle_{k,2}^{(A^{[j]})_j}\langle X|_{k,2}^{(A^{[j]})_j}\Big]\Big) = 2\ln(f_1^{(k)}(X)) \tag{3.37}$$

By successively minimizing $f_1^{(k)}$ for different cuts k through the system we hence not only minimize the projection error of each micro step but also allow for a more efficient approximation of the target state as $S_{1/2}$ over a given cut directly bounds the bond dimensions needed for approximating the given state by an MPS as discussed in Sec. 3.1.2. In addition to $f_1^{(k)}$, one can use in very high-dimensional problems $f_4^{(k)}(X) = -\|\Sigma(\text{mat}(X))\|_4^4$ as it has a simple representation as a tensor network and allows for an analytic computation of the gradient [102]. Note that although $f_4^{(k)}$ does not give rise to certified bounds on the needed bond dimensions the minimization of both cost functions has roughly the same effect as $\|\Sigma(\text{mat}(X))\|_2 = 1$: small singular values are suppressed and large singular values increased. With this we enforce a stronger decay of the Schmidt spectrum of a bipartiting cut separating the modes $[k]$ and $[k]^c$ which is the essential feature needed for an efficient approximation of a state by an MPS.

3.3.2. Optimization Set

In Eq. (3.33) the optimization is performed over all $U \in U(2p)$. However, the optimization set can and has to be adjusted to the cost functions chosen above and to symmetries present in the system if they are exploited by the DMRG. As both cost functions above measure the correlations of the regions $[k]$ and $[k]^c$ only, they are insensitive to operations that act only on one of the two regions, i.e., $f^{(k)}(X(U)) = f^{(k)}(X(U[U_1 \oplus U_2]))$ where $U_1, U_2 \in U(p)$. The optimization can therefore be restricted to the right coset

$$Gr(p, 2p) = U(p) \times U(p) \backslash U(2p), \tag{3.38}$$

Algorithm 3 – Algorithm for combining mode transformations and a two-site DMRG given an initial MPS and a maximal number of DMRG steps. W denotes the corresponding optimization set adapted to the symmetries of the system and we use the procedures introduced in Alg. 2.

procedure DMRG WITH MODE TRAFO($(A^{[j]})_{j\in[V]}$, max steps)
  $k \leftarrow 1$
  for $n = 1$ to max steps **do**
    direction $\leftarrow (-1)^{n//(V-2)}$
    $X_{\text{opt}} \leftarrow$ DMRG STEP$\Big((A^{[j]})_{j\in[V]}, k\Big)$
    $U_{\text{opt}}^{(k,2)} \leftarrow \underset{U\in W}{\arg\min}\, f^{(k)}(X_{\text{opt}}(U))$
    if $f^{(k)}(X_{\text{opt}}(U_{\text{opt}}^{(k,2)})) > f^{(k)}(X_{\text{opt}}(\mathbb{1}))$ **then**
      $U_{\text{opt}}^{(k,2)} \leftarrow \mathbb{1}$
    $A^{[k]}, A^{[k+1]} \leftarrow$ PROJECT$(X_{\text{opt}}(U_{\text{opt}}^{(k,2)}),$ direction$)$
    update couplings and presummed operators with U_{opt} (see App. B.2)
    $k \leftarrow k +$ direction

which is the Grassmann manifold of p-dimensional subspaces of a $2p$-dimensional space. In App. D we comment in general on $Gr(p, n)$ and describe the optimization algorithms we use to optimize over them.

If symmetries are exploited during the DMRG, the components $A^{[j]}$ come with a specific structure. In order to preserve this structure, only mode transformations that do not mix different symmetry sectors are admissible. That is, the mode transformation has to commute with the generators of the symmetry. Consider the generic scenario in quantum chemistry applications with $p = 2$, i.e., each site supports one spin-up and spin-down mode, the most frequently exploited symmetries are a global $U(1) \times U(1)$ or $SU(2)$. In these cases the local mode transformations commuting with the symmetries are of the form

$$U = \begin{pmatrix} U_{1,1}^{(1)} & 0 & U_{1,2}^{(1)} & 0 \\ 0 & U_{1,1}^{(2)} & 0 & U_{1,2}^{(2)} \\ U_{2,1}^{(1)} & 0 & U_{2,2}^{(1)} & 0 \\ 0 & U_{2,1}^{(2)} & 0 & U_{2,2}^{(2)} \end{pmatrix}, \tag{3.39}$$

so up to a reordering of sites $U = U^{(1)} \oplus U^{(2)}$ where $U^{(1)} \in Gr(1, 2)$ acts on the spin-up and $U^{(2)} \in Gr(1, 2)$ on the spin-down electrons on the sites k and $k + 1$ only[16], where in case of a global $SU(2)$ symmetry we in addition have to ensure $U^{(1)} = U^{(2)}$.

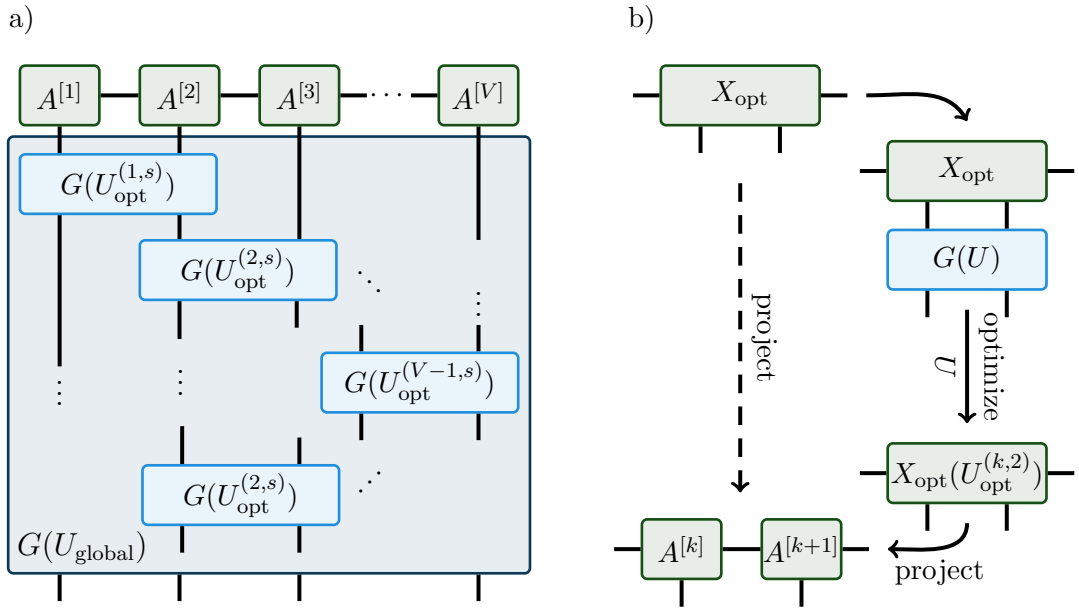

Figure 3.13. – a) displays the ansatz class we optimize over in the two-site DMRG with mode transformation. In addition to the MPS we allow for an additional change of the single-particle basis by the mode transformation U_{global} which is composed of a sequence of local mode transformations. In b) the additional steps needed for optimizing the single-particle basis in parallel to the MPS are shown. Instead of directly projecting a found solution X_{opt} in the micro step (dashed arrow), we optimize the basis locally in order and project the solution in the optimized single-particle basis.

3.3.3. Algorithm and Examples

The resulting scheme is summarized in Fig. 3.13 and Alg. 3. For a given $k \in [V-1]$ we calculate as usual the lowest eigenvector X_{opt} of $H^{(k,2)}$. Instead of directly projecting the solution X_{opt} we first solve (3.33) and then project $X_{\text{opt}}(U_{\text{opt}}^{(k,2)})$. In order to keep the energy expectation value constant, we also rotate all operators with the transformation, so for instance calculate the couplings of $H(\mathbb{1}_{pk} \oplus U_{\text{opt}}^{(k,2)} \oplus \mathbb{1}_{M-(2+k)p})$ using Eq. (2.37). In App. B.2 we show how the presummed operators needed for the DMRG have to be transformed in order to account for the changed couplings and that this can be done in a time scaling as $\mathcal{O}(M^3r^2)$. In total we therefore obtain an algorithm that performs a two-site DMRG in parallel to a single-particle basis optimization similar in spirit to the Hartree-Fock method. The resulting method is, as to be expected, again prone to getting stuck in local minima. This can happen already for a local optimization step as displayed in Fig. 3.14. Here we see that multiple local minima might exists for each local optimization problem. The check in line 7 of Alg. 3 ensures that the minimum found does not increase the projection error compared to not rotating the single-particle basis

[16] This result should be quite intuitive as we can only talk about a fixed number of spin-up and down electrons if we do not mix them into hybrids by a mode transformation.

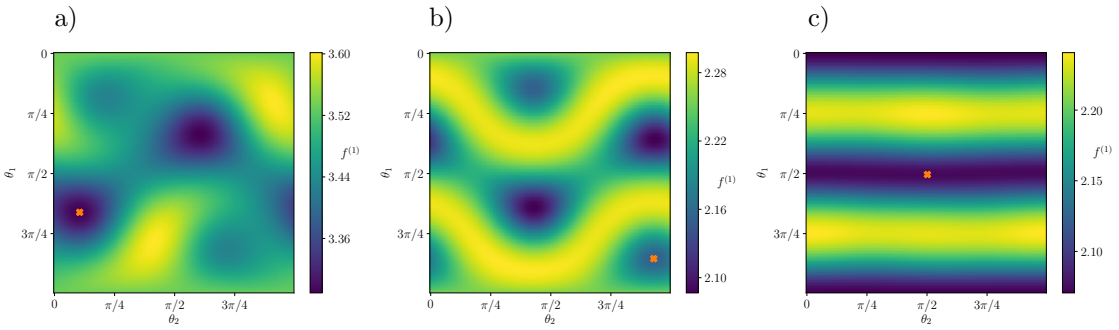

Figure 3.14. – Plot of three typical situations encountered during the minimization of the cost functions $f^{(1)}$. We run Alg. 3 for a system of 6 fermions in 12 modes with an interacting Hamiltonian as in Eq. (2.58) with random couplings t and v. We display the cost function over $Gr(1,2)$ using the parametrization explained in App. D. The orange cross indicates the minimum identified by a Nelder-Mead method. During the optimization of the single particle basis by local transformations, we typically encounter graphs of the cost function as displayed in a)-c). In a) and b), a non-trivial basis update can be found that reduces the Rényi-1/2 entropy over the considered cut. Note that in both, a) and b), local minima are present and the optimization scheme might actually converge only to a local minimum as shown in b). In c) we show an almost trivial case which is often encountered after several updates in situations where a no further local optimization of the basis is possible at the considered cut.

and is found to yield a stable scheme in practical applications. In addition we find that a scheme which combines different optimization methods for the single-particle basis is less sensitive to local minima. In order to explore the performance of the proposed method further we present three different examples below. First, a toy model is used in order to illustrate the underlying idea of the combination of TNS and mode transformations more concretely by showing that the correct single-particle basis does allow to reduce the entanglement present in the state to be approximated. Secondly, we test our method on a realistic example from quantum chemistry. Here, we discuss in more detail the potential of the individual and combined schemes and show that we can significantly reduce the resources needed or increase the accuracy of the obtained results. Thirdly, the insight that a real time evolution can be closely related to the two-site DMRG [91] allows us to formulate a time evolution scheme within an adaptive single-particle basis on the MPS manifold. We then show that for moderately interacting systems, the varying single-particle basis can account for a lot of the entanglement created during the time evolution such that the resources needed are again drastically reduced.

3.3.3.1. Toy Example: The Fermi-Hubbard Model

Let us illustrate the essential profit and idea behind the change of the single-particle basis during a DMRG by using a well understood toy model, the one-dimensional spinless Fermi-Hubbard model defined in Eq. (2.61). In the limit of no interactions the

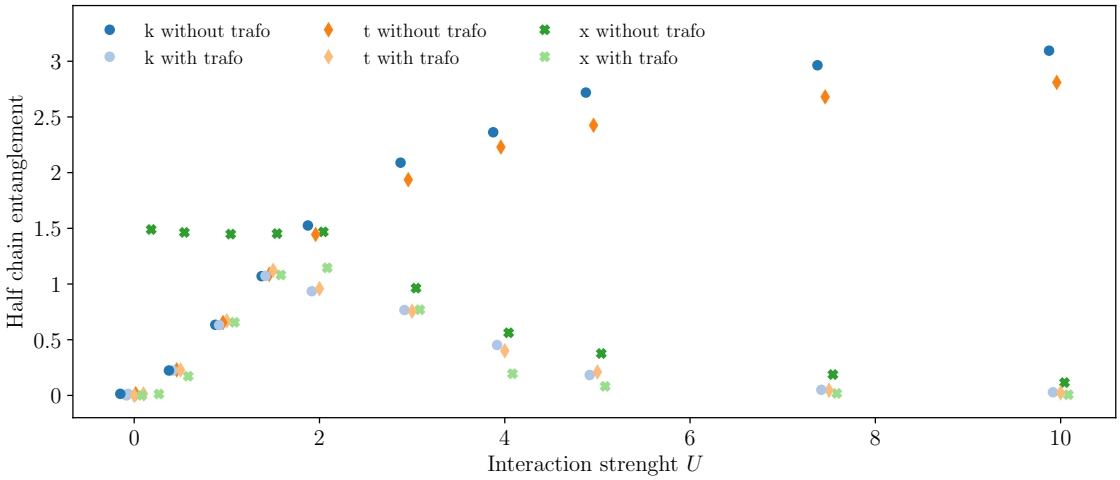

Figure 3.15. – Entanglement of the ground state of the one-dimensional spinless Fermi-Hubbard model defined in Eq. (2.61) on 32 sites with open boundary condition in different single-particle basis. By x, t, and k we denote the real space, the eigenbasis of the kinetic energy and the momentum space (obtained from a discrete Fourier transform of the real space), respectively. We plot the von Neumann entropy for a bipartiting cut at the center of the chain for different interaction strengths, where the hopping amplitude was fixed to 1. The data points are displaced in x direction for better visibility. In addition we show the results of applying adaptive mode transformations, starting out in all three different initial basis. In order to adapt the basis we use Alg. 3 and perform a global basis update minimizing the energy functional as in Eq. (3.32) every 4 sweeps. All calculations are performed with maximal bond dimension 128 at half filling with $U(1)$ symmetric MPS.

Hamiltonian is diagonalized by a Fourier, or in finite systems with open boundary condition a Fourier-like, transformation. Written in this basis the ground state is a single Slater determinant without entanglement between different modes. In the limit of infinite interactions however (or vanishing hopping amplitude), charge density wave states which in the real space occupation number representation take the form $|1, 0, 1, 0, \dots\rangle$ are states with minimal energy and possess no correlation between different modes for a real space formulation. However, written in the corresponding opposite basis, i.e., the real space basis for weak interactions or the momentum space for strong interactions, yields a representation with significant correlations. In Fig. 3.15 we show how the half chain cut von Neumann entropy of the ground state changes for different single-particle basis. We see a strong dependence of the correlations build up in the ground state on the interacting strength interpolating between the two easily understood regimes. In the real space basis we obtain a clear change of the half chain entropy around the phase transition at $U = 2t$ between the localized Mott phase for strong interactions and the delocalized metallic phase at weak interactions [112, Sec. 5.4.5]. Adapting the basis during the DMRG calculations allows to smoothly interpolate between the two regimes. Around the phase transition, correlations are built up, reflecting the criticality of the system while deep in the corresponding phase most correlations can be removed

using the appropriate single-particle basis. This nicely and intuitively illustrates that by choosing the correct initial basis we are able to reduce correlations present and can hope for more efficient approximations of the ground state. Note further that it can be advantageous to move away from a local formulation. At low interaction strengths we find that less correlations are built up for a formulation in momentum space which has a non-local Hamiltonian.

3.3.3.2. Interacting Fermions in Quantum Chemistry

As a more realistic proof of principle we want to consider an example involving non-local interactions and choose a ring of 6 beryllium atoms. This system was investigated in [108] were a strong dependence of the convergence behavior of a DMRG calculation on the single-particle basis was noted. It was observed, in agreement with other studies [109], that a basis build up from localized orbitals, obtained using a Foster-Boys localization scheme on the Hartree-Fock orbitals, yields a strong improvement of the overall convergence of a DMRG calculation compared to the use of plain Hartree-Fock orbitals. Furthermore, a ring of beryllium atoms features, just as the Hubbard model, a change in its correlation structure, here mediated by the interatomic distance. For short distances σ-bond like orbital structure emerges between pairs of atoms where in the limit of large distances individual atomic orbitals are the dominant contribution [108]. At an interatomic distance of about 2.6 Å the correlation effects are the strongest due to an avoided crossing of the ground and first excited state and the favored orbitals switch from molecular to atomic orbitals [108]. We therefore use four different initial basis configurations. The canonical Hartree-Fock and localized version of each of the two dominant contributions. *The corresponding integrals t and v were created and provided by Edoardo Fertitta; for details we refer to [108].* Starting in these different basis sets we use Alg. 3 in combination with additional basis optimization techniques. The results and corresponding routines used are presented and discussed in Fig. 3.16 and 3.17.

The drop of the bond dimension in the left panel of Fig. 3.16 by almost two orders of magnitude shows that Alg. 3 allows us to significantly reduce the resources needed during a ground state approximation when we employ a fixed accuracy in a realistic setting and start in an unfavorable initial basis. In the right panel it is shown that Alg. 3 allows us to improve the accuracy starting from both a non-optimal and a close to optimal initial basis in a calculation with limited resources. It is important to note that we obtain an improved accuracy in the ground state energy even when using localized orbitals which are often believed to be optimal. However also note that we do not find the globally optimal single particle basis. Starting from the canonical Hartree-Fock basis, in the right panel of Fig. 3.16 the optimized basis leads to a worse energy approximation than the localized basis (in an optimized ordering). In short, we get stuck in a local minimum which however is significantly better than the initial starting point.

In Fig. 3.17 we show that we can overcome certain convergence issues of the purely local scheme. There we show results for different optimization schemes which use various combinations of the three basic basis optimization approaches: the local mode updates via Alg. 3, the mutual information based reordering and the SCF-like energy based

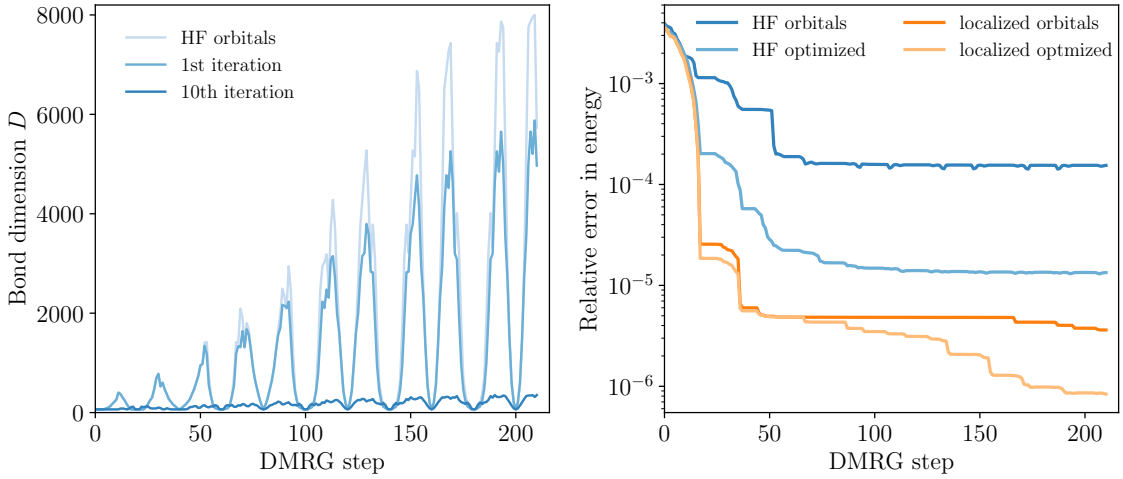

Figure 3.16. – Investigation of the convergence behavior of the two-site DMRG ground state approximation in a stretched configuration with interatomic distance of 3.3 Å. We only investigate the in this regime dominant canonical and localized configurations of the atomic orbitals. The results were obtained with the QC-DMRG code of Örs Legeza including additional convergence amplifiers such as the dynamical block state selection approach [95] and the configuration interaction based dynamically extended active space procedure [113]. The calculations were performed by Libor Veis. Starting in the different initial single-particle basis two ordinary DMRG sweeps have been performed followed by eight additional sweeps which optimize the single-particle basis using Alg. 3. The single-particle basis is then reordered based on the mutual information pattern of the approximated ground states. All calculations have been performed with open boundary $U(1) \times U(1)$ symmetric MPS and we use local mode transformation preserving the $SU(2)$ symmetry of the Hamiltonian. The left panel shows the bond dimension needed for a calculation starting in the canonical Hartree-Fock basis at bounded discarded weight $\epsilon \leq 10^{-6}$ with minimal bond dimension 64. The light blue line (HF orbitals) corresponds to 10 sweeps performed in the Hartree-Fock basis. In addition we display the bond dimension needed when adapting the single-particle basis according to the described scheme in its 1st and 10th iteration. The right panel shows the relative error of the approximated ground state energy at a bounded bond dimension of $r \leq 256$, where the reference value was obtained by a calculation with a bond dimension 2048 in the localized basis. The dark blue and dark orange line correspond to calculations in the canonical Hartree-Fock and localized orbitals respectively. In light blue and light orange we plot the results obtained from the adaptive single-particle basis scheme after the 15th and 10th iteration starting in the Hartree-Fock and localized basis respectively. In the left panel we find that adapting the single particle basis allows to reduce the resources needed by more than two orders of magnitude while keeping the accuracy of the calculation fixed. Furthermore, we are able to improve both the non-optimal Hartree-Fock and the localized basis substantially such that the ground state energy approximation improves by roughly an order of magnitude as shown in the right panel.

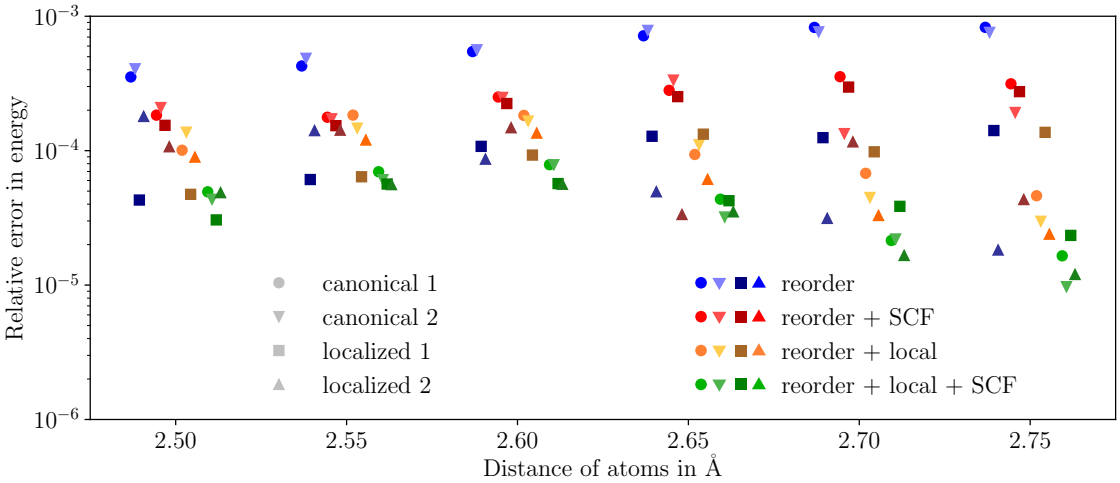

Figure 3.17. – Study of the ground state energy obtained with different DMRG schemes optimizing the single-particle basis at different configurations of the Be_6 ring. We display the minimal relative error obtained in each different optimization run for several interatomic distances around the avoided crossing were we displace the data in x-direction for better visibility. The different symbols indicate the different initial single-particle basis, the molecular canonical Hartree-Fock basis (canonical 1), the atomic Hartree-Fock basis (canonical 2) and the corresponding localized ones. The different colors denote the different basis optimization approaches. We display the minimal relative energy error encountered where as reference we use the best energy obtained from a calculation with bond dimension 1024 in the corresponding ideal localized basis provided by Örs Legeza using his QC-DMRG code. We performed 44 sweeps in every run (considerably less total steps than in Fig. 3.16) and used $U(1) \times U(1)$ symmetric open boundary MPS with a maximal bond dimension of 128. The local mode transformation respected only the $U(1) \times U(1)$ of the Hamiltonian. For the blue runs we use a plain two-site DMRG calculation with a reordering of the basis every 5 sweeps. In red we plot the result using a two-site DMRG with additional global basis updates every 2 sweeps and a reordering every 7 sweeps. The orange data was essentially obtained as the data in Fig. 3.16 with local mode updates after two initial convergence sweeps and a reordering every 7 sweeps (after each reordering we pause the local updates for one sweep). The green data was obtained by combining all methods where every 3 and 7 sweeps we performed a global energy based update or reorder the basis respectively. The blue data points obtained in the plain initial bases illustrate nicely the change of the favored configuration from 1 to 2 at around 2.60 Å. Using SCF-like updates alone (red data) yields an instable behavior in this example. Starting in localized single particle bases we find here that the obtained solution are often worse than using reorderings only. If we use local mode transformation only (orange data) we obtain more stable results but the converged energies can contain a strong dependence on the initial basis (at 2.75 Å they spread over an order of magnitude). The combined scheme on the other hand (green data) does provide a more robust method which yields results that are clearly less dependent on the initially chosen basis and reliably finds favorable single particle bases.

rotation obtained by solving the optimization problem in Eq. (3.32). Note that in the blue data, obtained in the different basis up to reorderings only, we see the transition from one preferred single-particle basis to the other at around 2.6 Å. Furthermore, we obtain that in most cases the different strategies lead to improvements in the approximability of the ground state energy. Note furthermore, that in this example the local mode transformation yield better and more stable results than the global energy based updates (orange vs. red data points). The only scheme that truly overcomes the strong memory effects from the initial bases is the combined scheme which yields in this example an up to two orders of magnitude improvement in the relative energy error. Even in this correlated setting this combined scheme identifies in almost all cases a single-particle basis which is superior to the best localized initial basis. Hence the combined scheme using local mode transformations, mutual information based reorderings and SCF-like global updates comes the closest to a black box scheme for optimizing the single-particle basis in conjunction with a DMRG based ground state approximation and allows to obtain highly accurate results with strongly restricted resources.

3.3.3.3. Time Evolution

All time evolution schemes based on tensor network algorithms are limited to short times as in a generic quantum system correlations spread over time rendering their approximability by MPS less efficient and in practice infeasible after certain times. As mentioned earlier, recently a connection between established time evolution schemes and DMRG has been derived [91]. As a result a heuristic algorithm was formulated which allows to perform a time evolution of non-local quantum systems by slightly adapting the routines of the two-site DMRG. This is in stark contrast to established tensor network time evolution schemes which are only able to simulate the time evolution generated by local Hamiltonians. As a consequence, the scheme presented in [91] can be applied irrespectively of the single-particle basis in which a fermionic problem is formulated such that we can design a heuristic hybrid method which tracks the time evolution generated from a Hamiltonian not only by inducing a flow on the MPS components but also by adapting the basis.

We will not review the details of the basic time evolution scheme here. Let us only note that given a Hamiltonian H and initial state $|\psi\rangle$ for which we have access to the reduced Hamiltonians $H^{(k,2)}$ we are able to design a symmetric first order integrator for a time step dt of the time dependent Schrödinger equation by performing a single sweep over the system [91]. Using a composed symmetric integrator which decomposes the step dt into smaller parts such that multiple runs of the integrator, i.e., sweeps, are employed we are able to reduce the error in each time step [114, Ch. 5]. In short, performing a preset number of sweeps through the system we are able to calculate the evolution of $|\psi(t)\rangle$ to $|\psi(t+dt)\rangle$. During this procedure, the bond dimension increases typically unboundedly in time, reflecting the spreading of present correlations in the system and rendering the approximation of the time evolved state by an MPS less accurate. Fixing the bond dimension to a finite value in order to regain efficiency then again leads to truncation errors.

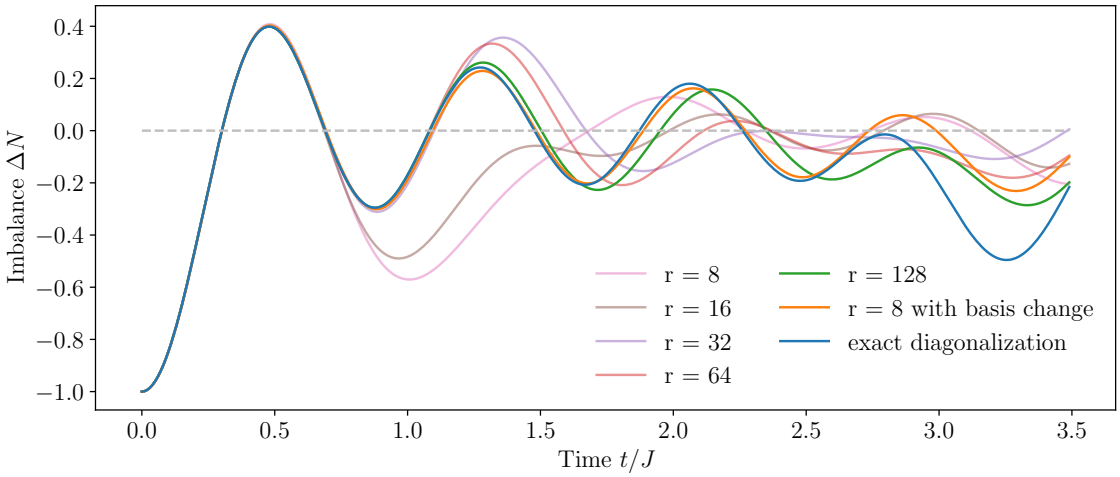

Figure 3.18. – Comparison of the time evolution in a fixed single-particle basis with increasing number of parameters to the described algorithm with a varying basis. We simulate the time evolution of the one-dimensional spinless Hubbard model defined in Eq. (2.61) with periodic boundary condition on 24 sites and interaction $U = 0.25J$. As initial state we choose a charge density wave which has the occupation number representation $|\psi(0)\rangle = |1, 0, 1, 0, \dots\rangle$ in the on-site basis. We track the imbalance of the state over time. We compare the exact solution in blue obtained with exact diagonalization with the MPS calculation in the fixed real space basis with bond dimensions $r = 8, 16, 32, 64, 128$ and a calculation in a varying basis with maximal bond dimension $r = 8$. For the time evolution we used a step size $dt = 0.01J$ and used a composite integrator of fourth order [114, Ch. 5] in all MPS calculations.

Taking this scheme for granted, we can perform individual time steps dt of the evolution. In order to allow for a varying basis we use a slight modification of Alg. 3. If we would use a scheme as intertwined as the one formulated for the two-site DMRG and adapt the single-particle basis during one global time step, we would lose the symmetry of the integrator which results in a loss of accuracy. Instead we adapt the basis after a fully performed time step dt which slightly changes the MPS. For this we perform a mode-sweep which only adapts the single-particle basis along the lines of Alg. 3 without performing a DMRG step in line 5 of Alg. 3 but instead using the simple contraction of the current MPS components as tensor X_{opt}, i.e., we alter line 5 to $X_{\text{opt}(i_k,i_{k+1}),\alpha_{k-1},\alpha_{k+1}} \leftarrow \sum_{\alpha_k} A^{[k]}_{i_k,\alpha_{k-1},\alpha_k} A^{[k+1]}_{i_{k+1},\alpha_k,\alpha_{k+1}}$. Interchanging between mode sweeps and time sweeps, i.e., adapting the single-particle basis to the current correlations in $|\psi(t)\rangle$ and performing then the next time step $|\psi(t)\rangle \to |\psi(t + dt)\rangle$, yields a scheme which allows for the time evolution of an MPS formulated within a time dependent single-particle basis. In this respect it can be seen as a heuristic MPS version of algorithms such as the multiconfigurational time dependent Hartree method which integrates the Schrödinger equation by varying the single-particle basis and adapting the CI expansion coefficients of the wave function simultaneously [115, 116]. In Fig. 3.18 we show an exemplary application of the described scheme to the one-dimensional Fermi-

Hubbard model. In reminiscence to previous numerical and experimental investigation of relaxation processes in the Hubbard model (see for instance [117]) we investigate the evolution of the imbalance ΔN defined by

$$\Delta N(t) = \frac{N_{even}(t) - N_{odd}(t)}{N}, \tag{3.40}$$

where $N_{even}(t) = \sum_{j\,\text{even}} \langle \psi(t)|n_j|\psi(t)\rangle$ denotes the expected particle number on all even sites, N_{odd} on all odd sites respectively. It is shown that if the interactions present in the system are not too strong, adapting the single-particle basis while performing a time evolution allows to significantly extend the validity of the time evolution and allows to save a tremendous amount of resources. The calculation performed at bond dimension $r = 8$ in a varying single-particle basis yields better results than the calculation in a fixed basis at bond dimension $r = 128$, saving more than an order of magnitude in the bond dimension (recall that memory cost and computation time scale roughly as $\mathcal{O}(r^2)$ and $\mathcal{O}(r^3)$). Compared to a calculation at the same bond dimension, the time for which we capture the evolution correctly extends significantly. Performing the same calculation with increasing interaction strength however, shows that using an adaptive basis yields similar computational advantages although the correctly simulated time window is less extended. This is due to more dominant two-particle effects which create correlations that can not be captured by a rotated single-particle basis. From this we expect that in the presence of moderate interactions, time dependent hybrid schemes which not only optimize the MPS components but also the single-particle basis to be significantly more efficient than comparable, established tensor network algorithms.

3.3.4. The Bird's-Eye View

Let us take a step back and consider the discussed schemes and obtained results on a more abstract level. From a technical and mathematical point of view we achieved to unite tensor network manifolds with the structure of an exterior product space. The optimization of the single-particle basis in conjunction with updates of the MPS components aims at best approximating a target vector by an MPS while rotating and deforming the underlying MPS manifold in order to find a more efficient parametrization and to reduce the distance to the target state. In doing so, we exploited algorithms and methods developed for the optimization on manifolds with unitary constraints, joining these two different branches of applied mathematics and numerical physics. The ultimate reason for why this is possible lies in the fact that we do not substantially change the structure of the cost function of the DMRG, the Rayleigh quotient, i.e., we do not create higher order coupling terms in the Hamiltonian.

From a physical point of view let us note that, from the perspective of non-interacting fermionic systems, MPS capture correlations in the wrong picture and do not allow for the most efficient approximation of a target state, e.g., a single Slater determinant can require an arbitrary high bond dimension to represent it. We resolve this by adding Gaussian transformations as an additional ingredient to the DMRG scheme. The mode transformations allow us to overcome the necessity of an area law or restricted correlation

of the target state in the for us natural basis for an efficient approximation. However, it is replaced with the need of suitable, i.e., area law like, correlations in some single-particle basis, not necessarily the one we would or can formulate the problem in. We therefore include efficiently parameterizable high entanglement effect into the tensor network description and extend its capabilities by this. Our results show that it can be advantageous to abandon locality in the description in order to increase the accuracy (compare the results discussed in Fig. 3.18).

On a more concrete level let us note that the local mode transformations have different advantages and disadvantages. Most importantly they can be closely intertwined with the DMRG updates and adapt the single-particle basis to the correlation structure of the state irrespective of the associated cost functions or deeper purpose of the used algorithm which yields flexibility and allows especially for their extension to time evolution schemes. In addition they do not yield much overhead in the computational costs and do not change the asymptotic scaling of the DMRG algorithm. However, for increasing system size they are only of limited use due to their local nature. Especially, the number of DMRG steps required in order to represent a generic global mode transformation does increase with the system size. In addition the scheme is, as other discussed methods, prone to get stuck in local minima. We can overcome both limitations by unifying the local updates of Alg. 3 with global energy based schemes and reorderings of the basis. The combination of these three methods provide a suitably robust black box method (compare Fig. 3.17) as the different optimization strategies are sensitive to different local minima.

The presented scheme is rather general. For fermionic systems it can be generalized to other tensor network decompositions; be it tree tensor networks or two-dimensional decompositions such as PEPS. Furthermore, based on the results found, it seems desirable to include possible structures in more general settings. When tensor network algorithms are employed for solving problems of applied mathematics, taking into account additional algebraic structure of the state space has the potential to improve their convergence rates and the approximability of the desired solution.

3.4. Summary

Over the past decades, a zoo of different numerical approaches for the simulation of non-local interacting fermionic systems has been developed. All these methods, invoke different approximations in order to make the computationally hard task of identifying the ground state of an interacting fermionic system tractable. As discussed, the approximations restrict the set of representable states, where such restrictions can be conveniently understood based on the correlation structures they allow us to resolve.

One, in terms of the representable correlation structure, very general variational set of states is given by tensor network states. They enable us to tackle highly correlated settings and allow for a heuristic error control by systematically increasing the number of parameters. With this, tensor network methods are found to be successfully applicable to non-local and strongly correlated settings. The necessity of an area law in all correlations

within the approximated state however, has to be compensated by a steadily increasing number of parameters. This renders tensor network methods usually less efficient in weakly correlated settings but a versatile and valuable tool for strong correlations in systems of moderate size. They allow for the approximation of ground and excited states as well as for simulating the time evolution of the given system using the DMRG algorithm and variants thereof.

In this chapter we carefully introduced and discussed the most important concepts of TNS, MPS and DMRG from a practical numerical perspective. We explained how a DMRG algorithm can be implemented which is capable of simulating non-local interacting fermionic systems efficiently while accounting for (Abelian) symmetries of the system. We further discussed a few structural insights into the DMRG scheme which allow us to infer the projection error of individual excited states and identify local minima of the DMRG routine originating from its concept of performing local updates only. The latter will be important when it comes to establishing a rigorous convergence analysis of the DMRG method which would allow to apply it far away from its original setting of local one-dimensional quantum systems as an independent eigensolver for large matrices.

Furthermore, we argued that MPS do not capture correlations in fermionic systems as efficiently as possible. As fermionic systems are usually simulated based on a lattice of individual modes within the Jordan-Wigner picture, the choice of the single-particle basis has a significant effect on the correlation structure of the final state. We presented a scheme which allows to adapt the single-particle basis according to the needs of the MPS. This method can be closely intertwined with the usual DMRG scheme and generalizes to methods of simulating the real time evolution of quantum systems. We showed that in practical applications we are indeed able to identify single-particle bases in which more efficient approximations of ground and time evolved states are possible and obtained in specific example a reduction of the needed parameters by several orders of magnitude. Combined with other established schemes for updating the single-particle basis we obtained a methods which performs close to black box tool and chooses the correct single-particle basis. We hence are able to obtain more accurate results with less resources and lift the ambiguity of the initial basis which usually has to be chosen based on intuition, experience or, in the worst case, convention. In what sense the obtained orbitals can be used beyond their technical advantages for more physical insights, for instance give rise to a different perspective on chemical bondings, is subject of current research.

4. Towards an Understanding of Mean Field Approaches via a Fermionic de Finetti Theorem

In the previous chapter we have seen how to numerically capture strongly correlated systems and approximate their properties based on refined schemes. In many instances however such an involved description is not necessary. Systems appearing in nature do in fact often possess an utterly simple structure compared to the complexity available in the full state space. In many cases effective descriptions of interacting systems in terms of very few parameters can be found. In the extreme case one finds that essential properties of the system are already captured by mean field approaches. Despite many issues such as spin contamination in open shell settings or the violation of other global symmetries [5], practical experience (see for instance [6, Sec. 3.7] and references therein) shows that the in Sec. 2.5.1 discussed Hartree-Fock approximation is often already able to account for many features of the system correctly and only small corrections are needed in cases where we want more than a rough and intuitive understanding of the system. It is however unclear, what the underlying structures are that a system needs in order to be well captured by mean field approaches. In fermionic systems we are not able to decide beforehand if a mean field approach is sufficient and if so, to which degree we can trust it. We hence know that in several systems a simple underlying structure exists and even know how to exploit it, but it seems fair to say that we do not properly understand it.

This chapter aims at shedding some light on this important question and take a small step into the direction of its solution. We present a setting where we are able to bound mean field approaches for a specific set of systems. Our main tool which we will employ for this is a fermionic de Finetti theorem which at the same time is the main technical result of this chapter. Our investigation is hereby motivated by the success of quantum versions of de Finetti's and related theorems in bounding and understanding mean field approaches for distinguishable particles.

Quantum de Finetti theorems allow us to link, in their easiest formulation, a strong symmetry constraint on a state to the correlations present in local reductions of the system. In detail, if a system is invariant under the permutation of its constituents the state is locally only classically correlated, i.e., it can be well approximated by a separable state. This directly connects to the intuition that mean field approaches should work best in systems in which all constituents are interacting equally with each other, which in the extreme can also be viewed as a system of infinite spatial dimension. In fermionic systems, such a symmetry constraint is found to impose further restrictions due to the

antisymmetric nature of the state. We will show here that a fermionic mode de Finetti theorem can be derived based on a more elaborate definition of permutation invariance which does not interfere with the canonical anti-commutation relations. The obtained theorem then allows us to approximate a permutation invariant state by a mode separable state, i.e., a state that can be written as a convex combination of fermionic product states in the sense of Sec. 2.3.1. In special cases these states can be linked to pure Gaussian states, which are the variational set of the Hartree-Fock approximation. Furthermore, we discuss how a de Finetti theorem is used to expand and generalize established results for fermionic systems on the example of fermionic central limit theorems which capture the static emergence of Gaussian states. Note however, that the investigations presented in the following only open a window towards a deeper understanding of realistic systems which are approximable by mean field methods. The assumptions needed on the system to be captured by our setting are rather strict such that much of the value of the following lies in conceptual insights as we will explain.

We will start our discussion by reviewing well understood and established de Finetti type theorems for distinguishable particles and how they help in bounding mean field approaches. Using the basic quantum de Finetti theorem for distinguishable particles we then derive a fermionic mode de Finetti theorem. We continue to discuss the structure of the obtained separable states in more details and show that we are in special cases able to bound fermionic mean field approaches.

4.1. Quantum de Finetti Theorems

De Finetti's theorem originally roots in probability theory and relates the joint distribution of multiple random variables which is invariant under permutation of the events to a combination of identical and independent processes. In detail it is shown that an extendable permutation invariant sequence of random events, meaning that their probability does not depend on the order of the events, can be well described by a mixture of independent identical copies of a common abstract model which depends on a single stochastic parameter (see for instance [118] for an introduction). A non-trivial and for instance also non-Markovian example of an extendable permutation invariant distribution would be the one resulting from Pólay's urn model, i.e. drawing balls from an urn with replacing the drawn ball by n copies. De Finetti's theorem was extended to a large number of different settings and cases for instance to exchangeable distributions of finitely many random variables, more general notions of permutation invariance, e.g. multiple sequences which are only internally permutation invariant, and to Markov processes (see [119] for a short introduction into different ramifications of the classical de Finetti theorem). In addition it allows to bridge gaps between subjective Baysian approaches versus more traditional frequentist interpretation of probability theory as it proves that one can assign an underlying common model to a sequence of exchangeable variables, an assumption which is natural in a frequentist interpretation but a priori denied within the Baysian view [118].

It was later realized that similar statements hold in quantum mechanical systems

[120–122]. In the context of quantum state tomography de Finetti type theorems help, just as in the case of classical probability distribution, to establish a foundation for an operational and epistemic interpretation and description of the tomography process, eliminating for instance the reference to unknown quantum states [122]. Furthermore, permutation invariance can often either be enforced by specific preparation protocols or naturally be assumed to hold for different systems. Along these lines, quantum versions of de Finetti's theorem allow to generalize and extend established schemes. Quantum state tomography is then able to construct a faithful model of the system if a permutation invariance of the measured subsystems is assumed instead of the ideal independence of copies [123]. In addition, quantum de Finetti theorems allow to extend the security proof of quantum key distribution protocols known for i.i.d. scenarios to general settings and by this to rigorously prove the practical security of such protocols [124, 125].

A quantum state is called permutation invariant, if it is invariant under an arbitrary permutation of the constituents. Consider for instance a state ρ on V distinguishable particles each described by an Hilbert space $\mathbb{C}^d$, i.e., $\rho \in \mathcal{D}(\mathbb{C}^{d^V})$, then ρ is permutation invariant if for any operators $A_j \in \mathcal{B}(\mathbb{C}^d)$ with $j \in [V]$ and permutation $\pi \in S_V$ we have

$$\mathrm{tr}(\rho A_1 \otimes A_2 \otimes \cdots \otimes A_V) = \mathrm{tr}(\rho A_{\pi(1)} \otimes A_{\pi(2)} \otimes \cdots \otimes A_{\pi(V)}), \tag{4.1}$$

or in short, if the permutation does not change ρ, i.e. $\pi\rho\pi^\dagger = \rho$. A state $\rho \in \mathcal{D}(\mathbb{C}^{d^V})$ is called extendable permutation invariant if for any $V' > 0$ there exists a state ρ' on $V+V'$ particles which is permutation invariant and $\mathrm{tr}_{[V]^c}\,\rho' = \rho$. Note that every extendable permutation invariant state is of course permutation invariant itself but the converse is not true. An extendable permutation invariant state can be shown to be separable, i.e. to contain no entanglement and all correlations present result from being a classical mixture [120–122]. Next to being able to capture extendable permutation invariant states, the reductions of a permutation invariant state of fixed size can be investigated [126–128]. Here one finds that the following theorem, which will be the backbone of our discussion further below.

Theorem 3 (De Finetti theorem for finite systems of distinguishable particles [127])**.** *Given a permutation invariant state ρ on V distinguishable particle of dimension d, i.e. $\rho \in \mathbb{C}^{d^V}$, then there exists an $r < \infty$ and sequences $\sigma_l \in \mathcal{D}(\mathbb{C}^d)$ and $a_l \in [0,1]$ for $l \in [r]$ with $\sum_{l=1}^r a_l = 1$ such that for any $k \in [V]$ we have*

$$\left\| \mathrm{tr}_{[k]^c}(\rho) - \sum_{l=1}^{r} a_l \sigma_l^{\otimes k} \right\|_1 \leq 2\frac{d^2 k}{V}. \tag{4.2}$$

The reduced states of a permutation invariant state are well approximated by a convex combination of i.i.d. product states and all remaining entanglement vanishes in the limit $V \to \infty$, i.e., of being extendable permutation invariant. Note that due to the permutation invariance Thm. 3 is of course insensitive to which subset of $[V]$ of size k is chosen and we choose the initial k sites out of convenience. Triggered by this finite size quantum de Finetti theorem different ramifications followed. If we change for instance the norm in Thm. 3 from the trace distance (distinguishing quantum states by

any operation or measurement) to the operational distinction using only local quantum operations and classical communication (LOCC norms) one obtains that the error will scale as $\log(d)$ instead of d [129–131]. This adaption yields more flexibility and results in quasipolynomial time algorithms for the detection of entanglement [129, 131]. Furthermore, one can relax the requirement of obtaining an i.i.d. product state. One can derive a de Finetti like theorem with a modified scaling of the error if one approximates the reductions $\mathrm{tr}_{[k]^c}\,\rho$ of a permutation invariant state with the convex combination and permutations of states which act as i.i.d. copies of local states on $l < k$ sites and are arbitrary on the remaining $k-l$ sites, i.e., states of the form $\sigma^{\otimes k} \otimes \rho'$. The error for such an approximation is exponentially suppressed in $k-l$, i.e., the size of the subsystem on which the states are not further specified [123]. Such an extension allows to generalize the hypothesis in hypothesis testing schemes [132] and to consider large subsystems with $k \in \mathcal{O}(V)$ [123].

Next to being useful in the context of quantum information theoretic considerations, quantum de Finetti theorems yield bounds on mean field approximations. Consider a system of V distinguishable particles with a (not necessarily geometrically) k-local Hamiltonian

$$H = \frac{1}{|P|} \sum_{S \in P} h_S, \tag{4.3}$$

where P is a collection of subsets of $[V]$ of maximal size k and all $\|h_S\| \leq 1$. Assume that H has a permutation invariant ground state space projector ρ. According to Thm. 3 there exist states σ_l and probabilities a_l such that we can approximate

$$\begin{aligned} 2\frac{d^2 k}{V} &\geq \frac{1}{|P|} \sum_{S \in P} \left| \mathrm{tr}\left(h_S \left[\sum_{l=1}^{r} a_l \sigma_l^{\otimes |S|} - \mathrm{tr}_{S^c}(\rho) \right] \right) \right| \geq \left| \mathrm{tr}\left(H \left[\sum_{l=1}^{r} a_l \sigma_l^{\otimes V} - \rho \right] \right) \right| \\ &\geq \min_{\sigma \in \mathcal{D}(\mathbb{C}^d)} \mathrm{tr}(H \sigma^{\otimes V}) - e_{\mathrm{GS}}, \end{aligned} \tag{4.4}$$

with e_{GS} denoting the ground state energy density of H. The quantum de Finetti theorem allows us to bound the error of a mean field approximation to the ground state, i.e. approximating the ground state by a product state in permutation invariant systems. It is however important to note that this insight is mostly of conceptual interest. Having a permutation invariant ground state space projector is a very restrictive assumption. Intuitively mean field approaches should be able to approximate the ground state of a system the better each part of the system is coupled to the rest. The permutation invariant systems are the extreme cases of this intuition. However, for distinguishable particles the gap between these extreme cases and realistic systems can be overcome. It can be shown that the strict assumption of being permutation invariant can be relaxed to having a well connected interaction graph [133]. The result gives rise to more general versions of de Finetti like theorems for non-permutation invariant states by considering the minimal local deviation of a general state from a separable state averaged over the full system.

Bosonic systems are intrinsically permutation invariant due to the constraint of being symmetric under the exchange of particles which can be expressed as $\pi\rho = \rho$ for all per-

mutations π and bosonic states ρ. Under this stronger version of permutation symmetry, slightly more powerful de Finetti theorems can be derived [127, 128, 134] and be employed in order to justify and bound the error of the use of the discrete Gross-Pitævskii equation (see for instance [134]). The fact that we can easily bound particle mean field ansatzes, i.e. approaches based on particle product states, in bosonic systems is based on the two properties that bosonic systems feature an intrinsic symmetry on the level of particles and can contain less physical modes than particles. The latter is necessary due to the dependence of de Finetti like theorems on the local Hilbert space dimension, here the size of the single particle Hilbert space, that needs to be countered by the systems size, i.e., particle number. In fermionic systems both features are absent. It is therefore unclear and remains subject of future research if fermionic particle de Finetti theorems can be formulated. When considering a permutation invariance with respect to the exchange of single particle modes, however, we find that fermionic systems possess a clearer structure. The set of fermionic states which are invariant under the permutation of single particle modes has been characterized in two different perspectives. On the on hand on the level of a C^* algebraic description [135] and on the other hand a parametrization based on second quantization and the investigation of lattice Hamiltonians was derived [136]. In both cases a full permutation invariance with respect to the permutation of the modes is assumed which, as we will discuss in detail below, interferes with the intrinsic antisymmetry of fermionic systems. Furthermore, the previous investigations focused on the thermodynamic limit $V \to \infty$ and concrete finite system bounds in the spirit of Thm. 3 are not available. The status for the fermionic de Finetti theorems discussed above is therefore comparable to the situation for distinguishable particles before the formulation of finite size theorems by [126, 127]. In the next section we partially fill the highlighted gaps. We derive a finite size fermionic de Finetti theorem in the spirit of Thm. 3 and especially show that for fermionic systems we can use a relaxed version of permutations invariance which does not conflict with the canonical anti-commutation relations.

4.2. A Fermionic Mode de Finetti Theorem

The following result was obtained in collaboration with Zoltán Zimborás and Jens Eisert and is published in [137]. Let us first introduce the needed machinery and notation. As already pointed out above, assuming a permutation invariance in fermionic systems on the level of modes in the spirit of permutation invariance for distinguishable particles interferes with the anti-commutation relations. Consider a fermionic state ρ on V sites with p modes per site that is fully permutation invariant, i.e. $\pi\rho\pi^\dagger = \rho$ for any permutation of modes π. Denote for $j \in [V]$ and $\sigma \in [2p]$ with $m_{j,\sigma}$ a Majorana operator supported on site j with σ labeling the local modes. Then the assumption of a full permutation invariance in combination with the canonical anti-commutation relations implies that for $j_1 \neq j_2$

$$\operatorname{tr}(\rho m_{j_1,\sigma} m_{j_2,\sigma}) = \operatorname{tr}(\rho m_{j_1,\sigma} m_{j_2,\sigma}) = -\operatorname{tr}(\rho m_{j_1,\sigma} m_{j_2,\sigma}) = 0, \tag{4.5}$$

i.e. expectation values of the above form vanish to 0 and we lose a part of the anti-commuting character of fermionic states by assumption. The first important point we want to stress here is that this assumption is not necessary for fermionic systems and we obtain a de Finetti theorem for a larger class of states. We will use the following definition for fermionic permutation invariant states:

Definition 2 (Permutation invariant fermionic states)**.** *Given a fermionic system with V sites, p modes per site and Majorana operators $m_{j,\sigma}$ with $j \in [V]$ and $\sigma \in [2p]$ we call a fermionic state permutation invariant if for all $r \in [Vp]$ we have*

(1) that all expectation values are invariant under order preserving redistribution of operators over the lattice, i.e. for all $(j_1, \sigma_1) < \cdots < (j_{2r}, \sigma_{2r})$ with $j_k \in [V]$ and $\sigma_k \in [2p]$ and permutations $\pi \in S_V$ which preserve that order, i.e. $(\pi(j_1), \sigma_1) < \cdots < (\pi(j_{2r}), \sigma_{2r})$ we get

$$\mathrm{tr}(\rho m_{j_1,\sigma_1} \dots m_{j_{2r},\sigma_{2r}}) = \mathrm{tr}(\rho m_{\pi(j_1),\sigma_1} \dots m_{\pi(j_{2r}),\sigma_{2r}}), \tag{4.6}$$

(2) that locally even operators can be swapped arbitrarily meaning that for all $(j_1, \sigma_1) < \cdots < (j_{2r}, \sigma_{2r})$ with $j_k \in [V]$ and $\sigma_k \in [2p]$ that fulfill $|\{l : j_l = j\}|$ is even for all $j \in [V]$ and all permutations $\pi \in S_V$ we have

$$\mathrm{tr}(\rho m_{j_1,\sigma_1} \dots m_{j_{2r},\sigma_{2r}}) = \mathrm{tr}(\rho m_{\pi(j_1),\sigma_1} \dots m_{\pi(j_{2r}),\sigma_{2r}}). \tag{4.7}$$

By not swapping odd operators we therefore avoid the interference between permutation invariance and antisymmetry of fermionic systems. Note that Def. 2 of course contains fully permutation invariant states. Concerning the general structure of permutation invariant states let us note that any permutation invariant state is obviously also translation invariant (when using the permutation invariance of Def. 2 only for open boundary conditions). Furthermore, a permutation invariant state displays no spatial dimensionality as all its expectation values do not depend on a metric, i.e. the distance of the operators. These states feature no decay of correlations and correspond in this sense, depending on the choice of jargon, to zero or V-dimensional systems. It is also important to note that Def. 2 not only contains the case of fully permutation invariant fermionic states but goes beyond it. Assume $p = 1$, the state

$$\rho = \frac{1}{2^V} \left(\mathbb{1} + i \tan\left(\frac{\pi}{2V}\right) \mu \sum_{\substack{j,l \in [V]: \\ j<l}} m_{j,1} m_{l,1} \right) \tag{4.8}$$

for $\mu \in [-1, 1]$[1] has the non-trivial expectation values

$$\mathrm{tr}(\rho\, m_{a,1} m_{b,1}) = \begin{cases} \tan\left(\frac{\pi}{2V}\right) \mu & \text{if } a > b \\ -\tan\left(\frac{\pi}{2V}\right) \mu & \text{if } b > a \end{cases} . \tag{4.9}$$

[1]Note that the factor $i \tan(\pi/2V)$ is needed in order to ensure positivity and Hermiticity. For large system sizes V it behaves approximately like $i\pi/2V$.

Hence it is permutation invariant according to Def. 2 but fails to be fully permutation invariant.

We will see below that for states which are permutation invariant according to Def. 2 all quantum correlations vanish locally upon increasing the system size including their antisymmetric character originating from the indistinguishability of fermions. In order to prove our final de Finetti theorem we will in fact first need to show that all notions of antisymmetry are suppressed in a permutation invariant state. To do this let us introduce the parity operators for each site $j \in [V]$

$$P_j = \prod_{\sigma=1}^{p} (1 - 2f_{j,\sigma}^\dagger f_{j,\sigma}) = (-i)^p \prod_{\sigma=1}^{p} m_{j,2\sigma-1} m_{j,2\sigma}. \tag{4.10}$$

Given an operator P we define for $s = \pm$ the maps $C_P^s : \mathcal{B}(\mathcal{F}_M) \to \mathcal{B}(\mathcal{F}_M)$ with

$$C_P^s(A) = \frac{1}{2}\big(A + sPAP\big). \tag{4.11}$$

Together with the parity operators we can define the symmetric projection of a permutation invariant state. First note that the maps $C_{P_j}^s$ for $s = +/-$ erase all terms with an odd/even number of Majorana operators on site j of an operator which can be verified by computing for $j \in [V]$

$$C_{P_j}^{+/-}(m_{j,\sigma_1} \dots m_{j,\sigma_r}) = \begin{cases} m_{j,\sigma_1} \dots m_{j,\sigma_r} & \text{if } r \text{ even/odd} \\ 0 & \text{if } r \text{ odd/even} \end{cases} . \tag{4.12}$$

Furthermore, consider the map

$$C = C_{P_V}^+ \circ \dots \circ C_{P_1}^+. \tag{4.13}$$

This map constitutes a quantum channel which can be verified easily by noting that it has a Kraus decomposition with the Kraus operators $2^{-V} P_{j_1} \dots P_{j_r}$ for all $r = 0, \dots, V$ and $j_1 > \dots > j_r \in [V]$. To be precise C is the channel which erases the part of a fermionic state that is sensitive to the intrinsic antisymmetry. This can be seen by considering a general operator $A \in \mathcal{B}(\mathcal{F}_M)$ and realizing that

$$\mathrm{tr}(C(\rho)A) = \mathrm{tr}(\rho\, C(A)) \tag{4.14}$$

by the cyclicity of the trace and fact that the Kraus operators of C are real. As $C_{P_j}^+$ filters all terms from A with an odd number of Majorana operators on site j we conclude that $C(\rho)$ has for operators that are even on all sites the same expectation values as ρ and all expectation values with operators which are odd on at least one site vanish. Hence, all expectation values of $C(\rho)$ that are sensitive to the order of operators and therefore display the antisymmetry of the system are 0. Equipped with this notation we are now able to formulate, understand and prove the suppression of quantum correlations between modes in fermionic permutation invariant states.

4.2.1. Suppression of Oddness and a Fermionic de Finetti Theorem

Let us first state the two essential results, the suppression of terms sensitive to the antisymmetry of a permutation invariant state followed by a mode de Finetti theorem and discuss them and their proofs on an intuitive basis. The full proofs are collected in the subsequent section for completeness. First we obtain that if ρ is permutation invariant then it is locally indistinguishable from $C(\rho)$ in the following sense:

Lemma 4. *Given a fermionic system on $V \geq 6$ sites with p modes per site and a permutation invariant state ρ then we find that for any $k \in [V]$*

$$\| \operatorname{tr}_{[k]^c}(\rho) - \operatorname{tr}_{[k]^c}(C(\rho)) \|_1 \leq \frac{2}{\sqrt{3}} \frac{2^{2p}(k-1)^{3/2}}{V}, \tag{4.15}$$

where $C = C^+_{P_1} \circ \cdots \circ C^+_{P_V}$ with P_j being the parity operator of site $j \in [V]$ as discussed above.

The essential intuition underlying the proof is very simple. Consider an expectation value which involves at some site a single Majorana operator $m_{j,\sigma}$. By the permutation invariance we can essentially freely shift this site over the lattice (if we forget for a moment the restriction of not swapping sites with odd operators for simplicity) such that instead of calculating one expectation value we can consider the average of all terms where the $m_{j,\sigma}$ is placed on any site. However, $\sum_{j\in[V]} m_{j,\sigma}$ has an operator norm of $\sqrt{V}$ only (which can be easily realized by noting that $\sqrt{V}^{-1} \sum_{j\in[V]} m_{j,\sigma}$ results from a Fourier transformation of the single-particle basis and constitutes a valid Majorana operator). Hence, the average over an odd operator is suppressed in norm as $\sqrt{V}^{-1}$. If we expand a physical operator in the Majorana basis, all terms are either even on all sites or odd on at least two sites such that we obtain a global suppression of the order V^{-1}. The full proof, laid out in the next section exploits this observation and shows in all rigor that it is true for more general situations than the one discussed here. Technically the lemma follows then from a combinatorial argument using the permutation invariance of the state and a version of the Cauchy-Schwarz inequality.

The above is a result of the global antisymmetry constraint on the state as it allows us to construct local non-commuting operators, which are needed for such a construction. In contrast, a similar line of argumentation does not hold in systems of distinguishable particles. Here, the operator norm of the average of any local operator will be constant and given by the norm of the local operator itself. We would hence not expect to be able to formulate a de Finetti theorem for distinguishable particles based on a notion of permutation invariance in the spirit of Def. 2. It is rather the additional antisymmetry constraint of fermionic systems which allows us to relax the definition of permutation invariance.

Secondly, using the suppression above it is clear that ρ can be well approximated by $C(\rho)$ which constitutes a permutation invariant state of distinguishable particles after a Jordan-Wigner transformation. Hence, we can apply the established de Finetti theorem easily and obtain:

Theorem 5 (A fermionic de Finetti theorem)**.** *Given a fermionic system on $V \geq 6$ sites with p modes per site and a permutation invariant state $\rho \in \mathcal{D}(\mathcal{F}_{Vp})$ there exists a $r \in \mathbb{N}$ and sequences of states $\xi_1, \ldots, \xi_r \in \mathcal{D}(\mathcal{F}_p)$ and weights $a_1, \ldots, a_r \in [0,1]$ with $\sum_l a_l = 1$ such that for all $k \in [V]$*

$$\left\| \mathrm{tr}_{[k]^c} \rho - \sum_{l=1}^{r} a_l \xi_l^{\otimes k} \right\|_1 \leq 2\frac{2^{2p}k}{V} + \frac{2}{\sqrt{3}} \frac{2^{2p}\sqrt{k-1}^3}{V} \tag{4.16}$$

where $\xi_l^{\otimes k} \in \mathcal{D}(\mathcal{F}_{kp})$ denotes the state which results from copying ξ_l to each of the k sites as explained in Sec. 2.3.1.

The proof follows essentially the above laid out intuition. We apply the de Finetti theorem of distinguishable particles to the Jordan-Wigner transformed permutation invariant state $C(\rho)$ and map the result back to a fermionic formulation.

4.2.2. Proof of Lem. 4 and Thm. 5

Before we prove Lem. 4 and Thm. 5 formulated in the previous section let us introduce two technical lemmata which we will need.

First, let us note that the maps C_P^s are contractive with respect to the operator norm if $\|P\| \leq 1$, i.e., we obtain:

Lemma 6. *For two general operators A and P with $A = A^\dagger$ and $\|P\| \leq 1$ we find that for $s = \pm$ the operators*

$$C_P^s(A) = \frac{A + sPAP}{2} \tag{4.17}$$

have an operator norm that is bounded by $\|C_P^s(A)\| \leq \|A\|$.

Proof. By direct calculation we find that

$$\|C_P^s(A)\| \leq \frac{\|A\|}{2} + \frac{\|PAP\|}{2} \leq \frac{\|A\|}{2} + \frac{\|P\|\,\|A\|\,\|P\|}{2} \leq \|A\|. \tag{4.18}$$

□

Secondly, as a final ingredient we need the following consequence of the Chauchy-Schwarz inequality:

Lemma 7. *For two operators ρ and A with $\rho = \rho^\dagger$, $\rho \geq 0$ and $\mathrm{tr}(\rho) = 1$ we find*

$$\|\,\mathrm{tr}(\rho A)\|^2 \leq \mathrm{tr}(\rho A A^\dagger). \tag{4.19}$$

Proof. As ρ is Hermitian and positive, $\sqrt{\rho}$ exists and is Hermitian. We obtain then from the Cauchy-Schwarz inequality

$$\|\,\mathrm{tr}(\sqrt{\rho}^\dagger \sqrt{\rho} A)\|^2 \leq |\,\mathrm{tr}(\sqrt{\rho}^\dagger \sqrt{\rho})|\,|\,\mathrm{tr}(A^\dagger \sqrt{\rho}^\dagger \sqrt{\rho} A)|. \tag{4.20}$$

The claim follows with the normalization and positivity of ρ and the cyclic invariance of the trace. □

4.2.2.1. Proof of Lem. 4

Let us repeat out of convenience the lemma:

Lemma. *Given a fermionic system on $V \geq 6$ sites with p modes per site and a permutation invariant state ρ then we find that for any $k \in [V]$*

$$\| \operatorname{tr}_{[k]^c}(\rho) - \operatorname{tr}_{[k]^c}(C(\rho))\|_1 \leq \frac{2}{\sqrt{3}} \frac{2^{2p}(k-1)^{3/2}}{V}, \tag{4.21}$$

where $C = C^+_{P_1} \circ \cdots \circ C^+_{P_V}$ with P_j being the parity operator of site $j \in [V]$ as discussed above.

Proof. The result above is trivial for $k = 1$ as the single site reduction of ρ and $C(\rho)$ are indeed the same as by Eq. (4.14) their expectation value agree for all physical (and therefore even) on-site observables. Set therefore $k > 1$ for the following. Using Eq. (4.14) we can rewrite the one-norm difference of the two states as

$$\big\|\operatorname{tr}_{[k]^c}(\rho) - \operatorname{tr}_{[k]^c}[C(\rho)]\big\|_1 = \sup_{\substack{A:\|A\|=1, A^\dagger = A \\ \operatorname{supp} A \subset [k]}} \big|\operatorname{tr}([A - C(A)]\rho)\big|. \tag{4.22}$$

Our first goal is now to decompose a general physical operator A into different components which are then bound individually in a second step exploiting the permutation invariance of the state ρ. Given any physical Hermitian operator A with support in $[k]$ and $\|A\| = 1$ we decompose A into the operators

$$A_l = C^-_{P_l} \circ C^+_{P_{l-1}} \circ C^+_{P_{l-2}} \circ \cdots \circ C^+_{P_1}(A) \tag{4.23}$$

for $l \in [k-1]$ and

$$A_k = C^+_{P_k} \circ \cdots \circ C^+_{P_1}(A) = C(A). \tag{4.24}$$

Note that by construction

$$A = \sum_{l=1}^{k} A_l \tag{4.25}$$

as A is overall even and for instance $A_k + A_{k-1} = C^+_{P_{k-2}} \circ \cdots \circ C^+_{P_1}(A)$ giving rise to a recursive simplification of the whole sum. In addition we have by Lem. 6 that the individual norms are bounded by $\|A_l\| \leq \|A\| \leq 1$ for all $j \in [k]$. Furthermore we know that for $l \in [k-1]$, if we decompose A_l in the Majorana operator basis, the decomposition will involve only terms that contain an even number of Majorana operators supported on every site in $[l-1]$ and an odd number of Majorana operators acting on site l itself.

Next we decompose the operators A_l for $l \in [k-1]$ further in order to have more control over the resulting components. For this step it is crucial to note that the Majorana operators themselves are Hermitian operators with eigenvalues ± 1 as $m^2_{j,\sigma} = \mathbb{1}$. We therefore can decompose A_l into the components $C^\pm_{m_{l,1}}(A_l)$ with $\|C^\pm_{m_{l,1}}(A_l)\| \leq 1$. One can convince oneself that when expanded in the Majorana basis $C^-_{m_{l,1}}(A_l)$ contains only

terms which involve the Majorana operator $m_{l,1}$ whereas $C^+_{m_{l,1}}(A_l)$ is the collection of all terms of A_l that do not contain $m_{l,1}$. Both operators are then split into the four components $C^{\pm}_{m_{l,2}}(C^+_{m_{l,1}}(A_l))$ and $C^{\pm}_{m_{l,2}}(C^-_{m_{l,1}}(A_l))$ for which we know which of them involves the operator $m_{l,1}$ or $m_{l,2}$ and which does not. Iterating this for all $m_{l,\sigma}$ at fixed l we obtain a decomposition of every A_l into

$$A_l = \sum_{r=1}^{p} \sum_{1\leq\sigma_1<\cdots<\sigma_{2r-1}\leq 2p} m_{l,\sigma_1}\dots m_{l,\sigma_{2r-1}} B_{l,(\sigma_h)_{h\in[2r-1]}} \tag{4.26}$$

with every $B_{l,(\sigma_h)_{h\in[2r-1]}}$ having a norm bounded by 1 and being even on the sites $1,\dots,l-1$, acting trivially on site l and being an overall odd operator.

In the next step we exploit the permutation invariance of the state ρ in order to show that every term of the decomposition of A_l in Eq. (4.26) is suppressed roughly as $\sqrt{k}/V$ in the system size. For this we assume $(k-1)\leq V/2$ and introduce a special subset of the permutations of the sites of the lattice. As the technical details will look rather convoluted let us discuss them first. For a fixed $l\in[k-1]$ we will decompose the set of available sites $[V]$ into three components, a left, a right and a central one where all sites in the left part have a smaller index than those in the central part and all in the central part have a smaller index than those in the right part. The permutations will then permute the site l into the central part of the lattice which has a size of about $V/2$. The sites $1,\dots,l-1$ are permuted as a block into the left part whereas the sites $l+1,\dots,k$ are permuted into the right part without changing the relative order of any pair of sites. In addition, consecutive sites will stay consecutive, expect of site l. The position of the site l permuted into the central part and position of the sites permuted to the other two is uncorrelated. Furthermore, the position of the sites permuted into the left part is uniquely defined by the ones that are permuted to the right part and for different permutations the sites permuted to the left and right part are either all permuted to the same sites or to disjoint sets. To make this concrete let us denote for $i,j\in[V]$ by $\tau_i^j\in S_V$ the swap of site i and j. In addition we introduce the abbreviations $n_k=\lfloor V/2(k-1)\rfloor$, for $x\in[n_k]$ $b_x^l = V-(x-1)(k-l)$, $c_x^l = n_k(l-1)-(x-1)(l-1)$ and $S_c^l=\{j\in[V]|n_k(l-1)<j\leq V-n_k(k-l)\}$. We then define for every $l\in[k-1]$, $a\in S_c^l$ and $x\in[n_k]$ the permutations $\pi^l_{a,x},\pi^l_x\in S_V$ by

$$\pi^l_{a,x} = \tau^1_{c_x^l-l+2}\circ\cdots\circ\tau^{l-1}_{c_x^l}\circ\tau^l_a\circ\tau^{l+1}_{b_x^l-k+l+1}\circ\cdots\circ\tau^k_{b_x^l} \tag{4.27}$$

$$\pi^l_{x} = \tau^1_{c_x^l-l+2}\circ\cdots\circ\tau^{l-1}_{c_x^l}\circ\tau^{l+1}_{b_x^l-k+l+1}\circ\cdots\circ\tau^k_{b_x^l}. \tag{4.28}$$

A visualization of the permutations is displayed in Fig. 4.1. Given that for any $l\in[k-1]$, $a\in S_c^l$ and $x\in[n_k]$ the permutation $\pi^l_{a,x}$ does not change the relative order of any site in $[k]$ we obtain from the permutation invariance of ρ that for any $r\in[p]$ and $1\leq\sigma_1<\cdots<\sigma_{2r-1}\leq 2p$

$$\operatorname{tr}\left(\rho\, m_{l,\sigma_1}\dots m_{l,\sigma_{2r-1}}B_{l,(\sigma_h)_{h\in[2r-1]}}\right) = \operatorname{tr}\left(\rho\frac{1}{|S_c^l|n_k}\sum_{a\in S_c^l}\sum_{x=1}^{n_k}\pi^l_{a,x}\left[m_{l,\vec{\sigma}}B_{l,\vec{\sigma}}\right]\right), \tag{4.29}$$

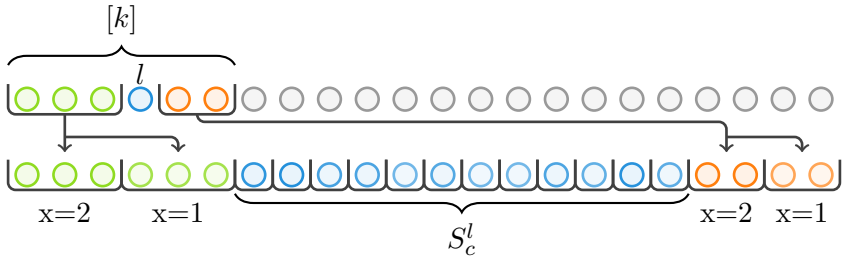

Figure 4.1. – Illustration of the permutations needed in the proof of lemma 4 for $V = 22$, $k = 6$ and $l = 4$. The above chain illustrates the starting condition where in green, blue and orange we highlight the sites $[l-1]$, l and the rest of $[k]$ respectively. Any permutation $\pi^l_{a,x}$ will then permute the system to one of the configurations depicted in the lower chain. The site l is permuted into the set S^l_c highlighted in blue, the sites to the left and right of l are permuted, depending on the choice of x, to the green and orange parts in the corresponding bins. Note that x defines the position for both parts and the target sites form disjoint sets for different x.

where we introduced the shorthand notation $m_{l,\vec{\sigma}} = m_{l,\sigma_1} \dots m_{l,\sigma_{2r-1}}$ for vectors of modes $\vec{\sigma} = (\sigma_h)_{h\in[2r-1]}$. Using lemma 7 we then directly obtain

$$\left|\mathrm{tr}\left(\rho\, m_{l,\vec{\sigma}} B_{l,\vec{\sigma}}\right)\right|^2 \leq \mathrm{tr}\left(\rho \frac{1}{|S^l_c|^2 n_k^2} \sum_{a,b\in S^l_c} \sum_{x,y=1}^{n_k} \pi^l_{a,x}\left[m_{l,\vec{\sigma}} B_{l,\vec{\sigma}}\right]\left(\pi^l_{b,y}\left[m_{l,\vec{\sigma}} B_{l,\vec{\sigma}}\right]\right)^\dagger\right) \tag{4.30}$$

$$= \mathrm{tr}\left(\rho \frac{1}{|S^l_c|^2 n_k^2} \sum_{a,b\in S^l_c} \sum_{x,y=1}^{n_k} m_{a,\vec{\sigma}} \pi^l_x\left[B_{l,\vec{\sigma}}\right] \pi^l_y\left[B_{l,\vec{\sigma}}{}^\dagger\right] m^\dagger_{b,\vec{\sigma}}\right). \tag{4.31}$$

The sum over x and y is symmetric where for $x \neq y$ the operators $\pi^l_x\left[B_{l,\vec{\sigma}}\right]$ and $\pi^l_y\left[B_{l,\vec{\sigma}}{}^\dagger\right]$ anti-commute due to the disjoint support and oddness of $B_{l,\vec{\sigma}}$. A similar argument holds for the sum over a and b which yields that only terms with $a = b$ and $x = y$ lead to a non vanishing contribution. Using that ρ is permutation invariant and that $\|m_{l,\vec{\sigma}} B_{l,\vec{\sigma}} B_{l,\vec{\sigma}}{}^\dagger m^\dagger_{l,\vec{\sigma}}\| \leq \|A\|^2 \leq 1$ we then conclude

$$\left|\mathrm{tr}\left(\rho\, m_{l,\vec{\sigma}} B_{l,\vec{\sigma}}\right)\right|^2 \leq \frac{1}{|S^l_c| n_k}. \tag{4.32}$$

where $n_k \approx V/2(k-1)$ and $|S^l_c| \approx V/2$. Applying this bound for all $l \in [k-1]$ to each of the $2^{2p}/2$ many terms in Eq. (4.26), we obtain

$$\|\,\mathrm{tr}_{[k]^c}(\rho) - \mathrm{tr}_{[k]^c}(C(\rho))\|_1 \leq \frac{2^{2p}(k-1)}{2\sqrt{|S^l_c| n_k}}, \tag{4.33}$$

as the $C(A)$ term in Eq. (4.22) is canceled by the A_k contribution.

In order to simplify the bound consider $k-1 \leq V/2$. In addition we need that for any $n \in \mathbb{N}$ the relative error of the floor function can be bounded as

$$\max_{w \in [n]} \frac{\frac{n}{w} - \left\lfloor \frac{n}{w} \right\rfloor}{\frac{n}{w}} \leq \frac{1}{2} \tag{4.34}$$

as the maximizer of the error is given by $w = \lfloor n/2 \rfloor + 1$. We obtain then

$$\frac{1}{|S_c^l| n_k} = \frac{1}{\frac{V^2}{2(k-1)}} \frac{1}{\frac{\left\lfloor \frac{V}{2(k-1)} \right\rfloor}{\frac{V}{2(k-1)}} - \frac{1}{2} \frac{\left\lfloor \frac{V}{2(k-1)} \right\rfloor}{\frac{V}{2(k-1)}} \frac{\left\lfloor \frac{V}{2(k-1)} \right\rfloor}{\frac{V}{2(k-1)}}} \leq \frac{2(k-1)}{V^2} \frac{1}{\frac{1}{2} - \frac{1}{2} \left(\max_{w \in [V]} \frac{\frac{V}{w} - \left\lfloor \frac{V}{w} \right\rfloor}{\frac{V}{w}} \right)^2} \tag{4.35}$$

$$\leq \frac{4(k-1)}{V^2} \frac{4}{3} \tag{4.36}$$

which yields the bound in the theorem. As the bound yields a value larger 2 for $k-1 > V/2$ for all $V \geq 6$ it applies trivially to all k. □

4.2.2.2. Proof of Thm. 5

Let us again, out of convenience, repeat the theorem:

Theorem (A fermionic de Finetti theorem)**.** *Given a fermionic system on $V \geq 6$ sites with p modes per site and a permutation invariant state $\rho \in \mathcal{D}(\mathcal{F}_{Vp})$ there exists a $r \in \mathbb{N}$ and sequences of states $\xi_1, \ldots, \xi_r \in \mathcal{D}(\mathcal{F}_p)$ and weights $a_1, \ldots, a_r \in [0,1]$ with $\sum_l a_l = 1$ such that for all $k \in [V]$*

$$\left\| \mathrm{tr}_{[k]^c} \rho - \sum_{l=1}^{r} a_l \xi_l^{\otimes k} \right\|_1 \leq 2 \frac{2^{2p} k}{V} + \frac{2}{\sqrt{3}} \frac{2^{2p} \sqrt{k-1}^3}{V} \tag{4.37}$$

where $\xi_l^{\otimes k} \in \mathcal{D}(\mathcal{F}_{kp})$ denotes the state which results from copying ξ_l to each of the k sites as explained in Sec. 2.3.1.

Proof. The proof of the theorem is now rather straightforward. By Lem. 1 we know that

$$\left\| \mathrm{tr}_{[k]^c} \rho - \mathrm{tr}_{[k]^c} C(\rho) \right\|_1 \leq \frac{2}{\sqrt{3}} \frac{2^{2p} \sqrt{k-1}^3}{V}. \tag{4.38}$$

In addition it is clear that $C(\rho)$ is a state that is locally even. Therefore it is fully permutation invariant by condition (2) of Def. 2 as all expectation values with operators that do not act as even operator on all sites vanish. Therefore $C(\rho)$ gives rise to a permutation invariant state on $(\mathbb{C}^{2^p})^{\otimes V}$ by a blocked Jordan-Wigner transformation. Applying the de Finetti theorem for distinguishable quantum systems yields immediately that there is a $r \in \mathbb{N}$, $a_1, \ldots, a_r \in [0,1]$ with $\sum_l a_l = 1$ and $\tilde{\xi}_1, \ldots, \tilde{\xi}_r \in \mathcal{D}(\mathbb{C}^{2^p})$ such that

$$\left\| \mathrm{tr}_{[k]^c} C(\rho) - \sum_{l=1}^{r} a_l \tilde{\xi}_l^{\otimes k} \right\|_1 \leq 2 \frac{2^{2p} k}{V}. \tag{4.39}$$

The states $\tilde{\xi}_l$ might under the Jordan-Wigner transformation not transform back to fermionic states as they could involve odd parts violating the superselection rule. However we can use a channel that is the reduction of the Jordan-Wigner transformed version of the channel C in order to project out these parts. For this, define the channel $\tilde{C}_k = C^+_{(Z^{\otimes p})_k} \circ \cdots \circ C^+_{(Z^{\otimes p})_1}$. Observe that $\tilde{C}(\mathrm{tr}_{[k]^c} C(\rho)) = \mathrm{tr}_{[k]^c} C(\rho)$ as under the Jordan-Wigner transformation we obtain $\tilde{C}_k = C^+_{P_k} \circ \cdots \circ C^+_{P_1}$. The contractiveness of the one-norm distance under channels gives us then directly

$$\left\| \mathrm{tr}_{[k]^c} C(\rho) - \sum_{l=1}^{r} a_l \tilde{C}(\tilde{\xi}_l^{\otimes k}) \right\|_1 \leq 2\frac{2^{2p}k}{V}. \tag{4.40}$$

Defining $\xi_l = C^+_{Z^{\otimes p}}(\tilde{\xi}_l)$ yields then the result. □

4.3. Structure of the Mode Product States and Applications

The mode product states and convex combinations thereof encountered in Thm. 5 are on first sight not very natural states for fermionic systems and especially in view of bounding mean field approaches such as the Hartree-Fock approximation we are interested in their underlying structure and potential connection to Gaussian states. It turns out that in certain limiting cases we can indeed connect mode product states to Gaussian states as we explain in the following. During this discussion we highlight two implications of the mode de Finetti theorem: The control of mean field approximations in special cases and the extension of the applicability of a fermionic central limit theorem which captures the convergence of the collection of i.i.d. copies of a state towards a Gaussian state.

4.3.1. Few Local Modes and Mean Field Approximations

First note that in view of the tensor product representation of quantum states, a mode product state corresponds to a matrix product operator of bond dimension one. Therefore part of our discussion directly applies to this set of states. The product states of Thm. 5 are of course special by their additional i.i.d. structure.

The first limiting case we want to discuss concerns the number of modes per sites p. Given a fermionic permutation invariant state ρ with corresponding r, states $\xi_l \in \mathcal{D}(\mathcal{F}_p)$ and weights $a_l \in [0,1]$ resulting from Thm. 5 we want to denote by $\rho^{(k)}$ the mode separable state $\rho^{(k)} = \sum_{l=1}^{r} a_l \xi_l^{\otimes k}$. Then, in the extreme case of $p = 1$ we obtain that all ξ_l are of the form $\xi_l = b_l|0\rangle\langle 0| + (1-b_l)|1\rangle\langle 1|$ for some $b_l \in [0,1]$ according to the parity superselection rule. Then $\rho^{(k)}$ takes the form

$$\rho^{(k)} = \sum_{i_1,\ldots,i_k=0}^{1} c^{(k)}_{i_1,\ldots,i_j} |i_1,\ldots,i_k\rangle\langle i_1,\ldots,i_k| \tag{4.41}$$

with $\sum c^{(k)}_{i_1,\ldots,i_j} = 1$ and $c^{(k)}_{i_1,\ldots,i_j} \geq 0$. Put differently, $\rho^{(k)}$ is the convex combination of Slater determinants. More general, for $p \leq 3$ we have that the ξ_l are necessarily convex

combinations of pure Gaussian states [138]. This property is then directly inherited by $\rho^{(k)}$.

This has the following implication on mean field approximations. Similar to the case of distinguishable particles, the mode de Finetti theorem allows us to bound the error of a ground state energy approximation using mode product states. To be precise let us denote by P a collection of subsets of maximal size k of the full system $[V]$. Consider a Hamiltonian of the form

$$H = \frac{1}{|P|} \sum_{S \in P} h_S, \tag{4.42}$$

with h_S being supported non-trivially on S only and being bounded by $\|h_S\| \leq 1$. Assuming a permutation invariant ground state space projector we can approximate the error of a mean field approximation along the lines of Eq. (4.4) using Thm. 5 and obtain

$$e_{\mathrm{GS}} - \min_{\xi \in \mathcal{D}(\mathcal{F}_p)} \mathrm{tr}(H \xi^{\otimes V}) \leq 2 \frac{2^{2p} k}{V} + \frac{2}{\sqrt{3}} \frac{2^{2p} \sqrt{k-1}^3}{V}. \tag{4.43}$$

For $p \leq 3$, $\xi^{\otimes V}$ will be the convex combination of pure Gaussian states such that we can bound by the linearity of the expectation value

$$e_{\mathrm{GS}} - \min_{\substack{\rho \in \mathcal{D}(\mathcal{F}_{pV}): \\ \rho \,\text{pure Gaussian}}} \mathrm{tr}(H \rho) \leq 2 \frac{2^{2p} k}{V} + \frac{2}{\sqrt{3}} \frac{2^{2p} \sqrt{k-1}^3}{V}. \tag{4.44}$$

Hence, in these systems we are able to bound the error of using a Hartree-Fock approximation in order to obtain the ground state energy. Note however, that this result and example has more of an illustrative character than being of practical physical interest. It shows that an insight into the correlation structure of fermionic state in the spirit of the mode de Finetti theorem does allow for bounding the error of mean field approximations based on mode or particle product states in special cases. The specific assumptions made on the setting are, however, more or less unnatural. It has for instance been argued that the normalization assumed in Eq. (4.42) leads to an unwanted suppression of terms in certain settings in fermionic systems [136] and more importantly the assumption of having a permutation invariant ground state in the first place is very restrictive and usually not matched in realistic systems. The result currently has to be seen as supporting the basic intuition that mean field type approaches are well justified in settings of high connectivity, i.e., systems in which everyone is coupled to everyone, and by this reveals one of the sufficient ingredients needed for mean field approaches. In order to bridge the gap between this intuition and realistic examples, generalizations of the above result in the spirit of [133] are needed, which are subject of ongoing research.

4.3.2. Extension of Hudson's Theorem

The second important limiting case we want to discuss concerns the size of the subsystem. For $k = 1$ the resulting mode product states can be any fermionic state $\xi \in \mathcal{D}(\mathcal{F}_p)$

which can be trivially seen from the fact that for any $\xi \in \mathcal{D}(\mathcal{F}_p)$ the state $\rho = \xi^{\otimes V}$ is permutation invariant and results in $\rho^{(1)} = \xi$. More interesting is the case of large subsystems. Here we obtain that mode product states behave like Gaussian states if we probe them only with modes that are averaged over the full subsystem. This can be made rigorous by an extension of Hudson's central limit theorem which shows in its basic version that the restriction to the 0 Fourier mode of states of the form $\rho^{\otimes k}$ converges for $k \to \infty$ to a Gaussian state [139]. This essentially means that for a state $\rho^{\otimes k}$ its moments with respect to the operators $1/\sqrt{k}\sum_{j=1}^{k} f_{j,\sigma}$ are the moments of a Gaussian state, i.e. decompose according to Wick's theorem up to an error decaying in k. Hence, if we probe the whole system equally weighted over large regions, we expect mode product states to appear Gaussian.

In order to make this notion more precise and to capture the convergence to a Gaussian state more conveniently we introduce the notation of cumulants. Given a fermionic system with M modes, Majorana operators m_j with $j \in [2M]$ and a state $\rho \in \mathcal{D}(\mathcal{F}_M)$ we define for $j_1, \ldots, j_w \in [2M]$ with w even the cumulants $K_w^\rho(m_{j_1}, \ldots, m_{j_w})$ via

$$\mathrm{tr}(m_{j_1} \ldots m_{j_w}\rho) = \sum_{P \in \mathcal{P}^e([w])} \mathrm{sign(P)} \prod_{p \in P} K_{|p|}^\rho((m_l)_{l \in p}), \tag{4.45}$$

where $\mathcal{P}^e([w])$ denotes the set of partitions of $[w]$ into parts of even size (recall Sec. 2.1.2). Note that the definition of the cumulant is multi-linear in the operators, which can be proven easily using a recursion, meaning that for instance

$$K_w^\rho\left(\sum_{k=1}^{2M} A_{k,j} m_j, m_{j_2}, \ldots, m_{j_w}\right) = \sum_{j=1}^{2M} A_{k,j} K_r^\rho(m_j, m_{j_2}, \ldots, m_{j_w}), \tag{4.46}$$

which means that the definition above is invariant under mode transformations and can also be applied to fermionic creation and annihilation operators as they are linear combinations of Majorana modes. For a Gaussian state ρ, all cumulants except of the second ones are zero as one would expect. This can be seen the following. For $w = 4, 6, 8, \ldots, 2M$ insert Wick's theorem in order to obtain

$$\mathrm{tr}(m_{j_1} \ldots m_{j_w}\rho) = \sum_{P \in \mathcal{P}_2([w])} \mathrm{sign(P)} \prod_{p \in P} K_{|p|}^\rho((m_l)_{l \in p}), \tag{4.47}$$

where we used that $K_2^\rho(m_{j_1}, m_{j_2}) = \mathrm{tr}(m_{j_1} m_{j_2}\rho)$. This cancels in the definition of the higher cumulants all terms which involve second moments only and shows upon iteration that the remaining part, the higher cumulants, are zero.

The established fermionic central limit theorem developed by Hudson [139] concerns in its original formulation only the Fourier 0-mode of a mode product state. Below we argue that this behavior can in fact be extended to all Fourier modes. We then obtain that the mode product state converges to a product of decoupled copies of Gaussian states, each of which is supported on conjugate pairs of momenta only. Let us note that if not the discrete Fourier transformation but a tensor product of Hadamard gates is used, the state can be fully decoupled such that ρ^k converges to $\rho_0^{\otimes k}$ with ρ_0 being Gaussian (this result is essentially contained in [140]).

Lemma 8 (Extension of Hudson's central limit theorem)**.** *Assume a fermionic system with V sites, p modes per site and a state $\rho \in \mathcal{D}(\mathcal{F}_p)$ on a single site. The product state $\rho^{\otimes V} \in \mathcal{D}(\mathcal{F}_{Vp})$ is then approximately described by a product $\bigotimes_q \rho_q$ with ρ_q Gaussian states to be specified in the following sense.*

Denote by $f^\dagger_{j,\sigma}$ and $f_{j,\sigma}$ with $j \in [V]$ and $\sigma \in [p]$ the creation and annihilation operators of the systems and introduce the shorthand notation $f^c_{j,\sigma}$ with $c = \pm 1$ and $f^{-1}_{j,\sigma} = f^\dagger_{j,\sigma}$ and $f^1_{j,\sigma} = f_{j,\sigma}$. We then introduce the Fourier modes of the lattice

$$a^c_{q,\sigma} = \frac{1}{\sqrt{V}} \sum_{j=1}^{V} e^{\frac{2\pi i}{V} jq} f^c_{j,\sigma} \tag{4.48}$$

for $q = -\lfloor (V-1)/2 \rfloor, \ldots, \lfloor V/2 \rfloor$. Furthermore, we denote by f^c_σ for $\sigma \in [p]$ and $c = \pm 1$ the creation and annihilation operators of a fermionic system with p modes.

We then find for $r \in [Vp]$, and triples $(q_1, c_1, \sigma_1) < \cdots < (q_{2r}, c_{2r}, \sigma_{2r})$ with $\sigma_l \in [p]$, $c_l = \pm 1$ and q_l as above that the cumulants of the Fourier modes with respect to the state $\rho^{\otimes V}$ are suppressed as

$$K^{\rho^{\otimes V}}_{2r}(a^{c_1}_{q_1,\sigma_1}, \ldots, a^{c_{2r}}_{q_{2r},\sigma_{2r}}) = \frac{1}{V^{r-1}} K^{\rho}_{2r}(f^{c_1}_{\sigma_1}, \ldots, f^{c_{2r}}_{\sigma_{2r}}) \delta_{\sum_{l=1}^{2r} c_l q_l \bmod V, 0} \tag{4.49}$$

Meaning that in the Fourier modes $\rho^{\otimes V}$ has up to an error decreasing in V the same moments as $\rho_0 \otimes \rho_{V/2} \otimes \bigotimes_{q=1}^{\lfloor (V-1)/2 \rfloor} \rho_{q,-q}$ with $\rho_0 = \rho_{V/2}$ and $\rho_{q,-q} = \rho_{q',-q'}$ all Gaussian and defined by the second moments of ρ where $\rho_{V/2}$ is only present if V is even.

Before we comment on the implication of Lem. 8 let us discuss the basic structure of the proof which can be found with all details in App. E. The result that the product state $\rho^{\otimes V}$ is Gaussian in the Fourier modes is rather intuitive. Consider a product of $2r$ fermionic operators in the Fourier basis. Expressed in the original operators this yields a sum over terms in which the original operators are distributed over the different copies of the system. Due to the state being an even operator we find that the leading contribution is given by terms with two or zero operators on each copy. Here the prefactor V^{-r} cancels with the sum over the r different copies the pairs are supported on and the resulting contribution is the same as for the Gaussified version of ρ as only second moments are probed. The next contribution comes from terms in which next to pairs of operators one copy supports four operators. Those terms are then suppressed by V^{-1} compared to the leading contribution as we have one less copy supporting operators to sum over when counting the number of such contributions, meaning that non-Gaussian correlation will be suppressed as V^{-1}. That the state decouples in the Fourier basis should also be not surprising due to the translational invariance of $\rho^{\otimes V}$. That the internal degrees of freedom and coefficients from the Fourier transform do not disturb this intuition is shown in App. E by proving the above lemma rigorously.

From Lem. 8 we can derive the intuition that if we probe $\rho^{\otimes k}$ in modes which are evenly smeared over all copies, the resulting moments will look very close to the ones of the Gaussified version of ρ – probing single copies on the other hand might of course reveal

non-Gaussian correlations as we could then restrict $\rho^{\otimes k}$ to fewer copies and increase the error in Lem. 8. In view of the structure of i.i.d. mode product states of Thm. 5 Lem. 8 implies that if we probe operators that can be written using a few Fourier mode operators of Lem. 8 only, then they will appear to be Gaussian up to an error that decays in the subsystem size as k^{-1}.

Hence, combining Thm. 5 and Lem. 8 yields a generalization of a result which is typical for de Finetti type theorems. Lem. 8 is restricted to i.i.d. product states. Using however the mode de Finetti theorem we conclude that its applicability can be easily extended to more general, namely permutation invariant, states. Combining Thm. 5 and Lem. 8 yields the immediate corollary that if a state $\rho \in \mathcal{D}(\mathcal{F}_{pV})$ is the reduction of a permutation invariant state on a much larger system (i.e. if ρ is extendable permutation invariant for some larger finite system size) then it appears to be the convex combination of Gaussian states if probed by the Fourier modes.

4.4. Summary

In this chapter we discussed the presence of a simple correlation structure imposed by symmetry in fermionic systems. We obtained for finite fermionic systems the mode de Finetti theorem in Thm. 5 much in the spirit of the case of distinguishable particles (compare Thm. 3). From the symmetry of a fermionic state under permutations of the single-particle modes we derived that all quantum correlations are suppressed with growing system size if we restrict our attention to local observables and combinations thereof. This includes in particular the anti-commuting character of the fermions which originates from their indistinguishability. Let us emphasize that local here does not refer to spatially local, a notion which is almost obsolete in the considered systems due to the permutation invariance, but only to being local in the sense of being supported on a limited number of modes. We extend previous results by considering finite size effects and more importantly by extending here the class of states considered by using a definition of permutation invariance tailored to fermionic systems which does not interfere with the antisymmetry of fermionic states. By this we are able to analyze in detail the suppression of the anti-commuting character which results in the possibility of employing simplifying descriptions of the corresponding quantum states.

We connect the result to two applications. First, we bound mean field approaches either based on product states in general or in special cases on pure Gaussian states to appropriate Hamiltonians. This allows to link our result in specific instances to the Hartree-Fock approximation and to find classes of models which are faithfully approximated within it. Furthermore, we extend the class of states to which we can apply structure theorems which are formulated for identical copies of fermionic state such as the fermionic central limit and extensions thereof. The combination of such theorems with versions of the quantum de Finetti theorem yields a generalization of these structural insights.

Note however that next to the implication on mean field approximations and the extensions of theorems such as the fermionic central limit theorem, a significant share of

the importance of Thm. 5 relies at the moment on technical aspects. The fact that a finite size de Finetti theorem can be formulated for fermionic system without demanding a full permutation invariance but from a symmetry restriction which does not conflict with the canonical anti-commutation relations is expected to be of even bigger importance when we aim to generalize the result to more generic settings. Direct implication for physical systems, i.e. bounding mean field approaches to realistic systems, need these generalizations of the above theorem. Either in the spirit of [133] based on the structure of the interaction graph of the system or generalizations to the particle picture by formulating a truly fermionic particle de Finetti theorem, based on antisymmetry instead of symmetry. Both ambitious goals will need and rely on an in depth discussion of the interplay of antisymmetry and a generalized notation of permutation invariance for which we view Thm. 5 as a first step.

5. Gaussification: Relaxation under Free Hamiltonians

One of the most general physical concepts which is present on all scales and touches essentially every branch of physics is the phenomenon of equilibration and thermalization. It is generically observed that physical systems, no matter if they are captured by classical or quantum mechanics, tend towards an equilibrium configuration. Furthermore, this equilibrium can often be well understood in terms of thermal ensembles which depend only on a few parameters. Very different initial conditions thermalize to the same configuration such that almost all memory on the specific initial condition is lost. We find the dynamical emergence of effective descriptions which, in the extreme case, only depend on a global energy constraint.

Modern experimental techniques allow to probe the relaxation or non-equilibrium processes in well controlled quantum systems [117, 141–145] and are able to resolve some of the details of equilibration and thermalization of closed quantum systems. Understanding however the emergence of these thermal ensembles based on a microscopic theory, is up to today, an open and challenging problem and the hope to obtain a derivation of statistical mechanics based on the principles of quantum mechanics is so far unmatched. Even worse, in the context of quantum mechanics further conceptual difficulties appear. The unitarity of the time evolution in closed quantum systems leads to an enormous amount of constants of motion which strongly restricts the global dynamics and the emergence of equilibrium ensembles becomes unclear on the first sight. We discuss approaches on how to resolve this apparent contradiction below. Moreover, although a certain understanding of the equilibration of closed quantum systems has been gained in recent years, we are not able to resolve this process properly in time. On a rigorous level, we are often only able to predict that a system will indeed equilibrate but can not control the time scale on which this happens such that the stability of the obtained predictions upon taking the thermodynamic limit is not guaranteed.

In order to advance with the above questions we investigate in this chapter relaxation processes in a restricted set of models. Considering the dynamics governed by free Hamiltonians we prove that in non-interacting systems a peculiar relaxation process appears. We find that after a short, meaning independent of the system size, relaxation time generic initial states will be locally indistinguishable from Gaussian states. In physical terms, the result is based on the assumption that the system supports a sufficient form of transport and that the initial states contains only appropriately decaying correlations. The Gaussification of free systems has two major implications. First, it explains how Gaussian states appear naturally and that in the description of the physics of free systems we can often safely restrict ourselves to the fully Gaussian setting. Secondly,

although the Gaussian state we find the system to converge to is in general time dependent, the result is reminiscent to the convergence towards a generalized Gibbs ensemble (GGE) [146] as we converge in any point in time to the maximum entropy state. If in addition the system equilibrates, as it will happen for instance in fully translation invariant settings, we obtain by this an actual convergence towards a GGE and hence explain the thermalization of a large class of free systems on controlled time scales.

Below we will first review different concepts for the description of equilibration and thermalization processes in closed quantum systems. Subsequently, we restrict ourselves to the setting of free fermionic systems and discuss the conditions under which we find these systems to Gaussify. After introducing the assumptions and presenting the intuition of the proof we conclude by discussing the physical implications in more detail.

5.1. Equilibration and Thermalization of Closed Quantum Systems

Let us first briefly review the current understanding of generic relaxation processes in quantum many-body systems obtained in recent years. We will restrict ourselves to the important concepts and results and in doing so we of course only scratch the surface of a rich and wide field. For very careful and formal reviews of the equilibration theory of closed quantum systems, which give also a more detailed account its historical development, we refer to [147, 148].

As already highlighted above, the unitarity of the time evolution of closed quantum mechanical systems poses a challenge towards explaining phenomena such as equilibration and thermalization in these system. In a global view, the time evolution generated from a time-independent Hamiltonian just leads to the rotation of a quantum state in the full Hilbert space with a constant rate of change. Hence, on the first sight it is unclear how this observation can be linked to our daily experience of finding systems in equilibrium.

Furthermore, from practical experience we know that the equilibrium configuration of a large physical system can conveniently be captured in terms of a thermodynamic ensemble which depends only on very few parameters, in the extreme case only on the global energy expectation value which fixes the temperature of the system. The notion of thermalization is even more surprising in the context of closed quantum systems. Any quantum system evolved under a time-independent Hamiltonian possesses for instance a number of independent integrals of motions which grows exponentially in the system size as the overlap of the initial state with every eigenspace of the systems Hamiltonian is constant in time. In addition, the unitary time evolution does not alter the global entropy of a quantum system, e.g. a pure initial state will stay pure at all times. Hence, common notions of ergodicity [147, 149] do not apply to quantum systems and have to be revised.

One solution to these problems is based on a change of perspective. Instead of considering the global system, the notion of equilibration and thermalization can be understood when considering the properties of finite subsystems instead of the global ones; a route

that also allows to overcome the ergodicity hypothesis in classical systems [149, 150]. Such a restriction is also encouraged by the fact that many quantities observable for closed quantum systems are composed of local parts. The total particle number is the sum of all local ones, currents are either measured as change of particle numbers or as local fluxes through a region etc. It is then at least conceivable that a system effectively serves as its own heat bath such that thermalization can generically occur on the level of subsystems, not interfering with the global constraints pointed out above.

To make this notion more precise let us introduce the concept of equilibration on average. Given a time-dependent signal $s(t)$ obtained from a physical system, e.g. the expectation value of an observable with respect to a time-dependent state, we are interested in its rough behavior over time. For this we consider its time average

$$\overline{s}^T = \frac{1}{T}\int_0^T s(t)dt \tag{5.1}$$

and define the infinite time average as $\overline{s}^\infty = \lim_{T\to\infty} \overline{s}^T$. Consider then an initial state $\rho = \rho(0)$ which is evolved under some Hamiltonian $H = \sum_k E_k P_k$, where E_k and P_k denote the eigenenergies of H and projectors on the eigenspaces respectively. One finds then, that the infinite time average of the expectation value of any observable A is captured by the expectation of A with respect to the state

$$\overline{\rho}^\infty = \lim_{T\to\infty}\int_0^T \rho(t)dt = \sum_k P_k \rho P_k, \tag{5.2}$$

meaning that $\overline{\mathrm{tr}(\rho(t)A)}^\infty = \mathrm{tr}(\overline{\rho}^\infty A)$ as can be verified easily by observing that any factor of the form $e^{it(E_k - E_{k'})}$ averages to zero. Assuming that all eigenenergies are non-degenerate such that $P_k = |k\rangle\langle k|$ one finds the simplified form $\overline{\rho}^\infty = \sum_k \langle k|\rho|k\rangle\ |k\rangle\langle k|$, i.e. $\overline{\rho}^\infty$ is diagonal in the eigenbasis of H with the same diagonal elements as ρ. Note that by this, $\overline{\rho}^\infty$ has the same overlaps with the eigenspaces of H and hence especially the same energy expectation value as the initial state ρ. In addition, one can show that $\overline{\rho}^\infty$ is the state which maximizes the von Neumann entropy, given these constraints (see for instance [147, Thm. 2.3.1]).

So far we have not argued that any signal received from a physical system might indeed equilibrate as its average could of course be a rather bad approximation of its overall behavior. However, one finds the general result that for any observable A given a Hamiltonian H with non-degenerate energy gaps, and an initial state ρ that has a non-vanishing overlap with a large number of eigenstates of H then the (infinite) time average is indeed a good description of the overall behavior of $\mathrm{tr}(A\rho(t))$ in the sense that the average deviation $\overline{[\mathrm{tr}(A\rho(t)) - \mathrm{tr}(A\overline{\rho}^\infty)]^2}^\infty$ is small [151, 152]. Note that the assumption of having non-degenerate energy gaps and overlap with many eigenstates of H is quite intuitive as it avoids the build up of resonances in the $e^{it(E_k - E_{k'})}$ factors and ensures that enough of them participate in order to effectively dephase the state. However,

the requirement of non-degenerate energy gaps excludes structured systems such as free models as here any energy gap will have a degeneracy which grows exponentially in the system size. Using the operational definition of the trace norm one finds that $\rho(t)$ and $\overline{\rho}^{\infty}$ are on average locally indistinguishable meaning that for any local region S of the system $\overline{\|\operatorname{tr}_{S^c}\rho(t)-\operatorname{tr}_{S^c}\overline{\rho}^{\infty}\|_1}^{\infty}$ is small. The error decreases algebraically with $(\sum_k\langle k|\rho|k\rangle^2)^{-1}$, i.e. the number of eigenstates $|k\rangle$ of H the initial state ρ has a significant overlap with, and increases exponentially with the size of the region S [152]. It is however important to note that by proving the smallness of the fluctuations of expectation values around their infinite time average we indeed find that the infinite time average does describe the configuration of the system at most times but obviously we loose any temporal information in the sense that we cannot say when a system will first look locally equilibrated. Furthermore, note that in any finite system, quantum or classical, versions of Poincaré's recurrence theorem enforce that a system leaves it equilibrium configuration as it has to return arbitrary closely to its initial values after a large but finite time [153] such that the equilibration has to set in well before that time. Moreover, a closer look at the details of the bound on the fluctuations around the time average leads to an equilibration time that grows exponentially with the system size. Hence, the relaxation times obtained from such an argumentation quickly diverge with the system size such that no statements can be made for the thermodynamic limit. One can argue that under a further smoothness assumption on the system, one would expect finite equilibration times in the infinite system based on the intuition of dephasing [154, 155] but the precise assumptions on H, A and $\rho(0)$ to be made are unclear. Note that in general, the assumptions leading to equilibration laid out above are rather technical in the sense that they can not be related to physical properties of the system which can be checked within an experimental setup, such as correlation properties etc. and are also hard to verify numerically and analytically. Hence, identifying settings where these assumptions can be cast in more physical terms, i.e. based on correlation structures in the initial state or transport properties of the system and deriving controlled relaxation times are the major goals of our upcoming investigation.

For this we will use a second, more controlled but at the same time more demanding approach to capture the equilibration of quantum systems. In this formulation we demand that a signal $s(t)$ is close to an equilibrium value after a short relaxation time t_{relax} for a whole time interval $[t_{\text{relax}}, t_{\text{rec}}]$ until recurrences occur after some recurrence time t_{rec}. One can show that free bosonic system exhibit such an equilibration during a time interval with a relaxation time t_{relax} independent of the system size and t_{rec} growing unboundedly with it under the assumption that the dynamics is governed by a local Hamiltonian supporting a sufficient transport and an initial state which has correlations that decay with the distance of the probed parts of the system [156–158]. Below we obtain a corresponding result for fermionic systems.

As noted above, the equilibrium state $\overline{\rho}^{\infty}$ maximizes the von Neumann entropy under the constraint that the desired state has the same form as ρ in each eigenspace of H. It hence respects all constants of motion in the system and depends on a number of parameters that grows exponentially with the system size. In contrast a thermalized

system is expected to be captured by a thermal ensemble that only depends on a few parameters, in the extreme case only the total energy. Under what precise conditions the equilibrium state $\overline{\rho}^{\infty}$ will be locally indistinguishable from a thermal state $\rho = e^{-\beta H}/\mathcal{Z}$ with $\mathcal{Z} = \mathrm{tr}(e^{-\beta H})$ is still under debate. One favored explanation for the thermalization of generic quantum systems is provided by the eigenstate thermalization hypothesis (ETH). Adopting the formulation of [159], the ETH assumes that the eigenstates $|k\rangle$ of a general interacting, non-degenerate Hamiltonian at an energy E_k in the bulk of the spectrum of H are already locally indistinguishable from the thermal state $e^{-\beta_k H}/\mathcal{Z}_k$. If the initial state now does have a significant overlap with many eigenstates of H but also only with eigenstates from a narrow energy window, then in systems fulfilling the ETH it will be locally indistinguishable from a thermal state essentially by assumption. Several numerical studies (see [160, Sec. 4] for a review) support the ETH for complex enough and generic systems. Furthermore, more rigorous links have been established which provide further evidence that the ETH does indeed capture essential aspects of thermalizing systems. It was shown that if one assumes thermalization to happen for all initial states which lead to equilibration and are supported in a narrow energy window only [147, Sec. 2.7.2] or all product states [161] then the ETH becomes in fact necessary, hence equivalent to thermalization.

The ETH attempts to explain thermalization of closed quantum systems on very general grounds. However, at the moment it corresponds to more of a rewriting of the whole problem than a solution and proving the ETH, either in general or in specific settings seems to be as challenging as the original task. A second possible route for approaching the problem of explaining thermalization is given by considering more specific settings. If we are willing to restrict the class of initial states for which thermalization occurs, a different argument can be employed. One can show that a special class of initial states thermalize if the subsystem under consideration it is only weakly coupled to the remaining system and if the density of states of the remaining system is exponentially decaying in energy [162]. Here a state ρ is an admissible initial state if its reduced state $\mathrm{tr}_S {}^c\overline{\rho}^{\infty}$ is close to the projector on an energy window $[E, E+\Delta]$ of the subsystem and coupling of subsystem and remaining part H_{SR} need to fulfill $\|H_{SR}\| \ll \beta^{-1} \ll \Delta$. It seems fair to say that both approaches towards explaining thermalization are again founded on rather technical assumptions which follow a certain physical intuition but have to be connected to numerically, analytically or experimentally testable properties of a specific system in future work.

In the following section we elaborate on the physical assumptions behind relaxation processes in closed and finite quantum systems. For this we consider the setting of free Hamiltonians as it gives rise to a simplified structure of the time evolution as explained in Sec. 2.4 and a clear and well defined notion of transport. We find that under the dynamics of a free Hamiltonian, a fundamental relaxation process besides equilibration and thermalization occurs namely that a general initial state will appear to be Gaussian after a short time. From this relaxation we are able to derive bounds on equilibration times of free models and provide a physical picture for mixing processes leading to the occurrence of maximal entropy states much in the spirit of thermalization.

5.2. Relaxation of Free Fermionic Systems

Understanding the non-equilibrium dynamics of generic interacting quantum systems and uncovering the details of relaxation processes such as equilibration and thermalization is an important and fundamental but also challenging objective of current research. Due to the complexity of the general task, we consider a restricted and technically easier to handle setting in the following. We want to assume that the dynamics of the system is governed by a non-interacting Hamiltonian. In Sec. 2.4 we explained that the dynamics of such systems is highly structured and can be captured in terms of mode transformations. It is this structure however that imposes additional constraints on the non-equilibrium dynamics. Most prominently, the covariance matrix and collections of higher moments will only be conjugated by an appropriate unitary in order to capture their time evolution such that their spectral properties are conserved. Hence, the eigenvalues of the covariance matrix and by this the occupation of the natural orbitals are constants of motion. Different from the general constants of motion in interacting systems which we expect to be irrelevant on local scales, free systems will even locally have a memory of these constants of motion. This leads to the observation that for free systems we never expect the system to fully thermalize but instead to relax to a generalized Gibbs ensemble (GGE) [146], meaning the maximal entropy state under the constraint that all second moments are determined by the covariance matrix of the initial state. As discussed in Sec. 2.4, the maximal entropy state with a predefined covariance matrix γ is the Gaussian state defined by γ. In free systems, thermalization is therefore expected to be replaced by the convergence towards a Gaussian equilibrium state which is fixed by the second moments of the initial state. In bosonic systems, this expectation can be rigorously confirmed and it is indeed found that under the time evolution generated by a non-interacting Hamiltonian, the system locally equilibrates towards a Gaussian state after a relaxation time which stays finite upon taking the thermodynamic limit [156–158].

Below we will argue that starting from any non-critical initial state the dynamics under a non-interacting local fermionic Hamiltonian that supports a sufficient form of transport, lets the system Gaussify in the sense that after short and finite time the state is locally indistinguishable from a Gaussian state. The assumptions imposed on the system are essentially equivalent fermionic versions of the ones needed in the bosonic setting. Gaussification is a relaxation process equally fundamental as equilibration itself and more general than thermalization (or more precisely convergence towards a GGE) in free systems as the Gaussian state the system approaches will in general be time-dependent and follow a non-equilibrium dynamic. We find that the system Gaussifies after a finite relaxation time t_{relax} and stays Gaussian until a recurrence time that grows unboundedly with the system size. The dynamics of the emerging Gaussian state under the non-interacting Hamiltonian is then easily captured by the evolution of the second moments. Hence, in settings where the second moments of the initial states equilibrate, we find that the full system equilibrates and also provably converges towards a GGE. The structure of free systems allows us to connect the technical assumptions needed for the

Gaussification of a system directly to its physical properties. The proof itself contains some cumbersome details and combinatorics in order to obtain a rigorous bound but the essential physical process leading to the decay of non-Gaussian correlation and by this the local loss of memory of the initial condition will become quite clear. By this we are able to capture a very general relaxation process in physical and intuitive terms. Before we discuss the physical implications of Gaussification further, we will first carefully explain the necessary assumptions on the system in the following section. Furthermore, we explain the structure of the proof, which can be found in the appendix in full detail. We then conclude by discussing some technical aspects and, most importantly, physical implications on systems of fermionic particles.

5.2.1. Conditions and Intuition for Gaussification

As mentioned above we find that under the assumption of transport a non-critical initial state will Gaussify under the dynamics created by a local non-interacting Hamiltonian. Let us make the notion of Gaussification and the assumptions more precise. For this let us consider an n-dimensional regular lattice of V sites with p fermionic modes per sites and denote by $M = Vp$ the total number of modes in the system. Furthermore, we denote for $j, k \in [M]$ by $d(j, k)$ the distance of two modes which is given by the distance of the two sites on which the corresponding modes are supported on the lattice. Moreover, we introduce the notation that for a general fermionic state ρ we denote by ρ_G the Gaussified version of ρ in the sense that ρ_G is the Gaussian state with the same second moments as ρ i.e. $\gamma(\rho) = \gamma(\rho_G)$. As it is typical for the investigation of non-equilibrium processes, the systems we want to consider are defined by some initial state $\rho = \rho(0)$ and a Hamiltonian governing the dynamics of the system.

First we discuss the assumption on the initial state. We want to assume that the initial state does not support long-range correlations but to only feature exponentially decaying correlations according to the following definition.

Definition 3 (Exponentially decaying correlations)**.** *Given a fermionic system defined on a lattice of V sites with natural distance $d(\cdot, \cdot)$ we define a state ρ to exhibit exponentially decaying correlations with constants c_{clust} and $\xi_{\text{clust}} > 0$ if for any observables A and B with $\|A\| = \|B\| = 1$ we find*

$$|\operatorname{tr}(\rho AB) - \operatorname{tr}(\rho A)\operatorname{tr}(\rho B)| \leq c_{\text{clust}} |\operatorname{supp}(A)||\operatorname{supp}(B)| e^{-\frac{d(A,B)}{\xi_{\text{clust}}}}, \tag{5.3}$$

where $d(A, B)$ denotes the minimal distance of the supports of A and B.

One very obvious candidate of states featuring exponentially decaying correlations are of course mode product states discussed extensively in Ch. 4 as different sites are fully uncorrelated. Less trivial, and hence more important examples of states with exponentially decaying correlations are states that can be written as injective matrix product state[1]. Note however that this does not generalize to more complex injective

[1]Motivated by results on translational invariant MPS [82] we call an MPS injective if the transfer

tensor network states as for instance critical injective PEPS with only algebraically decaying correlations can be constructed [163]. Furthermore, we will show below that if ρ has exponentially decaying correlations then its Gaussified version ρ_G will have exponentially decaying correlations as well however with a prefactor that depends in a factorial instead of linear fashion on the size of the support (see Lem. 14 in the App. F for details). This result is of course not surprising as any expectation value of ρ_G is via Wick's theorem related to the second moments of ρ which of course display an exponential decay. Beyond these technical candidates let us also mention some more physical ones. It can be rigorously proven that if a local Hamiltonian has a unique ground state and a stable spectral gap, meaning that the gap between ground states and first excited states can be uniformly bounded when the system size is increased, then this ground state has exponentially decaying correlations [164, 165]. Moreover, one can show that for any local Hamiltonian there exists a finite critical temperature from which on the thermal state of the Hamiltonian features exponentially clustering correlations [166].

These physical examples also illustrate how the following results can be tested experimentally. The ground state or thermal state of a suitable model could be prepared as an initial state followed by a sudden change of the systems Hamiltonian. Such a scenario of a quench would allow for preparing an appropriate non-equilibrium initial state independent of the non-interacting Hamiltonian whose dynamics we want to investigate.

The class of Hamiltonians we want to investigate are local and supposed to feature a kind of ergodic transport which allows to explore the system evenly. Assuming locality for the Hamiltonian has a profound implication. For any local Hamiltonian on a quantum lattice system one can show that the dynamics sets a general upper bound on the speed with which correlations can spread over the system. Such bounds are usually cast into the form of Lieb-Robinson bounds (see for instance [167] for a review) which state that the support of a local operator that is evolved under a local Hamiltonian spreads up to exponentially small error only linearly in time. These bounds capture essentially the intuition of a finite velocity of quasi-particles in solids but are more restrictive in the sense that they apply to all energy scales. For non-interacting fermionic systems with a local Hamiltonian one can show that:

Theorem 9 (Lieb-Robinson bounds of local quadratic fermionic Hamiltonians)**.** *Given an n-dimensional lattice of V sites with natural distance $d(\cdot,\cdot)$ and p fermionic modes per site. Denote by m_j the Majorana operators of the system with $j \in [M]$ for $M = Vp$. If the evolution of the system is given by a quadratic Hamiltonian as in Eq.* (2.50) *which*

operators $T^{[s]} \in \mathbb{C}^{r_{s-1}^2 \times r_s^2}$ of every site $s \in [V]$ defined as

$$T^{[s]}_{(\alpha,\alpha'),(\beta,\beta')} = \sum_{i=1}^{d} A^{[j]}_{i,\alpha,\beta} \overline{A^{[j]}_{i,\alpha',\beta'}} \tag{5.4}$$

has an unique eigenvalue of largest magnitude. It is then easy to see that any expectation value of the form $\mathrm{tr}(\rho AB)$ agrees with $\mathrm{tr}(\rho A)\,\mathrm{tr}(\rho B)$ up to an error that scales as $\gamma^{-d(A,B)}$ with γ being the smallest gap between the largest and second largest eigenvalue occurring for any $T^{[s]}$ on sites s between the supports of A and B.

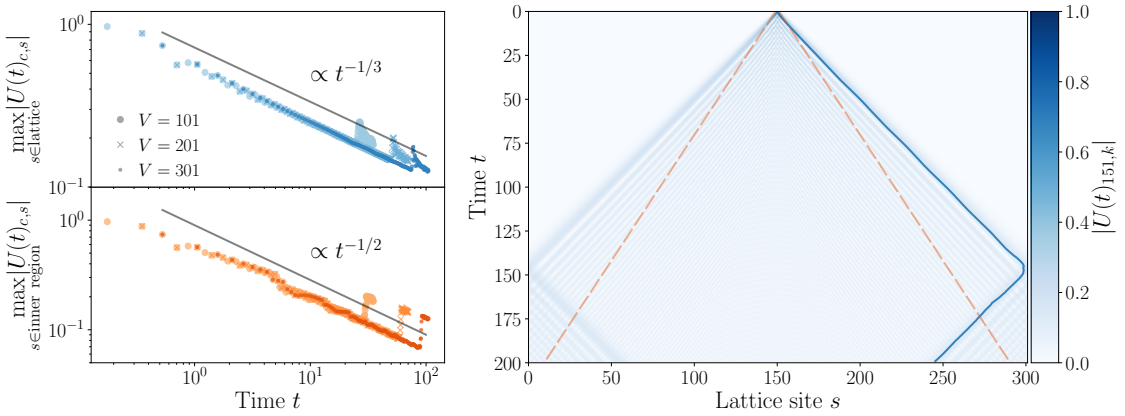

Figure 5.1. – Illustration of the Lieb-Robinson cone and the transport properties of a local non-interacting model. As Hamiltonian we choose the one-dimensional Hubbard model defined in Eq. (2.61) without interactions $U = 0$ and periodic boundary conditions. In the right panel we show the time evolution of the magnitude of one row of the mode transformation $U(t) = e^{ith^{(2)}}$ which governs the dynamics of the systems, where $h^{(2)}$ denotes the coupling matrix of the non-interacting Hamiltonian i.e. $H = \sum_{j,k} h^{(2)}_{j,k} f^\dagger_j f_k$, for $V = 301$. The maximal value is indicated by the smoothed blue curve and is near the wave front. In the upper left panel we show the evolution of the maximal amplitude near the wave front in time for different system sizes ($V = 101, 201, 301$ indicated by the corresponding symbols and we denote by c the corresponding central site index). The lower left panel displays for each considered system size the maximal value in the inner region of the Lieb-Robinson cone (defined by the two orange dashed lines in the right panel). In both plots we added lines that are proportional to $t^{-1/3}$ and $t^{-1/2}$ as guides to the eye. In both plots on the left side recurrences occur when the wave front starts to interfere with itself. The recurrence time for which this happens however grows with the system size.

is local, i.e. there exists an l independent of V such that $d(j,k) > l \Rightarrow h^{(2)}_{j,k} = 0$, then there are constants c_{lr}, v_{lr} and $\xi_{\text{lr}} > 0$ and a recurrence time t_{rec} such that $\forall t \in]0, t_{\text{rec}}]$ and $j, k \in [M]$

$$|e^{-h^{(2)}_{j,k}t}| \leq c_{\text{lr}} e^{-\frac{|d(j,k)-v_{\text{lr}}t|}{\xi_{\text{lr}}}}, \tag{5.5}$$

with $t_{\text{rec}} \to \infty$ for $V \to \infty$ and $d(j,k)$ denotes for modes $j, k \in [M]$ the distance of sites on which j and k are supported.

Proof. The proof follows directly from applying the more general Lieb-Robinson bounds stated in [168] for anti-commuting operators to $\{m_j(t), m_k\} = (e^{-h^{(2)}t})_{j,k}$. □

Lieb-Robinson bounds yield very generic upper bounds on the propagation of any correlation within a system but they especially do not guarantee it. Disordered systems as well as free hopping Hamiltonians fulfill the same Lieb-Robinson bounds despite showing distinct transport properties, the latter an ergodic exploration of the system while Anderson localization prevents transport in disordered ones [169]. It turns out that in order to find generic relaxation in free systems we also need to assume that the

systems features a certain minimal transport and that the Lieb-Robinson bounds are up to certain threshold tight. To be precise we use the following definition for a system which allows for transport.

Definition 4 (Delocalizing transport of quadratic Hamiltonians)**.** *Assume an n-dimensional lattice of V sites with p fermionic modes per site with Majorana operators m_j with $j \in [M]$ for $M = Vp$. A quadratic Hamiltonian as in Eq.* (2.50) *is said to exhibit delocalizing transport with constants c_{trans} and $d_{\text{trans}} > 0$ and recurrence time t_{rec} if $\forall t \in]0, t_{\text{rec}}]$ and $j, k \in [M]$*

$$|(e^{-h^{(2)}t})_{j,k}| \leq c_{\text{trans}} t^{-d_{\text{trans}}} \tag{5.6}$$

and $t_{\text{rec}} \to \infty$ for $V \to \infty$.

We show in App. G that a n-dimensional free nearest neighbor hopping Hamiltonian, implementing the usual kinetic energy on a regular lattice, exhibits delocalizing transport in the sense of Def. 4 with a decay of $d_{\text{trans}} = n/3$. In Fig. 5.1 we further illustrate the transport properties as well as the Lieb-Robinson cone for the one-dimensional nearest neighbor hopping Hamiltonian. The example shows effectively how the support of the annihilation operator $f_j(t) = \sum_k U_{j,k} fk$ spreads in time. The Lieb-Robinson cone is clearly visible and the maximal amplitude which decays as $t^{-1/3}$. We find in addition that the weight on the majority of sites is even more strongly suppressed in time; approximately as $t^{-1/2}$. Note that for a local Hamiltonian which possesses Lieb-Robinson bounds, the decay constant d_{trans} is bounded due to the orthogonality of the mode transformation governing the time evolution by the following argument. For a given mode $j \in MM$ the size of the Lieb-Robinson cone around the site on which the mode is supported on scales in time as t^n for a n-dimensional lattice. Hence, $(e^{-h^{(2)}t})_{j,k}$ can be assumed to vanish (up to an exponentially small error) for k outside of this cone. However, the orthogonality of $e^{-h^{(2)}t}$ demands that $\sum_k |e^{-h^{(2)}t}_{j,k}|^2 = 1$ such that if all elements $e^{-h^{(2)}t}_{j,k}$ for k inside the cone are equally suppressed we conclude that $d_{\text{trans}} \leq n/2$.

Prepared with these assumptions and definitions we can now formulate the following theorem.

Theorem 10. *Consider a fermionic system defined on an n-dimensional regular lattice of V sites where each site is equipped with p fermionic modes. Assume that the dynamics of the system is given by a non-interacting local Hamiltonian giving rise to delocalizing transport with decay constant $d_{\text{trans}} > n/4$. Given a region $S \subset [V]$ of the lattice we then find that any fermionic state ρ with exponentially decaying correlations and for all $0 < \epsilon < 4d_{\text{trans}} - n$ we can define a constant C independent of the system size V such that*

$$\|\text{tr}_{S^c} \rho(t) - \text{tr}_{S^c} \rho_G(t)\|_1 \leq C t^{-(4d_{\text{trans}} - n - \epsilon)} \tag{5.7}$$

for all $t \in]0, t_{\text{rec}}]$ where $\rho_G(t)$ denotes the time-dependent Gaussian state with the same second moments as $\rho(t)$ and $t_{\text{rec}} \to \infty$ for $V \to \infty$.

Under the stated assumptions we hence find that the initial state becomes locally indistinguishable from its Gaussified version, where the constant C depends only on the size of the region S and constants of the problem, so c_{trans}, c_{clust}, v_{lr}, etc. The constant is especially independent of the total systems size V such that we find that even in the infinite system, the state is essentially captured by its Gaussified version after a short time. Before we discuss the physical implications we sketch the proof in the remainder of this section and add a few technical comments in the following one.

The proof of the above theorem follows along rather intuitive lines. A complete and rigorous version of the proof can be found in the App. F. First we rewrite the problem using the operational definition of the one-norm into an expectation value of an operator supported on the subsystem S only. Every such operator can be expanded into the Majorana basis such that we can capture their time evolution easily using the mode transformation $e^{h^{(2)}t}$. The locality of the Hamiltonian allows us to restrict our attention to the Lieb-Robinson cone around the subsystem S and neglect any other part of the system by accepting an exponentially suppressed error. The exponential decay of correlations present in the initial state ρ and by this also in its Gaussified version ρ_G yields then that an expectation value that probes regions which are more than the correlation length apart factor into essentially independent expectation values, again up to an exponential small error. Here, the delocalizing transport ensures that the combinatorics works out in the sense that there are more and more contributions which probe far apart regions on the lattice at later times. Hence, we arrive at different combinations of terms which consist of products of individual local expectation values that are weighted by the corresponding matrix elements of the propagator $e^{h^{(2)}t}$ distributed over the Lieb-Robinson cone around S. The delocalizing transport ensures now that these contributions decay if fourth or higher moments are probed which follows roughly from the following combinatorial argument. The Lieb-Robinson cone has a typical size of t^n for an n-dimensional lattice while the prefactor of a fourth moment decays due to the delocalizing transport as $t^{-4d_{\mathrm{trans}}}$. We can therefore bound the different configuration of one fourth moment which is factored from the remaining terms roughly by $t^{-4d_{\mathrm{trans}}+n}$ such that for $d_{\mathrm{trans}} > n/4$ we obtain an algebraic decay. Showing that appearing second moments can be summarized in a large but finite additional factor and that the constraint that different factored expectation values are not distributed freely but have to have a minimal distance does not interfere with this intuition relies on more careful bookkeeping and more elaborate combinatorial estimates.

The key steps of the proof are hence exploiting Lieb-Robinson bounds in order to focus the attention on a subregion of the lattice which grows in time but is independent of the total system size. Then one shows that by the exponentially decaying correlations the expectation value of every time-evolved local operator can be written as the mixture of several local contributions. Delocalizing transport ensures that these local contribution decay fast enough if they involve at least one fourth moment. If a term consists of second moments only, ρ and ρ_G lead to the same contribution by construction.

5.2.2. Technical Comments on Gaussification

Let us add a few technical comments on the proof and the resulting theorem. First, let us note that the result is reminiscent to a central limit theorem. After exploiting the exponentially decaying correlations, we obtain that any expectation value can be written as the a mixture of products of numerous individual local contribution. From this we then find that in leading order the expectation value is captured by the Gaussified state which is the state that maximizes the von Neumann entropy when fixing the second moments.

Furthermore, note that the decay in time is governed by an algebraic law and hence does not yield a real time scale of relaxation (which would be present in an exponential decay in time). Still we find for a predefined accuracy δ that we can choose a time t_{relax} if the system size is large enough such that for all $t \in [t_{\mathrm{relax}}, t_{\mathrm{rec}}]$ $\rho(t)$ and $\rho_G(t)$ are locally indistinguishable up to an error δ, i.e. $\| \operatorname{tr}_S{}^c \rho(t) - \operatorname{tr}_S{}^c \rho_G(t) \|_1 \leq \delta$. The time t_{relax} is then determined by the prefactor C and error threshold δ.

When stating Thm. 10 we did not give a detailed account of the constant C. This is mostly due to the fact that the major goal of the theorem is to prove that there exists such a constant which is independent of the system size. A more detailed look into the constant by collecting the different contributions in App. F reveals that C depends factorially on the size of the subregion S and hence quickly leads to bounds which exceed any experimental relevant time scale and intuition. It is unclear which contributions to this strong subsystem size dependence are an artifact of our proof strategy and which are necessary. A more detailed investigation would be needed if one would want to derive a general more realistic bounds. However, our current proof relies on bounding the error made for individual products of Majorana or fermionic creation and annihilation operators (Ref. [170] contains the details of the latter). The final bound is then obtained as a worst case estimate based on this bound which increases the prefactor further. Having said that it is hence possible for specific observables to obtain a smaller, though still exponentially large in $|S|$, prefactor from the considerations in App. F.

5.2.3. Physical Implications and Relation to Thermalization

Let us now consider Thm. 10 from a more physical point of view. As we already pointed out before, Gaussification is not equivalent to equilibration or convergence to a GGE as the Gaussian state a system converges to is in general time-dependent. However, according to Thm. 10 we can describe the non-equilibrium dynamics of the system within the Gaussian setting only, although the initial state might be far from Gaussian. For practical applications this provides of course an immense simplification of the problem. The number of degrees of freedom we need to account for in order to capture all local properties of the system up to a controlled error is reduced such that the considered systems are more accessible for analytical as well as numerical investigations. Reminiscent to the loss of memory of the initial conditions in a thermalizing process, the information of all non-Gaussian correlations is lost and we converge to the state which at each point in time maximizes the von Neumann entropy. Gaussification can then be understood as

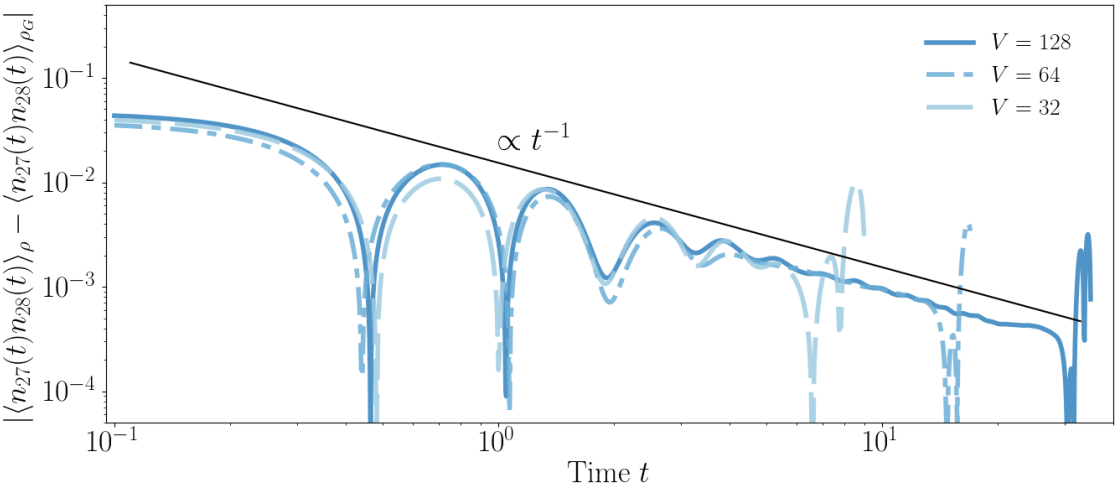

Figure 5.2. – Numerical investigation of the Gaussification of a nearest neighbor density-density correlator with increasing system sizes. As initial state we choose the ground state of the one-dimensional Hubbard model defined in Eq. (2.61) with interaction strength $U = 2$ and an additional on-site disorder potential $\sum_j h_j n_j$ with h_j drawn independently from a Gaussian distribution with zero mean and variance $1/4$. As Hamiltonian we choose the nearest neighbor hopping Hamiltonian (Eq. (2.61) with $U = 0$). The hopping amplitude was set to 1 for both Hamiltonians. All calculations are performed with periodic boundary conditions and for consistency we have drawn 128 random coefficients h_j once and used the first V of them to define the corresponding model. We show the evolution of the non-Gaussian contributions to the nearest neighbor density-density correlations on site 27 and 28. The black line indicates a t^{-1} decay as guide to the eye.

a convergence towards a time-dependent GGE and nicely displays a dynamical instance of Janyes' principle. It is interesting to note that Thm. 10 can be viewed as our second generalization of the fermionic central limit theorem (next to the one obtained in the previous chapter with the help of the derived de Finetti theorem) which this time emerges dynamically in time. Only in special cases does Gaussification imply equilibration. If we consider a system with a translation invariant Hamiltonian and translation invariant second moments then it is easy to verify that $\rho_G(t) = \rho_G(0)$ for all times t. Hence, by Thm. 10 we obtain that $\rho(t)$ equilibrates locally. In addition it proves that the system equilibrates to the expected GGE. This of course holds true in any setting where one can prove by additional means that the second moments equilibrate after some time.

Gaussification will be visible in the evolution of a higher moments and hence can be experimentally observed [171] with a signature as the one shown in Fig. 5.2. Here we show that the non-Gaussian contributions to a fourth moment decay algebraically in time until a recurrence occurs due to the collision of the two wave-fronts of the Lieb-Robinson cone. As expected, this recurrence is shifted to later times with increasing system size. It is interesting to note that we indeed find an algebraic decay which shows that the fact that our bound in Thm. 10 is only algebraically and not exponentially decaying in time is not an artifact of our proof but reflects the physical reality of free

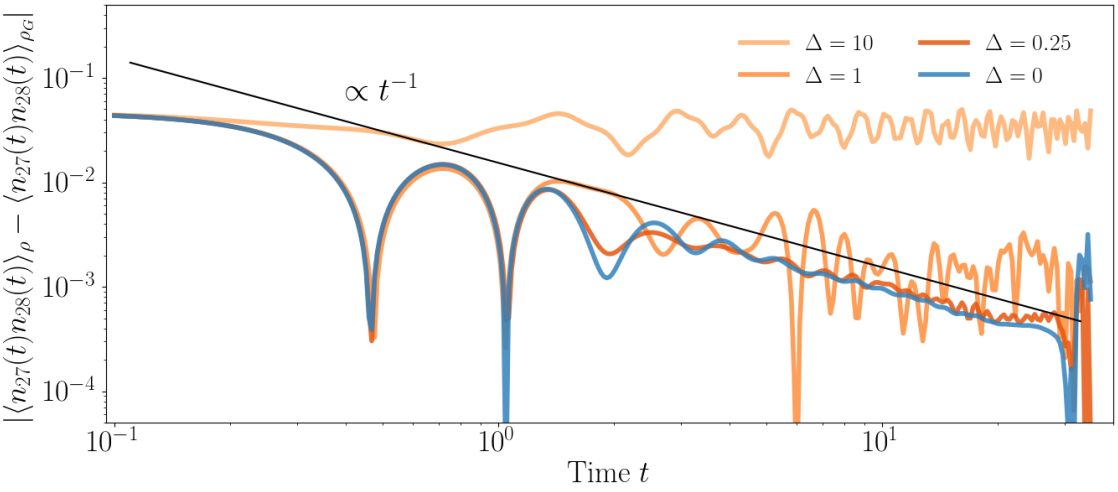

Figure 5.3. – Investigation of the Gaussification of a nearest neighbor density-density correlator in the presence of disorder. We use the same initial state as in Fig. 5.2 on $V = 128$ sites. As free Hamiltonian we choose the nearest neighbor hopping Hamiltonian (Eq. (2.61) with $U = 0$) with an additional on-site disorder potential $\Delta \sum_j h_j n_j$ with $h_j \in [0, 1]$ chosen uniformly at random. The black line indicates a t^{-1} decay as guide to the eye. With increasing disorder strength Δ, larger portions of the initial non-Gaussian correlations remain due to localization effects.

systems. Note however, that from our bound we would expect a decay as $t^{-1/3}$ as the nearest neighbor hopping Hamiltonian supports delocalizing transport with $d_{\text{trans}} = 1/3$ only (compare Fig. 5.1 and App. G). It is hence surprising on the first sight that we find an even stronger suppression for the correlator in Fig. 5.2. We can explain this behavior within our framework. As shown in Fig. 5.1 we find that only at the wave-front of the Lieb-Robinson cone weights are suppressed with $d_{\text{trans}} = 1/3$. Inside the cone and hence the majority of contributions is suppressed more strongly with $d_{\text{trans}} = 1/2$. Assuming an effective $d_{\text{trans}} = 1/2$ in our bound yields the observed t^{-1} decay.

Furthermore, let us argue in what sense the initial assumption to Thm. 10, transport and decaying correlation in the initial state, are indeed necessary and underline this with the corresponding physical pictures. The requirement of the system supporting delocalizing transport is rather intuitive. If the system would feature no transport at all, for instance due to Anderson localization effects in the presence of disorder [169], then non-Gaussian features will be preserved within small localized regions such that the system only Gaussifies up to the point in time where the dynamics explored these localized regions. With decreasing localization length, i.e. due to increasing disorder, this cutoff time is reached earlier as displayed in Fig. 5.3. On the other hand if the system contains only ballistically expanding particles with a linear dispersion relation $E(k) = vk$ and group velocity v, then we will find transport in the system which does not spread. It is then clear that at a given point in time t in a local region S we see only the local correlations from the regions vt away from S of the initial state translated to S and hence

expect to be able to resolve initial non-Gaussian features. Hence violating or weakening the assumption of delocalizing transport without assuming additional structure leads to non-Gaussifying settings.

The same holds true for the assumed (exponentially) decaying correlations. Intuitively they require that the initial state has enough overlap with several eigenstates of the system (which are delocalized in the presence of transport). If we drop the assumption of exponentially decaying correlations we can easily construct initial conditions which do not Gaussify. Consider for instance the one-dimensional nearest neighbor hopping Hamiltonian with periodic boundary conditions and a single mode per site. Denoting the real space creation and annihilation operators of the system by $f_j^\dagger$ and f_j respectively with $j \in [V]$ labeling the sites, we define the Fourier modes

$$a_q^\dagger = \frac{1}{\sqrt{V}} \sum_{j=1}^{V} e^{\frac{2\pi i}{V} jq} f_j^\dagger, \qquad a_q = \frac{1}{\sqrt{V}} \sum_{j=1}^{V} e^{-\frac{2\pi i}{V} jq} f_j, \tag{5.8}$$

for $q = 0, 1, \ldots, V-1$. We assume V to be dividable by 4 and choose the momentum eigenstates

$$|\psi_1\rangle = \prod_{q=\frac{V}{4}}^{3\frac{V}{4}-1} a_q^\dagger |0\rangle, \qquad |\psi_2\rangle = \prod_{q=0}^{\frac{V}{2}} a_{2q}^\dagger |0\rangle, \tag{5.9}$$

i.e. the ground state $|\psi_1\rangle$ in the half filling sector and a highly excited state, and consider the state $\rho = 1/2(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|)$. A lengthy but uneventful calculation shows that for $k, l \in [V]$ with $k \neq l$ we find

$$|\operatorname{tr}(\rho\, n_k n_l) - \operatorname{tr}(\rho\, n_k)\operatorname{tr}(\rho\, n_l)| = \frac{1}{V^2} \frac{\delta_{k-l,\text{odd}}}{1 - \cos\left[\frac{2\pi}{V}(k-l)\right]} + \frac{\delta_{|k-l|,\frac{V}{2}}}{8}, \tag{5.10}$$

$$|\operatorname{tr}(\rho\, n_k n_l) - \operatorname{tr}(\rho_G\, n_k n_l)| = \frac{1}{2V^2} \frac{\delta_{k-l,\text{odd}}}{1 - \cos\left[\frac{2\pi}{V}(k-l)\right]} + \frac{\delta_{|k-l|,\frac{V}{2}}}{16} \tag{5.11}$$

where $n_k = f_k^\dagger f_k$. For large system sizes we see that the correlations $|\operatorname{tr}(\rho\, n_k n_l) - \operatorname{tr}(\rho\, n_k)\operatorname{tr}(\rho\, n_l)|$ decay essentially as $1/(l-k)^2$ and are peaked at a distance $k - l = V/2$, hence, not exponentially decaying. Furthermore, the state is invariant under the time evolution, i.e. $\rho(t) = \rho(0)$ as the Fourier modes are the eigenmodes of the system and ρ is the projector on two different momentum eigenstates. Hence, we obtain that the non-Gaussian correlation $|\operatorname{tr}(\rho\, n_k n_{k+1}) - \operatorname{tr}(\rho_G\, n_k n_{k+1})| \approx 1/(8\pi^2)$ stays constant over time and is locally detectable.

Note that on the other hand the assumed locality of the Hamiltonian is most likely not strictly necessary. In our setting we are happy to assume it as many realistic models are well approximated by local Hamiltonians. Technically the locality implies by Lieb-Robinson bounds that we can reduce our attention to a subregion of the system independent of the system size. If we would want to abandon locality for some reasons,

we would need to replace it with some other form of upper bound on the speed of the spreading of correlations which is naturally expected in most systems.

In view of more general systems we expect from the intuition obtained here that relaxation processes and the loss of memory of initial conditions is tied to the correlation structure of the initial state and transport properties of the Hamiltonian. The intuition that many independent contributions start to mix quickly in time and by this allow for an effective ensemble description of the system is then expected to carry over to more general settings, i.e., interacting models. Note however, that this step is at the moment only schematically outlined. One important open problem in this area is to obtain a proper definition of transport in interacting systems where the picture of individual well defined particles exploited here has to be revised and there are most likely next to technical difficulties also conceptual challenges to be overcome.

5.3. Summary

In the above chapter we discussed different relaxation processes in closed quantum systems. During the expected thermalization of many-body quantum systems, a system loses its memory of almost all initial conditions after a short relaxation time and allows for an effective description of the state of the system in terms of thermal ensembles. It is however unclear how precisely these effective descriptions emerge from the unitary dynamics of closed quantum systems.

In order to shed some light on these old and fundamental questions of quantum mechanics, we invoke the study of a simpler setting which allows to relate technical assumptions to physical interpretations more easily – the setting of free Hamiltonians. If the interactions present in a system can be neglected, we find that a local Hamiltonian which features a spreading transport that slowly but surely explores large portions of the system leads to the for free systems genuine relaxation process of Gaussification. We find, as summarized in Thm. 10, that an initial state which features exponentially decaying correlations will be locally indistinguishable from a Gaussian state after a short time. All non-Gaussian correlations will delocalize over the system and be only accessible by global, specifically tailored measurements. It is important to emphasize that the resulting state is in general time dependent.

As a first important result we hence view that in such free systems the whole local non-equilibrium dynamics can up to a bounded error be well described within a purely Gaussian setting. As discussed in Sec. 2.4, this implies a tremendous simplification and an efficient description of the system. Furthermore, our theorem implies that Gaussian states appear naturally even from highly non-Gaussian initial states such that they can be assumed to be the initial state in the description of a non-interacting system if the system parameters have been kept fixed for some time. Secondly, in cases where the second moments of the initial state equilibrate, e.g. if the models and initial states are translation invariant, Gaussification directly implies equilibration and convergence of the state towards a GGE. For free systems we hence find that a suitable transport and correlation structure of the initial state are sufficient for the emergence of an ensemble

description after a well controlled relaxation time.

6. Outlook and Open Research Question

In this thesis we explored different directions of identifying and understanding the more detailed structure of fermionic systems. We first introduced in Ch. 2 a common general notation for finite fermionic systems and discussed in detail the structure of non-interacting models, the complexity of interacting systems and families of general kinematic constraints arising from the indistinguishable character of fermions. Subsequently we introduced various numerical schemes which approximate the ground state of interacting fermionic systems with a strong focus on tensor network methods in Ch. 3. Most notably, we combined tensor network states with mode transformations which leads to a larger variational class of states that include strong entanglement effects in order to approximate target states. We discussed multiple schemes and their corresponding advantages in order to adapt the single particle basis such that more efficient approximations can be found much in spirit of the Hartree-Fock approximation. In Ch. 4 and 5 we studied the emergence of natural structures. In Ch. 4 we formulated a mode de Finetti theorem which deduces the suppression of all quantum correlations between different fermionic modes from a permutation invariance. We argue that due to intrinsic antisymmetry of fermionic systems such a theorem can be derived under a relaxed version of permutation invariance and link our result conceptually to mean field approximations. In Ch. 5 we argued that under the evolution of a free Hamiltonian supporting transport a generic initial state will locally relax to a Gaussian state and described by this in detail a fundamental relaxation process in closed quantum systems on general grounds. Note that in both cases, the results are of rather general nature. Their major merit is the rigorous prediction of the approximability of energy densities and the appearance of separability in large permutation invariant systems and the explanation that and why we can expect Gaussians states to appear in nature and not the exact error for mean field approximations derived in a finite system or the relaxation time of a specific system. We summarized our findings in more detail at the end of the corresponding chapters and will not repeat these summaries here. Instead we want to conclude with an extended outlook and present further research questions around and emerging from the topics discussed in this thesis. The complexity and prospects vary strongly among the different questions and directions discussed below from rather small variations or modifications of existing approaches to long and involved research programs.

The finding that the antisymmetry of the fermionic wave function is manifested in a rich structure of the reduced density matrices and gives rise to numerous new kinematic constraints bears a great potential. On the one hand side these constraints could potentially explain the response of physical systems to the change of its parameters by the saturation of certain constraints and movement along facets resulting from these constraints. On the other hand the structure imposed by the constraints on the re-

duced one-body density matrix could be used within numerical methods by restricting optimization schemes to the states whose one-body reduced density matrix lies on or near certain facets. These hopes however are accompanied by challenges and problems. The sheer number of constraints is one major obstacle preventing progress along these lines. In order to overcome this limitation further research concerning the structure of the constraints is needed. Are there relaxed version of the constraints which suitably interpolate between the Pauli constraints and the full set of constraints obtained from Klyachko? Is there a way to efficiently predict for a model which constraints might be relevant in a certain parameter regime and which not, e.g., if an interacting model undergoes pinning [37], is there a way to predict the corner of the polytope we will find its ground state in using mean field or similar methods? Furthermore, it would be interesting to systematically investigate if typical variational sets used in numerical schemes which simulate interacting fermionic systems can be understood from the perspective of these additional constraints. The solution of a Hartree-Fock approximation will of course lie in a fixed corner of the polytope and CI or CC wave functions with only few excitations will be close to this point. It is however unclear if these states and states with higher excitations, will favor facets or be homogeneously smeared inside the polytope. The same question holds of course for MPS. For instance, do MPS with low bond dimensions aggregate inside the polytope or are these structures more or less uncorrelated. Fair sampling techniques over MPS manifolds are needed here in order to assess the generic structures appearing. In addition, it is only insufficiently understood how physical structures such as locality of correlations or translation invariance affect the structures found in the one or two-body reduced density matrix.

The success of tensor network based schemes outside the realm of approximating ground states of local gapped one-dimensional Hamiltonians shows that they provide an efficient tool in high dimensional settings and are in principle applicable to a wide range of numerical problems. Here, however the origins of DMRG and related methods weighs heavily as they are formulated for and tailored to eigenvalue or real time evolution problems only. Small changes in the DMRG micro iteration step however allow to formulate at least on paper more flexible schemes that are able to minimize more general functions than Rayleigh quotients over the MPS manifold. By this novel optimization methods exploiting and targeting low rank structures can be constructed. However, in order to build a general optimization toolbox we need a better understanding of the convergence behavior of local update schemes. As discussed in Sec. 3.2.2, an analysis of induced local minima and the structure of the tangent cones is needed to obtain a broader understanding of rank adaptive schemes. In the light of the generalizations of tensor network schemes via the incorporation of mode transformations presented in Ch. 3 the next important step would be the analysis of the geometry of the combined parameter manifold. Formulating variational methods for real or imaginary time evolution on the joint manifolds of MPS and mode transformations $\mathcal{A} \times U(M)$ would allow to formulate more rigorous versions of the methods discussed in Sec. 3.3 and will especially yield real time evolution schemes that outperform traditional MPS based methods as shown in Sec. 3.3.3.3. Applying the combined single particle optimization scheme in larger

settings could of course also help to choose a suitable active space[1] for the calculations of the electronic structure in systems from quantum chemistry. In addition if employed together with an MPS with low bond dimension, the scheme allows to generalize the Hartree-Fock approximation in such a way that the optimized single particle basis takes into account the most important corrections to the Hartree-Fock solution. The obtained single particle basis could then improve the performance of other approximation schemes such as CC, CI or DFT in more correlated settings.

Before we discuss the more concrete directions of future research let us note that by virtue of the Jordan Wigner transformation, the results presented in Ch. 4 and Ch. 5 can be translated to distinguishable particles. If a the dynamics of a system of qubits is governed by a Hamiltonian that maps to a non-interacting fermionic model supporting delocalizing transport, for instance the XX model, and possesses an initial state which maps to a fermionic state respecting the super selection rule, then the dynamic of the system is well captured by Thm. 10 and the system will Gaussify in the fermionic picture. Furthermore, our de Finetti theorem in Thm. 5 extends quantum de Finetti theorems of distinguishable particles as it applies directly to spin states which are not fully permutation invariant but only invariant in the weaker sense of Def. 2 and map again to a proper fermionic state under the Jordan-Wigner transformation. Beyond that, as already discussed in Ch. 4, in order to obtain bounds on mean field approximations of realistic models generalizations of the mode de Finetti theorem Thm. 5 are needed. A generalization along the lines of [133] based on the connectivity of the interaction graph seems possible at least if the norm in the desired bound is changed to LOCC norms which would allow for an large local dimension and to employ arguments based on the symmetry under the exchange of particles instead of modes. In order to overcome the limitations of the used norm and to obtain more general results a fully antisymmetric de Finetti theorem would be needed.

Concerning the Gaussification of free systems, next to obvious questions about how the bound might be tightened or if similar results would hold in continuous systems, one interesting connection to systems of solid state physics would be the question if our definition of transport can be brought into a more traditional view. Can we formulate a similar theorem for Gaussification based on the band structure of a given model? Furthermore, the current discussion leaves the question of equilibration of free models largely unsettled. It seems plausible that the combination of traditional tools of the equilibration theory of closed quantum systems with insights obtained in the context of Gaussification could lead to sufficient conditions for the equilibration of free models and allow for the formulation of equilibration times. Such a deeper understanding combined with our result could directly imply the convergence towards a GGE for many systems. A translation of the obtained results to interacting models is far more challenging. Here many-body localized systems with their local integrable structure might represent an

[1] In order to automatize the selection of a good finite subspace of $\bigwedge^N L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ often a large set of single particle orbitals is chosen. From them a smaller active space is chosen in which the ground state is approximated as accurate as possible using DMRG or other means and the passive orbitals are either fully occupied or empty.

intermediate stepping stone for the transition from the investigation of free models to fully interacting systems.

Bibliography

[1] Schuch, N. and Verstraete, F. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nat. Phys.*, **5**, 732 (2009).

[2] Liu, Y.-K., Christandl, M., and Verstraete, F. Quantum computational complexity of the n-representability problem: QMA complete. *Phys. Rev. Lett.*, **98**, 110503 (2007).

[3] Whitfield, J. D., Love, P. J., and Aspuru-Guzik, A. Computational complexity in electronic structure. *Phys. Chem. Chem. Phys.*, **15**, 397 (2013).

[4] Ashcroft, N. W. and Mermin, N. D. *Solid State Physics*. Brooks/Cole, Belmont (1976).

[5] Lykos, P. and Pratt, G. W. Discussion on the Hartree-Fock approximation. *Rev. Mod. Phys.*, **35**, 496 (1963).

[6] Szabo, A. and Ostlund, N. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Mineola (1996).

[7] Sholl, D. S. and Steckel, J. A. *Density Functional Theory*. John Wiley & Sons, Inc., Hoboken (2009).

[8] Schollwöck, U. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.*, **326**, 1, 96 (2011).

[9] Hastings, M. B. An area law for one-dimensional quantum systems. *J. Stat. Mech.*, **2007**, 08, P08024 (2007).

[10] Affleck, I., Kennedy, T., Lieb, E. H., and Tasaki, H. Rigorous results on valence-bond ground states in antiferromagnets. *Phys. Rev. Lett.*, **59**, 799 (1987).

[11] Pérez-García, D., Wolf, M. M., Sanz, M., Verstraete, F., and Cirac, J. I. String order and symmetries in quantum spin lattices. *Phys. Rev. Lett.*, **100**, 167202 (2008).

[12] Pollmann, F. and Turner, A. M. Detection of symmetry-protected topological phases in one dimension. *Phys. Rev. B*, **86**, 125441 (2012).

[13] Orús, R. Advances on tensor network theory: symmetries, fermions, entanglement, and holography. *The European Physical Journal B*, **87**, 11, 280 (2014).

[14] Nolting, W. *Grundkurs Theoretische Physik 5/2*. Springer, Berlin (2006).

[15] Fischer, G. *Lineare Algebra*. Vieweg, Wiesbaden (2005).

[16] Nielsen, M. A. and Chuang, I. L. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge (2000).

[17] Ciarlet, P. G., Lions, J. L., and Le Bris, C. *Special Volume: Computational Chemistry*, *Handbook of Numerical Analysis*, vol. 10. Elsevier, Amsterdam (2003).

[18] Cancès, E., Defranceschi, M., Kutzelnigg, W., Le Bris, C., and Maday, Y. Computational quantum chemistry: A primer. In P. G. Ciarlet, J. L. Lions, and C. Le Bris, (eds.) *Special Volume, Computational Chemistry*, *Handbook of Numerical Analysis*, vol. 10, 3. Elsevier (2003).

[19] Jordan, P. and Wigner, E. Über das Paulische Äquivalenzverbot. *Z. Physik*, **47**, 9, 631 (1928).

[20] Wick, G. C., Wightman, A. S., and Wigner, E. P. The intrinsic parity of elementary particles. *Phys. Rev.*, **88**, 101 (1952).

[21] Zimborás, Z., Zeier, R., Keyl, M., and Schulte-Herbrüggen, T. A dynamic systems approach to fermions and their relation to spins. *EPJ Quantum Tech.*, **1**, 1, 11 (2014).

[22] Johansson, M. Comment on 'reasonable fermionic quantum information theories require relativity' (2016). arXiv:1610.00539.

[23] Cahill, K. E. and Glauber, R. J. Density operators for fermions. *Phys. Rev. A*, **59**, 1538 (1999).

[24] Jones, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, **87**, 897 (2015).

[25] Bravyi, S. Lagrangian representation for fermionic linear optics. *Quantum Inf. Comp.*, **5**, 216 (2005).

[26] Friesdorf, M. *Closed Quantum Many-Body Systems out of Equilibrium*. Ph.D. thesis, Freie Universität Berlin (2016).

[27] Choi, J.-Y., Hild, S., Zeiher, J., Schauß, P., Rubio-Abadal, A., Yefsah, T., Khemani, V., Huse, D. A., Bloch, I., and Gross, C. Exploring the many-body localization transition in two dimensions. *Science*, **352**, 6293, 1547 (2016).

[28] De Roeck, W. and Imbrie, J. Z. Many-body localization: stability and instability (2017). arXiv:1705.00756.

[29] Van Voorhis, T. and Head-Gordon, M. A geometric approach to direct minimization. *Mol. Phys.*, **100**, 11, 1713 (2002).

[30] Whitfield, J. D. and Zimborás, Z. On the NP-completeness of the Hartree-Fock method for translationally invariant systems. *J. Chem. Phys.*, **141**, 234103 (2014).

[31] Lloyd, S. Universal quantum simulators. *Science*, **273**, 5278, 1073 (1996).

[32] Lanyon, B. P., Whitfield, J. D., Gillett, G. G., Goggin, M. E., Almeida, M. P., Kassal, I., Biamonte, J. D., Mohseni, M., Powell, B. J., Barbieri, M., Aspuru-Guzik, A., and White, A. G. Towards quantum chemistry on a quantum computer. *Nat. Chem.*, **2**, 106 (2010).

[33] Hastings, M. B., Wecker, D., Bauer, B., and Troyer, M. Improving quantum algorithms for quantum chemistry. *Quantum Info. Comput.*, **15**, 1-2, 1 (2015).

[34] Coleman, A. J. Structure of fermion density matrices. *Rev. Mod. Phys.*, **35**, 668 (1963).

[35] Klyachko, A. A. The Pauli exclusion principle and beyond (2009). arXiv:0904.2009.

[36] Klyachko, A. A. The Pauli principle and magnetism (2013). arXiv:1311.5999.

[37] Schilling, C., Gross, D., and Christandl, M. Pinning of fermionic occupation numbers. *Phys. Rev. Lett.*, **110**, 040404 (2013).

[38] Tennie, F., Vedral, V., and Schilling, C. Influence of the fermionic exchange symmetry beyond Pauli's exclusion principle. *Phys. Rev. A*, **95**, 022336 (2017).

[39] Borland, R. E. and Dennis, K. The conditions on the one-matrix for three-body fermion wavefunctions with one-rank equal to six. *J. Phys. B*, **5**, 1, 7 (1972).

[40] Ruskai, M. B. Connecting n-representability to Weyl's problem: the one-particle density matrix for n = 3 and r = 6. *J. Phys. A*, **40**, 45, F961 (2007).

[41] Klyachko, A. A. Quantum marginal problem and n-representability. *J. Phys. Conf. Ser.*, **36**, 1, 72 (2006).

[42] Altunbulak, M. and Klyachko, A. A. The Pauli principle revisited. *Commun. Math. Phys.*, **282**, 2, 287 (2008).

[43] Altunbulak, M. *The Pauli Principle, Representation Theory, and Geometry of Flag Varieties.* Ph.D. thesis, Bilkent University (2007).

[44] Mazziotti, D. A. Significant conditions for the two-electron reduced density matrix from the constructive solution of n representability. *Phys. Rev. A*, **85**, 062507 (2012).

[45] Erdahl, R. M. Representability. *Int. J. Quantum Chem.*, **13**, 6, 697 (1978).

[46] Mazziotti, D. A. Structure of fermionic density matrices: Complete n-representability conditions. *Phys. Rev. Lett.*, **108**, 263002 (2012).

[47] Mazziotti, D. A. Variational minimization of atomic and molecular ground-state energies via the two-particle reduced density matrix. *Phys. Rev. A*, **65**, 062511 (2002).

[48] Zhao, Z., Braams, B. J., Fukuda, M., Overton, M. L., and Percus, J. K. The reduced density matrix method for electronic structure calculations and the role of three-index representability conditions. *J. Chem. Phys.*, **120**, 5, 2095 (2004).

[49] Schur, I. Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie. *Sitzungsber. Berl. Math. Ges.*, **22**, 9 (1923).

[50] Horn, A. Doubly stochastic matrices and the diagonal of a rotation matrix. *Am. J. Math.*, **76**, 3, 620 (1954).

[51] Thompson, R. C. Principal submatrices of normal and Hermitian matrices. *Illinois J. Math.*, **10**, 2, 296 (1966).

[52] Hackbusch, W. *Tensor Spaces and Numerical Tensor Calculus.* Springer, Berlin (2012).

[53] Booth, G. H., Thom, A. J. W., and Alavi, A. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.*, **131**, 5, 054106 (2009).

[54] Marti, K. H. and Reiher, M. The density matrix renormalization group algorithm in quantum chemistry. *Z. Phys. Chem.*, **224**, 3-4, 583 (2010).

[55] Chan, G. K.-L. and Sharma, S. The density matrix renormalization group in quantum chemistry. *Annu. Rev. Phys. Chem.*, **62**, 1, 465 (2011).

[56] Wouters, S. and Van Neck, D. The density matrix renormalization group for ab initio quantum chemistry. *Eur. Phys. J. D*, **68**, 9, 272 (2014).

[57] Szalay, S., Pfeffer, M., Murg, V., Barcza, G., Verstraete, F., Schneider, R., and Legeza, O. Tensor product methods and entanglement optimization for ab initio quantum chemistry. *Int. J. Quantum Chem.*, **115**, 19, 1342 (2015).

[58] Landsburg, J. M., Qi, Y., and Ye, K. On the geometry of tensor network states. *Quantum Info. Comput.*, **12**, 3-4, 346 (2012).

[59] Verstraete, F., Murg, V., and Cirac, J. I. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.*, **57**, 2, 143 (2008).

[60] Eisert, J. Entanglement and tensor network states. In E. Pavarini, E. Koch, and U. Schollwöck, (eds.) *Emergent Phenomena in Correlated Matter*, *Modeling and Simulation*, vol. 3. Verlag Jülich (2013).

[61] Schuch, N. Condensed matter applications of entanglement theory. In D. P. DiVincenzo, (ed.) *Introduction to Quantum Information Science. Lecture Notes of the 44th IFF Spring School 2013, Schlüsseltechnologien / Key Technologies*, vol. 52. Verlag Jülich (2013).

[62] Orús, R. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Ann. Phys.*, **349**, 117 (2014).

[63] Bridgeman, J. C. and Chubb, C. T. Hand-waving and interpretive dance: An introductory course on tensor networks (2016). arXiv:1603.03039.

[64] Oseledets, I. V. Tensor-train decomposition. *SIAM J. Sci. Comput.*, **33**, 5, 2295 (2011).

[65] Vidal, G. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, **91**, 147902 (2003).

[66] Hillar, C. J. and Lim, L.-H. Most tensor problems are NP-hard. *J. ACM*, **60**, 6, 45:1 (2013).

[67] Uschmajew, A. and Vandereycken, B. The geometry of algorithms using hierarchical tensors. *Linear Algebra Appl.*, **439**, 1, 133 (2013).

[68] Haegeman, J., Mariën, M., Osborne, T. J., and Verstraete, F. Geometry of matrix product states: Metric, parallel transport, and curvature. *J. Math. Phys.*, **55**, 2, 021902 (2014).

[69] Verstraete, F. and Cirac, J. I. Renormalization algorithms for quantum-many body systems in two and higher dimensions (2004). arXiv:cond-mat/0407066.

[70] Verstraete, F., Porras, D., and Cirac, J. I. Density matrix renormalization group and periodic boundary conditions: A quantum information perspective. *Phys. Rev. Lett.*, **93**, 227205 (2004).

[71] Pippan, P., White, S. R., and Evertz, H. G. Efficient matrix-product state method for periodic boundary conditions. *Phys. Rev. B*, **81**, 081103 (2010).

[72] Vidal, G. Entanglement renormalization. *Phys. Rev. Lett.*, **99**, 220405 (2007).

[73] Evenbly, G. and Vidal, G. Tensor network states and geometry. *J. Stat. Phys.*, **145**, 4, 891 (2011).

[74] Page, D. N. Average entropy of a subsystem. *Phys. Rev. Lett.*, **71**, 1291 (1993).

[75] Foong, S. K. and Kanno, S. Proof of Page's conjecture on the average entropy of a subsystem. *Phys. Rev. Lett.*, **72**, 1148 (1994).

[76] Arad, I., Kitaev, A., Landau, Z., and Vazirani, U. An area law and sub-exponential algorithm for 1d systems (2013). arXiv:1301.1162.

[77] Verstraete, F. and Cirac, J. I. Matrix product states represent ground states faithfully. *Phys. Rev. B*, **73**, 094423 (2006).

[78] Landau, Z., Vazirani, U., and Vidick, T. A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians. *Nat. Phys.*, **11**, 566 (2015).

[79] Arad, I., Landau, Z., Vazirani, U., and Vidick, T. Rigorous RG algorithms and area laws for low energy eigenstates in 1d. *Commun. Math. Phys.*, **356**, 1, 65 (2017).

[80] Ge, Y. and Eisert, J. Area laws and efficient descriptions of quantum many-body states. *New J. Phys.*, **18**, 8, 083026 (2016).

[81] Holzhey, C., Larsen, F., and Wilczek, F. Geometric and renormalized entropy in conformal field theory. *Nuc. Phys.*, **B424**, 3, 443 (1994).

[82] Pérez-García, D., Verstraete, F., Wolf, M. M., and Cirac, J. I. Matrix product state representations. *Quantum Info. Comput.*, **7**, 5, 401 (2007).

[83] Singh, S., Pfeifer, R. N. C., and Vidal, G. Tensor network decompositions in the presence of a global symmetry. *Phys. Rev. A*, **82**, 050301 (2010).

[84] Weichselbaum, A. Non-abelian symmetries in tensor networks: A quantum symmetry space approach. *Ann. Phys.*, **327**, 12, 2972 (2012).

[85] Cornwell, J. F. *Group Theory in Physics.* Academic Press, San Diego (1997).

[86] Singh, S., Pfeifer, R. N. C., and Vidal, G. Tensor network states and algorithms in the presence of a global u(1) symmetry. *Phys. Rev. B*, **83**, 115125 (2011).

[87] White, S. R. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, **48**, 10345 (1993).

[88] Östlund, S. and Rommer, S. Thermodynamic limit of density matrix renormalization. *Phys. Rev. Lett.*, **75**, 3537 (1995).

[89] Dukelsky, J., Martín-Delgado, M. A., Nishino, T., and Sierra, G. Equivalence of the variational matrix product method and the density matrix renormalization group applied to spin chains. *Europhys. Lett.*, **43**, 4, 457 (1998).

[90] Haegeman, J., Cirac, J. I., Osborne, T. J., Pižorn, I., Verschelde, H., and Verstraete, F. Time-dependent variational principle for quantum lattices. *Phys. Rev. Lett.*, **107**, 070601 (2011).

[91] Haegeman, J., Lubich, C., Oseledets, I., Vandereycken, B., and Verstraete, F. Unifying time evolution and optimization with matrix product states. *Phys. Rev. B*, **94**, 165116 (2016).

[92] Sorber, L., Van Barel, M., and de Lathauwer, L. Unconstrained optimization of real functions in complex variables. *SIAM J. Optim.*, **22**, 3, 879 (2012).

[93] Eckart, C. and Young, G. The approximation of one matrix by another of lower rank. *Psychometrika*, **1**, 3, 211 (1936).

[94] Legeza, O. and Fáth, G. Accuracy of the density-matrix renormalization-group method. *Phys. Rev. B*, **53**, 14349.

[95] Legeza, O., Röder, J., and Hess, B. A. Controlling the accuracy of the density-matrix renormalization-group method: The dynamical block state selection approach. *Phys. Rev. B*, **67**, 125114 (2003).

[96] Chan, G. K.-L. and Head-Gordon, M. Highly correlated calculations with a polynomial cost algorithm: A study of the density matrix renormalization group. *J. Chem. Phys.*, **116**, 11, 4462 (2002).

[97] Uschmajew, A. Local convergence of the alternating least squar algorithm for canonical tensor approximation. *SIAM J. Matrix Anal. Appl.*, **33**, 2, 639 (2012).

[98] Khrulkov, V. and Oseledets, I. V. Desingularization of bounded-rank matrix sets (2016). arXiv:1612.03973.

[99] Kutschan, B. Tangent cones to TT varieties (2017). arXiv:1705.10152.

[100] Schneider, R. and Uschmajew, A. Convergence results for projected line-search methods on varieties of low-rank matrices via Łojasiewicz inequality. *SIAM J. Optim.*, **25**, 1, 622 (2015).

[101] Haegeman, J., Pirvu, B., Weir, D. J., Cirac, J. I., Osborne, T. J., Verschelde, H., and Verstraete, F. Variational matrix product ansatz for dispersion relations. *Phys. Rev. B*, **85**, 100408 (2012).

[102] Krumnow, C., Veis, L., Legeza, O., and Eisert, J. Fermionic orbital optimization in tensor network states. *Phys. Rev. Lett.*, **117**, 210402 (2016).

[103] Legeza, O., Veis, L., Poves, A., and Dukelsky, J. Advanced density matrix renormalization group method for nuclear structure calculations. *Phys. Rev. C*, **92**, 051303 (2015).

[104] Stoudenmire, E. M. and White, S. R. Studying two-dimensional systems with the density matrix renormalization group. *Annu. Rev. Condens. Matter Phys.*, **3**, 111 (2012).

[105] Ma, Y. and Ma, H. Assessment of various natural orbitals as the basis of large active space density-matrix renormalization group calculations. *J. Chem. Phys.*, **138**, 22, 224105 (2013).

[106] Rissler, J., Noack, R. M., and White, S. R. Measuring orbital interaction using quantum information theory. *Chem. Phys.*, **323**, 23, 519 (2006).

[107] Barcza, G., Legeza, O., Marti, K. H., and Reiher, M. Quantum-information analysis of electronic states of different molecular structures. *Phys. Rev. A*, **83**, 012508 (2011).

[108] Fertitta, E., Paulus, B., Barcza, G., and Legeza, O. Investigation of metal-insulator-like transition through the *ab initio* density matrix renormalization group approach. *Phys. Rev. B*, **90**, 245129 (2014).

[109] Olivares-Amaya, R., Hu, W., Nakatani, N., Sharma, S., Yang, J., and Chan, G. K.-L. The ab-initio density matrix renormalization group in practice. *J. Chem. Phys.*, **142**, 3, 034102 (2015).

[110] Ghosh, D., Hachmann, J., Yanai, T., and Chan, G. K.-L. Orbital optimization in the density matrix renormalization group, with applications to polyenes and β-carotene. *J. Chem. Phys.*, **128**, 14, 144117 (2008).

[111] Abrudan, T., Eriksson, J., and Koivunen, V. Conjugate gradient algorithm for optimization under unitary matrix constraint. *Signal Process.*, **89**, 9, 1704 (2009).

[112] Gebhard, F. *The Mott Metal-Insulator Transition: Models and Methods*, *Springer Tracts in Modern Physics*, vol. 137. Springer, Berlin (1997).

[113] Legeza, O. and Sólyom, J. Optimizing the density-matrix renormalization group method using quantum information entropy. *Phys. Rev. B*, **68**, 195116 (2003).

[114] Hairer, E., Wanner, G., and Lubich, C. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, *Springer Series in Computational Mathematics*, vol. 31. Springer, Berlin (2002).

[115] Meyer, H.-D., Manthe, U., and Cederbaum, L. S. The multi-configurational time-dependent Hartree approach. *Chem. Phys. Lett.*, **165**, 1, 73 (1990).

[116] Fasshauer, E. and Lode, A. U. J. Multiconfigurational time-dependent Hartree method for fermions: Implementation, exactness, and few-fermion tunneling to open space. *Phys. Rev. A*, **93**, 033635 (2016).

[117] Schreiber, M., Hodgman, S. S., Bordia, P., Lüschen, H. P., Fischer, M. H., Vosk, R., Altman, E., Schneider, U., and Bloch, I. Observation of many-body localization of interacting fermions in a quasirandom optical lattice. *Science*, **349**, 6250, 842 (2015).

[118] Bernardo, J. M. The concept of exchangeability and its applications. *Far East Journal of Mathematical Sciences*, **4**, 111 (1996).

[119] Diaconis, P. Recent progress on de Finettis notions of exchangeability. In *Bayesian statistics 3*, 111. Oxford Univ. Press (1988).

[120] Størmer, E. Symmetric states of infinite tensor products of C*-algebras. *J. Funct. Anal.*, **3**, 1, 48 (1969).

[121] Hudson, R. L. and Moody, G. R. Locally normal symmetric states and an analogue of de Finetti's theorem. *Z. Wahrscheinlichkeitstheor. Verwandte Geb.*, **33**, 4, 343 (1976).

[122] Caves, C. M., Fuchs, C. A., and Schack, R. Unknown quantum states: The quantum de Finetti representation. *J. Math. Phys.*, **43**, 9, 4537 (2002).

[123] Renner, R. Symmetry of large physical systems implies independence of subsystems. *Nat. Phys.*, **3**, 645 (2007).

[124] Renner, R. Security of quantum key distribution. *Int. J. Quantum Inf.*, **06**, 01, 1 (2008).

[125] Renner, R. and Cirac, J. I. de Finetti representation theorem for infinite-dimensional quantum systems and applications to quantum cryptography. *Phys. Rev. Lett.*, **102**, 110504 (2009).

[126] König, R. and Renner, R. A de Finetti representation for finite symmetric quantum states. *J. Math. Phys.*, **46**, 12, 122108 (2005).

[127] Christandl, M., König, R., Mitchison, G., and Renner, R. One-and-a-half quantum de Finetti theorems. *Commun. Math. Phys.*, **273**, 2, 473 (2007).

[128] König, R. and Mitchison, G. A most compendious and facile quantum de Finetti theorem. *J. Math. Phys.*, **50**, 1, 012105 (2009).

[129] Brandão, F. G. S. L., Christandl, M., and Yard, J. Faithful squashed entanglement. *Commun. Math. Phys.*, **306**, 3, 805 (2011).

[130] Li, K. and Smith, G. Quantum de Finetti theorem under fully-one-way adaptive measurements. *Phys. Rev. Lett.*, **114**, 160503 (2015).

[131] Brandão, F. G. S. L. and Harrow, A. W. Quantum de Finetti theorems under local measurements with applications. *Commun. Math. Phys.*, **353**, 2, 469 (2017).

[132] Brandão, F. G. S. L. and Plenio, M. B. A generalization of quantum Stein's lemma. *Commun. Math. Phys.*, **295**, 3, 791 (2010).

[133] Brandão, F. G. S. L. and Harrow, A. W. Product-state approximations to quantum states. *Commun. Math. Phys.*, **342**, 1, 47 (2016).

[134] Trimborn, F., Werner, R. F., and Witthaut, D. Quantum de Finetti theorems and mean-field theory from quantum phase space representations. *J. Phys. A*, **49**, 13, 135302 (2016).

[135] Crismale, V. and Fidaleo, F. De Finetti theorem on the CAR algebra. *Commun. Math. Phys.*, **315**, 1, 135 (2012).

[136] Kraus, C. V., Lewenstein, M., and Cirac, J. I. Ground states of fermionic lattice Hamiltonians with permutation symmetry. *Phys. Rev. A*, **88**, 022335 (2013).

[137] Krumnow, C., Zimborás, Z., and Eisert, J. A fermionic de Finetti theorem (2017). arXiv:1708.01266.

[138] De Melo, F., Ćwikliński, P., and Terhal, B. M. The power of noisy fermionic quantum computation. *New J. Phys.*, **15**, 1, 013015 (2013).

[139] Hudson, R. L. A quantum-mechanical central limit theorem for anti-commuting observables. *J. Appl. Probab.*, **10**, 3, 502 (1973).

[140] Greplová, E. *Quantum Information with Fermionic Gaussian States.* Master's thesis, LMU München (2013).

[141] Bloch, I., Dalibard, J., and Nascimbene, S. Quantum simulations with ultracold quantum gases. *Nat. Phys.*, **8**, 267 (2012).

[142] Cheneau, M., Barmettler, P., Poletti, D., Endres, M., Schausz, P., Fukuhara, T., Gross, C., Bloch, I., Kollath, C., and Kuhr, S. Light-cone-like spreading of correlations in a quantum many-body system. *Nature*, **481**, 484 (2012).

[143] Schneider, U., Hackermuller, L., Ronzheimer, J. P., Will, S., Braun, S., Best, T., Bloch, I., Demler, E., Mandt, S., Rasch, D., and Rosch, A. Fermionic transport and out-of-equilibrium dynamics in a homogeneous Hubbard model with ultracold atoms. *Nat. Phys.*, **8**, 213 (2012).

[144] Trotzky, S., Chen, Y.-A., Flesch, A., McCulloch, I. P., Schollwock, U., Eisert, J., and Bloch, I. Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas. *Nat. Phys.*, **8**, 325 (2012).

[145] Langen, T., Geiger, R., Kuhnert, M., Rauer, B., and Schmiedmayer, J. Local emergence of thermal correlations in an isolated quantum many-body system. *Nat. Phys.*, **9**, 640 (2013).

[146] Polkovnikov, A., Sengupta, K., Silva, A., and Vengalattore, M. *Colloquium:* nonequilibrium dynamics of closed interacting quantum systems. *Rev. Mod. Phys.*, **83**, 863 (2011).

[147] Gogolin, C. *Equilibration and thermalization in quantum systems.* Ph.D. thesis, Freie Universitt Berlin (2014).

[148] Gogolin, C. and Eisert, J. Equilibration, thermalisation, and the emergence of statistical mechanics in closed quantum systems. *Rep. Prog. Phys.*, **79**, 5, 056001 (2016).

[149] Singh, N. How and why does statistical mechanics work (2011). arXiv:1103.4003.

[150] Landau, L. D. and Lifschitz, E. M. *Statistische Physik Teil 1*. Harri Deutsch, Frankfurt (2008).

[151] Reimann, P. Foundation of statistical mechanics under experimentally realistic conditions. *Phys. Rev. Lett.*, **101**, 190403 (2008).

[152] Linden, N., Popescu, S., Short, A. J., and Winter, A. Quantum mechanical evolution towards thermal equilibrium. *Phys. Rev. E*, **79**, 061103 (2009).

[153] Bocchieri, P. and Loinger, A. Quantum recurrence theorem. *Phys. Rev.*, **107**, 337 (1957).

[154] Wilming, H., Goihl, M., Krumnow, C., and Eisert, J. Towards local equilibration in closed interacting quantum many-body systems (2017). arXiv:1704.06291.

[155] De Oliveira, T. R., Charalambous, C., Jonathan, D., Lewenstein, M., and Riera, A. Equilibration time scales in closed many-body quantum systems (2017). arXiv:1704.06646.

[156] Cramer, M., Dawson, C. M., Eisert, J., and Osborne, T. J. Exact relaxation in a class of nonequilibrium quantum lattice systems. *Phys. Rev. Lett.*, **100**, 030602 (2008).

[157] Flesch, A., Cramer, M., McCulloch, I. P., Schollwöck, U., and Eisert, J. Probing local relaxation of cold atoms in optical superlattices. *Phys. Rev. A*, **78**, 033608 (2008).

[158] Cramer, M. and Eisert, J. A quantum central limit theorem for non-equilibrium systems: exact local relaxation of correlated states. *New J. Phys.*, **12**, 5, 055020 (2010).

[159] Rigol, M., Dunjko, V., and Olshanii, M. Thermalization and its mechanism for generic isolated quantum systems. *Nature*, **452**, 854 (2008).

[160] D'Alessio, L., Kafri, Y., Polkovnikov, A., and Rigol, M. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Adv. Phys.*, **65**, 3, 239 (2016).

[161] De Palma, G., Serafini, A., Giovannetti, V., and Cramer, M. Necessity of eigenstate thermalization. *Phys. Rev. Lett.*, **115**, 220401 (2015).

[162] Riera, A., Gogolin, C., and Eisert, J. Thermalization in nature and on a quantum computer. *Phys. Rev. Lett.*, **108**, 080402 (2012).

[163] Pérez-García, D., Verstraete, F., Wolf, M. M., and Cirac, J. I. PEPS as unique ground states of local Hamiltonians. *Quantum Info. Comput.*, **8**, 6&7, 650 (2008).

[164] Hastings, M. B. and Koma, T. Spectral gap and exponential decay of correlations. *Commun. Math. Phys.*, **265**, 3, 781 (2006).

[165] Nachtergaele, B. and Sims, R. Lieb-Robinson bounds and the exponential clustering theorem. *Commun. Math. Phys.*, **265**, 1, 119 (2006).

[166] Kliesch, M., Gogolin, C., Kastoryano, M. J., Riera, A., and Eisert, J. Locality of temperature. *Phys. Rev. X*, **4**, 031019 (2014).

[167] Hastings, M. B. Locality in quantum systems (2010). arXiv:1008.5137.

[168] Hastings, M. B. Decay of correlations in Fermi systems at nonzero temperature. *Phys. Rev. Lett.*, **93**, 126402 (2004).

[169] Anderson, P. W. Absence of diffusion in certain random lattices. *Phys. Rev.*, **109**, 1492 (1958).

[170] Gluza, M., Krumnow, C., Friesdorf, M., Gogolin, C., and Eisert, J. Equilibration via Gaussification in fermionic lattice systems. *Phys. Rev. Lett.*, **117**, 190602 (2016).

[171] Haller, E., Hudson, J., Kelly, A., Cotta, D. A., Peaudecerf, B., Bruce, G. D., and Kuhr, S. Single-atom imaging of fermions in a quantum-gas microscope. *Nat. Phys.*, **11**, 738 (2015).

[172] Van der Walt, S., Colbert, S. C., and Varoquaux, G. The numpy array: A structure for efficient numerical computation. *IEEE Comput. Sci. Eng.*, **13**, 2, 22 (2011).

[173] Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.*, **17**, 87 (1975).

[174] Leininger, M. L., Sherrill, C. D., Allen, W. D., and Schaefer, H. F. Systematic study of selected diagonalization methods for configuration interaction matrices. *J. Comput. Chem.*, **22**, 13, 1574 (2001).

[175] Murnaghan, F. D. *The unitary and rotation groups.* Spartan Books, Washington (1962).

[176] Edelman, A., Arias, T. A., and Smith, S. T. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, **20**, 2, 303 (1998).

[177] Manton, J. H. Optimization algorithms exploiting unitary constraints. *IEEE Trans. Signal Process.*, **50**, 3, 635 (2002).

[178] Usevich, K. and Markovsky, I. Optimization on a Grassmann manifold with application to system identification. *Automatica*, **50**, 6, 1656 (2014).

[179] Knuth, D. E. Semi-optimal bases for linear dependencies. *Linear Multilinear Algebra*, **17**, 1, 1 (1985).

[180] Mehrmann, V. and Poloni, F. Doubling algorithms with permuted Lagrangian graph bases. *SIAM J. Matrix Anal. Appl.*, **33**, 3, 780 (2012).

A. Wick's Theorem and Equivalence of Definitions of Gaussian States in Finite Systems

In this appendix we present a proof of Wick's theorem for finite fermionic systems and comment on the equivalence of the different definitions of Gaussian states.

A.1. Wick's Theorem

Given $M \in \mathbb{N}$, let $f_j^\dagger$ and f_j for $j \in [M]$ denote fermionic creation and annihilation operators fulfilling the CAR and define the corresponding Majorana operators m_k for $k \in [2M]$. Let $\xi_j \in [-1, 1]$ for $j \in [M]$ and

$$\rho = \prod_{j=1}^{M} \left(\frac{1}{2}\mathbb{1} + \frac{i}{2}\xi_j m_{2j-1} m_{2j} \right) \tag{A.1}$$

be a Gaussian state. We can then calculate the expectation value of an arbitrary product of Majorana operators $m_{j_1} \dots m_{j_{2r}}$ according to Wick's theorem.

Theorem 11 (Wick's theorem)**.** *Given the system and Gaussian state ρ defined above we obtain for $r \in \mathbb{N}$*

$$\operatorname{tr}(m_{j_1} \dots m_{j_{2r}} \rho) = \operatorname{Pf}(\gamma^{(m)}[j_1, \dots, j_{2r}]) = \sum_{P \in \mathcal{P}_2([2r])} \operatorname{sign}(\mathrm{P}) \prod_{(p_1,p_2)\in P} \operatorname{tr}(m_{j_{p_1}} m_{j_{p_2}} \rho), \tag{A.2}$$

with the correlation matrix

$$\gamma^{(m)}[j_1, \dots, j_{2r}]_{a,b} = \begin{cases} \operatorname{tr}(m_{j_a} m_{j_b} \rho) & \text{if } a < b \\ -\operatorname{tr}(m_{j_b} m_{j_a} \rho) & \text{if } b < a \\ 0 & \text{else} \end{cases}, \tag{A.3}$$

for $a, b \in [2r]$.

Note that if all operators $m_{j_1} \dots m_{j_{2r}}$ are different, the entries of the correlation matrix are contained in the covariance matrix, as $\gamma^{(m)}[j_1, \dots, j_{2r}]_{a,b} = \operatorname{tr}([m_{j_a}, m_{j_b}]\rho)/2$.

Proof. We assume that the operators $m_{j_1}, \dots, m_{j_{2r}}$ are ordered such that $j_1 \leq j_2 \leq \dots \leq j_{2r}$. Define the clusters $C_k = \{l | j_l \in \{2k-1, 2k\}\}$ for $k \in [M]$. It is easy to see

that the expectation value decouples over the different clusters due to the structure of ρ such that

$$\mathrm{tr}(m_{j_1}\dots m_{j_{2r}}\rho)=\prod_{k=1}^{M}\mathrm{tr}\left(\rho\prod_{l\in C_k}m_{j_l}\right). \tag{A.4}$$

If any cluster C_k has an odd size, the expectation value yields zero, in accordance with Eq. (A.3) as for every permutation π there exists at least one factor with $j_{\pi(2l-1)}$ and $j_{\pi(2l)}$ belonging to different clusters which yields $\mathrm{tr}(m_{j_{\pi(2l-1)}}m_{j_{\pi(2l)}}\rho)=0$. Given any non-empty cluster C_k. We prove by induction that $\mathrm{tr}(\rho\prod_{l\in C_k}m_{j_l})=\mathrm{Pf}(\gamma^{(m)}[(j_l)_{l\in C_k}])$. If $C_k=\{l,l+1\}$ we obtain

$$\mathrm{tr}\left(m_{j_l}m_{j_{l+1}}\rho\right)=\mathrm{Pf}\begin{pmatrix}0 & \mathrm{tr}\left(m_{j_l}m_{j_{l+1}}\rho\right)\\ -\mathrm{tr}\left(m_{j_l}m_{j_{l+1}}\rho\right) & 0\end{pmatrix}, \tag{A.5}$$

which is true by the definition of the Pfaffian. Assuming that the claim is true for $|C_k|=2n$ we consider $|C_k|=2n+2$. Choose $a\in C_k$ such that $m_{j_a}=m_{j_{a+1}}$. Note that such an a exists as $|C_k|>2$. We obtain then from the anti-commutation relation of Majorana operators that

$$\mathrm{tr}\left(\rho\prod_{l\in C_k}m_{j_l}\right)=\mathrm{tr}\left(\rho\prod_{l\in C_k\setminus\{a,a+1\}}m_{j_l}\right). \tag{A.6}$$

Defining the matrix $B\in\mathbb{R}^{|C_k|\times|C_k|}$ with $B_{i,j}=\delta_{i,j}-\delta_{i,a}\delta_{j,a+1}$ we obtain that

$$(B^T\gamma^{(m)}[(j_l)_{l\in C_k}]B)_{x,y}=\begin{cases}\gamma^{(m)}[(j_l)_{l\in C_k}]_{x,y} & \text{if } x,y\neq a+1\\ 1 & \text{if } x=a,y=a+1\\ -1 & \text{if } x=a+1,y=a\\ 0 & \text{else}\end{cases}. \tag{A.7}$$

By this only terms with $(a,a+1)\in P$ contribute in the expansion of the Pfaffian and we obtain

$$\mathrm{Pf}(\gamma^{(m)}[(j_a)_{a\in C_k}])=\frac{1}{\det(B)}\mathrm{Pf}(B^T\gamma^{(m)}[(j_a)_{a\in C_k}]B) \tag{A.8}$$

$$=\mathrm{Pf}(\gamma^{(m)}[(j_a)_{a\in C_k\setminus\{l,l+1\}}]))=\mathrm{tr}\left(\rho\prod_{l\in\setminus C_k\{a,a+1\}}m_{j_l}\right), \tag{A.9}$$

as $\det(B)=1$ and the last equality corresponds to the assumption of the induction step. Using $\mathrm{Pf}(A\oplus B)=\mathrm{Pf}(A)\,\mathrm{Pf}(B)$ finishes the proof for ordered indices. Note that if $j_1,\dots j_{2r}$ are not ordered, we can always order them without changing the relative order of any j_a and j_b with $j_a=j_b$ which introduces a sign only. Changing the order of rows and columns of the correlation matrix accordingly, introduces the same sign for the Pfaffian. □

Lemma 12. *This result is stable under linear transformations, meaning that for any* $T \in \mathbb{C}^{2r\times 2M}$ *with*

$$c_k = \sum_{j=1}^{2M} T_{k,j} m_j \tag{A.10}$$

we obtain for a Gaussian state ρ

$$\mathrm{tr}(c_1 \dots c_{2r}\rho) = \mathrm{Pf}(\gamma^{(c)}[1,\dots,2r]). \tag{A.11}$$

Wick's theorem therefore applies to any set of Majorana and fermionic creation and annihilation operators.

Proof. The proof follows directly from inserting the definition of the c's in the expectation value, applying Wick's theorem there and reorganizing the terms appropriately. □

A.2. Definitions of Gaussian States

As highlighted in Sec. 2.4, fermionic Gaussian states have different equivalent definitions which highlight individual important aspects.

In the previous section we have seen that any state which can be brought into the form

$$\rho = \prod_{j=1}^{M} \left(\frac{1}{2}\mathbb{1} + \frac{i}{2}\xi_j m_{2j-1} m_{2j} \right), \tag{A.12}$$

using mode transformations only fulfills Wick's theorem. On the other hand let ρ be a state fulfilling Wick's theorem then ρ is obviously fully characterized by its covariance matrix. A state of the form (A.12) however can represent any covariance matrix possible. In addition it fulfills, as we have seen, Wick's theorem such that we can find for any state ρ fulfilling Wick's theorem a state ρ' of the form in Eq. (A.12) with the same expectation values. As the the collection of all expectation values uniquely defines a state, this proves the equivalence of defining a state to be Gaussian if it is of the form (A.12) or fulfills Wick's theorem.

As we do not use the Grassmann calculus in this thesis we only mention that Gaussian states as defined in Eq. (A.12) have a Gaussian Grassmann representation and vice versa (see for instance [25] for details).

Furthermore, Gaussian states maximize the von Neumann entropy if the second moments are kept fixed. Given a ρ and its Gaussified version ρ_G, meaning that ρ_G is Gaussian and $\gamma(\rho) = \gamma(\rho_G)$. Written in the correct basis, ρ_G is of the form (A.12). Denoting by $f_j^\dagger$ and f_j the fermionic creation and annihilation operators associated to the natural modes m_k of ρ_G we find that Eq. (A.12) takes the form

$$\rho_G = \prod_{j=1}^{M} \left(\frac{1-\xi_j}{2}(\mathbb{1} - n_j) + \frac{1+\xi_j}{2} n_j \right), \tag{A.13}$$

with $n_j = f_j^\dagger f_j$. Hence, ρ_G is diagonal in this basis and the product of commuting operators. Based in this observation we conclude

$$\operatorname{tr}\Big[\rho \ln(\rho_G)\Big] = \sum_{j=1}^{M} \operatorname{tr}\left[\rho \ln\left(\frac{1-\xi_j}{2}(\mathbb{1}-n_j) + \frac{1+\xi_j}{2} n_j\right)\right] \tag{A.14}$$

$$= \sum_{j=1}^{M} \left[\ln\left(\frac{1-\xi_j}{2}\right) \operatorname{tr}[\rho(\mathbb{1}-n_j)] + \ln\left(\frac{1+\xi_j}{2}\right) \operatorname{tr}[\rho(n_j)]\right] \tag{A.15}$$

$$= \sum_{j=1}^{M} \left[\frac{1-\xi_j}{2} \ln\left(\frac{1-\xi_j}{2}\right) + \frac{1+\xi_j}{2} \ln\left(\frac{1+\xi_j}{2}\right)\right] = \operatorname{tr}\Big[\rho_G \ln(\rho_G)\Big], \tag{A.16}$$

where from the first to the second line we exploited the fact that $\mathbb{1} - n_j$ and n_j are projectors on disjoint subspaces as well as that the expression stays valid in the limit $\xi_j = \pm 1$ as ρ has the same second moments as ρ_G and we use the convention $0 \ln 0 = 0$. The positivity of the relative entropy implies now

$$0 \leq -\operatorname{tr}\Big[\rho \ln(\rho_G)\Big] + \operatorname{tr}\Big[\rho \ln(\rho)\Big] = -\operatorname{tr}\Big[\rho_G \ln(\rho_G)\Big] + \operatorname{tr}\Big[\rho \ln(\rho)\Big] = S_{\mathrm{vN}}(\rho_G) - S_{\mathrm{vN}}(\rho). \tag{A.17}$$

As this inequality applies to any state ρ with the given covariance matrix, ρ_G does indeed maximize S_{vN}. As the relative entropy of ρ and ρ_G is zero if and only if $\rho = \rho_G$ [16, Thm. 11.7] we also find that this maximum is unique.

B. Details on the Implementation and Costs of the Two-Site DMRG with and without Mode Transformations

In this appendix we present details on the implementation of a standard two-site DMRG. In particular we explain how the matrices $H^{(m,2)}$ of a micro step are decomposed which is an essential step for the application of DMRG to realistic systems. Furthermore, we elaborate on the computational cost of DMRG and show that the mode transformations can be indeed incorporated for free in an existing two-site DMRG algorithm.

B.1. Presummed Operators and Costs of the Two Site DMRG

Assume we have a fermionic system of V sites with M modes. Let us focus on how to decompose $H^{(m,2)}$ as explained in Fig. 3.8 obtained from the generic Hamiltonian of interacting fermions with M modes

$$H = \sum_{i,j=1}^{M} t_{i,j} f_i^\dagger f_j + \sum_{i,j,k,l=1}^{M} v_{i,j,k,l} f_i^\dagger f_j^\dagger f_k f_l \tag{B.1}$$

for a two-site DMRG. Fix a site $m \in [V-1]$ and define the sets $S_L = [m-1]$, $S_C = \{m, m+1\}$ and $S_R = [m+1]^c$ and assume we have an initial MPS with components $(A^{[j]})_{j\in[V]}$. Let us denote by T_L and T_R the modes associated to the sites S_L and S_R respectively. The Jordan-Wigner transformed expression of each term $f_i^\dagger f_j^\dagger f_k f_l$ can be written as an MPO with bond dimension 1. We further define for an MPO of bond dimension 1 with components $(O^{[j]})_{j\in[V]}$ and $I = L, R, C$ the truncation maps

$$\Gamma^{(I)}((O^{[j]})_{j\in[V]}) = \bigotimes_{j\in S_I} \sum_{i_j,k_j} O^{[j]}_{i_j,k_j} |i_j\rangle\langle k_j| \tag{B.2}$$

which truncate the MPO down to the region S_L, S_C or S_R by discarding the remaining part where $|i_j\rangle$ for $j \in [d_j]$ denotes an orthonormal basis of $\mathcal{H}^{[j]}$. We further define for $I = L, R$ the partial contraction of truncated operators with the current MPS

$$O^{(I)}(\,\cdot\,) = \Big(\langle (A^{[j]})_{j\in S_I}|_{\alpha_1} \; \Gamma^{(I)}(\,\cdot\,) \; |(A^{[j]})_{j\in S_I}\rangle_{\alpha_2} \Big)_{\alpha_1,\alpha_2} \tag{B.3}$$

which is either an $r_{m-1} \times r_{m-1}$ or an $r_{m+1} \times r_{m+1}$ matrix, for $I = L, R$ respectively. Note that $O^{(I)}(\,\cdot\,)^\dagger = O^{(I)}(\,\cdot^\dagger\,)$. Last, we define for $I = L, R$ and $k, l \in T_I{}^c$ the partially

contracted operators which are presummed according to the weights of the Hamiltonian

$$P^{(I,0)} = \sum_{j,k\in T_I} \frac{1}{2} t_{j,k}\, O^{(I)}(f_j^\dagger f_k) + \sum_{i,j,k,l\in T_I} \frac{1}{2} v_{i,j,k,l}\, O^{(I)}(f_i^\dagger f_j^\dagger f_k f_l), \tag{B.4}$$

$$P_l^{(I,1)} = \sum_{j\in T_I} t_{j,l}\, O^{(I)}(f_j^\dagger f_l), \tag{B.5}$$

$$P_l^{(I,1')} = \sum_{i,j,k\in T_I} (v_{i,j,k,l} - v_{i,j,l,k})\, O^{(I)}(f_i^\dagger f_j^\dagger f_k f_l), \tag{B.6}$$

$$P_{k,l}^{(I,2)} = \sum_{i,j\in T_I} v_{i,j,k,l}\, O^{(I)}(f_i^\dagger f_j^\dagger f_k f_l), \tag{B.7}$$

$$P_{k,l}^{(I,2')} = \sum_{i,j\in T_I} (v_{i,k,j,l} - v_{k,i,j,l} - v_{i,k,l,j} + v_{k,i,l,j})\, O^{(I)}(f_i^\dagger f_k^\dagger f_j f_l). \tag{B.8}$$

Roughly speaking, $P^{(I,0)}$ contains all terms of H that are supported on the side $I = L, R$ only. $P^{(I,1)}$ and $P^{(I,1')}$ as well as $P^{(I,2)}$ and $P^{(I,2')}$ combine all terms that act the same on the side opposed to I, i.e. $P_j^{(L,1)}$ contains all terms that act in the center and on the right side of the system with f_j only etc. For simplicity of the notation let us assume that the modes are ordered such that $\{1,2\} \subset {T_R}^c$ and $\{M-1, M\} \subset {T_L}^c$. We obtain then the decomposition

$$\begin{aligned} H^{(k,2)} \to\ & (P^{(L,0)}, \mathbb{1}, \mathbb{1}) + (\mathbb{1}, \mathbb{1}, P^{(R,0)}) + \sum_{k\in {T_L}^c} (P_k^{(L,1)}, \Gamma^{(C)}(f_1^\dagger f_k), O^{(R)}(f_1^\dagger f_k)) \\ & + \sum_{k\in {T_L}^c} (P_k^{(L,1')}, \Gamma^{(C)}(f_1^\dagger f_2^\dagger f_k f_1), O^{(R)}(f_1^\dagger f_2^\dagger f_k f_1)) \\ & + \sum_{k\in {T_R}^c} (O^{(L)}(f_M^\dagger f_{M-1}^\dagger f_k f_M), \Gamma^{(C)}(f_M^\dagger f_{M-1}^\dagger f_k f_M), P_k^{(R,1')}) \\ & + \sum_{k,l\in {T_L}^c} (P_{k,l}^{(L,2)}, \Gamma^{(C)}(f_1^\dagger f_2^\dagger f_k f_l), O^{(R)}(f_1^\dagger f_2^\dagger f_k f_l)) \\ & + \sum_{k\in {T_L}^c} (P_{k,l}^{(L,2')}, \Gamma^{(C)}(f_1^\dagger f_k^\dagger f_1 f_l), O^{(R)}(f_1^\dagger f_k^\dagger f_1 f_l)) + \text{h.c.} \end{aligned} \tag{B.9}$$

where we denoted terms in the decomposition by triples $(H_L^{(k,2)}, H_C^{(k,2)}, H_R^{(k,2)})$ as explained in Fig. 3.8 and the +h.c. denotes here that we add the triples of the conjugated terms, i.e., $(H_L^{(k,2)\dagger}, H_C^{(k,2)\dagger}, H_R^{(k,2)\dagger})$. Note that in the above expression, the operators f_1, f_2, f_M and f_{M-1} are needed in order to guarantee the truncated operators to involve the correct signs and Z operators from the Jordan-Wigner transformation – we could have used any other fermionic annihilation operators supported on the corresponding part of the system instead.

For a fixed site m we then need the operators $P^{(I,0)}, P_l^{(I,1)}, P_l^{(I,1')}, P_{k,l}^{(I,2)}, P_{k,l}^{(I,2')}$ for $I \in \{L, R\}$ and $k, l \in {T_I}^c$ as well as $O^{(R)}(f_1^\dagger f_k)), O^{(R)}(f_j^\dagger f_k)), O^{(R)}(f_j f_k))$ for $j, k \in T_R$ and $O^{(L)}(f_M^\dagger f_k)), O^{(L)}(f_j^\dagger f_k)), O^{(L)}(f_j f_k))$ for $j, k \in T_L$. Note that from $O^{(R)}(f_j^\dagger f_k))$

we can reconstruct terms like $O^{(R)}(f_1^\dagger f_j^\dagger f_1 f_k)$). During a DMRG we store all described operators for all sites m fixed in the beginning on the disk which requires as disk storage of $\mathcal{O}(M^3 r^2)$. After updating the MPS based on $\mathbf{X}_{\text{opt}}$, we load half of the operators needed into the RAM and compute the remaining ones from the current presummed operators in a time scaling as $\mathcal{O}(M^2(dr^3+d^2r^2)+M^3r^2)$ with $r=\max(\{r_j|j\in[V-1]\})$ and $d=\max(\{d|j\in[V]\})$.[1] Note that if additional symmetries are present in the couplings t and v, the number of presummed operators needed can be reduced – for instance in many application t will not mix different types of particles etc.

B.2. Mode Transformation of Presummed Operators and Costs

In Alg. 3 we augment the micro iteration steps of a two-site DMRG with local mode transformation. For an update of the sites m and $m+1$ we find a transformation $U_{\text{opt}}^{(m,2)} \in U(2p)$ which gives rise to a global transformation $V^{(m)} = \mathbb{1}_{pm} \oplus U_{\text{opt}}^{(m,2)} \oplus \mathbb{1}_{M-(m+2)p}$. In addition to the rotation of the state by $V^{(m)}$ we also transform all operators needed such that expectation values are conserved

$$\langle X_{\text{opt}}|_{m,2}^{(A^{[j]})_j} H |X_{\text{opt}}\rangle_{m,2}^{(A^{[j]})_j} = \langle X_{\text{opt}}(U)|_{m,2}^{(A^{[j]})_j} G(V^{(m)}) H G(V^{(m)\dagger}) |X_{\text{opt}}(U)\rangle_{m,2}^{(A^{[j]})_j}, \quad \text{(B.10)}$$

as $|X_{\text{opt}}(U)\rangle_{m,2}^{(A^{[j]})_j} = G(V^{(m)})|X_{\text{opt}}\rangle_{m,2}^{(A^{[j]})_j}$, which allows for a consistent minimization of the energy over several steps. The transformation of H is achieved by rotating the coefficient tensors as in Eq. (2.37) which can be done in a time scaling as $\mathcal{O}(M^3)$ due to the locality of the transformation. However, the presummed operators from the precious section, are now constructed with respect to the wrong couplings t and v and need to be transformed as well. This can be done as during the construction of the presummed operators we never summed over a site that is affected by $U_{\text{opt}}^{(m,2)}$. We can therefore in principle update all presummed operators using

$$P_l^{(I,1)}(U_{\text{opt}}^{(m,2)}) = \sum_{j=T_C} P_j^{(I,1)} U_{\text{opt}\,j,l}^{(m,2)} \quad \text{(B.11)}$$

$$P_l^{(I,1')}(U_{\text{opt}}^{(m,2)}) = \sum_{j=T_C} P_j^{(I,1')} U_{\text{opt}\,j,l}^{(m,2)} \quad \text{(B.12)}$$

$$P_{k,l}^{(I,2)}(U_{\text{opt}}^{(m,2)}) = \sum_{i,j=T_C} P_{i,j}^{(I,2)} U_{\text{opt}\,i,k}^{(m,2)} U_{\text{opt}\,j,l}^{(m,2)} \quad \text{(B.13)}$$

$$P_{k,l}^{(I,2')}(U_{\text{opt}}^{(m,2)}) = \sum_{i,j=T_C} (U_{\text{opt}}^{(m,2)\dagger})_{k,i} P_{i,j}^{(I,2')} U_{\text{opt}\,j,l}^{(m,2)}. \quad \text{(B.14)}$$

However, in order to avoid to load all operators from the disk in each step, we only update the $I = L/R$ ones for a right/left sweep and update all loaded operators with

[1] Say we sweep to the right and updated the sites k and $k+1$. We can compute for instance $O^{(L)}(f_M^\dagger f_k^\dagger f_j f_M)$ from the previous $O^{(L)}(f_M^\dagger f_j)$ in a time scaling as $\mathcal{O}(dr^3+d^2r^2)$. We further obtain $P^{(L,1')}$ from the previous $P^{(L,2)}$ and $P^{(L,2')}$ etc. The M^3 scaling is due to the update of $P_{k,l}^{(I,2)}$ and $P_{k,l}^{(I,2')}$ as for instance $\mathcal{O}(M)$ many $O^{(I)}(f_i f_j)$ are added to $\mathcal{O}(M^2)$ many $P_{k,l}^{(I,2)}$

an accumulated mode transformation on $T_I \cup T_C$. This accumulated transformation is composed of all local transformations performed since the last time the corresponding loaded operator was needed. The update is then obtained by repeatedly applying the relations above for maximally $\mathcal{O}(M)$ many local rotations such that the transformation of the presummed operators can be calculated in a time scaling as $\mathcal{O}(M^3 r^2)$, i.e. without increasing the computational costs of the DMRG.

C. Comment on the Developed Code Structure for Tensor Network States

Let us in this appendix shortly comment on the code structure developed for simulating non-local quantum systems using the DMRG algorithm in the course of the research documented in this thesis. Due to its immense length (over 8000 lines of code in its current implementation) we will not append the code here and it is instead available on request. We here only highlight the different functionalities and outline the structure implemented. The code was designed and developed by the author of this thesis and implemented in python 2.7 and relies on numpy and scipy [172] routines for linear algebra.

First let us note that the developed code is for all applications less efficient than other state of the art implementations. The aim of the developed code is not superiority over existing implementations which are under development for more than a decade but to allow for maximal flexibility and adaptability to new ideas while being able to simulation relevant realistic systems.

The code allows to approximated ground and excited states by MPS using a k-site DMRG for arbitrary k by implementing the scheme discussed in Sec. 3.2 (the design of the algorithm $k > 2$ is obvious from the two-site case). The tractable Hamiltonians are of the form

$$H = \sum_{q} \sum_{i,j,k,l=1}^{V} v^{(q)}_{i,j,k,l} O^{(1,q)}_i O^{(2,q)}_j O^{(3,q)}_k O^{(4,q)}_l \tag{C.1}$$

with either $O^{(p,q)}_j = \mathbb{1}^{\otimes(j-1)} \otimes O^{(p,q)} \otimes \mathbb{1}^{V-j}$ or $O^{(p,q)}_j = Z^{\otimes(j-1)} \otimes O^{(p,q)} \otimes \mathbb{1}^{V-j}$ for some local operator $O^{(p,q)}$ and to all operators at fixed q common operator Z. This includes most prominently spin systems with general local dimension d, by the Jordan-Wigner transformation discussed in Sec. 2.2.1 spinless and spinfull interacting fermions as well as interacting bosonic systems with truncated local occupation numbers. The computational cost reduce if any $O^{(p,q)} = \mathbb{1}$. The Fermi-Hubbard model in Eq. (2.61) can therefore in real space be described by second order polynomials only by choosing $O^{(1,1)}_j = f^\dagger_j$, $O^{(2,1)}_k = f_k$ and $O^{(1,2)}_j = n_j$, $O^{(2,2)}_k = n_k$. The coefficients tensors $v^{(q)}$ are not restricted and we are therefore not constrained to any locality or spatial dimension. However, we of course have to keep in mind that we impose an MPS approximation which has the in Sec. 3.1.2 discussed limits to its representable correlation patterns. A potential one-dimensional spatial locality in the coefficients and relation of the operators such as $O^{(1,q)} = O^{(2,q)}$ or $O^{(1,q)} = O^{(2,q)\dagger}$ is exploited if present in order to reduce the computational resources needed. In addition we allow for all tractable Hamiltonians to simulate the generated time evolution of the time dependent Schrödinger equation by

implementing the two-site DMRG based scheme described in [68].

We allow for the incorporation of abelian symmetry by implementing symmetric MPS along the lines of Sec. 3.1.4 and are therefore able to perform calculation with a fixed particle number in fermionic systems or fixed magnetization in spin systems. We can then target the few lowest eigenstates of a Hamiltonian respecting this symmetry in any given symmetry sector.

Furthermore, the code contains the various algorithms and schemes presented in Sec. 3.3 for adapting the single-particle basis of fermionic systems while performing a DMRG calculation or time evolution. The resulting advantages are discussed and shown in Fig. 3.15, 3.16, 3.17 and 3.18.

Not all routines needed are provided by standard python implementations. We therefore extend the structure available from python 2.7 by implementing our own sparse tensor format for representing symmetric MPS components including the necessary linear algebra routines. In addition, we use our own implementation for an efficient Davidson eigensolver (adapting the version of [173, 174]) for computing the lowest lying eigenvalues and eigenvectors of a sparse matrix and include the optimization algorithms over Grassmannian manifolds discussed in App. D. Furthermore, we allow for a basic parallelization during the most time consuming parts of the micro step of the DMRG.

D. Optimizing with Unitary Constraints

In this appendix we review a few basics about the optimization under unitary constraints with special focus on Grassmannians. We first introduce the typical manifolds encountered and then review and reference possible optimization techniques with a strong focus on the ones we need for the combination of mode transformation and tensor network states.

D.1. Sets with Unitary Constraints

Unitary constraints in optimization problems can occur to different degrees. One prominent case is of course the optimization over the full unitary manifold

$$U(n) = \{U \in \mathbb{C}^{n\times n} | U^\dagger U = \mathbb{1}_n\}. \tag{D.1}$$

$U(n)$ is the collection of all orthonormal bases of $\mathbb{C}^n$ and has a real dimension of n^2 due to the constraints induced by the unitarity. Note that, in view of the local mode transformations used in Sec. 3.3, ever unitary $U \in U(n)$ can be decomposed into $\mathcal{O}(n^2)$ many unitaries of the form $\mathbb{1}_k \oplus U_{\text{local}} \oplus \mathbb{1}_{n-k-2}$ with $U_{\text{local}} \in U(2)$, where k starts at 0 and is increased or decreased by 1 just like during a DMRG sweep with mode transformation. The existence of such a decomposition follows from [175, p. 9ff] with additional swap gates in order to obtain the desired structure.

A second set we frequently encounter is the collection of isometries on $\mathbb{C}^N$, which form the Stiefel manifold

$$St(p,n) = \{V \in \mathbb{C}^{n\times p} | V^\dagger V = \mathbb{1}_p\}. \tag{D.2}$$

The Stiefel manifold $St(p,n)$ contains all possible bases of all p-dimensional subspaces of $\mathbb{C}^n$ and has a real dimension of $2np - p^2$. We can represent $St(p,n)$ as the quotient $St(p,n) = U(n)/U(n-p)$ which becomes clear by realizing that here $U(n-p)$ simply represents the freedom to choose any basis in the $(n-p)$-dimensional orthogonal complement to the linear space spanned by the columns of $V \in St(p,n)$. Finding a few eigenstates of a Hermitian matrix can for instance be viewed as an optimization problem over a Stiefel manifold.

The third construction we want to discuss here are Grassmannians

$$Gr(p,n) = U(n)/U(p) \times U(n-p). \tag{D.3}$$

The Grassmannian $Gr(p,n)$ collects all p-dimensional subspaces of $\mathbb{C}^n$, i.e., it consists of equivalence classes $[V]$ for $V \in St(p,n)$ with $V_1 \sim V_2$ if the columns of $V_1 \in St(p,n)$ span the same subspace of $\mathbb{C}^n$ as the columns of $V_2 \in St(p,n)$ and $Gr(p,n)$ has a real

dimension of $2(n-p)p$. Apart from our encounter of the Grassmannian in Sec. 3.3, they appear for instance in Hartree-Fock calculations as there internal basis changes in the occupied and in the unoccupied orbitals lead to no change in the energy. Put differently, the set of all Slater determinants with a particle number N in M modes, i.e. the corresponding set of pure Gaussian states, is isomorphic to the Grassmannian $Gr(N, M)$.

D.2. Optimization over Grassmannians

We will not describe all algorithms in full detail but rather point to the corresponding references from which they can be easily obtained and only describe the main concepts used.

The numerical solution of optimization problems gets easier the more information about the problem we can provide to the solver, e.g. an optimization set containing no or few redundant parameters, information about the cost function such as gradients or higher derivatives etc. For the problem in Eq. (3.32) for instance we can easily calculate the gradient analytically such that a gradient based optimzation scheme over the unitaries is most favorable, whereas the problem Eq. (3.33) typically does not allow for an analytic computation of the gradient, hence gradient-free methods are needed. The problem in Eq. (3.32) can therefore be solved using a conjugate gradient algorithm over $U(n)$ described in [111].

For optimizing over Grassmannians we choose different strategies although formulations of for instance the steepest descent method are available [176, 177]. For the special case $Gr(1, 2)$ we are able to find a global parametrization in terms two real parameters. Such a parameterization can be derived from the Euler angle decomposition of a general 2×2 unitary which states that $\forall U \in U(2)$ we find $\theta_1, \theta_2, \theta_3, \theta_4 \in [0, 2\pi]$ with

$$U = \exp(i\theta_1 Z) \exp(i\theta_2 Y) \exp(i\theta_3 Z) \exp(i\theta_4 \mathbb{1}_2). \tag{D.4}$$

It is then clear that two unitaries with the same θ_1 and θ_2 are equivalent in the sense of the equivalence relation of $U(2)/U(1) \times U(1)$. Instead of optimizing over $Gr(1, 2)$ we therefore choose to optimize over $\{\exp(i\theta_1 Z) \exp(i\theta_2 Y) | \theta_1, \theta_2 \in [0, 2\pi]\}$ which can be done easily using for instance the Nelder-Mead method for which numpy [172] provides a ready to use implementation.

In higher dimensional settings, Euler angle decompositions yield only an over-parametrization of the Grassmannian and the direct optimization becomes quickly inefficient and unstable such that we prefer to use a more elaborate scheme. As the gradient of our cost function can not be calculated analytically we found in our applications the scheme presented in [178] to be more flexible and efficient than a steepest descent method. Here one uses the fact [179] that any element of $Gr(p, n)$ can be represented by an element of

$$M = \{\begin{pmatrix} X & \mathbb{1}_p \end{pmatrix} O_\pi | \pi \in S_n, X \in \mathbb{C}^{p \times (n-p)}, |X_{j,k}| \leq 1 \forall j, k\}, \tag{D.5}$$

where $O_\pi \in O(n)$ denotes the orthogonal matrix implementing the permutation on $\mathbb{C}^n$. One can therefore optimize X for a fixed permutation $\pi \in S_n$ using a gradient-free

optimization method. If the optimum is found near the boundary of the allowed set, a permutation is chosen which allows to optimize in a next step over the corresponding "next" sector of $Gr(p, n)$ [180]. We find this scheme to be able to reliably identify local minima in the considered settings.

E. Proof of the Extension of Huddson's Central Limit Theorem

In this appendix we want to prove the extension to Hudson's central limit theorem used in Sec. 4.3.2. Let us repeat the theorem out of convenience:

Lemma. *Assume a fermionic system with V sites, p modes per site and a state $\rho \in \mathcal{D}(\mathcal{F}_p)$ on a single site. The product state $\rho^{\otimes V} \in \mathcal{D}(\mathcal{F}_{Vp})$ is then approximately described by a product $\bigotimes_q \rho_q$, with ρ_q being Gaussian states to be specified, in the following sense.*

Denote by $f^\dagger_{j,\sigma}$ and $f_{j,\sigma}$ with $j \in [V]$ and $\sigma \in [p]$ the creation and annihilation operators of the systems and introduce the shorthand notation $f^c_{j,\sigma}$ with $c = \pm 1$ and $f^{-1}_{j,\sigma} = f^\dagger_{j,\sigma}$ and $f^1_{j,\sigma} = f_{j,\sigma}$. We then introduce the Fourier modes of the lattice

$$a^c_{q,\sigma} = \frac{1}{\sqrt{V}} \sum_{j=1}^{V} e^{\frac{2\pi i}{V} jq c} f^c_{j,\sigma} \tag{E.1}$$

for $q = -\lfloor (V-1)/2 \rfloor, \ldots, \lfloor V/2 \rfloor$. Furthermore, we denote by f^c_σ for $\alpha \in [p]$ and $c = \pm 1$ the creation and annihilation operators of a fermionic system with p modes.

We then find for $r \in [Vp]$, and triples $(q_1, c_1, \sigma_1) < \cdots < (q_{2r}, c_{2r}, \sigma_{2r})$ with $\sigma_l \in [p]$, $c_l = \pm 1$ and q_l as above that the cumulants of the Fourier modes with respect to the state $\rho^{\otimes V}$ are suppressed as

$$K_{2r}^{\rho^{\otimes V}}(a^{c_1}_{q_1,\sigma_1}, \ldots, a^{c_{2r}}_{q_{2r},\sigma_{2r}}) = V^{1-r} K^\rho_{2r}(f^{c_1}_{\sigma_1}, \ldots, f^{c_{2r}}_{\sigma_{2r}}) \delta_{\sum_{l=1}^{2r} c_l q_l \bmod V, 0} \tag{E.2}$$

Put differently, in the Fourier modes $\rho^{\otimes V}$ has up to an error decreasing in V the same moments as $\rho_0 \otimes \rho_{V/2} \otimes \bigotimes_{q=1}^{\lfloor (V-1)/2 \rfloor} \rho_{q,-q}$ with $\rho_0 = \rho_{V/2}$ and $\rho_{q,-q} = \rho_{q',-q'}$ being Gaussian states and defined by the second moments of ρ where $\rho_{V/2}$ is only present if V is even.

Proof. In order to prove the extension of Hudson's central limit theorem we closely follow the original proof [139] and use an inductive argument in r. For $r = 1$ we explicitly calculate

$$K_2^{\rho^{\otimes V}}(a^{c_1}_{q_1,\sigma_1}, a^{c_2}_{q_2,\sigma_2}) = \frac{1}{V} \sum_{j_1,j_2=1}^{V} e^{\frac{2\pi i}{V}(c_1 q_1 j_1 + c_2 q_2 j_2)} \operatorname{tr}(f^{c_1}_{j_1,\sigma_1} f^{c_2}_{j_2,\sigma_2} \rho^{\otimes V}) \tag{E.3}$$

Using that $\rho^{\otimes V}$ is even on every site yields that terms with $j_1 \neq j_2$ vanish. In addition the resulting expectation values are independent of j_1, i.e. $\operatorname{tr}(f^{c_1}_{j,\sigma_1} f^{c_2}_{j,\sigma_2} \rho^{\otimes V}) = \operatorname{tr}(f^{c_1}_{\sigma_1} f^{c_2}_{\sigma_2} \rho)$.

Evaluating then the remaining sum over sites yields as claimed

$$K_2^{\rho^{\otimes V}}(a^{c_1}_{q_1,\sigma_1}, a^{c_2}_{q_2,\sigma_2}) = \delta_{(c_1 q_1 + c_2 q_2) \bmod V,0} \operatorname{tr}(f^{c_1}_{\sigma_1} f^{c_2}_{\sigma_2} \rho). \tag{E.4}$$

For $1 < r \leq Vp$ we use the definition of the cumulants

$$\sum_{P \in \mathcal{P}^e([2r])} \operatorname{sign}(\mathrm{P}) \prod_{p \in P} K_{|p|}^{\rho^{\otimes V}}((a^{c_l}_{q_l,\sigma_l})_{l \in p}) = \frac{1}{V^r} \operatorname{tr}\left(\prod_{l=1}^{2r} \left(\sum_{j=1}^{V} e^{\frac{2\pi i}{V} c_l q_l j} f^{c_l}_{j,\sigma_l} \right) \rho^{\otimes V} \right). \tag{E.5}$$

We expand the product over the independent sums on the right-hand side of the equation by using the partitions of the set $[2r]$ in order to account for all different distributions of the indices on different sites. Each partition selects then which indices are collect together on individual site. Furthermore we sum over all site configurations by summing over all increasing sequences of sites and permutations which shuffle the occupied sites. This yields

$$\begin{aligned} &\sum_{P \in \mathcal{P}^e([2r])} \operatorname{sign}(\mathrm{P}) \prod_{p \in P} K_{|p|}^{\rho^{\otimes V}}((a^{c_l}_{q_l,\sigma_l})_{l \in p}) \\ &= \frac{1}{V^r} \sum_{P \in \mathcal{P}^e([2r])} \operatorname{sign}(\mathrm{P}) \sum_{j_1 < \cdots < j_{|P|}} \sum_{\pi \in S_{|P|}} \prod_{w=1}^{|P|} \operatorname{tr}\left(\rho^{\otimes V} \prod_{l \in p_w} e^{\frac{2\pi i}{V} c_l q_l j_{\pi(w)}} f^{c_l}_{j_{\pi(w)},\sigma_l} \right) \end{aligned} \tag{E.6}$$

Inserting now the definition of the cumulants on the right-hand side leads to

$$\begin{aligned} &\sum_{P \in \mathcal{P}^e([2r])} \operatorname{sign}(\mathrm{P}) \prod_{p \in P} K_{|p|}^{\rho^{\otimes V}}((a^{c_l}_{q_l,\sigma_l})_{l \in p}) \\ &= \frac{1}{V^r} \sum_{P \in \mathcal{P}^e([2r])} \operatorname{sign}(\mathrm{P}) \sum_{j_1 < \cdots < j_{|P|}} \sum_{\pi \in S_{|P|}} \prod_{w=1}^{|P|} \sum_{Q \in \mathcal{P}^e(p_w)} \operatorname{sign}(\mathrm{Q}) \prod_{q \in Q} K^{\rho}_{|q|}((f^{c_l}_{\sigma_l})_{l \in q}) e^{\frac{2\pi i}{V}(\sum_{l \in q} c_l q_l) j_{\pi(w)}} \end{aligned} \tag{E.7}$$

It is now key to realize that we can write the last expression above as

$$\sum_{P \in \mathcal{P}^e([2r])} \operatorname{sign}(\mathrm{P}) \prod_{p \in P} K_{|p|}^{\rho^{\otimes V}}((a^{c_l}_{q_l,\sigma_l})_{l \in p}) = \frac{1}{V^r} \sum_{P \in \mathcal{P}([2r])} \operatorname{sign}(\mathrm{P}) \prod_{p \in P} K^{\rho}_{|p|}((f^{c_l}_{\sigma_l})_{l \in p}) \sum_{j=1}^{V} e^{\frac{2\pi i}{V} j \sum_{l \in p} c_l q_l} \tag{E.8}$$

which can be shown by noting that the combination of the sum over partitions P, sites $j_1, \ldots, j_{|P|}$ in combination with the permutations and partitions Q in Eq. (E.7) generates uniquely every configuration given by a partition P in Eq. (E.8) with the parts being distributed independently over the sites. In addition, the product of $\operatorname{sign}(\mathrm{P}) \prod \operatorname{sign}(\mathrm{Q})$ in Eq. (E.7) gives rise to the sign $\operatorname{sign}(\mathrm{P})$ in Eq. (E.8). We can then evaluate the sum over the sites, insert the induction assumption for all P with $|P| > 1$ and cancel the common terms on the left and right-hand side of the equation in order to obtain

$$K_{2r}^{\rho^{\otimes V}}((a^{c_l}_{q_l,\sigma_l})_{l \in [2r]}) = \frac{1}{V^{r-1}} K^{\rho}_{2r}((f^{c_l}_{\sigma_l})_{l \in [2r]}) \delta_{\sum_{l=1}^{2r} c_l q_l \bmod V,0}. \tag{E.9}$$

□

F. Proof of Gaussification

In this appendix we present the complete and rigorous proof for the relaxation of free systems in the form of Gaussification as stated in Thm. 10. We repeat the theorem out of convenience here.

Theorem 13. *Consider a fermionic system defined on an n-dimensional regular lattice of V sites where each site is equipped with p fermionic modes. Assume that the dynamics of the system is given by a non-interacting local Hamiltonian giving rise to delocalizing transport with decay constant $d_{\mathrm{trans}} > n/4$. Given a region $S \subset [V]$ of the lattice we then find that for any fermionic state ρ with exponentially decaying correlations and for all $0 < \epsilon < 4d_{\mathrm{trans}} - n$ we can define a constant C independent of the system size V such that*

$$\|\mathrm{tr}_{S^c}\,\rho(t) - \mathrm{tr}_{S^c}\,\rho_G(t)\|_1 \leq Ct^{-(4d_{\mathrm{trans}}-n-\epsilon)} \tag{F.1}$$

for all $t \in]0, t_{\mathrm{rec}}]$ where $\rho_G(t)$ denotes the time dependent Gaussian state with the same second moments as $\rho(t)$ and $t_{\mathrm{rec}} \to \infty$ for $V \to \infty$.

Proof. The proof of the theorem is performed essentially in four steps. First we reformulate the problem using the known results on the dynamics under quadratic Hamiltonians. Then we use Lieb-Robinson bounds in order to truncate the evolution of the system to a small region of interest. Using the exponential decay in the correlations of ρ we then find that the evolution is dominated by a mixture of independent contributions. Subsequently, the delocalizing transport allows us to obtain the final bound within a combinatorial argument.

F.1. Reformulation of the Problem

Let $m_{a,\sigma} = m_j$ denote the Majorana operators of the system with $j = (a, \sigma)$, $a \in [V]$ and $\sigma \in [2p]$ labeling the sites and modes per sites respectively. We define as before for $j = (a, \sigma)$ and $k = (b, \sigma')$ that $d(j, k) = d(a, b)$. For the region S we can write the one-norm difference of $\rho(t)$ and $\rho_G(t)$ using expectation values of operators supported on S. Switching into the Heisenberg picture yields then

$$\|\mathrm{tr}_{S^c}\,\rho(t) - \mathrm{tr}_{S^c}\,\rho_G(t)\|_1 = \sup_{\substack{A:\|A\|=1, A^\dagger=A \\ \mathrm{supp}\,A\subset S}} |\,\mathrm{tr}(A(t)[\rho - \rho_G])|, \tag{F.2}$$

where $\rho = \rho(0)$. Every observable A with $\mathrm{supp}\,A \subset S$ can be expanded in the Majorana operator basis such that we obtain for the time evolved operator

$$A(t) = \sum_{J\subset S\times[2p]} a_J \prod_{j\in J} m_j(t). \tag{F.3}$$

The coefficients are bounded by $|a_J| \leq 1$ as $\|A\| = 1$.[1] As the Hamiltonian is quadratic, its time evolution is given, as explained in Sec. 2.4, by a mode transformations such that

$$m_j(t) = \sum_{k \in [V]\times[2p]} O_{j,k}(t) m_k, \tag{F.4}$$

with $O(t) = e^{-ht}$ and h being the couplings of the Hamiltonian. We can then trivially bound

$$\|\operatorname{tr}_{S^c} \rho(t) - \operatorname{tr}_{S^c} \rho_G(t)\|_1 \leq 2^{2p|S|} \max_{J \subset S\times[2p]} \left| \sum_{k_1,\ldots,k_{|J|} \in [V]\times[2p]} \operatorname{tr}\left(\prod_{l=1}^{|J|} O_{j_l,k_l}(t) m_{k_l} [\rho - \rho_G] \right) \right|, \tag{F.5}$$

where $\{j_1, \ldots, j_{|J|}\} = J$ is assumed to be ordered. Set t_{rec} to be the minimum of the recurrence times provided by the delocalizing transport in Def. 4, the Lieb-Robinson bound in Thm. 9 and the exponential suppression of correlations in Def. 3. Fix from now on a time $t \leq t_{\mathrm{rec}}$ and a set $J \subset S \times [2p]$, where we assume w.l.o.g. $|J|$ to be even and we write $J = \{j_1, \ldots, j_{|J|}\}$ with $(j_l)_l$ ordered.

F.2. Restriction to the Lieb-Robinson Cone

For the given lattice we define for any mode $j \in [V] \times [2p]$ the ball of radius Δ around j as

$$B_\Delta(j) = \{k \in [V] \times [2p] \,|\, d(l,k) \leq \Delta\}. \tag{F.6}$$

Furthermore, we define by

$$|B_\Delta| = \max_{j \in [V]\times[2p]} |B_\Delta(j)| \tag{F.7}$$

the size of a ball of radius Δ, which scales for an n-dimensional lattice as Δ^n. For the set $J \subset [V] \times [2p]$ we define the Lieb-Robinson cone at time t to be

$$C_t = \bigcup_{j \in J} B_{2v_{\mathrm{lr}}t}(j). \tag{F.8}$$

Due to the Lieb-Robinson bound in Thm. 9 we can then restrict the sum over modes in Eq. (F.5) to the cone by the following argument. The sum over the first in index splits into components from the cone and the ones outside the cone such that

$$\begin{aligned} \left| \sum_{k_1,\ldots,k_{|J|} \in [V]\times[2p]} \operatorname{tr}\left(\prod_{l=1}^{|J|} O_{j_l,k_l}(t) m_{k_l} [\rho - \rho_G] \right) \right| \leq \sum_{k_1 \in {C_t}^c} |O_{j_1,k_1}(t)| \left| \operatorname{tr}\Big(m_{k_1} m_{j_2}(t) \ldots m_{j_{|J|}}(t) \right. \\ \left. \times [\rho - \rho_G] \Big) \right| + \left| \sum_{k_1 \in C_t} \sum_{k_2,\ldots,k_{|J|} \in [V]\times[2p]} \operatorname{tr}\left(\prod_{l=1}^{|J|} O_{j_l,k_l}(t) m_{k_l} [\rho - \rho_G] \right) \right|. \end{aligned} \tag{F.9}$$

[1] That $\|A\| = 1$ indeed implies $|a_J| \leq 1$ can be seen from the fact that $|a_J| \operatorname{tr} \mathbb{1} = |\operatorname{tr}(\prod_{j,\sigma \in J} m_{j,\sigma} A)| \leq \|\prod_{(j,\sigma)\in J} m_{j,\sigma}\|_1 \|A\| = \operatorname{tr} \mathbb{1}$.

The expectation value in the first term can be upper bounded by 2. Using the suppression from the Lieb-Robinson bound we obtain then

$$\sum_{k_1 \in {C_t}^c} |O_{j_1,k_1}(t)| \leq \sum_{l=0}^{\infty} |B_{2v_{\mathrm{lr}}t+l+1}(j_1) \backslash B_{2v_{\mathrm{lr}}t+l}(j_1)| e^{-\frac{|v_{\mathrm{lr}}t - 2v_{\mathrm{lr}}t - l|}{\xi_{\mathrm{lr}}}}. \tag{F.10}$$

Inserting that $|B_{2v_{\mathrm{lr}}t+l+1}(j_1) \backslash B_{2v_{\mathrm{lr}}t+l}(j_1)|$, the volume of the rings of width 1, can always be upper bounded as $\mathrm{vol}_n \times (2v_{\mathrm{lr}}t + l + 1)^{n-1}$, where

$$\mathrm{vol}_n = \max_{j \in [V] \times [2p]} \max_{l \in \mathbb{N}} |B_{l+1}(j) \backslash B_l(j)| / l^{n-1} \tag{F.11}$$

depends only on the lattice structure, yields then that we can define the constant

$$\mathrm{const}_{\mathrm{lr}} = \sup_{t \in \mathbb{R}^+} \sum_{l=0}^{\infty} \mathrm{vol}_n \, (2v_{\mathrm{lr}}t + l + 1)^{n-1} e^{-l/\xi_{\mathrm{lr}}} e^{-v_{\mathrm{lr}}t/2\xi_{\mathrm{lr}}}. \tag{F.12}$$

Note that it is clear that this constant is independent of the system size and time as well as finite due to the exponential suppression and depends only on vol_n, v_{lr} and ξ_{lr}. We can iterate this argument for all other k in Eq. (F.9) such that we obtain

$$\begin{aligned} \left| \sum_{k_1,\ldots,k_{|J|} \in [V] \times [2p]} \mathrm{tr} \left(\prod_{l=1}^{|J|} O_{j_l,k_l}(t) m_{k_l}[\rho - \rho_G] \right) \right| \leq & \left| \sum_{k_1,\ldots,k_{|J|} \in C_t} \mathrm{tr} \left(\prod_{l=1}^{|J|} O_{j_l,k_l}(t) m_{k_l}[\rho - \rho_G] \right) \right| \\ & + 2|J| \mathrm{const}_{\mathrm{lr}} e^{-v_{\mathrm{lr}}t/2\xi_{\mathrm{lr}}} \end{aligned} \tag{F.13}$$

proving that we can restrict our attention to the cone C_t up to an error that is exponentially suppressed in t.

F.3. Factorization by the Exponentially Suppressed Correlations

In the next step we split the expectation value in Eq. (F.13) into a product of independent parts, again up to an error that is exponentially suppressed in time using the exponentially decaying correlations of the state ρ. First we show that if ρ has exponentially decaying correlations in the sense of Def. 3 then the correlations of the Gaussified state ρ_G will decay as well.

Lemma 14. *Let ρ be a fermionic state exhibiting exponentially decaying correlations with constants c_{clust} and ξ_{clust} then we obtain for the Gaussified state ρ_G for all observables A and B with $\|A\| = \|B\| = 1$*

$$|\operatorname{tr}(AB\rho_G) - \operatorname{tr}(A\rho_G)\operatorname{tr}(B\rho_G)| \leq c_{\mathrm{clust}}^{(G)} e^{-\frac{d(A,B)}{\xi_{\mathrm{clust}}}}, \tag{F.14}$$

with $c_{\mathrm{clust}}^{(G)} = c_{\mathrm{clust}}[4p(|\operatorname{supp}(A)| + |\operatorname{supp}(B)|)]^{2p|\operatorname{supp}(A)|+|\operatorname{supp}(B)|}$

Proof. The result is rather intuitive as the correlations of the Gaussified state ρ_G are related to the ones of ρ by Wick's theorem. Using for $K \subset [V] \times [2p]$ the notation $m_K = \prod_{k \in K} m_k$ we can expand the normalized observables into $A = \sum_{K \subset \mathrm{supp}\,(A) \times [2p]} a_K m_K$ and $B = \sum_{K \subset \mathrm{supp}\,(B) \times [2p]} b_K m_K$ with $|b_K| \leq 1$ and $|a_K| \leq 1$ again. Using Wick's theorem we obtain

$$|\operatorname{tr}(\rho_G AB) - \operatorname{tr}(\rho_G A)\operatorname{tr}(\rho_G B)| \leq \sum_{K,L} |\operatorname{Pf}(\gamma^{(m)}[K,L]) - \operatorname{Pf}(\gamma^{(m)}[K] \oplus \gamma^{(m)}[L])| \quad \text{(F.15)}$$

where $\gamma^{(m)}[J,K]$ are the correlation matrices of the second moments. Note that every term of the second Pfaffian also appears in the first one and that all terms that solely appear in the first Pfaffian contain at least one expectation value of the form $\operatorname{tr}(\rho m_k m_l)$ for some $k \in K$ and $l \in L$. Those terms are exponentially suppressed in the distance of the support of A and B such that we obtain

$$|\operatorname{tr}(\rho_G AB) - \operatorname{tr}(\rho_G A)\operatorname{tr}(\rho_G B)| \leq \sum_{K,L} (2p[|\operatorname{supp}(A)| + |\operatorname{supp}(B)|])!! c_{\mathrm{clust}} e^{-\frac{d(A,B)}{\xi_{\mathrm{clust}}}}. \quad \text{(F.16)}$$

Inserting upper bounds for the factorial and number of different K and L yields the above result. □

In order to exploit the correlation structure of the states ρ and ρ_G we need to bookkeep the position and distances of the indices k inside the cone in more detail. We therefore define the notation of indices being arranged in clusters where different clusters have a minimal distance towards each other and indices inside a cluster should have a maximal distance. Formally we define the following. We introduce for a set $p \subset \mathbb{N}$ the function $i^p : p \to [|p|]$ where for $l \in p$ $i^p(l) = |\{k \in p : k \leq l\}|$ tells us that l is the $i^p(l)$-th element of p for p ordered increasingly. For $r \in \mathbb{N}$ let $K \subset [V] \times [2p]$ with $|K| = r$, $(k_l)_{l \in [r]} \in ([V] \times [2p])^{\times r}$ and $\Delta \in \mathbb{R}^+$. We then define the Δ-partition of $(k_l)_{l \in [r]}$ to be the partition $P^K_\Delta((k_l)_{l \in [r]}) \in \mathcal{P}(K)$ which fulfills

1) The elements of each part p of the partitions label indices $k_{i^p(l)}$ which are connected by a chain of indices corresponding to the same parts that have a maximal distance Δ from one element to the next, i.e.

$$\forall p \in P_\Delta((k_l)_{l \in [r]}) : \forall a, b \in p\, \exists c_1 = a, c_2 \ldots, c_{z-1}, c_z = b \in p : d(k_{i^K(c_l)}, k_{i^K(c_{l+1})}) \leq \Delta. \quad \text{(F.17)}$$

2) The indices that are labeled by elements of different parts of the partition are at least a distance Δ apart, i.e.

$$\forall p, q \in P_\Delta((k_l)_{l \in [r]}), p \neq q : \forall a \in p, b \in q : d(k_{i^K(a)}, k_{i^K(b)}) > \Delta. \quad \text{(F.18)}$$

For $P \in \mathcal{P}(K)$ and $I \subset [V] \times [2p]$ let us denote by $\mathcal{K}^I_P$ the set of indices that lie in I and are Δ partitioned according to the partition P, i.e.

$$\mathcal{K}^I_P = \{(k_l)_{l \in [r]} \in I^{\times r} | P^K_\Delta((k_l)_{l \in [r]}) = P\}. \quad \text{(F.19)}$$

We suppress any reference to the set K in the notation of $\mathcal{K}_P^I$ as K can be inferred from the partition P. Fix a $\Delta \in \mathbb{R}^+$ for the moment. For any part $p \subset [|J|]$ of a partition $P \in \mathcal{P}([|J|])$ and a sequence $(k_l)_{l\in[|J|]} \in [V] \times [2p]^{\times|J|}$ we then introduce the patch operators

$$\hat{p}_p^{(k_l)_{l\in[|p|]}} = \prod_{l\in p} O_{j_l,k_{i^{[|J|](l)}}}(t) m_{k_{i^{[|J|](l)}}}. \tag{F.20}$$

We can then rewrite the sum over indices $k \in C_t$ for $\omega = \rho, \rho_G$ and obtain

$$\sum_{k_1,\dots,k_{|J|}\in C_t} \operatorname{tr}\left(\omega \prod_{l=1}^{|J|} O_{j_l,k_l}(t) m_{k_l}\right) = \sum_{P\in\mathcal{P}([|J|])} \sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \operatorname{tr}\left(\omega \prod_{p\in P} \hat{p}_p^{(k_l)_{l\in p}}\right). \tag{F.21}$$

Using that both, ρ and ρ_G have exponentially decaying correlations we can factor the expectation values up to an error that is suppressed in Δ – scaling Δ in time will then yield a suppression exponentially in time in the end. For a given partition $P \in \mathcal{P}([|J|])$ with $|P| > 1$, we label the parts as $\{p_1, \dots, p_{|P|}\} = P$ in order to obtain

$$\begin{aligned}\left|\sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \operatorname{tr}\left(\rho \prod_{p\in P} \hat{p}_p^{(k_l)_{l\in p}}\right)\right| &= \left|\sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \operatorname{tr}\left(\rho\, \hat{p}_{p_1}^{(k_l)_{l\in p_1}}\right) \operatorname{tr}\left(\rho \prod_{p\in P\setminus p_1} \hat{p}_p^{(k_l)_{l\in p}}\right)\right| \\ &\quad + |p_1|(|J|-|p_1|) c_{\text{clust}} e^{-\frac{\Delta}{\xi_{\text{clust}}}}.\end{aligned} \tag{F.22}$$

Estimating $|p_1|(|J|-|p_1|) < |J|^2$ and iterating this process $|P| < |J|$ times we obtain

$$\left|\sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \operatorname{tr}\left(\rho \prod_{p\in P} \hat{p}_p^{(k_l)_{l\in p}}\right)\right| = \left|\sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \prod_{p\in P} \operatorname{tr}\left(\rho \hat{p}_p^{(k_l)_{l\in p}}\right)\right| + |J|^3 c_{\text{clust}} e^{-\frac{\Delta}{\xi_{\text{clust}}}}. \tag{F.23}$$

The same argumentation holds of course for ρ_G up to changing the prefactor $|p_1|(|J|-|p_1|)$ originating from the supports of $\hat{p}_{p_1}^{(k_l)_{l\in p_1}}$ and the rest to $(4p|J|)^{2p|J|}$ in every step. Using that $|\mathcal{P}([|J|])| \le |J|^{|J|}$ we conclude that

$$\begin{aligned}\sum_{P\in\mathcal{P}([|J|])} \left|\sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \operatorname{tr}\left((\rho-\rho_G)\prod_{p\in P} \hat{p}_p^{(k_l)_{l\in p}}\right)\right| &\le |J|^{|J|+1}(|J|^2 + (4p|J|)^{2p|J|}) c_{\text{clust}} e^{-\frac{\Delta}{\xi_{\text{clust}}}} \\ + \sum_{P\in\mathcal{P}^e([|J|])} \left|\sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \prod_{p\in P} \operatorname{tr}\left(\rho\, \hat{p}_p^{(k_l)_{l\in p}}\right) - \prod_{p\in P} \operatorname{tr}\left(\rho_G \hat{p}_p^{(k_l)_{l\in p}}\right)\right|,\end{aligned} \tag{F.24}$$

where we also used that on the right-hand side we can restrict ourselves to partitions with even parts only due to the eveness of ρ and ρ_G.

F.4. Suppression by Delocalizing Transport

The last part of the proof involves the delocalizing transport of the system. In essence we are going to see below that although the number of configurations of patches on the lattice grows with the size of the Lieb-Robinson cone in time, if the delocalizing transport is strong enough the weight contained in the Δ-regions is suppressed sufficiently such that we obtain an overall algebraic suppression. To be more precise, any patch to which at least 4 indices are associated, is spread over the lattice giving rise to $\approx t^n$ many terms which are each suppressed as $|O_{j,k}(t)| \approx t^{-4d_{\text{trans}}}$ hence, for strong enough transport, the combined contribution is suppressed algebraically. Note that according to the remark after Def. 4 d_{trans} is upper bounded by $n/2$ meaning that patches to which two indices are associated are suppressed by the same argument only if that bound is maximally saturated. We do not want to make this very strong assumption on the system and will bound the contributions originating from pairs of indices individually.

For any partitions $P \in \mathcal{P}^e([|J|])$ we introduce the notation

$$P_2 = \{p \in P : |p| = 2\} \qquad \text{and} \qquad P_{>2} = \{p \in P : |p| > 2\}, \tag{F.25}$$

collecting all parts of P of size 2 and all larger parts. We can then sort the sum over $\mathcal{P}^e([|J|])$ by the number of pairs and write for $\omega = \rho, \rho_G$

$$\sum_{P\in\mathcal{P}^e([|J|])} \left| \sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \prod_{p\in P} \operatorname{tr}\left(\omega\, \hat{p}_p^{(k_l)_{l\in p}}\right) \right| = \sum_{m=0}^{|J|/2-2} \sum_{\substack{P\in\mathcal{P}^e([|J|]):\\ |P_2|=m}} \left| \sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \prod_{p\in P} \operatorname{tr}\left(\omega\, \hat{p}_p^{(k_l)_{l\in p}}\right) \right|. \tag{F.26}$$

Note that ρ and ρ_G yield the same result if there are $|J|/2$ pairs.

We then split the sum over indices k into two parts, one part associated to patches of size larger than 2 and one part associated to the pairs. When splitting these contributions into two sums, it is important to note that the allowed configurations of the pairs will dependent on the configurations of the other patches as around each patch a forbidden region of size Δ is formed. In addition, the indices appearing in the pairs are the indices missing in the other patches. In order to completely decouple the two sums we can take the maximum over all regions of the cone as forbidden regions as well as the maximum over all sets of indices constituting the pairs and obtain

$$\sum_{m=0}^{|J|/2-2} \sum_{\substack{P\in\mathcal{P}^e([|J|]):\\ |P_2|=m}} \left| \sum_{(k_l)_{l\in[|J|]}\in\mathcal{K}_P^{C_t}} \prod_{p\in P} \operatorname{tr}\left(\omega\, \hat{p}_p^{(k_l)_{l\in p}}\right) \right| \leq \sum_{m=0}^{|J|/2-2} \sum_{\substack{P\in\mathcal{P}^e([|J|]):\\ |P_2|=m}} \sum_{(k_l)_{l\in[|J|-2m]}\in\mathcal{K}_{P_{>2}}^{C_t}}$$
$$\times (c_{\text{trans}} t^{-d_{\text{trans}}})^{|J|-2m} \left(\max_{I\subset C_t} \max_{\substack{K\subset[|J|]:\\ |K|=2m}} \max_{Q\in\mathcal{P}_2(K)} \left| \sum_{(k_l)_{l\in[2m]}\in\mathcal{K}_Q^{C_t\setminus I}} \prod_{q\in Q} \operatorname{tr}(\omega \hat{p}_q^{(k_{i^q(l)})_{l\in q}}) \right| \right) \tag{F.27}$$

where we inserted the bound $|O_{j,k}(t)| \leq c_{\text{trans}} t^{-d_{\text{trans}}}$ for all patches in $P_{>2}$. Defining the abbreviation

$$f(m,t) = \max_{I \subset C_t} \max_{\substack{K \subset [|J|]: \\ |K|=2m}} \max_{Q \in \mathcal{P}_2(K)} \left| \sum_{(k_l)_{l \in [2m]} \in \mathcal{K}_Q^{C_t \setminus I}} \prod_{q \in Q} \operatorname{tr}(\omega \hat{p}_q^{(k_{iq(l)})_{l \in q}}) \right| \tag{F.28}$$

and using the fact that $|\{P \in \mathcal{P}^e([|J|]) : |P_2| = m\}| \leq |J|^{|J|}$ we then obtain the bound

$$\begin{aligned} \sum_{m=0}^{|J|/2-2} \sum_{\substack{P \in \mathcal{P}^e([|J|]): \\ |P_2|=m}} \left| \sum_{(k_l)_{l \in [|J|]} \in \mathcal{K}_P^{C_t}} \prod_{p \in P} \operatorname{tr}\left(\omega\, \hat{p}_p^{(k_l)_{l \in p}}\right) \right| &\leq \sum_{m=0}^{|J|/2-2} (c_{\text{trans}} t^{-d_{\text{trans}}} |C_t|^{\frac{1}{4}} |B_{|J|\Delta}|)^{|J|-2m} \\ &\times |J|^{|J|} f(m,t) \end{aligned} \tag{F.29}$$

and we are left to bound the pair contribution $f(m,t)$ for which we employ a recursive argument. Note that trivially we set $f(0,t) = 1$. For $m = 1$ we need that for any linear transformation $A \in \mathbb{C}^{M \times M}$ we can bound the norm of $m_j^{(A)} = \sum_{l=1}^M A_{j,l} m_l$ by $\|m_j^{(A)}\| \leq \|A\|^2$.[2] Rewriting the sum over indices withing a distance Δ in a sum over all indices and subtracting the contributions which are at least Δ apart we can use the exponential decay of correlation in order to obtain

$$f(1,t) \leq \max_{I \subset C_t} \max_{\substack{K \subset J \\ |K|=2}} \left(\left| \sum_{k_1,k_2 \in C_t \setminus I} \operatorname{tr}(\rho\, \hat{p}_K^{(k_l)_{l \in [2]}}) \right| + \left| \sum_{\substack{k_1,k_2 \in C_t \setminus I: \\ d(k_1,k_2) > \Delta}} \operatorname{tr}(\rho\, \hat{p}_K^{(k_l)_{l \in [2]}}) \right| \right) \tag{F.31}$$

$$\leq 1 + c_{\text{clust}} c_{\text{trans}}^2 |C_t|^2 t^{-2d_{\text{trans}}} e^{-\frac{\Delta}{\xi_{\text{clust}}}}. \tag{F.32}$$

In order to access terms with higher m note that if the indices k would be distributed independently (without the constraints imposed by the Δ partitioning) we can use the bound for $f(1,t)$ and obtain

$$\max_{I \subset C_t} \max_{\substack{K \in [|J|]: \\ |K|=2m}} \max_{Q \in \mathcal{P}_2(K)} \prod_{q \in Q} \left| \sum_{k_1,k_2 \in \mathcal{K}_{\{q\}}^{C_t \setminus I}} \operatorname{tr}(\rho\, \hat{p}_q^{(k_l)_{l \in [2]}}) \right| \leq f(1,t)^m. \tag{F.33}$$

We therefore need to bound the difference of the restricted sum and the unrestricted sum. Terms that are contained in the latter but not in the first however will have to

[2] One can verify this bound using that

$$\left\| \sum_{l=1}^M A_{j,l} m_l \right\| = \sup_{\substack{|\psi\rangle: \\ |\psi\rangle\langle\psi|=1}} \left| \sum_{k,l=1}^M A_{j,l} \overline{A}_{j,k} \langle\psi| m_l m_k |\psi\rangle \right| = \langle j| A \left(\operatorname{tr}(|\psi\rangle\langle\psi| m_a m_b) \right)_{a,b \in [M]} A^\dagger |j\rangle, \tag{F.30}$$

where $|j\rangle$ denotes the j-th unite vector of $\mathbb{C}^M$. Using that the covariance matrix of second moments has a maximal eigenvalue of 1 we obtain the bound.

have at least a pair of pairs close to each other. This yields the basic insight needed in order to set up the recursion as now a pair is effectively merged into a larger patch and we have less pairs remaining. Along these lines we obtain

$$f(m,t) \leq f(1,t)^m + \max_{I\subset C_t} \max_{\substack{K\subset[|J|]\\|K|=2m}} \max_{P\in\mathcal{P}_2(K)} \Bigg| \sum_{(k_l)_{k\in[2m]}\in\mathcal{K}_Q^{C_t\setminus I}} \prod_{q\in Q} \mathrm{tr}(\rho\, \hat{p}_q^{(k_l)_{l\in q}}) - \prod_{q\in Q} \sum_{(k_l)_{k\in[2]}\in\mathcal{K}_{\{q\}}^{C_t\setminus I}} \mathrm{tr}(\rho\, \hat{p}_q^{(k_l)_{l\in[2]}}) \Bigg| \tag{F.34}$$

$$\leq f(1,t)^m + \max_{I\subset C_t} \max_{\substack{K\subset[|J|]\\|K|=2m}} \sum_{Q\in\mathcal{P}^e(K)\setminus\mathcal{P}_2(K)} \Bigg| \sum_{(k_l)_{l\in[2m]}\in\mathcal{K}_Q^{C_t\setminus I}} \prod_{q\in Q} \mathrm{tr}(\rho\, \hat{p}_q^{(k_l)_{l\in q}}) \Bigg|. \tag{F.35}$$

But this is exactly the same expression which we started out from such that we obtain directly

$$f(m,t) \leq f(1,t)^m + \sum_{h=0}^{m-2} |J|^{|J|} (c_{\mathrm{trans}} |C_t|^{\frac{1}{4}} t^{-d_{\mathrm{trans}}} |B_{|J|\Delta}|)^{2m-2h} f(h,t). \tag{F.36}$$

Defining the constants

$$z = 1 + c_{\mathrm{clust}} c_{\mathrm{trans}}^2 |C_t|^2 t^{-2d_{\mathrm{trans}}} e^{-\frac{\Delta}{\xi_{\mathrm{clust}}}} \tag{F.37}$$

$$y = |J|^{|J|} (c_{\mathrm{trans}} |C_t|^{\frac{1}{4}} t^{-d_{\mathrm{trans}}} |B_{|J|\Delta}|)^4 \frac{1 - (c_{\mathrm{trans}} |C_t|^{\frac{1}{4}} t^{-d_{\mathrm{trans}}} |B_{|J|\Delta}|)^{|J|}}{1 - (c_{\mathrm{trans}} |C_t|^{\frac{1}{4}} t^{-d_{\mathrm{trans}}} |B_{|J|\Delta}|)} \tag{F.38}$$

we can define a recursion for an upper bound $b(m,t)$ of $f(m,t)$ by

$$b(0,t) = 1, \tag{F.39}$$

$$b(m,t) = z^m + y \sum_{h=0}^{m-1} b(h,t), \tag{F.40}$$

where we also added the $m-1$ term to the sum over h which only increases the bound as all terms are positive and we have $b(m,t) \geq f(m,t)$ for all m. We resolve this recursion by using that $b(m,t) > b(m-1,t)$ which can be seen as follows. We claim that $b(m,t) \geq z b(m-1,t) \geq b(m-1,t)$ (where the second inequality is obvious as $z \geq 1$) which is shown by induction

$$b(1,t) = z + y b(0,t) \geq z \geq 1 = b(0,t) \tag{F.41}$$

and using the assumption we get

$$z b(m-1,t) = z^m + y \sum_{h=0}^{m-2} z b(h,t) \geq z^m + \sum_{h=1}^{m-1} b(h,t) \geq z^m + \sum_{h=0}^{m-1} b(h,t) = b(m,t). \tag{F.42}$$

With this we can upper bound the sum and obtain

$$b(m,t) = z^m + y\sum_{h=0}^{m-1} b(h,t) \leq z^m + ymb(m-1,t) \tag{F.43}$$

Inserting the bound repeatedly we obtain

$$b(m,t) \leq \sum_{h=0}^{m} z^h y^{m-k} \frac{m!}{h!} \leq \sum_{h=0}^{m} z^h y^{m-k} \frac{m!}{h!} \frac{|J|^{m-h}}{(m-h)!} = (z+|J|y)^m. \tag{F.44}$$

F.5. Putting Everything Together

With this we have all components for the proof. Collecting all different errors in Eq. (F.13), (F.24) and in the previous section yields for the bound

$$\begin{aligned}
\left\|\mathrm{tr}_{S^c}\,\rho(t) - \mathrm{tr}_{S^c}\,\rho_G(t)\right\|_1 \leq\ & 2^{2p|S|+1}(2p|S|)\mathrm{const}_{\mathrm{lr}} e^{-v_{\mathrm{lr}}t/2\xi_{\mathrm{lr}}} + 2^{2p|S|}(2p|S|)^{2p|S|+1} c_{\mathrm{clust}} \\
&\times (4p^2|S|^2 + (8p^2|S|)^{4p^2|S|}) e^{-\frac{\Delta}{\xi_{\mathrm{clust}}}} + 2^{2p|S|}(2p|S|)^{2p|S|}\left(c_{\mathrm{trans}} t^{-d_{\mathrm{trans}}} |C_t|^{\frac{1}{4}} |B_{2p|S|\Delta}|\right)^4 \\
&\times \frac{1-(c_{\mathrm{trans}} t^{-d_{\mathrm{trans}}} |C_t|^{\frac{1}{4}} |B_{2p|S|\Delta}|)^{2p|S|}}{1-(c_{\mathrm{trans}} t^{-d_{\mathrm{trans}}} |C_t|^{\frac{1}{4}} |B_{2p|S|\Delta}|)} \Bigg[1 + c_{\mathrm{clust}} c_{\mathrm{trans}}^2 t^{-2d_{\mathrm{trans}}} |C_t|^2 e^{-\frac{\Delta}{\xi_{\mathrm{clust}}}} \\
&+ (2p|S|)^{2p|S|+1} (c_{\mathrm{trans}} t^{-d_{\mathrm{trans}}} |C_t|^{\frac{1}{4}} |B_{2p|S|\Delta}|)^4 \frac{1-(c_{\mathrm{trans}} t^{-d_{\mathrm{trans}}} |C_t|^{\frac{1}{4}} |B_{2p|S|\Delta}|)^{2p|S|}}{1-(c_{\mathrm{trans}} t^{-d_{\mathrm{trans}}} |C_t|^{\frac{1}{4}} |B_{2p|S|\Delta}|)} \Bigg]
\end{aligned} \tag{F.45}$$

where we inserted the upper bound $|J| \leq 2p|S|$. Note that as $|C_t| \propto t^n$ and $d_{\mathrm{trans}} > n/4$ there exists by assumption an ϵ such that for $\Delta \propto t^{\frac{\epsilon}{4n}}$ the term $|B_{2p|S|\Delta}||C_t|^{\frac{1}{4}} t^{-d_{\mathrm{trans}}} \propto t^{-(d_{\mathrm{trans}} - \frac{n+\epsilon}{4})}$ is still suppressed in time. Hence, we can scale Δ in time such that the first two terms of the bound are suppressed exponentially in t. The last contribution is then at least suppressed as $t^{-(4d_{\mathrm{trans}}-n-\epsilon)}$ where the term in square brackets tends to 1 for large t. It is then clear that we can define a constants C such that

$$\left\|\mathrm{tr}_{S^c}\,\rho(t) - \mathrm{tr}_{S^c}\,\rho_G(t)\right\|_1 \leq C t^{-(4d_{\mathrm{trans}}-n-\epsilon)} \tag{F.46}$$

for any $0 < \epsilon < 4d_{\mathrm{trans}} - n$ as long as $t \leq t_{\mathrm{rec}}$ where C will depend on the parameters c_{trans}, d_{trans}, c_{clust}, ξ_{clust}, $\mathrm{const}_{\mathrm{lr}}$, n, p and the subsystem size $|S|$ but will be independent of the total system size V. □

G. Delocalizing Transport in Free Models

In the following appendix we want to show that the prototypical free model exhibiting transport, the free nearest neighbor hopping model, does indeed possess delocalizing transport as defined in Def. 4. To be more precise we show that the free nearest neighbor hopping model on an n-dimensional cubic lattice exhibits delocalizing transport with $c_{\text{trans}} = 25^n$, $t_{\text{rec}} = V^{\frac{6}{7n}}$ and $d_{\text{trans}} = n/3$.

Let us first consider the one-dimensional case. The nearest neighbor hopping Hamiltonian with hopping amplitude 1, a single mode per site and periodic boundary conditions takes then the form

$$H = \sum_{j,k} t^{\text{nn}}_{j,k} f^\dagger_j f_k = \sum_j \Big(f^\dagger_j f_{j+1} + f^\dagger_{j+1} f_j \Big). \tag{G.1}$$

The discrete Fourier transform $U^{\text{FT}} \in U(V)$ defined by

$$U^{\text{FT}}_{q,j} = \frac{1}{\sqrt{V}} e^{\frac{2\pi i}{V} jq} \tag{G.2}$$

with $q = 0, 1, \ldots, V-1$ diagonalizes t^{nn} with $\text{spec}(t^{\text{nn}}) = (2\cos(2\pi q/V))_{q=0,\ldots,V-1}$ and we set $\Lambda = \text{diag}(\text{spec}(t^{\text{nn}}))$. Expressed in terms of Majorana modes, the Hamiltonian is given by

$$H = \sum_{j,k} \frac{i}{4} h^{(2)} m_j m_k, \tag{G.3}$$

with $h^{(2)} = t^{\text{nn}} \otimes iY$. The coupling matrix $h^{(2)}$ is diagonalized by

$$U = \frac{1}{\sqrt{2}} U^{\text{FT}} \otimes \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \tag{G.4}$$

with the spectrum $i[\Lambda \oplus (-\Lambda)]$. Let us decompose Majorana mode indices $j \in [2V]$ into a tuple (a, σ) with $a \in [V]$ labeling the lattice site and $\sigma \in \{0, 1\}$ the mode, i.e., $j = 2a - \sigma$. Then we obtain for the propagator

$$\left| e^{-h^{(2)} t}_{(a,\sigma),(b,\sigma')} \right| \leq \sum_{\kappa=-1,1} \frac{1}{2V} \left| \sum_{q=0}^{V-1} e^{\frac{2\pi i}{V} q(a-b)} e^{i\kappa 2t \cos(\frac{2\pi}{V} q)} \right|. \tag{G.5}$$

The proof of the suppression follows from the argumentation in [158, App. A.1] which we repeat here for completeness. Define the function $f_\kappa(\phi) = e^{2i\kappa t \cos(\phi)}$. We obtain then that its Fourier modes

$$f_\kappa(n) = \frac{1}{2\pi} \int_0^{2\pi} e^{-in\phi} f_\kappa(\phi) d\phi \tag{G.6}$$

are given by $f_\kappa(n) = J_n(2t)(-i)^n(-1)^{\frac{\kappa+1}{2}n}$ with $J_n(2t)$ denoting the Bessel functions of first kind. Using partial integration, we can relate the Fourier modes $f_\kappa(n)$ to the integral

$$|f_\kappa(n)| = \frac{1}{2\pi n^2} \int_0^{2\pi} e^{-in\phi} \frac{d^2}{d\phi^2} f_\kappa(\phi) d\phi. \tag{G.7}$$

Evaluating the derivatives and estimating the resulting integrals yields

$$|f_\kappa(n)| \leq 2\frac{t^2+t}{n^2}. \tag{G.8}$$

Inserting the Fourier expansion of f_κ leads to

$$\sum_{q=0}^{V-1} e^{\frac{2\pi i}{V}q(a-b)} e^{i\kappa 2t\cos(\frac{2\pi}{V}q)} = \sum_{q=0}^{V-1} e^{\frac{2\pi i}{V}q(a-b)} \sum_{n=-\infty}^{\infty} e^{in\frac{2\pi}{V}q} f_\kappa(n) = V \sum_{p=-\infty}^{\infty} f_\kappa(pV+b-a). \tag{G.9}$$

Splitting off the $p=0$ term, exploiting the relation of $f_\kappa(n)$ and the Bessel function and inserting the bound derived above we can bound

$$\left| \frac{1}{V} \sum_{q=0}^{V-1} e^{\frac{2\pi i}{V}q(a-b)} e^{i\kappa 2t\cos(\frac{2\pi}{V}q)} - (-i)^{b-a}(-1)^{\frac{\kappa+1}{2}(b-a)} J_{b-a}(2t) \right| = \left| \sum_{z=1}^{\infty} \Big[f_\kappa(zV+b-a) \right. \tag{G.10}$$

$$\left. + f_\kappa(-zV+b-a) \Big] \right| \leq 2\frac{t^2+t}{V^2} \sum_{z=1}^{\infty} \left[\frac{1}{(\frac{b-a}{V}+z)^2} + \frac{1}{(\frac{b-a}{V}-z)^2} \right]. \tag{G.11}$$

Due to the translation invariance and reflection symmetry of the model we are only interested in the case $|b-a| \leq V/2$ such that we can bound

$$\left| \frac{1}{V} \sum_{q=0}^{V-1} e^{\frac{2\pi i}{V}q(a-b)} e^{i\kappa 2t\cos(\frac{2\pi}{V}q)} - (-i)^{b-a}(-1)^{\frac{\kappa+1}{2}(b-a)} J_{b-a}(2t) \right| \leq 2(\pi^2-4)\frac{t^2+t}{V^2}. \tag{G.12}$$

Inserting furthermore the bound $J_n(x) \leq x^{-1/3}$ for the Bessel function of first kind yields

$$\left| \frac{1}{V} \sum_{q=0}^{V-1} e^{\frac{2\pi i}{V}q(a-b)} e^{i\kappa 2t\cos(\frac{2\pi}{V}q)} \right| \leq 2(\pi^2-4)\frac{t^2+t}{V^2} + \frac{1}{(2t)^{\frac{1}{3}}}. \tag{G.13}$$

For $t \leq t_{\text{rec}} = V^{\frac{6}{7}}$ we conclude that

$$\left| e^{-h^{(2)}t}_{(a,\sigma),(b,\sigma')} \right| \leq 25 t^{-\frac{1}{3}}. \tag{G.14}$$

In higher dimensions we want to restrict ourselves to the case of a cubic lattice of $V = L^n$ sites with L denoting the linear expansion in all directions. Then, we find that the coupling $t^{\mathrm{nn,n}}$ of the hopping Hamiltonian decomposes as

$$t^{\mathrm{nn,n}} = \sum_{l=0}^{n-1} \mathbb{1}_L^{\otimes l} \otimes t^{\mathrm{nn,1}} \otimes \mathbb{1}_L^{\otimes n-l-1}, \tag{G.15}$$

with $t^{\mathrm{nn,1}}$ denoting the coupling of the one-dimensional chain. The mode transformation governing the time evolution of the system takes then the form

$$e^{th^{(2),n}} = \left(e^{th^{(2),1}}\right)^{\otimes n}, \tag{G.16}$$

such that

$$|e^{th^{(2),n}}_{j,k}| \leq 25^n t^{-\frac{n}{3}} \tag{G.17}$$

follows immediately from the above for all $t \leq t_{\mathrm{rec}} = L^{\frac{6}{7}} = V^{\frac{6}{7n}}$.

H. Backmatter

H.1. Publications of the Author of the Thesis

This thesis contains among other results the content of the following publications:

- C. Krumnow, L. Veis, Ö. Legeza, and J. Eisert. *Fermionic orbital optimization in tensor network states.* Phys. Rev. Lett., **117**, 210402, (2016).

- M. Gluza, C. Krumnow, M. Friesdorf, C. Gogolin, and J. Eisert. *Equilibration via Gaussification in fermionic lattice systems.* Phys. Rev. Lett., **117**, 190602, (2016).

- C. Krumnow, Z. Zimborás, and J. Eisert. *A fermionic de Finetti theorem.* arXiv:1708.01266, (2017), accepted in J. Math. Phys.

In parallel to the work on this thesis the author has further contributed to the following publications which are partially mentioned in this thesis but not a central part of it:

- A. Nietner, C. Krumnow, E. J. Bergholtz, and J. Eisert. *Composite symmetry protected topological order and effective models.* arxiv:1704.02992, (2017).

- H. Wilming, M. Goihl, C. Krumnow, and J. Eisert. *Towards local equilibration in closed interacting quantum many-body systems.* arxiv:1704.06291, (2017).

- M. Goihl, M. Gluza, C. Krumnow, and J. Eisert. *Construction of exact constants of motion and effective models for many-body localized systems.* arXiv:1707.05181, (2017).

H.2. Supervision of Bachelor and Master Students

In parallel to the work on this thesis the author co-supervised different research projects related to the completion of bachelor or master theses:

Bachelor Students

- Leon Klein, *Exploration of the Interplay of Dissipation and Disorder with Tensor Network Methods*, (2016), shared supervision with Albert Werner.
- Luis Herrman, *Local Inversion of Matrix Product Operators: An Analysis of Several Numerical Approaches*, (2016), shared supervision with Albert Werner.
- Hauke Rasch, *The Lieb-Robinson Cone in the Presence of Disorder*, (2017), shared supervision with Marcel Goihl and Marek Gluza.
- Zeno Schätzle, *Investigating measures of many-body localization*, (2017), shared supervision with Marcel Goihl and Marek Gluza.

Master Students

- Alexander Nietner, *Composite Topological Order and the Bilayer Δ-Chain*, (2015), shared supervision with Emil Bergholtz and Jens Eisert.
- Marek Gluza, *Relaxation of correlated fermionic states upon sudden quenches to free evolution*, (2016), shared supervision with Mathis Friesdorf and Jens Eisert.

H.3. Acknowledgment

The completion of every larger project, such as the endeavor of attaining a PhD, builds on many small steps. Therefore I first of all would like to thank my many teachers and professors who motivated, challenged and inspired me during my school time and studies. Especially I would like to thank Dr. Neubert for preparing me remarkably well for my studies.

I would like to thank Jens Eisert for his constant enthusiasm and trust he has given me. During my PhD he introduced me to several interesting and challenging topics and questions and gave me at the same time the freedom and opportunities to pursue own ideas. Furthermore, I want to thank all further collaborators who I enjoyed working with during the past years. I am most grateful for the many discussions with Mathis Friesdorf, Max Pfeffer and Zoltán Zimborás on mathematics, physics "and everything". Furthermore, I thank Örs Legeza and Libor Veis for multiple discussions and the joint work on implementing the novel ideas presented in this thesis in their code structure – and especially Örs for his repeated hospitality in Budapest. I thank Marek Gluza for paving with us a route through the mist of months of mathematical discussions and Christian Gogolin for helping to straighten it.

Even in the presence of a stable private life and bundle of helpful collaborators does the personal sanity depend a lot on an enjoyable working atmosphere. I therefore want to thank, beyond my collaborators of course, the members of the qmio group, especially Janina, Adrian, Alex, Carlos, Dominik, Emilio, Frederik, Henrik, Ingo, and Martin as well as Jörg and Christian. I am thankful for all discussions, meetings, jokes and anecdotes that we shared.

Many thanks to Marek and Marcel for prove-reading this thesis and helping me with their comments to improve the presentation at different places. Furthermore, I thank the Studienstiftung des deutschen Volkes for the financial support during my PhD.

Most importantly however I thank the people surrounding me who are not enlisted on any paper in this thesis or the physic institutes webpage. I thank my friends and relatives who kept me sane in good and in bad times during my studies and PhD. I am by far most grateful to my parents Doris and Detlef, my brother Stefan and my wife Kathi. For being there. Always.

H.4. Abstract

Understanding fermionic systems and creating more powerful tools for their simulation has been a focus of modern theoretical physics since the first days of the formulation of quantum mechanics. Due to their intrinsic hardness we are not able to design universal schemes which compute and extract the dynamic or static properties of generic interacting fermionic systems efficiently and the distinct structures of more specific settings have to be exploited in order to obtain efficient methods. In this thesis we set out to identify and understand some of these structures in finite fermionic systems in more detail from both, a practical and conceptual point of view.

From a practical application point of view, we extend tensor network methods such that they are able to resolve Gaussian fermionic correlations. By combining tensor network states (TNS) and mode transformations we overcome the defect of TNS of not being able to approximate independent fermions efficiently and incorporate thus structured high entanglement effects into TNS. The obtained schemes adapt established TNS methods to fermionic systems and allow in specific cases to significantly reduce the amount of resources needed for a ground state search and a real time evolution of a non-local fermionic system. By this we construct TNS methods in the spirit of other multi-configuration schemes and allow TNS, which are able to resolve complex mode correlations, to detect close-to-product structures in a particle picture.

Conceptually we investigate the emergence of efficient structures in different classes of models. We formulate a fermionic mode de Finetti theorem which deduces a separability of a fermionic state, i.e., the suppression of all quantum correlation between different fermionic modes, from an underlying permutation invariance of the state and by this restricts the correlation structure of a state based on its symmetries. This insight directly relates to the certification of mean field approaches such as the Hartree-Fock method as we discuss and provides a new perspective towards the understanding of why these rough approximations provide surprisingly accurate results in certain systems. Furthermore, we prove that under the evolution of a free Hamiltonian supporting a sufficient form of transport generic non-critical initial states become Gaussian after a short time. This result links to general relaxation processes of closed quantum systems such as equilibration and thermalization and is in general reminiscent to the convergence towards a generalized Gibbs ensemble (and provides a rigorous proof of such a convergence in special systems). In both cases, Gaussification and the fermionic de Finetti theorem, we carefully discuss the initial assumptions and special role of the fermionic antisymmetry constraint.

H.5. Zusammenfassung

Das Verständnis von fermionischen Systemen sowie die Entwicklung von Methoden für deren Beschreibung ist einer der Schwerpunkte moderner theoretischer Physik seit der Entwicklung der Quantenmechanik. Die allgemeinen Systemen inhärente Komplexität verwehrt es universelle Methoden zu entwickeln, die im allgemeinen Fall statische und dynamische Eigenschaften wechselwirkender Fermionen bestimmen. Um dennoch Einblicke in ein spezifisches System zu erhalten, müssen in der Regel tiefere Strukturen des vorliegenden Systems ausgenutzt werden. Das praktische Auffinden sowie das theoretische Verständnis solcher Strukturen sind der Fokus dieser Arbeit.

Um Strukturen, die eine effizientere Simulation fermionischer Systeme ermöglichen, zu finden, etablieren wir eine Kombination von Tensornetzwerkzuständen (TNS) und fermionischen Modentransformationen. Diese Kombination erlaubt es etablierte TNS Algorithmen und Methoden auf fermionische Systeme anzupassen und durch die adäquate Wahl der Einteilchenbasis strukturierte starke Verschränkungseffekte in diesen zu erfassen. Die resultierenden Multikonfigurations TNS Methoden erlauben es in praktischen Anwendungen die benötigten Ressourcen zum Teil immens zu verringern, da sie die gewählte Einteilchenbasis an die entsprechende Korrelationsstruktur des Systems anpassen können.

Auf einer theoretischen Ebene betrachten wir den Ursprung und das Entstehen verschiedener Strukturen in fermionischen Systemen. Wir formulieren ein Moden de Finetti Theorem, welches die Separabilität eines fermionischen Zustandes aus einer Permutationssymmetrie des Systems folgert. Solch eine Unterdrückung von Quantenkorrelationen zwischen einzelnen Einteilchenmoden steht im direkten Zusammenhang mit der Anwendbarkeit von Molekularfeldnäherungen, wie etwa der Hartree-Fock Näherung, an das gegebene System und ermöglicht es die Genauigkeit dieser Näherungen zu zertifizieren. Darüber hinaus zeigen wir, dass in einem freien fermionischen System, welches einen genügenden Transport erlaubt, ein allgemeiner, nicht kritischer Zustand nach einer kurzen Zeit zu einem Gaußschen Zustand relaxiert. Strukturell ähnelt und trägt das Resultat zum Verständnis von anderen allgemeinen Relaxationsprozessen in geschlossenen Quantensystemen, wie der Equilibrierung und der Thermalisierung, bei und beweist schlüssig, dass freie fermionische Systeme zu generalisierten Gibbs-Ensembles equilibrieren und nicht thermalisieren. Sowohl für das de Finetti Theorem als auch für die Gaußifizierung freier fermionischer Systeme diskutieren wir insbesondere die Rolle der nötigen hinreichenden Annahmen an das System und die der kanonischen, fermionischen Antisymmetriebedingungen.

H.6. Eigenständigkeitserklärung

Ich bestätige hiermit, dass ich die vorliegende Arbeit selbstständig und nur mit Hilfe der angegebenen Hilfsmittel angefertigt habe. Alle Stellen der Arbeit, die wörtlich oder sinngemäß aus Veröffentlichungen oder anderen fremden Quellen entnommen wurden, sind als solche kenntlich gemacht. Ich habe die Arbeit nicht in einem früheren Promotionsverfahren eingereicht.

Berlin, den 24.11.2017 _____

Christian Detlef Krumnow