# Effective One-Dimensional Models from Matrix Product States

Masterarbeit zur Erlangung des akademischen Grades Master of Science

> vorgelegt von Frederik Keim geboren in Tettnang

Lehrstuhl für Theoretische Physik I Fakultät Physik Technische Universität Dortmund 2012

- 1. Gutachter : Prof. Dr. Uhrig
- 2. Gutachter : Prof. Dr. Anders

Datum des Einreichens der Arbeit: 28. September 2012

## Abstract

In this thesis a method for deriving effective one-dimensional models based on the matrix product state formalism is introduced. It exploits tranlational invariance to to work directly in the thermodynamic limit. The method is tested on the analytically solvable Ising model in a transverse magnetic field. Results for ground state energy and dispersion are given as well as a way to find a real space representation for the local creation operator. From this, the one particle contribution to the spectral weight is calculated.

## Kurzfassung

In dieser Arbeit wird eine variationelle Methode zur Ableitung effektiver eindimensionaler Modelle vorgestellt, die auf dem Formalismus der Matrixproduktzustände basiert. Durch Ausnutzung von Translationsinvaranz kann direkt im thermodynamische Limes gearbeitet werden. Die Methode wird anhand des Ising Modells in einem transversalen Magnetfeld getestet, das exakt lösbar ist. Es werden Ergebnisse für die Grundzustandsenergie und die Einteilchen-Dispersion angegeben, sowie ein Weg den lokalen Erzeuger im Ortsraum zu konstruieren. Damit wird der Einteilchen-Beitrag zum spektralen Gewicht berechnet.

# Contents

Contents	V
List of Figures	VI
List of Tables	VII
1 Introduction         1.1 Motivation         1.2 Existing methods         1.3 The approach         1.4 Structure of the thesis	1 1 1 3 3
2 Model and general approach         2.1 The transverse field Ising model         2.2 Introduction to matrix product states         2.3 Effective models	<b>5</b> 5 9 21
3 Ground state energy         3.1 Matrix product states for infinite systems         3.2 Variational ground state search         3.3 Results         3.4 Existing methods	<b>23</b> 23 30 37 38
4 Dispersion         4.1 Description of excitations         4.2 Algorithm         4.3 Results         4.4 Existing methods	<b>39</b> 39 46 47 51
5Local creation operator and spectral weight5.1Local creation operator5.2Spectral weight	<b>53</b> 53 55
<ul> <li>6 Conclusions and outlook</li> <li>6.1 Summary of method</li> <li>6.2 Summary of results</li> <li>6.3 Outlook</li> <li>Bibliography</li> </ul>	63 63 63 64 65

# List of Figures

2.2Graphical representation of local matrices $A^{\sigma_\ell}$ at the edges and in the bulk of the chain. The matrcies at the edges are simply vectors and therefore have only one matrix index $\alpha_1$ and $\alpha_{L-1}$ respectively. Complex conjugates that arise in the description of bra-vectors are depicted with a downward pointing physical index.122.4Application of matrix product operator165.5System buildup in numerical renormalization groups172.6Superblock in standard DMRG203.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to $M$ matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 1.0$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 50
2.3Tensor network representation of local matrices $A^{\sigma_\ell}$ at the edges and in the bulk of the chain. The matrices at the edges are simply vectors and therefore have only one matrix index $\alpha_1$ and $\alpha_{L-1}$ respectively. Complex conjugates that arise in the description of bra-vectors are depicted with a downward pointing physical index.122.4Application of matrix product operator162.5System buildup in numerical renormalization groups172.6Superblock in standard DMRG182.7DMRG sweep203.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to $M$ matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 50
bulk of the chain. The matrices at the edges are simply vectors and therefore have only one matrix index $\alpha_1$ and $\alpha_{L-1}$ respectively. Complex conjugates that arise in the description of bra-vectors are depicted with a downward pointing physical index
have only one matrix index $\alpha_1$ and $\alpha_{L-1}$ respectively. Complex conjugates that arise in the description of bra-vectors are depicted with a downward pointing physical index
that arise in the description of bra-vectors are depicted with a downward pointing physical index
$\begin{array}{llllllllllllllllllllllllllllllllllll$
2.4Application of matrix product operator162.5System buildup in numerical renormalization groups172.6Superblock in standard DMRG182.7DMRG sweep203.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to $M$ matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 1.0$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
2.5System buildup in numerical renormalization groups172.6Superblock in standard DMRG182.7DMRG sweep203.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to M matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 1.0$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
2.6Superblock in standard DMRG182.7DMRG sweep203.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to $M$ matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 50
2.7DMRG sweep203.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to M matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
3.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to M matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 1.0$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 50
3.1Splitting scheme for infinite systems233.2Diagram for local MPO293.3Central contributions to $M$ matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 50
3.2Diagram for local MPO293.3Central contributions to $M$ matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 1.0$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
3.3Central contributions to M matrix343.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
3.4Ground state energy per lattice site $E_0/L$ 373.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
3.5Deviation of ground state energy from exact result384.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.1Results from naive dispersion approach414.2Eigenvalues of the norm matrix $N_q$ 434.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 444.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.1Eigenvalues of the norm matrix $N_q$ 4.34.2Eigenvalues of the norm matrix $N_q$ 4.34.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 434.5Dispersion for $\lambda = 0.5$ 444.6Dispersion for $\lambda = 0.9$ 484.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.2Engentiates of the norm internal $N_q^{\alpha\beta}$ 4.34.3Tensor network for norm element $N_j^{\alpha\beta}$ 434.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 444.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.3Tensor network for horm element $N_j$ 4.34.4Representative contributions to the matrix element $h_j^{\alpha\beta}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.4Representative contributions to the matrix element $h_j^{\gamma}$ 444.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.5Dispersion for $\lambda = 0.5$ 484.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.6Dispersion for $\lambda = 0.9$ 494.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.7Dispersion for $\lambda = 1.0$ 494.8Excitation gap energy $\Delta$ 504.9Deviation of $\Delta$ from exact result50
4.8 Excitation gap energy $\Delta$
4.9 Deviation of $\Delta$ from exact result
5.1 Eigenvector $v_a$ of momentum space EVP
5.2 Coefficients $v_i$ of creation operator (linear)
5.3 Coefficients $v_i$ of creation operator (log.)
5.4 Septral weight $S_{1\pi}^{xx}$ for $\lambda = 0.5$
5.5 Septral weight $S_{1x}^{xx}$ for $\lambda = 0.9$
5.6 Septral weight $S_{1x}^{xx}$ for $\lambda = 0.999$
5.7 Correlation length

# List of Tables

2.1	Examples on tensor network notation	12
3.1 3.2	Local variation algorithm	36 36
$4.1 \\ 4.2$	Alorithm for calculating the energy dispersion $\ldots \ldots \ldots$	$\frac{46}{47}$

# Chapter 1

# Introduction

## 1.1 Motivation

One of the main fields of research in quantum many body physics are systems of strongly correlated electrons. Sometimes, mostly for one-dimensional models, special properties of a given model allow for an analytical solution. However, the number of such models is small and in general one will have to resort to numerical calculations. The exponential growth of the Hilbert space dimension with the number of particles strongly limits the size of a system that can be analyzed by exact diagonalization.

There are various approaches to work around these limitations. Renormalization group methods try to concentrate the computational power of a classical computer on a fraction of the Hilbert space in which the physics takes place that one is interested in. Different methods mainly differ in the criterion that is used to decide what information to keep and what to discard. A very prominent example is the density matrix renormalization group method (DMRG) [1].

These methods usually directly produce the results of simulated quantum measurements. A different approach is to map a given Hamiltonian onto an effective Hamiltonian that is diagonal in the subspace of interest and can be used to derive further physical properties of the system. A group of such methods are e.g. continuous unitary transformations (CUT), which come in different variants depending on the model and the specific goal. Examples are perturbative CUT (pCUT) [2, 3] and graph based CUT (gCUT) [4]. A problem of these methods is however, that the interaction range accessable is strongly limited by the computational ressources of today's classical computers.

In this thesis a method will be presentend, that combines ideas from both approaches: A variational ansatz is used to obtain not only the ground state energy, but also the dispersion relation and a real space representation of the local creation and annihilation operators. This provides an effective model for the one particle space that can be used in further studies.

### 1.2 Existing methods

Although the main goal of the method is the derivation of an effective model in the spirit of CUTs, its algorithms are that of a renormalization method and it has to be seen in the context of DMRG and matrix product state related methods. The first breakthrough of renormalization group methods was Wilsons solution of the Kondo problem in a single impurity Anderson model [5] in 1975. His Numerical Renormalization Group (NRG) very successfully used the energy as a criterion to select the correct part of the Hilbert space. Later it became evident, that energy alone is not always the relevant criterion. This led to S.R. White's DMRG method [1]. He could show that using the eigenvalues of a density matrix is in a certain sense optimal (cf. Sect. 2.2.5). The DMRG is still one of the most powerful methods to obtain the properties of ground states and low lying excited states of one-dimensional systems

Today, there are a multitude of extensions to DMRG, some of which are closely related to the method presented here. Extensions to standard DMRG cover the calculation of dynamical properties in both frequency space (DDMRG) [6] and real-time (tDMRG) [7, 8]. High precision DMRG results for ground state energies and excitation gaps are often used as benchmark today. However, calculating an energy dispersion is rather extensive [8, 9].

Since the original formulation of DMRG in inherently one-dimensional, the poor performance in two or more spatial dimensions has always been a drawback. This is due to the interactions becoming longranged when mapping a two dimensional system onto a onedimensional chain. Efforts in overcoming this resulted in a momentum space formulation of DMRG [10, 9] that also provied a method of calculating dispersions. The 2D performance however was still moderate.

The concept of matrix product states (MPS) is also a very powerful tool that predates DMRG and has been introduced under different circumstances by different people, e.g. in Refs. [11, 12, 13]. Östlund and Rommer discovered in 1995 [14], that in a translationally invariant system the DMRG automatically leads to a MPS form in the thermodynamic limit. The MPS can be rederived purely variationally, without any reference to DMRG. Although they did not take the thermodynamic limit in their calculations [14, 15], their work is the basis for infinite systems DMRG (iDMRG) [16, 17] and related methods.

The reformulation of the successful DMRG method in terms of MPS or, more general, tensor networks created much interest and resulted in serveral related methods based on tensor networks. G. Vidal's infinite time-evolving block decimation (iTEBD) algorithm [18] exploits translational invariance to very efficiently simulate the time evolution of infinite one-dimensional systems using the Suzuki-Trotter decomposition of the time evolution operator. Evolution in imaginary time effectively cools down the system and provides a good ground state approximation.

Other MPS based approaches for infinite chain systems that use transfer matrices were developed by Bañuls et al. [19] and Ueda et al. [20]

The MPS formulation also paved the way for successful extension to higher spatial dimensions of the DMRG. Projected entangled pair states (PEPS) [21, 22] replace every physical lattice site with a number of virtual spin-1/2 systems, corresponding to the number of nearest neighbours a site has. These auxillary spins form maximally entangled states across every bond. The physical state in form of an MPS is the obtained by projecting the auxillary systems onto the Hilbert space of the physical sites.

Combining conscepts of PEPS and iTEBD allows the simulation of infinite two-dimensional systems using iPEPS [23].

Another approach that is closely related to the method presented here was recently proposed by Pirvu et al. [24]. It uses a momentum eigenstate ansatz for a MPS and is well suited to obtain accurate dispersion relations.

## 1.3 The approach

In this thesis a variational method to derive effective models is introduced. To our knowledge, this has not been done before. The method uses the matrix product formalism, which can also be used to describe a variety of other variational methods e.g. Wilson's NRG [25]. If translational invariance is assumed, matrix porduct states present a very efficient way to work in the thermodynamic limit, thus ridding the results of finite size errors.

The ground state search algorithm is closely related to the above mentioned DMRG method [26]. Unlike DMRG however, the method can use intermediate results from this ground state search to obtain the energy dispersion of the elementary excitations with about the same precision as the ground state energy itself. Moreover, as a byproduct of the dispersion relation, a real space representation of the local second quantization creation operator is found.

One of the key elements of the new method is that often high-dimensional minimization can be replaced by more robust iterated diagonalization. At the moment, it allows to compute static and momentum dependent properties at zero temperature.

The results are to be considered as proof of concept only. The current implementation is in GNU octave script, which presents a severe limitation of efficiency. Due to a lack of time, some parts of the algoritm are rather crude. A lot of optimizations are to be made in the future in order to understand if the remaining problems arise from the method itself or from the model that is investigated.

An implementation in C++ and the use of more effecient algorithms such as the Lanczos algorithm for diagonalization should considerably boos the accuracy of the results.

The calculations were done with double precision, most of them on workstation computers.

## **1.4** Structure of the thesis

The thesis is structured as follows: In Chap. 2 the model that is used to test the new method is introduced and a general overview of the idea of matrix product states (MPS) is given.

In Chap. 3 to 5, the method is developed and the results are presented and compared to the exact solution and to some results obtained from other methods. Chapter 3 shows how the ground state energy per lattice site for infinite systems can be calculated with an MPS representation of the ground state. In Chap. 4 a way to describe local excitations is intoduced and results for the one-particle dispersion are given. In Chap. 5 a method to derive a real space representations of the local creation operator is described, that allows for the computation of the one-particle spectral weight and further studies of one particle properties.

Finally, in Chap. 6 the method and the results are summed up and an outlook on future investigations is given.

## Chapter 2

# Model and general approach

#### 2.1 The transverse field Ising model

#### 2.1.1 Exact solution

In this thesis, the presented method is tested on the one-dimensional quantum Ising model in a transverse magnetic field (ITF) as used, e.g., by P. G. De Gennes to describe tunneling in ferroelectric crystals [27]. In this section a quick reminder of the analytic solution and the closed expressions for ground state energy, dispersion relation, and the one-particle spectral weight are given.

The model is defined by the Hamiltonian

$$H = -\Gamma \sum_{j} S_{j}^{z} - J \sum_{j} S_{j}^{x} S_{j+1}^{x}, \qquad \Gamma, J > 0$$
(2.1)

where  $S^{z}$  and  $S^{x}$  are the common spin- $\frac{1}{2}$  operators. The model describes a spin chain with nearest neighbour interaction along the x-axis in a perpendicular external field. It is analytically solvable and well understood. The ratio of the coupling constants

$$\lambda := \frac{J}{2\Gamma} , \qquad \lambda \in [0,\infty)$$
(2.2)

serves as control parameter that defines the system behaviour. In the strong field (or free spin) limit (J = 0), the ground state aligns all spins along the external field and elementary excitations are spin flips. In the weak field (or Ising) limit ( $\Gamma = 0$ ), the ground state is ferromagnetic and twofold degenerate. Then, elementary excitations are domain walls, separating sections with different ground state realizations. As  $\lambda$  approaches 1, the correlation length diverges and a quantum phase transition occours at  $\lambda = 1$ .

In the Ising regime, the ground state has an intrinsic long range order, wherefore this regime is also called the ordered phase. This order disappears for  $\lambda < 1$ , so that the strong field regime is also referred to as disordered phase.

As Pfeuty has shown in Ref. [28], the model can be solved analytically by mapping the spins to spinless fermions.

By introducing the spin ladder operators

$$S_j^{\pm} := S_j^{\mathrm{x}} \pm i S_j^{\mathrm{y}} , \qquad (2.3)$$

in terms of which the spin operators read

$$S_{j}^{x} = \frac{1}{2}(S_{j}^{+} + S_{j}^{-}), \qquad S_{j}^{y} = \frac{1}{2i}(S_{j}^{+} - S_{j}^{-}), \qquad S_{j}^{z} = S_{j}^{+}S_{j}^{-} - \frac{1}{2}, \qquad (2.4)$$

the Hamiltonian becomes

$$H = \sum_{j} \frac{\Gamma}{2} - \Gamma \sum_{j} S_{j}^{+} S_{j}^{-} - \frac{J}{4} \sum_{j} \left( S_{j}^{+} + S_{j}^{-} \right) \left( S_{j+1}^{+} + S_{j+1}^{-} \right) .$$
(2.5)

In the strong field limit, this can be interpreted as a quasiparticle model. Since a spin can only be flipped once, only one excitation can exist on a site, which is a fermionic property. On the other hand, all spin operators on different sites commute. This results in mixed commutation and anticommutation relations

$$\{S_i^+, S_i^-\} = 1, \qquad [S_i^+, S_j^-] = 0 \quad j \neq i .$$
(2.6)

Excitations with these properties are commonly called hardcore bosons. Note, however that in the strong field limit, the ground state aligns all spins "upwards". Hence  $S^+$  does not create an excitation but annihilates one. Therefore, another set of operators is defined by

$$\alpha_j := S_j^+, \qquad \alpha_j^\dagger := S_j^- , \qquad (2.7)$$

so that  $\alpha_j^{\dagger}$  creates and  $\alpha_j$  annihilates a quasiparticle. This transformation preserves the hardcore properties. In normal order H now reads

$$H = -\sum_{j} \frac{\Gamma}{2} + \Gamma \sum_{j} \alpha_{j}^{\dagger} \alpha_{j} - \frac{J}{4} \sum_{j} \left( \alpha_{j}^{\dagger} + \alpha_{j} \right) \left( \alpha_{j+1}^{\dagger} + \alpha_{j+1} \right) .$$
(2.8)

In one dimension a Jordan-Wigner transformation [29, 30]

$$c_{j} = \exp\left(i\pi \sum_{i < j} \alpha_{i}^{\dagger} \alpha_{i}\right) \alpha_{j} \qquad c_{j}^{\dagger} = \alpha_{j}^{\dagger} \exp\left(-i\pi \sum_{i < j} \alpha_{i}^{\dagger} \alpha_{i}\right)$$
$$\alpha_{j} = \exp\left(-i\pi \sum_{i < j} c_{i}^{\dagger} c_{i}\right) c_{j} \qquad \alpha_{j}^{\dagger} = c_{j}^{\dagger} \exp\left(i\pi \sum_{i < j} c_{i}^{\dagger} c_{i}\right)$$

can be used to map the hardcore bosons to spinless fermions. The  $c_j$  satisfy the canonical anticommutation relations [30]

$$\{c_i^{\dagger}, c_j\} = \delta_{ij}, \qquad \{c_i^{\dagger}, c_j^{\dagger}\} = \{c_i, c_j\} = 0.$$
 (2.9)

In the case of open boundary conditions this results in

$$H = -\frac{L\Gamma}{2} + \Gamma \sum_{j}^{L} c_{j}^{\dagger} c_{j} - \frac{J}{4} \sum_{j}^{L-1} (c_{j}^{\dagger} - c_{j}) (c_{j+1}^{\dagger} + c_{j+1}) , \qquad (2.10)$$

where L is the number of lattice sites in the chain. For periodic boundary conditions, the second sum runs from 1 to L with L + 1 := 1, giving rise to an additional subextensive term, coupling the last and the first site of the chain. For both open and periodic boundary conditions, the error in letting both sums run to L and neglecting the corrections becomes small if L is large. Thus in the thermodynamic limit H reads

$$H = -\frac{L\Gamma}{2} + \Gamma \sum_{j}^{L} c_{j}^{\dagger} c_{j} - \frac{J}{4} \sum_{j}^{L} (c_{j}^{\dagger} - c_{j}) (c_{j+1}^{\dagger} + c_{j+1}) . \qquad (2.11)$$

By a Fourier transform

$$\phi_q = \frac{1}{\sqrt{L}} \sum_j \exp(iqr_j)c_j \qquad \qquad \phi_q^{\dagger} = c_j^{\dagger} \frac{1}{\sqrt{L}} \sum_j \exp(-iqr_j)$$
$$c_j = \frac{1}{\sqrt{L}} \sum_q \exp(-iqr_j)\phi_q \qquad \qquad c_j^{\dagger} = \frac{1}{\sqrt{L}} \sum_q \phi_q^{\dagger} \exp(iqr_j)$$

one arrives at

$$H = \sum_{q} \left( -\frac{\Gamma}{2} + \frac{J}{4} e^{iqa} \right) + \sum_{q} \left( \Gamma - \frac{J}{4} \cos(qa) \right) \phi_q^{\dagger} \phi_q - \frac{J}{4} \sum_{q} (\phi_{-q} \phi_q e^{iqa} - \phi_q^{\dagger} \phi_{-q}^{\dagger} e^{-iqa})$$

$$\tag{2.12}$$

$$= -\Gamma \sum_{q>0} \begin{pmatrix} \phi_q^{\dagger} & \phi_{-q} \end{pmatrix} \begin{pmatrix} A_q & iB_q \\ -iB_q & -A_q \end{pmatrix} \begin{pmatrix} \phi_q \\ \phi_{-q}^{\dagger} \end{pmatrix}$$
(2.13)

Note that the constant terms in the first line are cancelled out by the anti-commutator required to obtain the matrix form in the second line. In this, a is the lattice constant so that  $r_j = aj$  and

$$A_q := 1 - \lambda \cos(qa), \qquad B_q := \lambda \sin(qa) . \tag{2.14}$$

For simplicity a will be normalized to 1. Finally a Bogolyubov transform [31]

$$\begin{pmatrix} \phi_q \\ \phi_{-q}^{\dagger} \end{pmatrix} = \begin{pmatrix} i\cos\Theta_q & -\sin\Theta_q \\ \sin\Theta_q & -i\cos\Theta_q \end{pmatrix} \begin{pmatrix} \eta_q \\ \eta_{-q}^{\dagger} \end{pmatrix}$$
(2.15)

leads to the diagonal form

$$H = \Gamma \sum_{q>0} \begin{pmatrix} \eta_q^{\dagger} & \eta_{-q} \end{pmatrix} \begin{pmatrix} \Lambda_q & 0\\ 0 & -\Lambda_q \end{pmatrix} \begin{pmatrix} \eta_q\\ \eta_{-q}^{\dagger} \end{pmatrix}$$
(2.16a)

$$=\Gamma\sum_{q}\Lambda_{q}\eta_{q}^{\dagger}\eta_{q} - \frac{\Gamma}{2}\sum_{q}\Lambda_{q} . \qquad (2.16b)$$

From the requirement that the off-diagonal elements vanish, the conditions

$$\tan(2\Theta_q) = \frac{B_q}{A_q} \tag{2.17a}$$

$$\Lambda_q = \sqrt{1 + \lambda^2 - 2\lambda \cos q} \tag{2.17b}$$

follow. This yields the one-particle energy dispersion and the ground state energy per lattice site in the thermodynamic limit, cf. Ref. [28]:

$$\omega_q = \Gamma \Lambda_q = \Gamma \sqrt{1 + \lambda^2 - 2\lambda \cos q}$$
(2.18a)

$$\frac{E_0}{L} = -\frac{\Gamma}{2L} \sum_q \Lambda_q = -\frac{\Gamma(1+\lambda)}{\pi} \int_0^{\frac{\pi}{2}} \sqrt{1 - \frac{4\lambda}{(1+\lambda)^2} \sin^2(q) \,\mathrm{d}q} \,. \tag{2.18b}$$

It is easy to see that the excitation energy gap  $\Delta$  follows as

$$\Delta = \min_{q} \omega_q = \Gamma |1 - \lambda| \tag{2.19}$$

and vanishes at the quantum critical point  $\lambda = 1$ .

The excitations created by  $\eta_q^{\dagger}$  are nonlocal quasiparticles called magnons since they are disturbances in the magnetic configuration of the ground state. For convenience these magnons will simply be referred to as particles.

The lower boundary of the two-particle energy continuum is given by

$$\Omega_q := \min_{q_1} \left( \omega_{q_1} + \omega_{q_2} \right) \Big|_{q_1 + q_2 = q}$$
(2.20)

which is strictly greater than  $\omega_q$  except at criticality where  $\Omega_q = \omega_q$ .

#### 2.1.2 Spectral weight

An important quantity in comparing theoretical models to experiments is the so called dynamical structure factor (DSF) [32]

$$S^{\alpha\beta}(\omega,q) := \frac{1}{2\pi L} \sum_{ij} \int_{-\infty}^{\infty} \mathrm{d}t \, e^{i[\omega t + q(r_i - r_j)]} \left\langle S_j^{\alpha}(t) S_i^{\beta}(0) \right\rangle \tag{2.21}$$

with  $\alpha, \beta \in \{x, y, z, +, -\}$ . The DSF describes the intensity distribution in inelastic neutron scattering. Angular brackets denote the ground state expectation value in the zero temperature case.

At the moment, the presented method does not provide frequency or time dependent quantities. However, since the energy spectrum is discrete (except for  $\lambda = 1$ ) the one-particle DSF can be obtained from the "spectral form" of the DSF in Ref. [33] as

$$S_{1p}^{\alpha\beta}(\omega,q) = \delta(\omega - \omega_q) S_{1p}^{\alpha\beta}(q) . \qquad (2.22)$$

This is expected to describe the low energy physics adequately. The one-partice spectral weights  $S_{1p}^{\alpha\beta}(q)$  are given by

$$S_{1p}^{\alpha\beta}(q) := \Omega^{\alpha*}(q) \,\Omega^{\beta}(q) \tag{2.23a}$$

with 
$$\Omega^{\alpha}(q) = \frac{1}{\sqrt{L}} \sum_{j} \langle \psi_q | S_j^{\alpha} | \psi_0 \rangle e^{iqr_j}$$
. (2.23b)

These depend only on the wave vector q and can be calculated as shown in Sect. 5.2. In Eq. (2.23)  $|\psi_0\rangle$  denotes the ground state and  $|\psi_q\rangle := a_q^{\dagger}|0\rangle$  a state with one particle of momentum q created by a yet to be defined creation operator  $a_q^{\dagger}$ . From the exact solution above for the ITF follows  $a_q^{\dagger} = \eta_q^{\dagger}$ . As an example the spectral weight  $S_{1p}^{xx}(q)$  will be calculated. From Eq. (2.23) it follows as

$$S_{1p}^{xx}(q) = \frac{1}{L} \sum_{ij} \langle \psi_0 | S_i^{x\dagger} | \psi_1 \rangle \langle \psi_1 | S_j^x | \psi_0 \rangle e^{iq(r_j - r_i)} .$$
(2.24)

In Ref. [33], Hamer et al. conjecture an exact result for the ITF, extrapolated from high order series expansions

$$S_{1p}^{xx}(q) = \frac{(1-\lambda^2)^{1/4}}{4\,\omega(q,\lambda)} \ . \tag{2.25}$$

Though not rigorously proven, it is in very good agreement with our results and is therefore used as a reference.

As will be shown later, the present method is not yet able to correctly handle ground state degeneracy. Therefore, studies are concentrated on the disordered regime, i.e.  $\lambda \in [0,1]$ .

## 2.2 Introduction to matrix product states

In this section a short introduction to the idea of matrix product states (MPS) is given. General properties are shown and the notation used in this thesis is presented. As the name suggests, MPS are a matrix based formulation of Schrödinger picture quantum mechanics. The concept as such was introduced in different contexts. Baxter used it 1968 to calculate dimerisation of spins on a plain [11]. In the context of quantum spin chains it recieved a lot of attention through the work of Fannes et al. [12, 13]. And much research has been done in this field since Östlund and Rommer found that MPS provide a powerful mathematical framework for renormalization methods.

The approach to the topic is based on a review by Schollwöeck, Ref. [26].

#### 2.2.1 Definition and construction

The key to the concept of MPS is the singular value decomposition (SVD) of a matrix.

**Theorem 2.1 (Singular value decomposition)** For every  $m \times n$  complex (or real) matrix  $\Psi$  there is a unique decomposition

$$\Psi = USV^{\dagger} , \qquad (2.26)$$

such that U is a  $m \times \min(m,n)$  column-orthogonal matrix, V is a  $n \times \min(m,n)$  columnorthogonal matrix and S is a  $\min(m,n) \times \min(m,n)$  real diagonal matrix with  $S_{ii} \ge 0$ ,  $S_i \ge S_j$  for i > j, holding the so called singular values of  $\Psi$ .

Note that  $U^{\dagger}U = 1$  and  $V^{\dagger}V = 1$ . Either U or V is a square matrix and therefore unitary. If  $\Psi$  is square, then both U and V are unitary and  $U^{\dagger}U = UU^{\dagger} = V^{\dagger}V = VV^{\dagger} = 1$ .

Now consider an arbitrary state  $|\psi\rangle$  of a quantum system:

$$|\psi\rangle = \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} c_{\sigma_1, \sigma_2, \dots, \sigma_L} |\sigma_1, \sigma_2, \dots, \sigma_L\rangle .$$
 (2.27)

The  $\sigma_i$  can be any quantum numbers characterizing a state of the system. Because the ITF described above is a linear chain model of spatially fixed spins, the  $\sigma_i$  will from here on just be the z-component of spin *i*. Therefore, all  $\sigma_i$  take *d* possible values where *d* is the dimension of the local Hilbert space of a single spin<sup>1</sup>.

Thus, there are  $d^L$  coefficients  $c_{\sigma_1,\sigma_2,\ldots,\sigma_L}$ , L being the number of spins in the chain. As these coefficients are (possibly time-dependent)  $\mathbb{C}$ -numbers, they can be understood as the components of a  $d^L$ -dimensional vector or as the elements of a  $d \times d^{L-1}$  matrix  $\Psi$ 

$$c_{\sigma_1,\sigma_2,...,\sigma_L} = \Psi_{(\sigma_1),(\sigma_2,...,\sigma_L)}$$
 (2.28)

Now the SVD is applied to this  $\Psi$  yielding

$$c_{\sigma_{1},\sigma_{2},...,\sigma_{L}} = \Psi_{(\sigma_{1}),(\sigma_{2},...,\sigma_{L})}^{[1]} = \left( U^{[1]}S^{[1]}V^{[1]\dagger} \right)_{(\sigma_{1}),(\sigma_{2},...,\sigma_{L})}$$
$$= \sum_{\alpha_{1}=1}^{d} U_{\sigma_{1},\alpha_{1}}^{[1]}S^{[1]}_{\alpha_{1},\alpha_{1}}V^{[1]\dagger}_{\alpha_{1},(\sigma_{2},...,\sigma_{L})} .$$
(2.29)

<sup>&</sup>lt;sup>1</sup>In case of the ITF there are d = 2 possible states. However, the MPS formalism and also the method developed in this thesis can be applied to other models wherefore they are introduced more generally.

Theorem 2.1 states that  $U^{[1]}$  and  $S^{[1]}$  are of dimension  $d \times d$  and  $V^{[1]\dagger}$  is of dimension  $d \times d^{L-1}$ .

 $U^{[1]}$  has d rows addressed by the index  $\sigma_1$ , each row corresponding to a physical state of lattice site (quantum number) 1. The index  $\sigma_1$  is therefore called the physical index.  $U^{[1]}$  can also be interpreted as a set of d matrices of dimension  $1 \times d$ , cf. Fig. 2.2

$$U^{[1]} = \begin{pmatrix} U^{[1]}_{1,1} & \dots & U^{[1]}_{1,d} \\ \vdots & & \vdots \\ U^{[1]}_{d,1} & \dots & U^{[1]}_{d,d} \end{pmatrix} \to \begin{cases} (U^{[1]}_{1,1}, \dots, U^{[1]}_{1,d}) \\ \vdots \\ (U^{[1]}_{d,1}, \dots, U^{[1]}_{d,d}) \end{cases} =: \begin{cases} A^{\sigma_1 = 1} \\ \vdots \\ A^{\sigma_1 = d} \end{cases} .$$
(2.30)

Next, the product  $S^{[1]}V^{[1]\dagger}$  is defined as a new  $d^2 \times d^{L-2}$  dimensional matrix  $\Psi^{[2]}$ 

$$\left(S^{[1]}V^{[1]\dagger}\right)_{\alpha_1,(\sigma_2,\dots,\sigma_L)} := \Psi^{[2]}_{(\alpha_1,\sigma_2),(\sigma_3,\dots,\sigma_L)}$$
(2.31)

that can again be decomposed by SVD. The index in square brackets marks both the step in the decomposition process and the lattice site that the leftover  $U^{[i]}$  is associated with. Carried out over all  $\sigma_i$  this results in

$$c_{\sigma_1,\sigma_2,\dots,\sigma_L} = \Psi^{[1]}_{(\sigma_1),(\sigma_2,\dots,\sigma_L)} = \left[ U^{[1]} S^{[1]} V^{[1]\dagger} \right]_{(\sigma_1),(\sigma_2,\dots,\sigma_L)}$$
(2.32a)

$$=\sum_{\alpha_1=1}^{a} \frac{U_{\sigma_1,\alpha_1}^{[1]}}{V_{\alpha_1,\alpha_1}^{[1]}} S_{\alpha_1,\alpha_1}^{[1]\dagger} V_{\alpha_1,(\sigma_2,\dots,\sigma_L)}^{[1]\dagger}$$
(2.32b)

$$=: \sum_{\alpha_1} A^{\sigma_1}_{1,\alpha_1} \Psi^{[2]}_{(\alpha_1,\sigma_2),(\sigma_3,\dots,\sigma_L)}$$
(2.32c)

$$= \sum_{\alpha_1} \sum_{\alpha_2} A^{\sigma_1}_{1,\alpha_1} U^{[2]}_{(\alpha_1,\sigma_2),\alpha_2} S^{[2]}_{\alpha_2,\alpha_2} V^{[2]\dagger}_{\alpha_2,(\sigma_3,\dots,\sigma_L)}$$
(2.32d)

$$=: \sum_{\alpha_1,\alpha_2} A^{\sigma_1}_{1,\alpha_1} A^{\sigma_2}_{\alpha_1,\alpha_2} \Psi^{[3]}_{(\alpha_2,\sigma_3),(\sigma_4,\dots,\sigma_L)} = \cdots$$
(2.32e)

$$= \sum_{\alpha_1, \alpha_2, \dots, \alpha_{L-1}} A^{\sigma_1}_{1, \alpha_1} A^{\sigma_2}_{\alpha_1, \alpha_2} \cdots A^{\sigma_{\ell}}_{\alpha_{\ell-1}, \alpha_{\ell}} \cdots A^{\sigma_{L-1}}_{\alpha_{L-2}, \alpha_{L-1}} A^{\sigma_L}_{\sigma_{L-1}, 1}$$
(2.32f)

Figure 2.1 shows how the big "blob"  $c_{\sigma_1,\sigma_2,\ldots,\sigma_L}$  is decomposed into sets of local matrices  $A^{\sigma_i}$  that are associated with one lattice site each.



Figure 2.1: Left to right decomposition of the coefficient vector  $c_{\sigma_1,\sigma_2,...,\sigma_L}$  into local sets of matrices  $A^{\sigma_i}$ . The colors indicate which matrices in Eq. (2.32) are associated with which lattice site.

The set of equalities (2.32) shows that every coefficient  $c_{\sigma_1,\sigma_2,...,\sigma_L}$  for a given physical configuration  $\{\sigma_1,\sigma_2,\ldots,\sigma_L\}$  of the system can be obtained by choosing the right matrix  $A^{\sigma_i}$  for every site and carrying out the matrix product. This leads to the

**Definition: Matrix product state (MPS):** Each state  $|\psi\rangle$  of a quantum mechanical system can be written as

$$|\psi\rangle = \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} A^{\sigma_1} \cdot A^{\sigma_2} \cdots A^{\sigma_L} |\sigma_1, \sigma_2, \dots, \sigma_L\rangle$$
(2.33)

where  $A^{\sigma_i}$  is a set of local matrices with one element for each possible state of the quantum number  $\sigma_i$ .

Note that no explicit knowledge about the basis  $|\sigma_1, \sigma_2, \ldots, \sigma_L\rangle$  is required other than existence and ortho normality (which will be assumed as given).

By construction the dimension of the matrices  $A^{\sigma_i}$  is  $d^{i-1} \times d^i$ , i.e. the maximum dimension in at  $i = \frac{L}{2}$ , growing exponentially with L. In some cases the Schmidt rank  $r_i$  (number of non-zero singular values of  $\Psi^{[i]}$ ) can be smaller than the full dimension of  $S^{[i]}$ , which leads to somewhat smaller matrix sizes.

The true potential of the MPS formalism is however, that by choosing a fixed maximum matrix size of D, the number of parameters for a variational description can be reduced from  $\mathcal{O}(d^L)$  to  $\mathcal{O}(LdD^2)$ . This happens in a systematic fashion, because all the sites in the bulk of the chain are influenced by this truncation in the same way.

An intuitive way of truncating the matrix size is to keep only the D largest singular values  $S_i$  in each step. This is also optimal in a certain sense as will be shown in Sect. 2.2.5. Obviously, the approximation is the better, the faster the decrease in the  $S_i$  is. This approach is however inherently asymmetric as apparent from the construction method shown above. The truncation on bond  $(\ell) - (\ell + 1)$  influences the sites to the right but not those to the left whose matrices have already been truncated.

#### 2.2.2 Tensor network notation

As the formulation of the required matrix products in terms of sums over multiple indices, e.g. Eq. in (2.32), is rather cumbersome, a graphical, more intuitive representation is introduced.

The tensor network notation describes n-dimensional tensors as objects (e.g. circles, squares) with n "legs" sticking out, one for each free index.

An element of a local matrix  $A_{\alpha_{\ell-1},\alpha_{\ell}}^{\sigma_{\ell}}$  is addressed by three indices  $\sigma_{\ell}$ ,  $\alpha_{\ell-1}$  and  $\alpha_{\ell}$ . This can also be understood as a tensor of rank 3, giving rise to the pictogram for local matrices  $A^{\sigma_{\ell}}$  in Fig. 2.3.

In a tensor network solid lines connecting two objects represent indices that are contracted, i.e. summed over.

The example in Tab. 2.1 shows the computation of the trace Tr(AB) of a matrix product in tensor network notation where A and B are matrices of suiteable dimensions. The result is an object with zero free indices, i.e. a scalar.

In this notation, the expansion coefficient in Eq. (2.33) looks like

$$c_{\sigma_1,\sigma_2,\dots,\sigma_L} = \bigcup_{\bullet} \bigcup_$$



Figure 2.2: Graphical representation of re-casting the left SVD factor  $U^{[\ell]}$  into a set of d local matrices  $A^{\sigma_{\ell}}$ .

$$A^{\sigma_{\ell}}: \bigoplus_{\ell=1}^{\sigma_{1}} \alpha_{1} \quad \alpha_{\ell-1} \bigoplus_{\ell=1}^{\sigma_{\ell}} \alpha_{\ell} \quad \alpha_{L-1} \bigoplus_{\ell=1}^{\sigma_{L}} \alpha_{\ell} \quad \alpha_{\ell-1} \bigoplus_{\sigma_{\ell}} \alpha_{\ell}$$

Figure 2.3: Tensor network representation of local matrices  $A^{\sigma_{\ell}}$  at the edges and in the bulk of the chain. The matrices at the edges are simply vectors and therefore have only one matrix index  $\alpha_1$  and  $\alpha_{L-1}$ respectively. Complex conjugates that arise in the description of bra-vectors are depicted with a downward pointing physical index.



#### 2.2.3 General properties

The form of the MPS constructed in (2.32) is called "left-canonical", as it is constructed from left to right. By construction this form is also left-normalized, i.e. at each site *i* the matrices  $A^{\sigma_i}$  satisfy

$$\sum_{\sigma_i=1}^d A^{\sigma_i \dagger} A^{\sigma_i} = U^{[i]\dagger} U^{[i]} = \mathbb{1} \qquad \forall \ i \ , \tag{2.35}$$

but in general

$$\sum_{\sigma_i} A^{\sigma_i} A^{\sigma_i^{\dagger}} = U U^{\dagger} \neq \mathbb{1} .$$
(2.36)

Equation (2.35) implies that

$$\langle \psi | \psi \rangle = 1 \ . \tag{2.37}$$

While the left-canonical construction is intuitive, it is by no means the only possibility. One can as well start the decomposition from right to left and in this case interpret the  $V^{[i]\dagger}$  as a set of d local matrices  $B^{\sigma_i}$ 

$$c_{\sigma_1,...,\sigma_L} = \Psi^{[L]}_{(\sigma_1,...,\sigma_{L-1}),(\sigma_L)}$$
(2.38a)

$$= \sum_{\alpha_{L-1}} U^{[L]}_{(\sigma_1,\dots,\sigma_{L-1}),\alpha_L} S^{[L]}_{\alpha_{L-1},\alpha_{L-1}} V^{[L]\dagger}_{\alpha_{L-1},\sigma_L}$$
(2.38b)

$$= \sum_{\alpha_{L-1}} \Psi^{[L-1]}_{(\sigma_1,\dots,\sigma_{L-2}),(\sigma_{L-1},\alpha_{L-1})} B^{\sigma_L}_{\alpha_{L-1},1} = \cdots$$
(2.38c)

$$= \sum_{\alpha_1,...,\alpha_L} B^{\sigma_1}_{1,\alpha_1} B^{\sigma_2}_{\alpha_1,\alpha_2} \cdots B^{\sigma_{L-1}}_{\alpha_{L-1},\alpha_L} B^{\sigma_L}_{\alpha_L,1}$$
(2.38d)

The representation obtained this way is then right-normalized:

$$\sum_{\sigma_i=1}^d B^{\sigma_i} B^{\sigma_i\dagger} = V^{[i]\dagger} V^{[i]} = 1 \qquad \forall \ i \ . \tag{2.39}$$

A third possibility is the mixed-canonical representation, where the decomposition is carried out form both the left and the right side. In this case, there is a leftover matrix  $S^{[\ell]}$  containing the singular values on the bond between the left-canonical and the rightcanonical parts.

$$c_{\sigma_1,\sigma_2,\dots,\sigma_L} = \bigoplus_{j=1}^{\sigma_1} \bigoplus_{j=1}^{\sigma_\ell} \bigoplus_{j=1}^{\sigma_\ell} \bigoplus_{j=1}^{\sigma_{\ell+1}} \bigoplus_{j=1}^{\sigma_L} \dots \quad (2.40)$$

To see how Eq. (2.35) and (2.39) imply normalization of the state consider the norm  $\langle \psi | \psi \rangle$ 

$$\langle \psi | \psi \rangle = \begin{cases} \langle \psi | & & & \\ | \psi \rangle & & & \\ | \psi \rangle & & & \\ \end{pmatrix}$$
(2.41a)

$$= \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} c^*_{\sigma_1, \sigma_2, \dots, \sigma_L} c_{\sigma_1, \sigma_2, \dots, \sigma_L} = c_{\sigma_1, \sigma_2, \dots, \sigma_L} c^*_{\sigma_1, \sigma_2, \dots, \sigma_L}$$
(2.41b)

$$= \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} (A^{\sigma_L \dagger} \cdots A^{\sigma_1 \dagger}) (A^{\sigma_1} \cdots A^{\sigma_L})$$
(2.41c)

$$=\sum_{\sigma_L} A^{\sigma_L \dagger} \cdots \underbrace{\left(\sum_{\sigma_1} A^{\sigma_1 \dagger} A^{\sigma_1}\right)}_{\mathbb{I}} \cdots A^{\sigma_L} = 1$$
(2.41d)

and for right-canonical MPS analogously. For mixed-canonical MPS the norm is

$$\langle \psi | \psi \rangle = \operatorname{Tr} \left[ \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} B^{\sigma_L \dagger} \cdots B^{\sigma_{\ell+1}} S^{[\ell] \dagger} A^{\sigma_\ell \dagger} \cdots A^{\sigma_1 \dagger} A^{\sigma_1} \cdots A^{\sigma_\ell} S^{[\ell]} B^{\ell+1} \cdots B^{\sigma_L} \right]$$
(2.42a)

From here on the local matrices will be labelled  $M^{\sigma_i}$  if the left- or right-canonical properties are not used explicitly.

This shows, that the matrix product representation is not at all unique. And there are still many more gauge degrees of freedom which change the representation but not the state  $|\psi\rangle$ . On each bond an invertible matrix  $X^{[i]}$  can be introduced and the transformation

$$M^{\sigma_i} \to M^{\sigma_i} X^{[i]}, \qquad M^{\sigma_{i+1}} \to (X^{[i]})^{-1} M^{\sigma_{i+1}}$$
 (2.43)

leaves the MPS invariant. Fixing all  $X^{[i]}$  and the boundary conditions makes the state unique.

So far, only a chain with open boundary conditions (OBC) has been discussed. In this case, the matrices at the edges of the chain were simply vectors, thus making the complete product a scalar value  $c_{\sigma_1,\sigma_2,...,\sigma_L}$ . Since the matrices at each site *i* carry information about the interaction with all sites to the left (or to the right), in a chain with periodic boundary conditions (PBC) all matrices must have dimensions greater than 1. For translationally invariant systems all matrices are of the same size. This will make the product itself a matrix instead of a scalar. The solution to obtain a scalar again is rather intuitive, looking at the tensor network

The example on tensor networks in Tab. 2.1 shows that the long line coupling the end of the chain to its first site corresponds to the trace operation. Thus for PBC, the coefficients are

$$c_{\sigma_1,\sigma_2,\dots,\sigma_L} = \sum_{\sigma_1,\sigma_2,\dots,\sigma_L} \operatorname{Tr}(M^{\sigma_1}\cdots M^{\sigma_L}) .$$
 (2.45)

This form also holds for OBC, since the trace of a scalar is still the same scalar.

Thus for OBC the coefficients are automatically scalars and PBC can be expressed by a trace operation which can be seen from the tensor network. There are however, other possible boundary conditions e.g. antiperiodic, fixed or linear combinations of any of these. Particularly for variational algorithms it is desireable to have the same fixed matrix size D on each lattice site independently of the boundary conditions. This leads to a more general ansatz. **Definition:** General ansatz for variational MPS: Each physical state  $|\psi\rangle$  of a one-

dimensional system can be approximated variationally by

р

$$|\psi_{\text{var}}\rangle = \sum_{\alpha,\beta=1}^{D} \sum_{\sigma_1,\sigma_2,\dots,\sigma_L} a_{\alpha}^T M^{\sigma_1} \cdot M^{\sigma_2} \cdots M^{\sigma_L} b_{\beta} |\sigma_1,\sigma_2,\dots,\sigma_L\rangle$$
(2.46)

where the  $M^{\sigma_i}$  are  $D \times D$  matrices and the  $a_{\alpha}$  and  $b_{\beta}$  are D-dimensional column vectors.

The choice of the vectors  $a_{\alpha}$  and  $b_{\beta}$ , which may depend on each other, provides the necessary degrees of freedom to implement various boundary conditions. For instance, the trace operation for PBC follows from a choice of, e.g.,  $a_{\alpha} = e_{\alpha}$  and  $b_{\beta} = \delta_{\alpha\beta}e_{\beta}$  where the  $e_i$  are Cartesian unit vectors. A concrete method to derive the  $a_{\alpha}$  and  $b_{\beta}$  for particular boundary conditions will not be elaborated at this point, because it is not required for the presented method.

As an example let us consider the overlap  $\langle \phi | \psi \rangle$  of two different states of the system, where the state  $|\psi\rangle$  is described by local matrices  $M^{\sigma_i}$  and  $|\phi\rangle$  is described by  $\tilde{M}^{\sigma_i}$ . Then

$$\langle \phi | \psi \rangle = \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} c^{\phi *}_{\sigma_1, \sigma_2, \dots, \sigma_L} c^{\psi}_{\sigma_1, \sigma_2, \dots, \sigma_L}$$
(2.47a)

$$= \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} \operatorname{Tr}(\tilde{M}^{\sigma_1 *} \cdots \tilde{M}^{\sigma_L *}) \operatorname{Tr}(M^{\sigma_1} \cdots M^{\sigma_L}) , \qquad (2.47b)$$

which is true for both OBC and PBC. In this form, there are  $d^L$  products of 2L matrices each, so that the overall computational effort is  $\mathcal{O}(Ld^L)$  which is exponentially expensive. However, most of the operations are unnecessary, because only two matrices change in each product. Therefore, a better way is to evaluate the expression as

$$\langle \phi | \psi \rangle = \frac{\langle \phi |}{|\psi \rangle} \underbrace{\mathsf{Tr}}(\tilde{M}^{\sigma_1 *} \dots \tilde{M}^{\sigma_L *}) \operatorname{Tr}(M^{\sigma_1} \dots M^{\sigma_L})$$
(2.48a)

$$= \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} \operatorname{Tr}(\tilde{M}^{\sigma_1 *} \cdots \tilde{M}^{\sigma_L *}) \operatorname{Tr}(M^{\sigma_1} \cdots M^{\sigma_L})$$
(2.48b)

$$= \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} \operatorname{Tr} \left[ (\tilde{M}^{\sigma_1 *} \cdots \tilde{M}^{\sigma_L *}) \otimes (M^{\sigma_1} \cdots M^{\sigma_L}) \right]$$
(2.48c)

$$= \operatorname{Tr}\left[\left(\sum_{\sigma_1} \tilde{M}^{\sigma_1 *} \otimes M^{\sigma_1 *}\right) \cdots \left(\sum_{\sigma_L} \tilde{M}^{\sigma_L} \otimes M^{\sigma_L}\right)\right] , \qquad (2.48d)$$

where  $\otimes$  denotes the tensor product.

In this way, there is one product of L matrices of dimension  $D^2 \times D^2$ , which results in a total computational effort of  $\mathcal{O}(dLD^6)$  and the actual growth is only linear in the system size.

While the form in Eq. (2.48d) follows naturally from  $\text{Tr}(A)\text{Tr}(B) = \text{Tr}(A \otimes B)$ , this shows that the tensor product is the correct form of product to use when contractig over physical indices.

#### 2.2.4 Matrix product operators (MPO)

Now that the matrix product form for quantum states has been defined, a compatible definition for operators is needed. States are defined by their expansion coefficients

$$c_{\sigma_1,\sigma_2,\dots,\sigma_L} = \langle \sigma_1,\sigma_2,\dots,\sigma_L | \psi \rangle = M^{\sigma_1} \cdots M^{\sigma_L} , \qquad (2.49)$$

where the  $M^{\sigma_i}$  are tensors of order 3. Correspondingly, operators are defined by their matrix elements which for a product of local operators are given by

$$\langle \sigma_1, \dots, \sigma_L | \hat{O} | \sigma'_1, \dots, \sigma'_L \rangle = W^{\sigma_1 \sigma'_1} \cdots W^{\sigma_L \sigma'_L} ,$$
 (2.50)

and are products of order 4 tensors  $W^{\sigma_i \sigma'_i}$ . Thus if  $\hat{O}$  is a product of local operators, its MPO form is

$$\hat{O} = \sum_{\sigma_1, \dots, \sigma_L, \sigma'_1, \dots, \sigma'_L} W^{\sigma_1 \sigma'_1} \cdots W^{\sigma_L \sigma'_L} |\sigma_1, \dots, \sigma_L\rangle \langle \sigma'_1, \dots, \sigma'_L|$$
(2.51)

and as a tensor network

$$\hat{O} \triangleq \begin{array}{c} \sigma_1 & \sigma_\ell & \sigma_L \\ \hline \bullet & \bullet & \bullet & \bullet \\ \sigma'_1 & \sigma'_\ell & \sigma'_L \end{array}$$
(2.52)

The tensor network representation also shows how to apply a MPO to a MPS.



Figure 2.4: Application of a matrix product operator to a matrix product state

As shown in Fig. 2.4, the network is contracted over the physical indices  $\sigma'_i$ , where the product form to be used is the direct matrix product

$$\begin{aligned} |\phi\rangle &= \hat{O} |\psi\rangle = \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} \left( \sum_{\sigma_1'} W^{\sigma_1 \sigma_1'} \otimes M^{\sigma_1'} \right) \cdots \left( \sum_{\sigma_L'} W^{\sigma_L \sigma_L'} \otimes M^{\sigma_L'} \right) |\sigma_1, \sigma_2, \dots, \sigma_L \rangle \\ &= \sum_{\sigma_1, \sigma_2, \dots, \sigma_L} N^{\sigma_1} \cdots N^{\sigma_L} |\sigma_1, \sigma_2, \dots, \sigma_L \rangle , \end{aligned}$$
(2.53)

in analogy to the overlap in Eq. (2.48). This means that the dimension of the local matrices  $N^{\sigma_i}$  describing the new state  $|\phi\rangle$  are of dimension  $D \cdot D_W \times D \cdot D_W$ , where  $D_W$  is the dimension of the operator matrices  $W^{\sigma_i \sigma'_i}$ . This multiplication of matrix dimensions is denoted by the thicker lines in the diagram.

At this point one usually has to solve two issues: First, how to explicitly construct the matrices  $W^{\sigma_i \sigma'_i}$  and sencond, how to reduce the matrix dimension of  $N^{\sigma_i}$  to  $D \times D$ . There are systemic ways to do this. The construction and use of general MPOs is known as MPO formalism and the latter as MPS compression. Both are discussed in Ref. [26].

However, this general case is not really applicable for our method, as it deals with infinite systems. Also, as will be shown in detail in Chap. 3, it is not necessary.

As a simple example, take the case of a matrix element  $\langle \phi | \hat{O} | \psi \rangle$ , where  $\hat{O}$  is nontrivial only on a single lattice site *i*. In this case the MPO form is simple

$$\hat{O} = \sum_{\sigma_1, \dots, \sigma_L, \sigma'_1, \dots, \sigma'_L} \delta_{\sigma_1, \sigma'_1} \cdots \delta_{\sigma_{i-1}, \sigma'_{i-1}} W^{\sigma_i \sigma'_i} \delta_{\sigma_{i+1}, \sigma'_{i+1}} \cdots \delta_{\sigma_L \sigma'_L} |\sigma_1, \dots, \sigma_L\rangle \langle \sigma'_1, \dots, \sigma'_L|$$

$$(2.54a)$$

$$= \sum_{\sigma_1,\dots,\sigma_L,\sigma'_i} W^{\sigma_i \sigma'_i} |\sigma_1,\dots,\sigma_i,\dots,\sigma_L\rangle \langle \sigma_i,\dots,\sigma'_i,\dots,\sigma_L|$$
(2.54b)

In the local Hilbert space of a single site the operator  $\hat{O}$  and therefore  $W^{\sigma_i \sigma'_i}$  is just a  $d \times d$  matrix O. Thus

$$\langle \phi | \hat{O}_i | \psi \rangle = \begin{cases} \langle \phi | & & \\ | \psi \rangle \end{cases}$$

$$= \operatorname{Tr} \left[ \left( \sum_{\sigma_1} \tilde{M}^{\sigma_1 *} \otimes M^{\sigma_1} \right) \cdots \left( \sum_{\sigma_i \sigma'_i} O_i^{\sigma_i \sigma'_i} \tilde{M}^{\sigma_i *} \otimes M^{\sigma_i} \right) \cdots \left( \sum_{\sigma_L} \tilde{M}^{\sigma_L *} \otimes M^{\sigma_L} \right) \right]$$

$$(2.55b)$$

Because the Hamiltonian (2.1) only consists of products of local operators, this is all that is required for our method at this point.

#### 2.2.5 Connection to DMRG

For readers familiar with the DMRG method it is insctructive, to see how the two concepts connect. This section also shows, that DMRG's density matrix criterion is equivalent to truncating the singular values on each bond in a MPS construction and that this approach is optimal in a certain sense.

In his paper from 1992 [1], White started from two main ideas: First, in truncating the dimension of a Hilbert space, it is best not to keep the lowest energy states but the most probable ones. Secondly, an approximate state  $|\tilde{\psi}\rangle$  optimally represents the expectation values, i.e., the physics of the real state  $|\psi\rangle$ , if the deviation

$$S := \left\| \left| \psi \right\rangle - \left| \tilde{\psi} \right\rangle \right\|^2 \tag{2.56}$$

between them is minimal. Both lead to the same DMRG formalism that can be elegantly formulated using MPS.

We recall the way chain models are usually handled in numerical renormalization methods. A given system of length  $\ell$  is recursively expanded to size  $\ell + 1$  by adding a single lattice site as shown in Fig. 2.5.

$$\begin{array}{cccc} A: & 1, \dots, \ell \end{array} \longrightarrow \\ & \ell+1 \end{array} \rightarrow \qquad A': & 1, \dots, \ell+1 \end{array}$$

Figure 2.5: Handling of chain models in numerical renormalization methods: The system is built up by adding one site at a time.

Once the dimension of the Hilbert space reaches a certain limit, the basis of the newly formed block A' is truncated using a suiteable criterion. A major drawback in Wilson's NRG for chain models was the poor handling of the boundary conditions in the build up of such blocks. White solved this by means of his superblock ansatz.

In addition to increasing the system size in each step, he added another block called environment. System and environment together form the superblock (see Fig. 2.6). Then, the Hamiltonian of this superblock is diagonalized and its states are projected onto the system block. This has the advantage that at the time site  $\ell + 1$  it is added to the system, is not a free end but a site in the bulk of a larger system.



superblock

Figure 2.6: Expansion scheme in standard DMRG: In each step one (or two) sites are added. System and environment are combined into the superblock.

Note that, as Fig. 2.6 shows, one can also add an additional site to the environment block at each step. A comment on terminology is in order at this point. In applications, the physical system of interest is the superblock (sometimes also called "world"). The splitting in "system" and "environment" is methodical, not physical.

The question is now how to project the superblock states onto the extended system block and how to truncate the system block basis to avoid exponential growth. Both questions lead to the use of the density matrix, more precisely to the reduced density matrix of the system block (of length  $\ell + 1$ ).

For simplicity the superblock will be assumed to be in a pure state  $|\psi\rangle$ , but the argument also holds for mixed states [34]. Let  $\{|i\rangle\}$  with  $i = 1, \ldots, d_S$  be a complete orthonormal basis of the system's Hilbert space and  $\{|j\rangle\}$  with  $1, \ldots, d_E$  one of the environment's Hilbert space. Then the superblock state is given by the product state

$$|\psi\rangle = \sum_{i}^{d_{\rm S}} \sum_{j}^{d_{\rm E}} \psi_{ij} |i\rangle \otimes |j\rangle = \sum_{ij} \psi_{ij} |i\rangle |j\rangle \quad \text{with } \psi_{ij} = \langle j|\langle i|\psi\rangle .$$
 (2.57)

Consider an observable operator that acts only on the system block and is the identity on the environment. Its spectral representation is

$$A^{\rm SB} = A^{\rm S} \otimes \mathbb{1}^{\rm E} = \sum_{ii'j} A_{ii'} |i\rangle |j\rangle \langle j|\langle i'| . \qquad (2.58)$$

Its expectation value with respect to  $|\psi\rangle$  is

$$\langle A \rangle = \langle \psi | A^{\mathrm{S}} \otimes \mathbb{1}^{\mathrm{E}} | \psi \rangle = \langle \psi | \left( \sum_{ii'j} A_{ii'} | i \rangle | j \rangle \langle j | \langle i' | \right) | \psi \rangle$$
(2.59)

$$=\sum_{i''j''}\sum_{i''j''}\sum_{ii'j}\psi_{i''j''}^{*}\psi_{i''j''}\psi_{i''j''}A_{ii'}\langle j''|\langle i''|i\rangle|j\rangle\langle j|\langle i'|i''\rangle|j'''\rangle$$
(2.60)

$$=\sum_{ii'j} A_{ii'} \psi_{ij}^* \psi_{i'j} = \sum_{ii'} A_{ii'} \sum_j \psi_{ij}^* \psi_{i'j}$$
(2.61)

$$= \operatorname{Tr}(\rho A) . \tag{2.62}$$

The last equality defines the object

$$\rho_{ii'} = \sum_{j} \psi_{ij}^* \psi_{i'j}$$
(2.63)

as the reduced density matrix of the system block. This matrix holds all information of the superblock state  $|\psi\rangle$  needed to compute the expectation value of any A that acts only on the system.

Let the eigenvalues of  $\rho$  be  $w_{\alpha}$  and the corresponding eigenvectors  $|u^{\alpha}\rangle$ . The  $|u^{\alpha}\rangle$  form a valid basis of the system block Hilbert space because as a density matrix  $\rho$  is hermitian and positife-semindefinite. The  $w_{\alpha}$  are assumed to be ordered  $w_1 \geq \cdots \geq w_{d_s}$ . By definition of the density matrix, the eigenstates with the largest  $w_{\alpha}$  are the most probable states. If, in accordance with the initial idea, the D most probable states<sup>2</sup> of the system block are kept, the approximate superblock state can be rewritten as

$$|\tilde{\psi}\rangle = \sum_{j}^{d_{\rm E}} \sum_{\alpha}^{D < d_{\rm S}} a_{j\alpha} |u^{\alpha}\rangle |j\rangle = \sum_{\alpha} |u^{\alpha}\rangle \sum_{j} a_{j\alpha} |j\rangle = \sum_{\alpha} a_{\alpha} |u^{\alpha}\rangle |v^{\alpha}\rangle .$$
(2.64)

Thus, the expansion coefficient  $\psi_{ij}$  with respect to the complete basis  $|i\rangle |j\rangle$  is given by

$$\tilde{\psi}_{ij} := \langle j | \langle i | \tilde{\psi} \rangle = \sum_{\alpha} a_{\alpha} \langle i | u^{\alpha} \rangle \langle j | v^{\alpha} \rangle = \sum_{\alpha} a_{\alpha} u_i^{\alpha} v_j^{\alpha} .$$
(2.65)

Note that this has the form of an element in a product of three matrices  $(UAV^{\dagger})_{ij}$ , where A is a diagonal matrix with  $A_{\alpha\alpha} = a_{\alpha}$ . In terms of the expansion coefficients the deviation S becomes

$$S = \sum_{ij} \left| \psi_{ij} - \tilde{\psi}_{ij} \right|^2 = \sum_{ij} \left| \psi_{ij} - \sum_{\alpha} u_i^{\alpha} a_{\alpha} v_j^{\alpha} \right|^2$$
(2.66)

or on a matrix level

$$S = \left\| \Psi - \tilde{U}\tilde{A}\tilde{V}^{\dagger} \right\|_{\mathrm{F}}^{2} \tag{2.67}$$

where the F denotes the Frobenius norm. This form of  $\tilde{\Psi}$  looks very much like the SVD from theorem 2.1. Indeed, linear algebra proves S to be minimal if  $\tilde{U}\tilde{A}\tilde{V}^{\dagger}$  is chosen as the SVD of  $\tilde{\psi}_{ij}$  interpreted as  $d_{\rm S} \times d_{\rm E}$  matrix. Details can, e.g., be found in Ref. [35].

By construction,  $\tilde{V}^{\dagger}$  describes the environment block and  $\tilde{U}$  the system block. The  $a_{\alpha}$  are the singular values on the bond connecting the two parts. Also, by the properties of the SVD  $UU^{\dagger} = 1$  and  $V^{\dagger}V = 1$ . This is exactly the form of a mixed canonical MPS as seen

<sup>&</sup>lt;sup>2</sup>In DMRG literature D is referred to as m most times.

in Eq. (2.42). This means that any state occurring in a block buildup algorithm can be decomposed into a mixed-canonical MPS. From Eq. (2.42) also follows that

$$\langle \tilde{\psi} | \tilde{\psi} \rangle = \operatorname{Tr}(S^{\dagger}S) = \sum_{\alpha} S_{\alpha\alpha}^2 = \operatorname{Tr}(\rho) = \sum_{\alpha} w_{\alpha}$$
 (2.68)

$$\Rightarrow \quad w_{\alpha} = S_{\alpha\alpha}^2 \ . \tag{2.69}$$

So the eigenvalues of the reduced density matrix  $\rho$  are the squares of the singular values on the bond between system and environment. Moreover, keeping the D most probable states is therefore equivalent to the aforementioned MPS truncation scheme of keeping the D largest singular values on each bond.

The DMRG method proves especially powerful for finite systems. When the superblock has reached the desired system size, the result can be further improved, because during the buildup the truncated basis is not chosen optimally with respect to the target system of fixed size.



Figure 2.7: Scheme of a DMRG sweep. In each phase, the shading marks the block that is expanded.

The optimization is done by iteratively moving the boundary between system and environment blocks through the chain. At every step, one site of the system and possibly one site of the environment, too, is reintroduced with its full local Hilbert space dimension. The physical state of the rest of the chain is kept fixed. It can be described using block basis' that are already available from previous calculations. Then the superblock is built again, the Hamiltonian diagonalized, its target state(s) projected onto the system block and the basis is truncated according to the density matrix criterion.

The process of moving the boundary from the initial position at the center of the chain to the right edge, then to the left edge and back to the center is called a sweep. It is illustrated in Fig. 2.7. Several sweeps can be necessary to reach convergence. The limiting factor in the accuracy that can be achieved finally is the number of states D kept in each step.

### 2.3 Effective models

The spin language is very intuitive to describe magnetic properties of electronic systems. However, it is not well suited to understand dynamic properties of the excitations. Therefore, a controlled, systematic way is needed to map the spin-picture Hamiltonian to

an effective Hamiltonian in the quasiparticle-picture

$$H^{\text{eff}} = E_0 + \sum_{ij} t_{ij} c_i^{\dagger} c_j + [\text{higher terms}] . \qquad (2.70)$$

The low energy sector at zero temperature is mostly dominated by states with a single quasiparticle. The "higher terms" in (2.70) describe any multi particle interaction. Twoparticle states can be handled in much the same way shown for one particle below. But this is beyond the scope of the present thesis.

The effective one-quasiparticle model then reads

$$H^{\text{eff}} = E_0 + \sum_{ij} t_{ij} c_i^{\dagger} c_j . \qquad (2.71)$$

This model describes a single excitation moves through the system.

In a translationally invariant system, the effective hopping element  $t_{ji}$  only depends on the relative coordinate j - i but not on the absolute positions. Therefore it is simplyfied to

$$t_{ij} = t_{j-i} = t_{\delta} \quad \text{with } \delta := j - i . \tag{2.72}$$

As an example consider the strong field limit (J = 0) of the ITF. Let  $|\psi_i\rangle$  be a state with one spin flipped from the polarized ground state on site *i* which directly corresponds to a one-particle state. The hopping element is then defined by

$$t_{j-i} = \langle \psi_j | (H - E_0) | \psi_i \rangle . \qquad (2.73)$$

In a translationally invariant system a momentum space basis also proves very useful

$$|\psi_q\rangle := \frac{1}{\sqrt{L}} \sum_j e^{iqr_j} |\psi_j\rangle . \qquad (2.74)$$

The effective Hamiltonian acts on such a state in the following way

$$H^{\text{eff}} |\psi_q\rangle = \left(E_0 + \sum_{ij} t_j c_{i+j}^{\dagger} c_i\right) \frac{1}{\sqrt{L}} \sum_n e^{iqr_n} |\psi_n\rangle$$
(2.75a)

$$= E_0 |\psi_q\rangle + \frac{1}{\sqrt{L}} \sum_{ijn} e^{iqr_n} t_j \underbrace{c_{i+j}^{\dagger} c_i |\psi_n\rangle}_{\delta_{n,i} |\psi_{i+j}\rangle}$$
(2.75b)

$$= E_0 |\psi_q\rangle + \sum_j t_j e^{-iqr_j} \underbrace{\sum_i \frac{1}{\sqrt{L}} e^{iq(r_i + r_j)} |\psi_{i+j}\rangle}_{|\psi_q\rangle}$$
(2.75c)

$$= \left(E_0 + \sum_j t_j e^{-iqr_j}\right) |\psi_q\rangle , \qquad (2.75d)$$

assuming the thermodynamic limit. Thus  $|\psi_q\rangle$  is an eigenstate of the effective Hamiltonian with the eigenvalue

$$E_q = E_0 + \sum_j t_j e^{-iqr_j} =: E_0 + \omega_q .$$
 (2.76)

Therefore, the one particle energy dispersion is just the Fourier transform of the hopping elements  $t_j$  and can be obtained as

$$\omega_q = \langle \psi_q | (H - E_0) | \psi_q \rangle = \left( \frac{1}{\sqrt{L}} \sum_j e^{-iqr_j} \langle \psi_i | \right) (H - E_0) \left( \frac{1}{\sqrt{L}} \sum_i e^{iqr_i} | \psi_i \rangle \right) \quad (2.77a)$$

$$= \frac{1}{L} \sum_{ji} e^{iq(r_i - r_j)} \underbrace{\langle \psi_j | (H - E_0) | \psi_i \rangle}_{=:H_{j-i}} = \sum_j e^{iqr_j} H_j .$$
(2.77b)

If the momentum eigenstates  $|\psi_q
angle$  are not normalized, this can be solved by setting

$$|\tilde{\psi}_q\rangle := \frac{1}{\sqrt{N_q}} \frac{1}{\sqrt{L}} \sum_j e^{iqr_j} |\psi_j\rangle . \qquad (2.78)$$

The normalization constant  $N_q$  follows from

$$1 \stackrel{!}{=} \langle \tilde{\psi}_q | \tilde{\psi}_q \rangle = \frac{1}{N_q} \frac{1}{L} \sum_{ij} e^{iq(r_i - r_j)} \underbrace{\langle \psi_j | \psi_i \rangle}_{=:N_{j-i}} = \frac{1}{N_q} \sum_j e^{iqr_j} N_j$$
(2.79a)

$$\Leftrightarrow \quad N_q = \sum_j e^{iqr_j} N_j \ . \tag{2.79b}$$

Then the dispersion reads

$$\omega_q = \frac{\langle \psi_q | (H - E_0) | \psi_q \rangle}{\langle \psi_q | \psi_q \rangle} = \frac{\sum_j e^{iqr_j} H_j}{\sum_j e^{iqr_j} N_j} =: \frac{H_q}{N_q} .$$
(2.80)

All this is also true in the general case. However, in general

$$|\psi_i\rangle = a_i^{\dagger} |\psi\rangle \tag{2.81}$$

where  $|\psi\rangle$  is the ground state and  $a_i^{\dagger}$  is the creation operator for a more complex particle and is yet to be specified. To do this for cases that are not as easy as the strong field limit of the ITF is precisely the goal of the method presented here.

# Chapter 3

# Ground state energy

In this chapter an adaptation of the concepts in Sect. 2.2 for infinite, translationally invariant systems is shown. Based on that, a variational algorithm to find an MPS representation of the ground state for such systems is presented. In the last section, numercial results are compared to the exact result.

### 3.1 Matrix product states for infinite systems

Like for all variational methods, the first goal is to minimize the energy functional

$$E = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \tag{3.1}$$

by varying a set of parameters. Here these parameters are the elements of the local matrices  $A^{\sigma_i}$ . To compute (3.1), expressions of the form  $\langle \psi | \hat{O} | \psi \rangle$  and  $\langle \psi | \psi \rangle$  need to be evaluated. They result in tensor networks those given in Eq. (2.48a) and (2.55a).

The method is designed for translationally invariant systems. The Hamiltonian operator of such a system is given by a local term  $h_i$  that acts on a finite number n of lattice sites. The full Hamiltonian is just the sum over L such terms, one for each lattice site (cf. Eq. (2.1))

$$H = \sum_{i} h_i . aga{3.2}$$

As the method is designed to explicitly handle infinite systems, the lattice sites will be labelled  $j = -\infty, \ldots, \infty$ , with site j = 0 in the middle. For most purposes, the chain or the tensor network is split into three sections: A left half-infinite part, a right half-infinite part and a finite section, close to or around site j = 0.



Figure 3.1: Splitting of the tensor network of an infinite system into three parts

The main idea behind working directly in the thermodynamic limit is that there are very simple expressions for the right and left half infinite parts of the the chain, obtained via a transfer matrix. The idea dates back at least to the solution of the Ising model by transfer matrices [36] and it was adopted for MPS, e.g., in [11]. It is also used in iPEPS [23] and the MPS approaches for infinite systems in Ref. [19, 20]. The latter is very similar to the methods in this section.

Clearly, in general only a variational ansatz is possible, since matrices for an exact description of the system would have to be of infinite dimension. However, assuming translational invariance, the problem becomes much less complicated. In this case, the local matrices for all lattice sites are equal in the ground state, and only one set of local matrices mus be optimized. In some cases even an exact description may be possible with a matix dimension as low as D = 2 (see e.g. the examples in Ref. [26]).

From hereon, the local matrices describing the ground state will be labelled  $A^s$  with  $s \in \{1, \ldots, d\}^1$ . Matrices associated with excited states will be labelled  $B^s$ . They all are of dimension  $D \times D$  where D is finite and fixed throughout the calculation.

#### 3.1.1 Calculating the norm

First, the norm  $\langle \psi | \psi \rangle$  will be calculated. For non-canonical MPS it follows from the overlap in (2.48) for  $\langle \phi | = \langle \psi |$ . The handling of the boundary conditions is moved to two auxiliary systems of local dimension D with states  $|\alpha\rangle$  and  $|\beta\rangle$ , which leads to a slightly modified ansatz

$$|\psi\rangle = \sum_{\alpha,\beta} \sum_{s_1,\dots,s_L} \operatorname{Tr} \left( a_{\alpha}^T A^{s_1} \cdots A^{s_L} b_{\beta} \right) |s_1,\dots,s_L\rangle |\alpha\rangle |\beta\rangle .$$
(3.3)

Assuming translational invariance, the norm becomes

$$\langle \psi | \psi \rangle = \sum_{\alpha \beta, \alpha' \beta'} \sum_{s_1, \dots, s_L, s'_1, \dots, s'_L} \operatorname{Tr} \left( a_{\alpha}^{T*} A^{s_1*} \cdots A^{s_L*} b_{\beta}^* \right) \operatorname{Tr} \left( a_{\alpha'}^T A^{s'_1} \cdots A^{s'_L} b_{\beta'} \right)$$

$$\times \underbrace{\langle \alpha | \alpha' \rangle}_{\delta_{\alpha \alpha'}} \underbrace{\langle \beta | \beta' \rangle}_{\delta_{\beta \beta'}} \underbrace{\langle s_1, \dots, s_L | s'_1, \dots, s'_L \rangle}_{\delta_{s_1 s'_1} \cdots \delta_{s_L s'_L}}$$

$$= \operatorname{Tr} \left[ \left( \sum_{\alpha} a_{\alpha}^{T*} \otimes a_{\alpha}^T \right) \left( \sum_{s} A^{s*} \otimes A^s \right) \cdots \left( \sum_{s} A^{s*} \otimes A^s \right) \left( \sum_{\beta} b_{\beta}^* \otimes b_{\beta} \right) \right]$$

$$(3.4a)$$

$$(3.4b)$$

From (3.4) a new object is defined, the transfer matrix (or transfer operator)

$$T := \sum_{s} A^{s*} \otimes A^{s} . \tag{3.5}$$

It corresponds to one rung in the ladder-like diagram in Fig. 3.1 and is a  $D^2 \times D^2$  matrix. The name derives from the idea, that if the right (left) part of the diagram is known, applying T ( $T^{\dagger}$ ) to it transfers the end one site to the left (right).

The objects built from  $a_{\alpha}$ ,  $b_{\beta}$  and their complex conjugates are  $D^2$ -dimensional vectors.

<sup>&</sup>lt;sup>1</sup>For the transverse Ising model  $s \in \{\uparrow, \downarrow\}$ 

They encode the boundary conditions and also turn the expression into a scalar, so that the trace operation can be omitted<sup>2</sup>.

In terms of T, the norm becomes

$$\langle \psi | \psi \rangle = \operatorname{Tr}(\vec{a}^{\dagger} T^L \vec{b}) = \vec{a}^{\dagger} T^L \vec{b} = (T^{\dagger \frac{L}{2}} \vec{a})^{\dagger} (T^{\frac{L}{2}} \vec{b}) , \qquad (3.6)$$

where L is again the number of lattice sites. Eventually only the thermodynamic limit  $L \to \infty$  will be considered.

Independently of its complete diagonalizability, T has an eigenvalue  $\Lambda$  with largest absolute value (which for the moment will be assumed to be unique) and a corresponding right-eigenvector  $\vec{v}_0$ . The power method for finding eigenvalues shows that

$$\lim_{L \to \infty} \frac{T^L}{\Lambda^L} \vec{b} = \beta_0 \vec{v}_0, \quad \text{with } \beta_0 \in \mathbb{C} .$$
(3.7)

The hermitian conjugate  $T^{\dagger}$  of T has the same eigenvalues as T, especially its eigenvalue with largest absolute value is also  $\Lambda$ . The corresponding eigenvector is labelled  $\vec{u}_0$ 

$$\lim_{L \to \infty} \frac{T^{\dagger L}}{\Lambda^L} \vec{a} = \alpha_0 \vec{u}_0, \quad \text{with } \alpha_0 \in \mathbb{C} .$$
(3.8)

Therefore, under the assumption that  $\Lambda$  is unique  $\alpha$  and  $\beta$  are finite, the norm is dominated by

$$\langle \psi | \psi \rangle = \Lambda^L \ \vec{u}_0^{\dagger} \vec{v}_0 \ \alpha_0^* \beta_0 \ . \tag{3.9}$$

Expanded in the - not necessarily orthogonal - respective eigenvectors  $\vec{v}_i$  of T and  $\vec{u}_i$  of  $T^{\dagger}$  the boundary vectors  $\vec{a}$  and  $\vec{b}$  are given by

$$\vec{b} = \beta_0 \vec{v}_0 + \sum_{i=1}^{D^2 - 1} \beta_i \vec{v}_i, \qquad \vec{u} = \alpha_0 \vec{u}_0 + \sum_{i=1}^{D^2 - 1} \alpha_i \vec{u}_i .$$
(3.10)

Note, that if  $\vec{v}_0$  is an eigenvector of T, then so is  $\beta_0 \vec{v}_0$ . Therefore, the eigenvectors  $\vec{v}_0$  and  $\vec{u}_0$  will be assumed normalized such, that  $\alpha_0 = \beta_0 = 1$ . These renormalized eigenvectors are labelled  $\vec{v}$  and  $\vec{u}$ , whereby the norm becomes

$$\langle \psi | \psi \rangle = \Lambda^L \ \vec{u}^{\dagger} \vec{v} \ . \tag{3.11}$$

As long as the system has a unique ground state, it is plausible to assume that  $\Lambda$  is unique and that the boundary conditions are irrelevant for an infinite system. Therefore, the norm always takes the form in Eq. (3.11) independently of the actual form of  $\vec{a}$  and  $\vec{b}$  and the boundary conditions they describe. Thus,  $\vec{a}$  and  $\vec{b}$  can be chosen in a way convenient for the further discussion of the method.

Often, it is advantageous to adopt an alternative interpretation of T and  $\vec{v}$ . Instead of a  $D^2$ -vector,  $\vec{v}$  can be seen as a  $D \times D$  square matrix v, and T as a superoperator that acts on such matrices. This leads to the following definition.

<sup>&</sup>lt;sup>2</sup>In other notation, e.g., in Ref. [24] the trace operation is kept and  $\vec{a}$  and  $\vec{b}$  are combined in a so called boundary matrix  $Q := \vec{b}\vec{a}^{\dagger}$ .

**Definition: Vectorization of a matrix:** Let A be a  $m \times n$  matrix, which is a  $1 \times n$  block matrix of its columns  $A = (\vec{a}_1 \cdots \vec{a}_n)$ . Then the vectorization of this matrix is defined as

$$\operatorname{vec}(A) = \vec{A} := \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_n \end{pmatrix} .$$
(3.12)

With a combined index  $(\alpha, \alpha')$  where  $\alpha$  designates the column-vector and  $\alpha'$  the element within the vector, the relations

$$\operatorname{vec}(A)_{\alpha\alpha'} = A_{\alpha\alpha'}^T = A_{\alpha'\alpha} \tag{3.13}$$

follow. A compatible indexing scheme for  $mn \times mn$  block matrices is

$$B_{\alpha\alpha',\beta\beta'} \quad \text{with } \alpha,\beta \in \{1,\ldots,n\}, \ \alpha',\beta' \in \{1,\ldots,m\}$$
(3.14)

so that  $(\alpha,\beta)$  designates the block and  $(\alpha',\beta')$  the element within the block.

A scalar product for matrices compatible with the above vectorization is defined as follows:

**Definition: Scalar product for matrices:** For two  $m \times n$  matrices u and v, a scalar product is defined by the hermitian form

$$(u,v) := \operatorname{Tr}(u^{\dagger}v) = \sum_{\alpha\beta} u^*_{\beta\alpha} v_{\beta\alpha} = \operatorname{vec}(u)^{\dagger} \cdot \operatorname{vec}(v) , \qquad (3.15)$$

which is equivalent to the standard inner product of  $\mathbb{C}^{mn}$ .

From its definition in Eq. (3.5) an element of the transfer matrix is given by

$$T_{\alpha\alpha',\beta\beta'} = \sum_{s} A^{s*}_{\alpha\beta} A^{s}_{\alpha'\beta'} . \qquad (3.16)$$

Applying T to a vector  $\vec{v}$  yields

$$(T\vec{v})_{\alpha\alpha'} = \sum_{\beta,\beta'} T_{\alpha\alpha',\beta\beta'} \vec{v}_{\beta\beta'} = \sum_{\beta,\beta'} \sum_{s} A^{s*}_{\alpha\beta} A^{s}_{\alpha'\beta'} v^{T}_{\beta\beta'}$$
(3.17a)

$$=\sum_{s}\sum_{\beta\beta'}A^{s}_{\alpha'\beta'}v_{\beta'\beta}A^{s\dagger}_{\beta\alpha} = \left(\sum_{s}A^{s}vA^{s\dagger}\right)_{\alpha'\alpha} .$$
(3.17b)

From Eq. (3.17) follows for T in the superoperator interpretation

$$T(v) = \sum_{s} A^{s} v A^{s\dagger} . \qquad (3.18)$$

This form of applying T has two advantages, especially for larger D. First, it eliminates the necessity of actually computing and storing a  $D^2 \times D^2$  matrix, as the Lanczos algorithm can be used to obtain  $\Lambda$ , v and u. Secondly, multiplying a  $D^2 \times D^2$  matrix with a  $D^2$  vector requires  $\mathcal{O}(D^4)$  operations, whereas the form (3.18) takes  $\mathcal{O}(2d^2D^3)$  operations. This is a clear speed up if  $2d^2 < D$ . By definition of the hermitian conjugate

$$(u,Tv) = (T^{\dagger}u,v) \tag{3.19}$$

$$(u,Tv) = \operatorname{Tr}\left[u^{\dagger}\left(\sum_{s} A^{s}vA^{s\dagger}\right)\right] = \operatorname{Tr}\left[\left(\sum_{s} A^{s\dagger}uA^{s}\right)^{\dagger}v\right] = \operatorname{Tr}\left[(T^{\dagger}u)^{\dagger}v\right] \qquad (3.20)$$

$$\Rightarrow \quad T^{\dagger}u = \sum_{s} A^{s\dagger} \, u \, A^{s} \tag{3.21}$$

$$\Rightarrow \quad T^{\dagger} = \sum_{s} A^{s^{\dagger}*} \otimes As^{\dagger} , \qquad (3.22)$$

which means that  $T^{\dagger}$  is constructed from  $A^{s\dagger}$  the same way that T is constructed from  $A^s$ .

To ensure that the norm is well defined, i.e.,  $\langle \psi | \psi \rangle \geq 0$ , both factors in Eq. (3.11) have to be non-negative. To see that  $(u,v) \geq 0$ , it will be shown that T is an endomorphism of the (anti-)hermitian  $D \times D$  matrices and also of the positive-semidefinite  $D \times D$  matrices.

$$(Tu)_{\alpha\beta} = \left(\sum_{s} A^{s} u A^{s\dagger}\right)_{\alpha\beta} = \sum_{s} \sum_{\mu\nu} A^{s}_{\alpha\mu} u_{\mu\nu} A^{s\dagger}_{\nu\beta}$$
$$(Tu)_{\alpha\beta}^{\dagger} = \left(\sum_{s} A^{s} u A^{s\dagger}\right)_{\alpha\beta}^{\dagger} = \left(\sum_{s} A^{s} u j^{\dagger} A^{s\dagger}\right)_{\alpha\beta}^{\dagger} = \sum_{s} \sum_{\mu\nu} A^{s}_{\alpha\mu} u^{\dagger}_{\mu\nu} A^{s\dagger}_{\nu\beta}$$
$$= \begin{cases} (Tu)_{\alpha\beta} & \text{if } u^{\dagger} = u \\ -(Tu)_{\alpha\beta} & \text{if } u^{\dagger} = -u \end{cases}.$$
(3.23)

As for the positive-semidefinite matrices, it needs to be shown that

$$v^{\dagger}T(u)v \ge 0 \quad \forall v \quad \text{if} \quad \tilde{v}^{\dagger}u\tilde{v} \ge 0 \quad \forall \ \tilde{v} \ .$$

First, note that the dyadic product formed from a  $\mathbb{C}$ -vector  $\vec{v}$  and its hermitian conjugate is always positive-semidefinite

$$\vec{u}^{\dagger}(\vec{v}\vec{v}^{\dagger})\vec{u} = (\vec{u}^{\dagger}\vec{v})(\vec{v}^{\dagger}\vec{u}) = (\vec{u}^{\dagger}\vec{v})(\vec{u}^{\dagger}\vec{v})^{*} = |\vec{u}^{\dagger}\vec{v}|^{2} \ge 0 \quad \forall \ \vec{u}$$
(3.24)

with equality, if at least one of the vectors is the nullvector of if  $\vec{u} \perp \vec{v}$ . In the eigenbasis of  $vv^{\dagger}$ , this means all eigenvalues  $\lambda_i$  are greater than or equal to zero. Therefore, in the respective eigenbasis'  $\{ |\alpha \rangle \}$  of u and  $\{ |\beta \rangle \}$  of  $vv^{\dagger}$ ,

$$v^{\dagger}T(u)v = \operatorname{Tr}(T(u)vv^{\dagger}) = \operatorname{Tr}\left[\sum_{s} A^{s}\left(\sum_{\alpha} \mu_{\alpha} \left|\alpha\right\rangle \left\langle\alpha\right|\right) A^{s\dagger}\left(\sum_{\beta} \lambda_{\beta} \left|\beta\right\rangle \left\langle\beta\right|\right)\right] \quad (3.25a)$$

$$= \operatorname{Tr}\left[\sum_{\alpha\beta} \mu_{\alpha} \lambda_{\beta} \sum_{s} \langle \beta | A^{s} | \alpha \rangle \langle \alpha | A^{s\dagger} | \beta \rangle\right] = \sum_{s} \sum_{\alpha\beta} \mu_{\alpha} \lambda_{\beta} A^{s}_{\beta\alpha} A^{s\dagger}_{\alpha\beta} \qquad (3.25b)$$

$$=\sum_{\alpha\beta}\mu_{\alpha}\lambda_{\beta}\sum_{s}A^{s}_{\beta\alpha}A^{s*}_{\beta\alpha}$$
(3.25c)

$$= \sum_{\alpha\beta} \sum_{s} \mu_{\alpha} \lambda_{\beta} |A^{s}_{\alpha\beta}|^{2} \begin{cases} \geq 0 & \text{if } \mu_{\alpha} \geq 0 \quad \forall \alpha \\ \leq 0 & \text{if } \mu_{\alpha} \leq 0 \quad \forall \alpha \end{cases},$$
(3.25d)

where  $\mu_{\alpha}$  are the eigenvalues of u and the  $\lambda_{\beta}$  those of  $vv^{\dagger}$ . This also proves that T is an endomorphism of the negative-semidefinite  $D \times D$  matrices.

It follows that the number of nonzero eigenvalues of T(u) is equal to the number of nonzero eigenvalues of u if the  $A^s$  do not have a common nontrivial nullspace. In this case especially the following holds

$$u > 0 \quad \Rightarrow \quad Tu > 0 \ . \tag{3.26}$$

The chain of equalities in (3.25) also shows that if fixed, positive-definite matrices u and v are given, the expression

$$(B|A) := \operatorname{Tr}\left(\sum_{s} v B^{s\dagger} u^{\dagger} A^{s}\right)$$
(3.27)

defines as scalar product (inner product) for two sets of local matrices. Positivity is proven in (3.25), sesquilinearity and hermitianity are apparent from Eq. (3.25c).

As established earlier, the vectors  $\vec{a}$  and  $\vec{b}$  in the expression for the norm  $\langle \psi | \psi \rangle$  in Eq. (3.6) can be chosen rather freely. A reasonable and conveniently simple choice is the identity matrix

$$\vec{a} = \vec{b} = \operatorname{vec}(\mathbb{1}_{D \times D}) , \qquad (3.28)$$

as it is obviously hermitian and positive-definite. Therefore, it will be assumed, that the eigenmatrices v and u of T and  $T^{\dagger}$  belonging to the eigenvalue with largest absolute value,  $\Lambda$ , have nonzero coefficients in the expansion of the identity according to Eq.  $(3.10)^3$ . Then v and u are both hermitian and positive-definite.

Let again  $\{ |\alpha\rangle \}$  be the eigenbasis of u and  $\{ |\beta\rangle \}$  the eigenbasis of v. Let further  $\{\mu_{\alpha}\}$  and  $\{\lambda_{\beta}\}$  be the respective eigenvalues of u and v. It follows then that

$$(u,v) = \operatorname{Tr}(u^{\dagger}v) = \operatorname{Tr}\left[\left(\sum_{\alpha} |\alpha\rangle\langle\alpha|\right) \left(\sum_{\beta} \lambda_{\beta} |\beta\rangle\langle\beta|\right)\right]$$
(3.29a)

$$=\sum_{\alpha\beta}\mu_{\alpha}\lambda_{\beta}\mathrm{Tr}(|\alpha\rangle\langle\alpha||\beta\rangle\langle\beta|) = \sum_{\alpha\beta}\mu_{\alpha}\lambda_{\beta}(\langle\beta|\alpha\rangle\langle\alpha|\beta\rangle)$$
(3.29b)

$$=\sum_{\alpha\beta}\mu_{\alpha}\lambda_{\beta} |\langle\alpha|\beta\rangle|^{2} > 0. \qquad (3.29c)$$

Now it is easy to see from the definitions, that

$$\Lambda(u,v) = (u,Tv) = \operatorname{Tr}\left(\sum_{s} u^{\dagger} A^{s} v A^{s\dagger}\right) = (A|A) > 0$$
(3.30)

$$\Rightarrow \quad \Lambda \in \mathbb{R}^+ \tag{3.31}$$

$$\Rightarrow \quad \langle \psi | \psi \rangle \in \mathbb{R}^+ \ . \tag{3.32}$$

Thus, under the assumptions that  $\Lambda$  is unique - also in its absolute value - and the correct boundary conditions have finite overlap with 1, the norm (3.4) is well defined.

This also constitues the key to handling systems in the thermodynamic limit. Every norm element or expectation value requires the computation of a right and a left half-infinite tensor network as seen in Fig. 3.1. This section shows, that computing these parts comes down to finding the eigenvalue  $\Lambda$  of the transfer matrix T with the largest absolute value, and the corresponding eigenvectors  $\vec{v}$  of T and  $\vec{u}$  of its hermitian conjugate  $T^{\dagger}$ .

<sup>&</sup>lt;sup>3</sup>If this is not the case, the ground state explicitly depends on the boundary conditions. This case will be delt with in future studies.
In other words, one can work on the finite central part with moderate effort, while the effects of the system being infinite are completely contained in  $\Lambda$ , u and v.

Note, that formally infinite powers of  $\Lambda$  lead to divergence if  $|\Lambda| \neq 1$ . However, since any quantities that are actually computed are of the form in Eq. (3.1), the norm and any expectation values can be formally devided by  $\Lambda^L$ . This results in a factor of  $\Lambda^{-m}$  in expectation values, where m is the number of lattice sites spanned by the central part of the network. See also Sect. 3.2.2 on the normalization of  $\Lambda$ . This division by  $\Lambda^L$  will be implied in all matrix elements and overlaps from here on.

#### 3.1.2 Local operators

Next, a suitable expression for oparators is needed, so that expectation values and matrix elements can be calculated as well.

Since all operators considered in this thesis consist of products of a finite number of local operators, the tensor networks are of the form (2.55a). Let  $\hat{O}_j$  be a local operator acting on site j and O its local matrix representation, which is of the size  $d \times d$ . For example for the z-component of the spin operator

$$\hat{O}_j = S_j^z \qquad \to \qquad O = \frac{1}{2}\sigma_z \ . \tag{3.33}$$

As seen in Eq. (2.55b), an operator tensor object of the form shown in Fig. 3.2 is given by the expression

$$O_{j}^{(\tilde{M},M)} := \sum_{ss'} O_{ss'} \tilde{M}^{s*} \otimes M^{s'} .$$

$$\alpha - - \sigma \beta$$

$$\alpha' - \sigma \beta$$

$$\alpha' - \sigma \beta'$$

Figure 3.2: Diagram for a local matrix product operator in a matrix element calculation, cf. Eq. (2.55)

Again, it is instructive to look at the tensor network, as it suggests to apply objects of the type (3.34) to one end of the tensor network in much the same way as T and  $T^{\dagger}$  are applied. Let  $\hat{O}_i$  be the identity operation 1 and  $\tilde{M}^s = M^s = A^s$ . Then (3.34) becomes

$$\mathbb{1}_{j}^{(A,A)} = \sum_{ss'} \mathbb{1}_{ss'} A^{s*} \otimes A^{s'} = \sum_{s} A^{s*} \otimes A^{s} = T , \qquad (3.35)$$

i.e., the transfer operator T is nothing but an identity operation between two  $A^s$  matrix sets. Therefore, local operators can be applied as (cf. Eq. (3.18))

$$O_{j}^{(A,A)}v := \sum_{ss'} O_{ss'} A^{s'} v A^{s\dagger} , \qquad (3.36)$$

which can be proven in analogy to Eq. (3.17). If the states providing the matrices  $\tilde{M}^s$  and  $M^s$  are not the ground state, this generalizes to

$$O_{j}^{(B,A)}v = \sum_{ss'} O_{ss'}A^{s'}v B^{s\dagger}$$
(3.37a)

$$O_{j}^{\dagger(B,A)}v = \sum_{ss'} O_{ss'}^{\dagger} A^{s\dagger} v B^{s'}$$
(3.37b)

where  $B^s$  labels the set of local matrices taken from the bra-side of the diagram and  $A^s$  the matrix set taken from the ket-side.

In this way, products of local oparators can be applied by successively applying objects of type (3.34) to one end of the network.

Let v be the eigenmatrix of T that represents (up to powers of  $\Lambda$ ) the right end of the chain, and  $\hat{O}$  and  $\hat{P}$  be operators acting on a local site. Then the expectation value of  $\hat{O}_{j}\hat{P}_{j+1}$  is given by

$$\langle \psi | \hat{O}_{j} \hat{P}_{j+1} | \psi \rangle = (u, O_{j}^{(A,A)}(P_{j+1}^{(A,A)}(v))) \Lambda^{-2} = (O_{j}^{\dagger(A,A)}(u), P_{j+1}^{(A,A)}(v)) \Lambda^{-2} , \qquad (3.38)$$

where the sites are labelled for book-keeping only. The way the operators are applied is independent of their actual location. The factor of  $\Lambda^{-2}$  results from the fact that T is applied to all but the two sites where the local operators act and division by  $\Lambda^{L}$  is implied. The expectation value of  $\hat{O}_{j}\hat{P}_{j}$  where both operators act on the same site, can be calculated by defining  $Q_{j} = O_{j}P_{j}$  and applying

$$(u, Q_j^{(A,A)}v)\Lambda^{-1} = \left(u, \sum_{ss'} Q_{ss'}A^{s'}vA^{s\dagger}\right)\Lambda^{-1}.$$
 (3.39)

Note that if  $\hat{O}$  is hermitian,  $O_j^{(A,A)}$  - like T - is an endomorphism of the (anti-)hermitian  $D \times D$ -matrices

$$(O_j^{(A,A)}v)^{\dagger} = \left(\sum_{ss'} O^{ss'} A^{s'} v A^{s\dagger}\right)^{\dagger} = \sum_{ss'} O^{ss'\dagger} A^s v^{\dagger} A^{s'\dagger} = O_j^{(A,A)} v , \qquad (3.40)$$

where the last equality holds if O and v are hermitian.

## 3.2 Variational ground state search

### 3.2.1 Local variation

If the Hamiltonian consists of local terms  $h_i$  that are the same for each lattice site, as it is the case for the ITF, the variational ground state energy per site is given by

$$\frac{E_{0,\text{var}}}{L} = \frac{\langle \psi(A^s) | h_i | \psi(A^s) \rangle}{\langle \psi(A^s) | \psi(A^s) \rangle} . \tag{3.41}$$

By the methods of the previous section, this is

$$e_0(A^s) := \frac{E_{0,\text{var}}}{L} = \frac{(u(A^s), h_i^{(A,A)} v(A^s))}{\Lambda(A^s)^n (u(A^s), v(A^s))} , \qquad (3.42)$$

where n is the number of sites that the local Hamiltonian  $h_i$  acts upon, and all of  $\Lambda$ , v and u are determined by the local matrix set  $A^s$ . This is a highly non-linear function in the coefficients of  $A^s$ . Interpreting these as a single vector,  $e_0(A^s)$  can be treated as a minimization problem on  $\mathbb{C}^{dD^2}$  using standard algorithms. But, since the derivative of  $e_0(A^s)$  cannot be computed and the energy landscape may be of arbitrary shape, it is hard to ensure convergence in reasonable time. This approach can be called "global variation" as the matrices on all sites are optimized simultaneously. Another method is similar to DMRG: Minimize  $e_0$  by varying the matrices of only one lattice site. Then, adopt this result on all other sites and start the optimization for the single site again. Repeat this, until no further improvement can be reached. This "local variation" approach will now be explained in detail, because it is also the basis for the subsequent treatment of excitations.

Let  $|\psi_j\rangle$  be the state that has  $A^s$  matrices at each site except for site j, where  $A^s$  is replaced by a different set  $B^s$ , that will be depicted by triangular symbols. Then the norm of this state is

$$\langle \psi_j | \psi_j \rangle = \begin{cases} \langle \psi_j | \bullet \bullet \bullet \bullet \bullet \bullet \bullet \\ | \psi_j \rangle \bullet \bullet \bullet \bullet \bullet \bullet \bullet \bullet \\ = (u, \ \mathbb{1}_j^{(B,B)} v) \Lambda^{-1} = (B|B) \Lambda^{-1} > 0 .$$
 (3.43)

The factor  $\Lambda^{-1}$  arises, because there is one transfer operator T on every rung but j. There, instead of T is an object that has the same structure as T, but it is built from the  $B^s$ . The series of equalities in (3.35) shows that such an object is equal to an identity operator on site j.

The energy expectation value is bit more complicated, because the exchanged matrix set  $B^s$  at site j breaks the translational invariance of the system and the relative location of the local Hamiltonian  $h_i$  to site j becomes important. Instead of the local term  $h_i$ , one has to look at the full Hamiltonian  $H = \sum_i h_i$  now. For an infinite system, H will always produce an infinite constant ground state energy  $E_0 = Le_0$  that has to be subtracted in order to obtain a well defined result. In each step the current best estimate for the exact  $e_0$ , i.e., the  $e_0(A^s)$  computed from the last available  $A^s$  is used.

Therefore, the function to be minimized in the local variation process is

$$E(A^{s}, B^{s}) := \frac{\langle \psi_{j}(A^{s}, B^{s}) | \left[ \sum_{i} \{h_{i} - e_{0}(A^{s})\} \right] |\psi_{j}(A^{s}, B^{s}) \rangle}{\langle \psi_{j}(A^{s}, B^{s}) | \psi_{j}(A^{s}, B^{s}) \rangle} .$$
(3.44)

The explicit dependence on  $A^s$  will be dropped in further notation, as the coefficients are fixed during one variational step.

As an example, let  $h_i$  consist of only one local operator  $\hat{O}_j$ . Then the energy as function of  $B^s$  is given by

$$E(B^{s}) = \sum_{j} \left( \frac{E_{j}}{(B|B)\Lambda^{-1}} - e_{0} \right)$$
(3.45a)  
with 
$$E_{j} = \begin{cases} \Lambda^{-|j|-1} \left( u, O_{j}^{(A,A)} T^{|j|-1} \mathbb{1}_{0}^{(B,B)} v \right) & \text{if } j < 0 \\ \Lambda^{-1} \left( u, O_{0}^{(B,B)} v \right) & \text{if } j = 0 \\ \Lambda^{-j-1} \left( u, \mathbb{1}_{0}^{(B,B)} T^{j-1} O_{j}^{(A,A)} v \right) & \text{if } j > 0 \end{cases}$$
(3.45b)

It is to be expected that the terms  $E_j - e_0(B|B)\Lambda^{-1}$  will converge to zero sufficiently quickly for  $|j| \gg 1$ . When interpreting the coefficients of  $B^s$  as a single vector

$$\vec{B} := \begin{pmatrix} \operatorname{vec}(B^1) \\ \vdots \\ \operatorname{vec}(B^d) \end{pmatrix} , \qquad (3.46)$$

the minimization of  $E(B^s)$  subject to the constraint of  $|\psi_j\rangle$  being normalized yields the bilinear form

$$\vec{B}^{\dagger}M\vec{B} - \mu\vec{B}^{\dagger}N\vec{B} = 0 , \qquad (3.47)$$

where residual powers of  $\Lambda$  are absorbed into the scalar lagrange multiplier  $\mu$ . The matrix N represents the normalization constraint for  $|\psi_j\rangle$ . Therefore, the variation leads to a generalized eigenvalue problem (EVP) of dimension  $dD^2$ 

$$\frac{\partial}{\partial \vec{B}^{\dagger}} (\vec{B}^{\dagger} M \vec{B} - \mu \vec{B}^{\dagger} N \vec{B}) = 0 \quad \Leftrightarrow \quad M \vec{B} - \mu N \vec{B} = 0 .$$
(3.48)

Note that due to the properties of the vectorization operator  $\operatorname{vec}(\cdot)$  the derivative of (3.47) with respect to  $\vec{B}^{\dagger}$  corresponds to the derivative of (3.45a) with respect to the coefficients of  $B^{s*}$  (without transposing).

To make the clouds of indices that will arise in defining M and N a little easier to understand, it will first be shown how the indexing scheme (3.14) extends by another level where sets of local matrices are concerned

$$A\vec{B} = \begin{pmatrix} A^{11} & \cdots & A^{1d} \\ \vdots & \ddots & \vdots \\ A^{d1} & \cdots & A^{dd} \end{pmatrix} \begin{pmatrix} \vec{B}^1 \\ \vdots \\ \vec{B}^d \end{pmatrix} .$$
(3.49)

The additional block-level coming from the physical index s will be labelled by superscript indices. Thus, on this physical block level, an element of the resulting vector reads

$$(A\vec{B})^s = \sum_{s'} A^{ss'} \vec{B}^{s'} .$$
(3.50)

Each of these blocks  $A^{ss'}$  fits the indexing scheme (3.14) again, so that an element of A is identified by a triple row index  $(s,\alpha,\alpha')$  and by an analogous column index  $(s',\beta,\beta')$ . An element of the product  $A\vec{B}$  is given by

$$(A\vec{B})^{s}_{\alpha\alpha'} = \sum_{s',\beta,\beta'} A^{ss'}_{\alpha\alpha',\beta\beta'} B^{s'}_{\beta\beta'} .$$
(3.51)

Starting with the normalization constraint, the matrix N is defined by

$$\frac{\partial}{\partial B^{s*}_{\alpha\alpha'}}(B|B) \stackrel{!}{=} \left(N\vec{B}\right)^{s}_{\alpha\alpha'} \tag{3.52a}$$

$$\frac{\partial}{\partial B^{s*}_{\alpha\alpha'}} \operatorname{Tr}\left(\sum_{s} B^{s\dagger} u^{\dagger} B^{s} v\right) = \frac{\partial}{\partial B^{s*}_{\alpha\alpha'}} \sum_{s} \sum_{\alpha\alpha'\beta\beta'} B^{s\dagger}_{\alpha'\alpha} u^{\dagger}_{\alpha\beta} B^{s}_{\beta\beta'} v_{\beta'\alpha'}$$
(3.52b)

$$=\sum_{s}\sum_{\beta\beta'}u^{\dagger}_{\alpha\beta}v_{\beta'\alpha'}B^{s}_{\beta\beta'}=\sum_{ss'}\sum_{\beta\beta'}\delta_{ss'}u^{\dagger}_{\alpha\beta}v^{T}_{\alpha'\beta'}B^{s}_{\beta\beta'} \qquad (3.52c)$$

$$=\sum_{s}\sum_{\beta\beta'}\left(\sum_{s'}\delta_{ss'}u^{\dagger}_{\alpha\beta}v^{T}_{\alpha'\beta'}\right)B^{s}_{\beta\beta'}=:N^{ss'}_{\alpha\alpha',\beta\beta'}B^{s}_{\beta\beta'} \quad (3.52d)$$

$$\Rightarrow \quad N_{\alpha\alpha',\beta\beta'}^{ss'} = \sum_{s} \delta_{ss'} u_{\alpha\beta}^{\dagger} v_{\alpha'\beta'}^{T} \tag{3.52e}$$

$$\Rightarrow \quad N = \mathbb{1}_d \otimes u^{\dagger} \otimes v^T . \tag{3.52f}$$

As an example on how to construct M, take the j = 0 term from (3.45b)

$$\frac{\partial}{\partial B^{s*}_{\alpha\alpha'}}(u, O^{(B,B)}_{0}v) \stackrel{!}{=} \left(M^{[j=0]}\vec{B}\right)^{s}_{\alpha\alpha'} \tag{3.53a}$$

$$\frac{\partial}{\partial B^{s*}_{\alpha\alpha'}}\operatorname{Tr}\left(u^{\dagger}\sum_{ss'}O^{ss'}B^{s'}vB^{s\dagger}\right) = \frac{\partial}{\partial B^{s*}_{\alpha\alpha'}}\operatorname{Tr}\left(\sum_{ss'}O^{ss'}B^{s\dagger}u^{\dagger}B^{s'}v\right)$$

$$= \frac{\partial}{\partial B^{s*}_{\alpha\alpha'}}\sum_{ss'}O^{ss'}\sum_{\alpha\alpha'\beta\beta'}B^{s\dagger}_{\alpha'\alpha}u^{\dagger}_{\alpha\beta}v_{\beta'\alpha'}B^{s'}_{\beta\beta'}$$

$$= \sum_{s'}\sum_{\beta\beta'}\left(\sum_{s}O^{ss'}u^{\dagger}_{\alpha\beta}v^{T}_{\alpha'\beta'}\right)B^{s'}_{\beta\beta'} \tag{3.53b}$$

$$\Rightarrow M^{[j=0]ss'}_{\alpha\alpha'} = \sum_{s'}O^{ss'}u^{\dagger}_{\alpha\beta'}v^{T}_{\alpha'\beta'} \tag{3.53c}$$

$$\Rightarrow \quad M^{[j=0]ss'}_{\alpha\alpha',\beta\beta'} = \sum_{s} O^{ss'} u^{\dagger}_{\alpha\beta} v^{T}_{\alpha'\beta'} \tag{3.53c}$$

$$\Rightarrow \quad M^{[j=0]} = O \otimes u^{\dagger} \otimes v^T . \tag{3.53d}$$

For  $j \neq 0$  the contribution  $M^{[j]}$  is constructed in exactly the same way with some minor modifications

$$M^{[j\neq 0]} = \mathbb{1}_d \otimes u^{\prime\dagger} \otimes v^{\prime T} \tag{3.54a}$$

with 
$$v' = \begin{cases} T^{j-1}O_j^{(A,A)}v & \text{if } j > 0\\ v & \text{if } j < 0 \end{cases}$$
 (3.54b)

and 
$$u' = \begin{cases} u & \text{if } j > 0 \\ T^{\dagger |j| - 1} O^{\dagger (A, A)} u & \text{if } j < 0 \end{cases}$$
 (3.54c)

The full matrix M is then given by

$$M = \sum_{j=-\infty}^{\infty} (M^{[j]} - e_0 N) .$$
 (3.55)

Obviously, for numerical calculations, the sum has to be cut off at some point. For finite correlation length, this can be done when

$$\left\| M^{[i]} - e_0 N \right\| < \epsilon \tag{3.56}$$

for some threshold value  $\epsilon$ . If the system shows critical behaviour, i.e. diverging correlation length, this point will never be reached and a hard cutoff at some  $j_{\text{max}}$  has to be made. This leads to truncation errors.

Now the case of a more complex local Hamiltonian  $h_i$  remains to be covered. If  $h_i$  is a sum of local operators acting on a single site i

$$h_i = \sum_{k=1}^{n} O_i^k \tag{3.57}$$

it easy to see that

$$E_{j=0} = \sum_{k=1}^{n} (u, O_0^{k, (B, B)} v)$$
(3.58)

and for  $j \neq 0$  analogously. If  $h_i$  is a product for local operators acting on adjacent sites

$$h_i = O_j^1 O_{j+1}^2 \cdots O_{j+n-1}^n , \qquad (3.59)$$

the single term for j = 0 in (3.45b) is replaced with a sum of n terms, as can be seen easily from the tensor networks in Fig. 3.3.

$$E_{\text{center}} = \sum_{k} E_{k} = \sum_{k} (\underbrace{u, O^{1, (A, A)} \cdots O^{k-1, (A, A)}}_{u'} O^{k, (B, B)} \underbrace{O^{k+1, (A, A)} \cdots O^{n, (A, A)}}_{v'} v) \quad (3.60)$$

These  $E_k$  result in contributions to M of the form

$$M^{[k]} = O^k \otimes u'^{\dagger} \otimes v'^T - e_0 N . \qquad (3.61)$$



Figure 3.3: Example for the tensor networks arising from the local Hamiltonian  $h_i$  acting on site 0 where  $A^s$  is replaced with  $B^s$ . Here  $h_i$  acts on n = 3 lattice sites. This could e.g. be a next-nearest neighbour interaction  $h_i = S_i^x \mathbb{1}_{i+1} S_{i+2}^x$ .

Note that N is always hermitian if v and u are hermitian

$$N^{\dagger} = (\mathbb{1} \otimes u^{\dagger} \otimes v^{T})^{\dagger} = \mathbb{1}^{\dagger} \otimes u \otimes (v^{T})^{\dagger} = \mathbb{1} \otimes u^{\dagger} \otimes v^{T} = N .$$

$$(3.62)$$

If additionally all the local operators appearing in H are hermitian, so is M, as local operators map hermitian matrices to hermitian matrices, and therefore all v' and u' are hermitian. Then, for any contribution to M the relation

$$M^{[j]\dagger} = (O \otimes u^{\prime\dagger} \otimes v^{\prime T})^{\dagger} = O^{\dagger} \otimes u^{\prime} \otimes (v^{\prime T})^{\dagger} = O \otimes u^{\prime\dagger} \otimes v^{\prime T} = M^{[j]}$$
(3.63)

holds.

The basic idea of the local variation approach is to solve the generalized EVP

$$M\vec{B} = \mu N\vec{B} \tag{3.64}$$

for the eigenvalues  $\mu_{\alpha}$  and the corresponding eigenvectors  $\vec{B}_{\alpha}$ . The  $\mu_{\alpha}$  will be assumed to be orderd  $\mu_0 \leq \mu_1 \leq \cdots \leq \mu_{dD^2}$  in the subsection discussions. The eigenvector  $\vec{B}_0$  with the lowest lying eigenvalue is used as the new local matrix set  $A^s$  in the next variation step. This can be repeated until convergence is reached i.e.  $\vec{B}_0 \equiv \text{vec}(A^s)$ .

There is no rigorous proof that using the eigenvector with the lowest local energy (the lowest eigenvalue  $\mu_0$ ) on every site in the next step will improve the result in this iterated process. But this is found to be the case, given that the generalized EVP *can* be solved.

#### 3.2.2 Notes on implementation

In some cases, when  $h_i$  is of simple form, the matrices M and N are sparse and the EVP can become ill-conditioned. The problem can be stabilized for numerical calculations by transforming it into a standard EVP. Let  $N = Q^{\dagger} D_N Q$  so that  $D_N$  is diagonal. Because N is hermitian, Q is unitary i.e.  $Q^{-1} = Q^{\dagger}$ .

$$M\vec{B} = \mu NB \quad \Leftrightarrow \quad M\vec{B} = \mu Q^{\dagger} D_N Q\vec{B}$$
 (3.65a)

$$\Leftrightarrow \quad M\vec{B} = \mu Q^{\dagger} \sqrt{D_N} \mathbb{1} \sqrt{D_N} Q\vec{B} \tag{3.65b}$$

$$\Leftrightarrow \underbrace{\sqrt{D_N}^{-1}QMQ^{\dagger}\sqrt{D_N}^{-1}}_{=:M'}\underbrace{\sqrt{D_N}Q\vec{B}}_{=:\vec{B}'} = \mu\sqrt{D_N}\vec{B}'$$
(3.65c)

$$\Leftrightarrow M'\vec{B}' = \mu\vec{B}' \quad \Rightarrow \quad \vec{B} = Q^{\dagger}\sqrt{D_N}^{-1}\vec{B}' . \tag{3.65d}$$

The last form is a hermitian standard EVP (in case M and N are real it is even symmetric) and it is better suited for numerical solution.

In some cases, e.g., in the strong field limit of the ITF, convergence of the local variation algorithm is very slow. It can then be advantageous to use the global variation method instead if the energy landscape is of simple engough shape.

Also, to make calculations easier, it is advantageous to have  $\Lambda = 1$ , which for unique  $\Lambda$  can always be achived by the transformation

$$A^s \to \tilde{A}^s := \frac{A^s}{\sqrt{\Lambda}}$$
 (3.66)

If  $\Lambda$  is degenerate in its absolute value, different  $\Lambda_i$  will be of absolute value 1. This avoids introducing new errors when multiplying or deviding by large powers of  $\Lambda_i$ . Especially division is important, since every explicitly applied operator - including T - comes with a factor of  $\Lambda^{-1}$ . The easiest way to assure the correct powers of  $\Lambda$  is to implement this division directly into the application of operators.

The scalar product of the eigenvectors v and u can also be renormalized to 1 by setting

$$v \rightarrow \tilde{v} := \frac{v}{\sqrt{(u,v)}}, \quad u \rightarrow \tilde{u} := \frac{u}{\sqrt{(u,v)}}.$$
 (3.67)

To check for convergence in the  $M^{[i]}$  instead of using the criterion Eq. (3.56) the sum can be cut when

$$T^{j-1}h_i^{(A,A)}v \cong T^{j-2}h_i^{(A,A)}v$$
(3.68)

within a reasonable tolerance. Equation (3.56) holds, if  $h_i$  acts far away from the perturbation  $B^s$  and thus its influence is negligible. The same is also true, if moving one site further away from  $B^s$  does not change the result anymore which is reflected in Eq. (3.68). This has the advantage, that the matrices in Eq. (3.68) are much smaller than those in Eq. (3.56).

The convergence behaviour of the ground state search can be stabilized further if not just  $\vec{B}_0$  is used as the new  $A^s$  in each step, but a linear combination of the given  $A^s$  and the resulting  $\vec{B}_0$ . This linear combination is determined by another linear minimization

$$\min_{x} \left[ e_0(\cos(x)\operatorname{vec}(A^s) + \sin(x)\vec{B}_0) \right] .$$
(3.69)

This variant ensures, that the energy decreases monotonously. According to our results it is, however, more prone to get stuck in areas of the energy landscape with very little slope.

#### 3.2.3 Algorithm

From the above sections, a variational algorithm to find an approximation of the ground state in a given MPS class can be defined. It is given in Tab. 3.1.

The actual criterion used to detect convergence is the change in the ground state energy

$$\Delta e_0 := e_0(A_n^s) - e_0(A_{n+1}^s) \tag{3.70}$$

from step n to step n + 1. If  $\Delta e_0$  is smaller than some threshold value  $\epsilon_1$ , the algorithm breaks. As mentioned above, this can also happen when it gets in a near flat part of the energy landscape. In this case however, usually  $\vec{B}_0 \neq \text{vec}(A^s)$ . Therefore, the additional criterion

$$\vec{B}_0 \cong \operatorname{vec}(A^s) \quad \Leftrightarrow \quad \left\| \frac{T_A}{\sqrt{\operatorname{Tr}(T_A^{\dagger}T_A)}} - \frac{T_B}{\sqrt{\operatorname{Tr}(T_B^{\dagger}T_B)}} \right\| < \epsilon_2$$
 (3.71)

has to be satisfied, where  $T_A$  and  $T_B$  are the transfer operators build from  $A^s$  and  $B_0$  respectively and  $\epsilon_2$  is another threshold value that has to be chosen several orders of magnitude larger than  $\epsilon_1$ .

#### Table 3.1: Local variation algorithm

- 0. Initialization: Set an initial value for  $A^s$ . If nothing better is available, use random matrices.
- 1. Compute  $\Lambda$ , v, u,  $e_0$  for the given  $A^s$ .
- 2. Construct matrices M and N for generized EVP. Solve  $M\vec{B} = \mu N\vec{B}$ .
- 3. If  $\Delta e_0 < \epsilon_1$  and  $A^s \cong \vec{B}_0$  convergence is reached. Break. If  $\Delta e_0 < \epsilon_1$  but  $A^s \neq \vec{B}_0$  start from step 0.
- 4. Else set  $A^s := \min_{e_0(x)} (\cos(x) \operatorname{vec}(A^s) + \sin(x) \overline{B}_0)$  and start from step 1

As mentioned above, in some cases the generalized EVP is numerically unstable. Also, if the initial value of  $A^s$  is chosen poorly, convergence may be very slow. Therefore, the current implementation uses a slightly more complex version of the algorithm which is given in Tab. 3.2 where n is a positive integer with  $n \approx 50$ .

#### Table 3.2: Extended ground state search algorithm

- 1. For a maximum of n times try the local variation algorithm in Tab. 3.1. If it does not converge within n iterations, break and use a new random initial guess.
- 2. If the local variation algorithm fails to produce a result after n runs, switch to the global variation method, i.e. minimize  $e_0$  using a standard minimization algorithm.

## 3.3 Results

This section shows the results for the ground state energy per site obtained from the method in comparison to the exact solution (2.18b).



Figure 3.4: Ground state energy per lattice site  $E_0/L$  as function of parameter  $\lambda$ 

Figure 3.4 shows the ground state energy as function of the parameter  $\lambda$ . Even for a value of D as small as 3 the agreement is remarkably good. Rather more telling, however, is the deviation from the exact result

$$\Delta E := \frac{E_{0,\text{var}}}{L} - \frac{E_{0,\text{exact}}}{L} \tag{3.72}$$

which is shown in Fig. 3.5. There, wild oscillations for small  $\lambda$  can be observed. This is due to the fact that close to the free spin limit, especially the norm matrix N is very sparse, because u and v contain many near-zero values, cf. Eq. (3.52f). This results in many near-zero eigenvalues of N, making the generalized EVP ill-conditioned and amplifying numercial noise.

Also the quality of the results from the global variation method strongly depends on the starting vector in the current implementation.

For larger  $\lambda$  the deviation increases rather smoothly, peaks at a value close to but smaller than 1 and decreases again. Although the result for each value of  $\lambda$  is used as a starting vector for the next higher one, this form of the deviation curve is *not* a hysteresis effect. The peak occurs at the same position if the sweep is done from higher to lower values of  $\lambda$ . So the method picks up the critical point, marked by the peak in the deviation, but does so at a too low value of  $\lambda$ . However, for increasing D this point moves towards  $\lambda = 1$ .

Note that although the ground state shows a twofold degeneracy for  $\lambda > 1$ , the method is able to find one of the realizations with increasing precision for higher  $\lambda$ .



Figure 3.5: Deviation of the variational results for the ground state energy form exact solution (2.18b)

## 3.4 Existing methods

The ground state energy of an infinite one-dimensional system can also be obtained from a multitude of other methods, including those mentioned in Chap. 1.

Especially finite size scaling with by now highly optimized real space DMRG algorithms is known to yield very accurate results for a large variety of models.

Vidal's iTEBD [18] and the MPS approach in Ref. [19] work directly in the thermodynamic limit, but rely on imaginary-time evolution to find the ground state, which is inefficient [16].

McCullochs iDMRG method, that is conceptionally similar to our approach, also yields results for infinite systems. The MPS algorithm in Ref. [20] uses a slightly different MPS format, but is basically equivalent to the global variation ground state search.

And finally, finite size scaling from the methods in Ref. [22, 24] are of course possible although, at the cost of several runs.

## Chapter 4

# Dispersion

This chapter shows how the matrices B from the ground state search algorithm in Sect. 3.2 describe excitations so that the one-particle dispersion  $\omega_q$  is obtained. The numerical results will be compared to the analytical expression (2.18a). In the last section existing methods to achive this will be discussed, especially MPS based approaches.

## 4.1 Description of excitations

#### 4.1.1 Momentum space variation

When a ground state approximation  $A^s$  for given D has been found, the eigenvector  $\vec{B}_0$ of the generalized EVP in Eq. (3.64) with the lowest eigenvalue  $\mu_0$  describes the ground state again. Since (3.64) arises from the energy functional (3.1), it is not far fetched to assume that solutions with higher lying eigenvalues  $\mu_{\alpha>0}$  describe excitations in the system.

Let  $|\psi_j^{\alpha}\rangle$  be the state that has the ground state matrix set  $A^s$  everywhere but on site j, where the set formed from the eigenvector  $\vec{B}_{\alpha}$  ( $\alpha > 0$ ) is inserted instead.

Especially, if an elementary excitation is localized to a single lattice site, one would expect the eigenvector  $\vec{B}_1$  with the second lowest eigenvalue  $\mu_1$  to represent this excitation. The one-particle dispersion could then be obtained as shown in Sect. 2.3

$$\omega_q^{(\alpha)} = \frac{\langle \psi_q^{\alpha} | (H - E_0) | \psi_q^{\alpha} \rangle}{\langle \psi_q^{\alpha} | \psi_q^{\alpha} \rangle} \bigg|_{\alpha = 1}$$
(4.1a)

with 
$$|\psi_q^{\alpha}\rangle := \frac{1}{\sqrt{L}} \sum_j e^{iqr_j} |\psi_j^{\alpha}\rangle$$
 (4.1b)

Note that the normalization  $\frac{1}{\sqrt{L}}$  is actually never computed in practice, as it always cancels out when evaluating expectation values in a translationally invariant system.

At the free spin limit  $\lambda = 0$ , there are no dispersive effects and any excitation has the energy  $\omega \equiv \Gamma$ , independently of the wave vector q. This is found to be correctly represented by  $\vec{B}_1$  up to deviations that can be regarded as numerical artefacts. But, for increased spin-spin coupling J, the assumption no longer holds. This is not surprising, since for  $\lambda > 0$  a flipped spin is expected to create a polarization cloud, so the elementary excitation is no longer completely localized.

By construction any  $|\psi_i^{\alpha>0}\rangle$  is orthogonal to the ground state  $|\psi_0\rangle$ 

$$\langle \psi_j^{\alpha>0} | \psi_0 \rangle = \vec{B}_{\alpha>0} N \vec{B}_0 = 0 \quad \forall \ j \ .$$

$$\tag{4.2}$$

Therefore, assuming the ground state was calculated sufficiently exactly, the dispersion can be obtained by variation in its orthogonal complement, i.e., the subspace spanned by the  $\vec{B}_{\alpha>0}$ .

For two eigenvectors  $\vec{B}_{\alpha}$  and  $\vec{B}_{\beta}$  at the lattice sites *i* and *j*, the overlap and the matrix element of the Hamiltonian (corrected by the ground state energy) are defined as

$$N_{ji}^{\alpha\beta} := \langle \psi_j^{\alpha} | \psi_i^{\beta} \rangle \tag{4.3a}$$

$$H_{ji}^{\alpha\beta} := \langle \psi_j^{\alpha} | (H - E_0) | \psi_i^{\beta} \rangle .$$
(4.3b)

Since the system is infinite and translationally invariant but for the local perturbations  $B^s_{\alpha}$  and  $B^s_{\beta}$ , the matrix elements  $H^{\alpha\beta}_{ji}$  and  $N^{\alpha\beta}_{ji}$  depend only on the relative coordinate j - i, but not on the actual locations j and i. Therefore, from now on i = 0 will be assumed. This leads to the simplified expressions

$$N_j^{\alpha\beta} := \langle \psi_j^{\alpha} | \psi_0^{\beta} \rangle \tag{4.4a}$$

$$H_j^{\alpha\beta} := \langle \psi_j^{\alpha} | (H - E_0) | \psi_0^{\beta} \rangle .$$
(4.4b)

See Sect. 4.1.2 on how to compute these matrix elements.

Unfortunately, the  $|\psi_i^{\alpha}\rangle$  do not form an orthogonal basis, i.e.,

$$\langle \psi_i^{\alpha} | \psi_j^{\alpha} \rangle \neq \delta_{ij} . \tag{4.5}$$

This can be solved by using the momentum space basis defined in Eq. (4.1b) which is orthogonal with respect to q

$$\langle \psi_q^{\alpha} | \psi_{q'}^{\beta} \rangle \propto \delta_{qq'}$$
 (4.6)

Figure 4.1 shows the dispersion computed from

$$\omega_q^{(\alpha)} = \frac{\langle \psi_q^{\alpha} | (H - E_0) | \psi_q^{\alpha} \rangle}{\langle \psi_q^{\alpha} | \psi_q^{\alpha} \rangle} = \frac{H_q^{\alpha \alpha}}{N_q^{\alpha \alpha}}$$
(4.7a)

with 
$$H_q^{\alpha\alpha} := \sum_{j=-\infty}^{\infty} e^{iqj} H_j^{\alpha\alpha}, \qquad N_q^{\alpha\alpha} := \sum_{j=-\infty}^{\infty} e^{iqj} N_j^{\alpha\alpha}$$
(4.7b)

for several  $\alpha$ . The plot shows two important things. First, not only  $\vec{B}_1$  contributes to the dispersion, but serveral other  $\vec{B}_{\alpha}$  too. However, no single curve alone is a good approximation. Secondly, there are distinct peaks where the Fourier transform of the norm,  $N_q^{\alpha\alpha}$ , has near-zeros.

This indicates, that the  $|\psi_q^{\alpha}\rangle$  are still not orthogonal with respect to  $\alpha$ . Also, some of them can hold redundant information.

The best estimate for the one-particle dispersion  $\omega_q$  can now be determined by minimizing

$$\omega_q = \frac{\langle \phi_q | (H - E_0) | \phi_q \rangle}{\langle \phi_q | \phi_q \rangle} \tag{4.8a}$$

with 
$$|\phi_q\rangle := \sum_{\alpha} v_q^{(\alpha)} |\psi_q^{\alpha}\rangle$$
 (4.8b)



Figure 4.1: Results from naive dispersion approach Eq. (4.7) for various eigenvectors  $\vec{B}_{\alpha}$ . Matrix dimension is D = 5, parameter is  $\lambda = 0.8$ . Note from the curve  $\alpha = 4$  that the eigenvectors with the lowest local energy  $\mu_{\alpha}$  are *not* always those with the most significant contributions to the dispersion. The remaining 9 curves for  $\alpha > 8$  have been left out, as they are all well above the lower boundary of the two-quasiparticle continuum (dashed red line).

This leads to a bilinear form in the vector  $v_q$  and subsequently to the generalized EVP

$$H_q v_q = \omega_q N_q v_q , \qquad (4.9)$$

in which the elements of the matrices  $H_q$  and  $N_q$  are given by

$$H_q^{\alpha\beta} := \sum_j e^{ikj} H_j^{\alpha\beta}, \qquad N_q^{\alpha\beta} := \sum_j e^{ikj} N_j^{\alpha\beta} . \tag{4.10}$$

Computing these coefficients  $H_j^{\alpha\beta}$  is by far the most time consuming part of the calculation. Therefore, it is desirable to restrict the dimension of the momentum space EVP to a dimension  $F < D^2d - 1$ .

Note, that  $H_q$  and  $N_q$  are hermitian and therefore an orthonormal eigensystem exsits. This follows from Eq. (4.21). Instead of solving the generalized EVP (4.9) directly, again the approach of transforming it into a standard EVP is chosen. To do this, first the norm matrix  $N_q$  is diagonalized.

Figure 4.2 shows the first 8 eigenvalues  $n_q^{(\alpha)}$  of  $N_q$ . It is easy to see that there is one eigenvalue that is very small and likely to cause numerical problems. This was also noted by Verstraete et al. in Ref. [24]. It is caused by the afore-mentioned redundancy in the  $\vec{B}_{\alpha}$ 

which means, they are not orthogonal in momentum space. Due to this over-counting, the dimension of the space effectively spanned by the  $\vec{B}_{\alpha}$  is less than  $dD^2 - 1$ , causing close to zero eigenvalues of the norm matrix.

However, if these lowest eigenvalues are well separated from the others, as it is the case in Fig. 4.2, their eigenspaces can be discarded without significant loss of precision.

The energy matrix  $H_q$  then has to be projected onto the subspace effectively spanned by the  $\vec{B}_{\alpha}$ , where the norm is well behaved. For larger D, the number of such problematic norm eigenvalues is expected to grow. Let f be the number of eigenvalues of  $N_q$  that are smaller than some threshold value for all q in the Brillouin zone and which are well separated from the rest. The  $n_q^{(\alpha)}$  are assumed to be in ascending order  $n_q^{(1)} \leq n_q^{(2)} \leq \cdots \leq n_q^{(F)}$ . The corresponding eigenvectors are labelled  $\vec{\kappa}_q^{(\alpha)}$ . The energy matrix is then transformed as

$$\tilde{H}_q = K^{\dagger} H_q K, \quad \text{with} \quad K = \left(\frac{\vec{\kappa}_q^{(f+1)}}{\sqrt{n_q^{(f+1)}}} \quad \cdots \quad \frac{\vec{\kappa}_q^{(F)}}{\sqrt{n_q^{(F)}}}\right) .$$
(4.11)

This leads to the standard EVP

$$H_q \tilde{v}_q = \omega_q \tilde{v}_q , \qquad (4.12)$$

whose lowest eigenvalue is the best estimate for the dispersion relation. The meaning of the corresponding eigenvector  $v_q$  will be discussed in Chap. 5.

Figure 4.1 also shows remaining major problem. It is clear that only a few  $\vec{B}_{\alpha}$  significantly contribute to the dispersion and most of them do not. But other than manually looking at the diagonal elements in *q*-space, no failsafe way could be found to determine which  $\alpha$  are the relevant ones.

Increasing the EVP dimension F increases the chance of including all relevant  $\vec{B}_{\alpha}$ . But it also increases the chance of the low-lying  $n_q^{(\alpha)}$  intersecting with the higher ones. This makes it difficult to find and discard superfluous eigenspaces of the norm matrix.

At this point, it is not fully understood how exactly the information is distributed amoung the  $\vec{B}_{\alpha}$ . At low matrix dimensions and low values of  $\lambda$  the relevant information appears to be contained in the eigenvectors with the lowest local energy. Increasing F beyond these relevant contributions does not significantly improve the result.

For increased parameter  $\lambda$ , there is a higher probability to find non-relevant eigenvectors inbetween the important ones (see the  $\alpha = 4$  curve in Fig. 4.1), so that local energy alone is not a sufficient criterion. Also, it appears that for larger D the information becomes more widely scattered across the  $D^2d - 1$  eigenvectors.

The exact properties of the eigenvectors and the way the quality of the result depends on the choice of them will be subject of detailed investigations in the future.

#### 4.1.2 Computing the matrix elements

The tensor networks for the norm elements are of the form shown in Fig. 4.3 and lead to the following expressions

$$N_{j}^{\alpha\beta} = \begin{cases} \Lambda^{-|j-i|-1} \left( u, \, \mathbb{1}_{j}^{(B_{\alpha},A)} T^{|j-i|-1} \mathbb{1}_{0}^{(A,B_{\beta})} v \right) & \text{if } j < 0 \\ \Lambda^{-1} \left( u, \, \mathbb{1}_{0}^{(B_{\alpha},B_{\beta})} v \right) & \text{if } j = 0 \quad (\text{cf. Eq. (3.43)}) \quad . \quad (4.13) \\ \Lambda^{-|j-i|-1} \left( u, \, \mathbb{1}_{0}^{(A,B_{\beta})} T^{j-i-1} \mathbb{1}^{(B_{\alpha},A)} v \right) & \text{if } j > 0 \end{cases}$$



Figure 4.2: Eigenvalues  $n_q^{(\alpha)}$  of the norm matrix  $N_q$  from Eq. (4.9). Note that the lowest eigenvalue  $n_q^{(1)}$  is smaller than  $10^{-5}$  in the whole Brillouin zone and well seperated from the others. Its subspace can be discarded in calculating  $\omega_q$ . Also note that there are intersections between the higher eigenvalues, e.g. between  $n_q^{(3)}$  and  $n_q^{(4)}$ . Discarding all eigenvalues up to  $n_q^{(3)}$  would lead to the consideration of different Hilbert spaces on each side of an intersection and thus give unpredictable results (cf. Eq. (5.3) an comments).



**Figure 4.3:** Tensor network for norm element  $N_i^{\alpha\beta}$ 

The energy matrix elements are more complex. Since again translational invariance is broken by the local perturbations, the full Hamiltonian has to be taken into account. Thus

$$H_j^{\alpha\beta} = \sum_i \left( \langle \psi_j^{\alpha} | h_i | \psi_0^{\beta} \rangle - e_0 N_j^{\alpha\beta} \right) =: \sum_i \left( h_i^{\alpha\beta} - e_0 N_j^{\alpha\beta} \right) . \tag{4.14}$$

Again, it is to be expected that if both  $B^s$  are far away from the site *i* the local Hamiltonian acts on, the influence is negligible. This means, that the summands should converge to zero sufficiently quickly. At and close to criticality, again truncation errors occur.

The Hamiltonian matrix element is given by a sum of tensor networks of the kind shown in Fig. 4.4.



Figure 4.4: Representative contributions to the matrix element  $h_{j}^{\alpha\beta}$ 

Three possible situations need to be distinguished.

(i) The term  $h_i$  acts outside the network block enclosed by the two  $B^s$  matrix sets (diagram Fig. 4.4a). This leads to scalar products of the form

$$h_{j}^{[i]\alpha\beta} = \begin{cases} (u', \mathbb{1}_{j}^{(B_{\alpha},A)}T^{|j|-1}\mathbb{1}_{0}^{(A,B_{\beta})}v') & \text{if } j < 0\\ (u', \mathbb{1}_{0}^{(A,B_{\beta})}T^{j-1}\mathbb{1}_{j}^{(B_{\alpha},A)}v') & \text{if } j > 0 \end{cases}$$
(4.15)

with 
$$v' = \begin{cases} v & \text{if } i < 0\\ T^{i-\max(j,0)-1}h_i^{(A,A)}v & \text{if } i > 0 \end{cases}$$
 (4.16)

and 
$$u' = \begin{cases} T^{\dagger |i| - |\min(j,0)| - n - 1} h_i^{\dagger (A,A)} u & \text{if } i < 0 \\ u & \text{if } i > 0 \end{cases}$$
 (4.17)

(ii)  $h_i$  acts on site 0 or site j (diagram Fig. 4.4b) or both (diagram Fig. 4.4c). This works in close analogy to the local term (j = 0) covered in the previous chapter. These "transition terms" yield expressions such is

$$h_{j}^{[i]\alpha\beta} = \begin{cases} (u, O^{1,(A,A)} \cdots O_{i=j}^{k,(B_{\alpha},A)} \cdots O^{n,(A,A)} T^{|j|-1-(n-k)} \mathbb{1}_{0}^{(A,B_{\beta})} v) & \text{if } j < 0\\ (u, O^{1,(A,A)} \cdots O_{i=0}^{k,(A,B_{\beta})} \cdots O^{n,(A,A)} T^{|j|-1-(n-k)} \mathbb{1}_{j}^{(B_{\alpha},A)} v) & \text{if } j > 0 \end{cases}$$
with  $k \in \{1, \dots, n\}$ 

$$(4.18)$$

for the left side perturbation at site min(j,0), and analogously for the right side. If |j| < n, another type of addition appears, where  $h_i$  acts on both  $B^s$ 

$$h_{j}^{[i]\alpha\beta} = \begin{cases} (u, O^{1,(A,A)} \cdots O^{k,(B_{\alpha},A)}_{i+k-1=j} \cdots O^{k+|j|-1,(A,B_{\beta})}_{0} \cdots O^{n,(A,A)}v) & \text{if } j < 0\\ (u, O^{1,(A,A)} \cdots O^{k,(A,B_{\beta})}_{i+k-1=0} \cdots O^{k+|j|-1,(B_{\alpha},A)}_{j} \cdots O^{n,(A,A)}v) & \text{if } j > 0 \end{cases}$$
with  $k \in \{1, \dots, n-|j|\}$ .

(iii) for |j| > n a different type occurs, where  $h_i$  is completely enclosed between both  $B^s$  (diagram Fig. 4.4d). This results in contributions of the form

$$h_{j}^{[i]\alpha\beta} = \begin{cases} (u, \mathbb{1}_{j}^{(B_{\alpha},A)} T^{m} h_{i}^{(A,A)} T^{|j|-1-n-m} \mathbb{1}_{0}^{(A,B_{\beta})} v) & \text{if } j < 0\\ (u, \mathbb{1}_{0}^{(A,B_{\beta})} T^{m} h_{i}^{(A,A)} T^{|j|-1-n-m} \mathbb{1}_{j}^{(B_{\alpha},A)} v) & \text{if } j > 0\\ \text{with } m \in \{0, \dots, |j|-1-n\} . \end{cases}$$

$$(4.20)$$

All  $h_j^{\alpha\beta}$  acquire an additional factor of  $\Lambda^m$  where m is the number of lattice sites spanned by the central part of the network.

Computing these  $H_j^{\alpha\beta}$  is by far the most time consuming part of the calculation. Therefore it is very useful that because of translational invariance the following relation holds

$$H_j^{\alpha\beta} = \langle \psi_j^{\alpha} | (H - E_0) | \psi_0^{\beta} \rangle$$
(4.21a)

$$= (\langle \psi_0^{\beta} | (H^{\dagger} - E_0^*) | \psi_j^{\alpha} \rangle)^*$$
(4.21b)

$$= (\langle \psi_{-j}^{\beta} | (H - E_0) | \psi_0^{\alpha} \rangle)^* = H_{-j}^{\beta \alpha *}$$
(4.21c)

and analogously for  $N_j^{\alpha\beta}$ . Interpreting  $H_j^{\alpha\beta}$  and  $N_j^{\alpha\beta}$  as elements of matrices  $H_j$  and  $N_j$  this means

$$H_{-j} = H_j^{\dagger} \tag{4.22a}$$

$$\mathbf{N}_{-j} = N_j^{\dagger} , \qquad (4.22b)$$

which reduces the computational cost by a factor of 2.

#### 4.1.3 Notes on implementation

A look at Eq. (4.13) suggests that  $N_j^{\alpha\beta}$  can be computed more efficiently if keeping, e.g., the right part of the tensor network given by the  $D \times D$  matrix

$$v' = T^{j-1} \mathbb{1}_{j}^{(B_{\alpha},A)} v \qquad (j > 0)$$
(4.23)

in memory. Then the next Fourier coefficient  $N_{i+1}^{\alpha\beta}$  is easily computed as

$$N_{j+1}^{\alpha\beta} = (\mathbb{1}^{(A,B_{\beta})}u, Tv') .$$
(4.24)

For the energy coefficients  $H_j^{\alpha\beta}$ , a similar approach can be used. But, since for each value of j the local Hamiltonian  $h_i$  is "moved through the network", all occurring powers of Tand  $T^{\dagger}$  need to be stored. This procedure reduces CPU time from  $\mathcal{O}(j_{\max}^2)$  to  $\mathcal{O}(j_{\max})$ , but at the cost of increasing memory consumption from a constant to  $\mathcal{O}(j_{\max})$ .

We found that the matrix elements  $H_q^{\alpha\beta}$  and  $N_q^{\alpha\beta}$  are either purely real or purely imaginary up to numerical noise from the Fourier transform. Also, if  $H_q^{\alpha\beta}$  is real, then  $N_q^{\alpha\beta}$  is so as well and correspondingly for  $H_q^{\alpha\beta}$  being imaginary.

It is not clear if this is a property of the method or the model. We still use this property to numerically stabilize the diagonalization process. A diagonal transformation of the form

$$U := \begin{pmatrix} u_1 & 0 \\ & \ddots & \\ 0 & & u_F \end{pmatrix} \quad \text{with} \quad u_{\alpha} = \begin{cases} 1 & \text{if } N_q^{1\alpha} \in \mathbb{R} \\ i & \text{if } N_q^{1\alpha} \in \mathbb{C} \end{cases}$$
(4.25)

transforms  $H_q$  and  $N_q$  into real symmetric matrices

$$H'_q := U^{\dagger} H_q U \tag{4.26a}$$

$$N'_q := U^{\dagger} N_q U , \qquad (4.26b)$$

from where the calculation proceeds as sketched above. Since such a transformation is unitary, it does not change the eigenvalues and  $\omega_a$  remains the same.

The corresponding eigenvector  $v_q$ , however, has to be transformed back into the original basis by

$$v_q = U^{\dagger} v_q' \tag{4.27}$$

for further use. This also avoids relative phase shifts in  $v_q$  from one value of q to the next, which would be misleading in the analysis of its Fourier transform.

#### 4.2Algorithm

Although the basic idea in calculating the dispersion is simple, the above-mentioned problems require a more sophisticated algorithm. It is listed in Tab. 4.1.

The algorithm is partially "heuristic", i.e., it is based on the observations of the previous sections. It does not always yield the optimal result possible for a given combination of Dand F, but it balances the use of computational resources, the need for human review at intermediate stages and the quality of the results.

Note, that the number f of small norm eigenvalues  $n_q^{(\alpha)}$  varies with the model parameters and also with D. Also, the threshold value below which problems occur may vary. Therefore, f from now on labels the number of norm eingenvalues actually discarded in the computation.

In the current implementation the F eigenvalues  $\vec{B}_{\alpha}$  with the lowest local energy are used as a starting point. But, this does not mean the same  $\alpha$  are utilized to compute the actual result. Observation shows, that for  $F \leq 10$  there is always at least one very small and well separated eigenvalue of  $N_q$ . Therefore, the basic algorithm discards the lowest eigenvector  $n_q^{(\alpha)}$ . However, in the calculations for Fig. 4.6 and 4.7 higher, manually determined values of f were used.

The ordering procedure in step 2 is based on the observation in Fig. 4.1. In most cases, the integral over the squared absolute value of the derivative is a good indicator for the probability that an eigenvector  $\vec{B}_{\alpha}$  significantly contributes to the dispersion. In this sense, if the consistency check in step 4 detects a problem, the algorithm will drop the eigenvector with the smallest probability.

**Table 4.1:** Alorithm for calculating the energy dispersion

- 1.
- Compute the Fourier coefficients  $H_j^{\alpha\beta}$  and  $N_j^{\alpha\beta}$  for  $j = 1, \ldots, j_{\text{max}}$ Compute the Fourier transform  $H_q^{\alpha\alpha}$  of the diagonal elements, order  $H_j$  and  $N_j$ 2. descendingly by  $\int_0^{\pi} \left| \frac{\mathrm{d}}{\mathrm{d}q} H_q^{\alpha \alpha} \right|^2 \mathrm{d}q$
- For  $q \in [0,\pi]$  compute and solve  $\tilde{H}_q \tilde{v}_q = \omega_q \tilde{v}_q$ , compute  $v_q$ 3.
- If  $|\max \omega_q \min \omega_q| > 10$ , reduce dimension of  $H_q$  and  $N_q$  by 1. 4. Start step 3. again

## 4.3 Results

In this section the results for the energy dispersion of the elementary excitations obtained from our method are presented. Figures 4.5 through 4.7 show the dispersion curve as function of wave vector q for one half of the first Brillouin zone, because  $\omega_q$  is an even function in q. The lower part of each plot shows the deviation of the variational results from the exact solution (2.18a). It is plotted as absolute value in a logarithmic scale for clarity. In this representation, the downward spikes indicate intersections with the exact solution. These mean, that the ground states was not found exactly. Otherwise, the variational principle would guarantee  $\omega_q$  to be a stict upper boundary.

The calculations were done with F = 20 except for D = 3 where  $F = dD^2 - 1 = 2 \cdot 9 - 1 = 17$  which is the maximum number possible.

As expected, on one hand higher local dimension D yields higher accuracy. On the other hand, parameter values  $\lambda$  closer to 1 decrease precision as the correlations become longer ranged. Of course, the failure to capture the closing energy gap at criticality is important to note. However, the results improve for increasing D. A detailed analysis of the convergence properties will require a more efficient implementation. But at least extrapolations of critical quantities should prove possible.

Table 4.2 compares the deviation in the ground state energy to the maximum and to the average deviation in the dispersion. Over a large part of the disordered phase ( $\lambda \leq 0.9$ ) the dispersion can be obtained to the same precision on average as the ground state energy. At the critical point, the precision deviates by about 3 orders of magnitude. The deviation in the dispersion is of course increased when the gap closes. Also, one should compare these deviations with the maximum deviation in the ground state energy, as the critical point is detected at a too low parameter value for small D.

Note, that the model shows no bound states of multiple particles in the thermodynamic limit. Therefore, knowing the one-particle dispersion  $\omega_q$  also provides the boundaries of multi-particle continua by means of Eq. (2.20) and similar expressions.

$\lambda$	D	$ \Delta E_0 $	$\max  \Delta \omega $	$ \langle \Delta \omega \rangle $
0.5	3	$3.73\cdot 10^{-7}$	$3.03\cdot 10^{-6}$	$1.47\cdot 10^{-6}$
	4	$3.01\cdot 10^{-9}$	$2.12\cdot 10^{-8}$	$5.53\cdot 10^{-9}$
	5	$1.22 \cdot 10^{-10}$	$5.60 \cdot 10^{-10}$	$1.50 \cdot 10^{-10}$
0.9	3	$1.62 \cdot 10^{-4}$	$7.16 \cdot 10^{-4}$	$3.94 \cdot 10^{-4}$
	4	$6.20 \cdot 10^{-6}$	$1.42 \cdot 10^{-4}$	$1.26 \cdot 10^{-5}$
	5	$1.68\cdot 10^{-6}$	$4.98\cdot 10^{-5}$	$6.13\cdot 10^{-6}$
1.0	3	$8.95\cdot10^{-5}$	$1.34 \cdot 10^{-1}$	$3.37 \cdot 10^{-2}$
	4	$2.34 \cdot 10^{-5}$	$5.55 \cdot 10^{-2}$	$1.29 \cdot 10^{-2}$
	5	$8.40\cdot10^{-6}$	$2.34\cdot10^{-2}$	$4.89\cdot10^{-3}$

**Table 4.2:** Comparison of the quality of the results for the ground state energyand for the one-particle dispersion

Figure 4.8 shows the energy gap  $\Delta$  as a function of the parameter  $\lambda$ . Up to  $\lambda = 0.9$ , the agreement is quite good. The inset shows a magnification of the region around the critical point. Each curve shows a kink, but each of them is at a value  $\lambda < 1$ . The position is consistent with the maximum in the deviation of the ground state energy, see Fig. 3.5.



Figure 4.5: Energy dispersion for parameter value  $\lambda = 0.5$ . In this calculation as well as in Fig. 4.6 and 4.7 the F = 20 lowest eigenvectors were used except for D = 3, where the maximum number of 17 was used.

The shading for  $\lambda > 1$  emphasizes the fact that the method in its current implementation cannot be expected to give accurate results for excitations in the ordered phase. The approach as described in Chap. 3 inherently assumes a single ground state realization at both ends of the chain. Since elementary excitations are domain walls in the Ising regime, this implies that there has to be an even number of excitations in any state that can be described. Indeed, there is roughly a factor of 2 between the curve for D = 5 and the exact solution.

Figure 4.9 shows the deviation of the numerically determined energy gap from the exact one. Again, consistent with the deviations in the ground state energy, there are spurious oscillations at low parameter values.

All calculations were done with F = 8. As mentioned above, for larger D the required information is not always contained in the lowest eigenvectors. This is reflected in the black curve for D = 5 showing larger deviations than the blue one for D = 5 and even some above the green line for D = 3.



Figure 4.6: Energy dispersion for parameter value  $\lambda = 0.9$ 



**Figure 4.7:** Energy dispersion for parameter value  $\lambda = 1.0$ 



Figure 4.8: Energy gap  $\Delta$  as a function of the parameter  $\lambda$ . The shaded area indicates the Ising regime, where due to the ground state degeneracy the method cannot give accurate results.



Figure 4.9: Deviation of the results for the energy gap from exact solution

## 4.4 Existing methods

This section describes how some of the methods mentioned in Sect. 1.2 that are related to our approach can be used to obtain a dispersion relation.

### 4.4.1 Momentum Space DMRG

The conventional real space DMRG method is very good at finding ground states and their energies. In gapped systems the quasiparticle gap is given by the energy difference between the ground state and the lowest lying excited state. Since that is the ground state in the space of excited states, its energy can also be obtained with high precision. Another possibility is to target several states simultaniously. However, at a fixed number of kept states D, this always comes with a loss of accuracy.

While real space DMRG is not suited very well to derive momentum dependent quantities, it can fairly easily be adapted to work on the reciprocal lattice i.e. in momentum space directly, which was originally done in search for an extension of DMRG to 2 spatial dimensions [10, 9]. In momentum space, every bloch state with momentum q is a lattice site. However, considerable care has to be taken in the order in which these momentum lattice sites are integrated into the system. This ordering scheme may depend on the physical model investigated.

In the momentum space DMRG method, overall momentum is a good (conserverd) quantum number. Therefore, a quasiparticle dispersion relation can be evaluated for a given q as

$$\omega_q = \min_q E_q - E_0 \tag{4.28}$$

where  $E_0$  is the ground state energy and  $E_q$  the lowest energy of an excited state with fixed overall momentum q.

While this approach utilizes the full power of the DMRG method, there are two major drawbacks in performance. Since this method works only in finite size systems, results have to be extrapolated to reach the thermodynamic limit. And, every single data point in the dispersion curve is obtained by at least one complete DMRG run. To reach full accuracy, even two runs are required.

A more elegant way using a real-time evolution variant of DMRG was proposed by White in Ref. [8]. In this approach the dispersion can be obtained as the maximum of the Fourier transform of the time and position dependent correlation function  $\langle S_i^-(t)S_0^+(0)\rangle$ . This requires only a single DMRG run. However, this method introduces an additional Suzuki-Trotter error in the time evolution. Also, tDRMG requires larger D and more sweeps. And it still requires finite size scaling analysis to derive results for the thermodynamic limit.

#### 4.4.2 Momentum space MPS

In Ref. [24] Privu, Haegeman and Verstraete recently proposed an MPS based algorithm for translationally invariant spin systems with PBC that produces very accurate results for dispersion relations.

The approach is quite similar to our method. It also exploits the fact that in translationally invariant systems all local matrices of an MPS representation of the ground state are the same. However, they work strictly on finite chains of length L and approximate the transfer matrix T by a number of m > 1 eigenvectors only if it occours in high powers. Also, they work directly with the momentum eigenstates  $|\psi_q\rangle$ . To this end, the Fourier transformation is done on the level of the matrices M and N which then yields a generalized EVP whose eigenvectors represent states with definite (quasi-) momentum q. In this picture, the energy of  $|\psi_q^1\rangle$  directly corresponds to the lowest branch of the dispersion. In a finite chain, this is however not always a single particle state. Therefore, in this form the method cannot easily be extended to the derivation of effective models in the way shown in chapter 5.

The precision to which the lowest but also higher branches of the dispersion can be obtained even at criticality is impressive. Relative errors for the lowest dispersion branch in a critical ITF chain are of order  $10^{-6}$  at D = 8 compared to our  $10^{-2}$  at D = 5) [24]. The drawback is however, that, to calculate the dispersion relations for variable q the Fourier coefficients have to be kept in momory. And here these are a set of  $L^2$  matrices of dimension  $dD^2$ . This currently poses a severe limitation on the chain length and / or the local matrix size D.

### 4.4.3 Projected entangled multipartite states (PEMS)

The PEPS method mention in chapter 1 still suffers from the implicit locality of correlations inherent in DMRG-like approaches, since entanglement is only fully encoded for nearest neighbours.

The idea of projected entangled multipartite states (PEMS) [37] is to overcome this limitation by adding an additional *L*-dimensional auxiliary system on each physical site, that encodes multipartite entanglement. Here *L* is the number of sites in a one-dimensional system. However, like PEPS, the approach readily extends to higher spatial dimension  $D_s$ , in which case the additional auxiliary systems are of dimension  $L^{D_s}$ . The multipartite state can also convey other properties into the PEPS, e.g. give the resulting physical state a definite linear momentum (see Ref. [37] for details).

An iterative DMRG-like algorithm can then be used to obtain the spectrum of lowest lying eigenstates with definite momentum q. This yields the one particle dispersion as the energy gap for given q as

$$\omega_q = E_q^{[1]} - E_0 \tag{4.29}$$

where  $E_0$  is the ground state energy and  $E_q^{[1]}$  the energy of the lowst excited state with momentum q.

This method may be better suited to describe certain critical systems than conventional DMRG and MPS based methods. However, it is limited to finite systems, as it requires the auxiliary systems of size L.

Therefore, this method, too, relies on extrapolation to obtain results in the thermodynamic limit. Also, a complete run is required for every given value of q

## Chapter 5

# Local creation operator and spectral weight

In this chapter it will be shown how a real space representation of the local creation operator  $a^{\dagger}$  for the hardcore boson excitations in the system can be obtained from the results of the previous chapter.

As one application of the effective model implied by these results, the one-particle contribution to the spectral weight  $S_{1p}^{xx}(q)$  is calculated and compared to the closed expression proposed by Hamer et al. in Ref [33].

## 5.1 Local creation operator

Consider the eigenvector  $v_q$  of the momentum space EVP defined in Eq. (4.9). For each q its components  $v_q^{(\alpha)}$  describe how the states  $|\psi_q^{\alpha}\rangle$  are mixed to form a state  $|\phi_q\rangle$ , so that

$$|\phi_q\rangle = \sum_{\alpha} v_q^{(\alpha)} |\psi_q^{\alpha}\rangle = a_q^{\dagger} |\psi_0\rangle . \qquad (5.1)$$

follows. Therefore, the Fourier transform of  $|\phi_q\rangle$  yields an expression for the action of the local creation operator on the ground state in real space

$$a_{i}^{\dagger} |\psi_{0}\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iqr_{i}} |\phi_{q}\rangle \, \mathrm{d}q = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iqr_{i}} \sum_{\alpha j} v_{q}^{(\alpha)} e^{iq(r_{i}+r_{j})} |\psi_{i+j}^{\alpha}\rangle \, \mathrm{d}q \qquad (5.2a)$$

$$= \sum_{\alpha j} |\psi_{i+j}^{\alpha}\rangle \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{iqr_j} v_q^{(\alpha)} \, \mathrm{d}q = \sum_{\alpha j} v_j^{(\alpha)} |\psi_{i+j}^{\alpha}\rangle \,. \tag{5.2b}$$

The unitary transformation in Eq. (4.25) implies that the components of  $v_q$  are either purely real or purely imaginary. Figure 5.1 shows an example of such an eigenvector  $v_q$ , where for each component  $v_q^{(\alpha)}$  the non-vanishing (real or imaginary) part is plotted.

Since for each q the vector  $v_q$  is the result of an independent diagonalization, the smoothness of the components has to be ensured by setting

$$v_q = \frac{v_{q-\Delta q}^{\dagger} v_q}{|v_{q-\Delta q}^{\dagger} v_q|} v_q , \qquad (5.3)$$

where  $\Delta q$  is the sampling interval in q space.

We found, however, that the components of  $v_q$  show discontinuities at the locations of intersections of the lowest  $n_q^{(\alpha)}$  kept with lower ones that were discarded. These discontinuities cannot be repaired by the phase correction in Eq. (5.3).



Figure 5.1: Example of the components of eigenvector  $v_q$  of the EVP (4.9) belonging to the energy dispersion  $\omega_q$ . The components are either purely real or imaginary. The plot shows the non-zero part of each component. Note that the imaginary components are symmetric, whereas the real ones are antisymmetric.

The Fourier coefficients  $v_i^{(\alpha)}$  in real space are given by

$$v_j^{(\alpha)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} v_q^{(\alpha)} e^{-iqj} \, \mathrm{d}q \;.$$
 (5.4)

Figure 5.2 shows them on a linear scale. They exhibit a quick decrease with distance j and an alternating sign. A better impression of the characteristics of their decrease is given by the logscale plot of the absolute values  $|v_j^{(\alpha)}|$  in Fig. 5.3. This shows exponential decay for all components. A slight increase at the right boundary can be attributed to a residual periodicity induced by the cutoff of the Fourier series in Eq. (4.10) at some finite  $j_{\text{max}}$ .

This yields the central result of this thesis: The local creation operator  $a_i^{\dagger}$  of elementary excitations can be constructed in real space

$$a_i^{\dagger} |\psi_0\rangle := \sum_{j,\alpha} v_j^{(\alpha)} |\psi_{i+j}^{\alpha}\rangle .$$
(5.5)

Due to the exponential decrease of the coefficients  $v_j^{(\alpha)}$  this expression can be well approximated with a finite number of coefficients. We expect a power-law decrease of the coefficients at criticality, which was not found in the present results at small D. How well it is described by larger matrices will be subject to future studies.

Note, that in the one-particle space from Eq. (5.5) also the action of the annihilation operator  $a_i$  in expressions of the type  $\langle \psi_0 | a_i a_j^{\dagger} | \psi_0 \rangle$  follows, as it acts on  $\langle \psi_0 |$  as a creation operator.

As a measure for the non-locality of the excitations, the quantity  $\zeta$  is defined by

$$V_j \propto \exp\left(-\frac{|j|}{\zeta}\right)$$
 (5.6a)

with 
$$V_j := \sum_{\alpha} |v_j^{(\alpha)}|$$
. (5.6b)

It can be determined by linear regression over the pairs of values  $(j, \log V_j)$ .



Figure 5.2: Fourier transform of the vector components in Fig. 5.1

## 5.2 Spectral weight

Using the definition (5.5), expressions such as

$$m_j := \langle \psi_0 | a_j S_0^{x} | \psi_0 \rangle = (\langle \psi_0 | S_0^{x\dagger} a_j^{\dagger} | \psi_0 \rangle)^*$$
(5.7)

can be evaluated. They are required for the calculation of the spectral weight  $S_{1p}^{xx}$ , cf. Eq. (2.24). Noting the Fourier transform of the creation operator  $a_j^{\dagger}$  is given by

$$a_q^{\dagger} := \frac{1}{L} \sum_j e^{iqr_j} a_j^{\dagger} , \qquad (5.8)$$



Figure 5.3: Absolute value of Fourier transform in Fig. 5.2. In the logarithmic plot it is evident that the decrease is exponential.

and using translational invariance Eq. (2.24) yields the one-particle contribution to the spectral weight

$$S_{1p}^{xx}(q) = \left(\sum_{i} \langle \psi_{0} | S_{i}^{x\dagger} \frac{1}{\sqrt{L}} | \psi_{q} \rangle e^{-iqr_{i}}\right) \left(\sum_{j} \frac{1}{\sqrt{L}} \langle \psi_{q} | S_{j}^{x} | \psi_{0} \rangle e^{iqr_{j}}\right)$$

$$\left(\sum_{i} \langle \psi_{i} | S_{i}^{x\dagger} e^{-iqr_{i}} \frac{1}{\sqrt{L}} \sum_{i} e^{iqr_{k}} e^{\frac{1}{2}} | \psi_{0} \rangle\right) \left(\sum_{j} e^{-iqr_{j}} \sum_{i} e^{-iqr_{k}} e^{\frac{1}{2}} | \psi_{0} \rangle\right)$$
(5.9a)

$$\left(\sum_{i} \langle \psi_{0} | S_{i}^{\mathsf{x}^{\mathsf{T}}} e^{-iqr_{i}} \frac{1}{L} \sum_{k} e^{iqr_{k}} a_{k-i}^{\mathsf{T}} | \psi_{0} \rangle \right) \left(\frac{1}{L} \sum_{j} e^{iqr_{j}} \sum_{\ell} e^{-iqr_{\ell}} \langle \psi_{0} | a_{j-\ell} S_{j}^{\mathsf{x}} | \psi_{0} \rangle \right)$$

$$(5.9b)$$

$$= \left(\frac{1}{L}\sum_{ik} e^{iq(r_k - r_i)} \langle \psi_0 | S_i^{\mathsf{x}^{\dagger}} a_{k-i}^{\dagger} | \psi_0 \rangle \right) \left(\frac{1}{L}\sum_{j\ell} e^{iq(r_j - r_\ell)} \langle \psi_0 | a_{j-\ell} S_j^{\mathsf{x}} | \psi_0 \rangle \right)$$
(5.9c)

$$= \left(\sum_{i} \langle \psi_0 | S_0^{\mathbf{x}} a_i^{\dagger} | \psi_0 \rangle e^{iqr_i} \right) \left(\sum_{j} \langle \psi_0 | a_j S_0^{\mathbf{x}} | \psi_0 \rangle e^{-iqr_j} \right)$$
(5.9d)

$$= \left(\sum_{i} m_{i}^{*} e^{iqr_{i}}\right) \left(\sum_{j} m_{j} e^{-iqr_{j}}\right)$$
(5.9e)

$$= |m_q|^2 . (5.9f)$$

Thus, the spectral weight  $S_{1p}^{xx}$  is the square of the absolute value of the Fourier transform of the above defined matrix element  $m_j$ .

From Eq. (5.5) and  $S^{x\dagger} = S^x$  follows

$$m_{j}^{*} = \sum_{i,\alpha} v_{i}^{(\alpha)} \langle \psi_{0} | S_{0}^{x} | \psi_{j+i}^{\alpha} \rangle .$$
(5.10)

In terms of the matrix product formalism of the previous chapters, these matrix elements are given by

$$m_{j}^{*} = \sum_{i\alpha} v_{i}^{(\alpha)} S_{j+i}^{\mathbf{x}(\alpha)}$$

$$\left( (u, \mathbf{1}_{i+i}^{(A,B_{\alpha})} T^{|j+i|-1} S_{\alpha}^{\mathbf{x}(A,A)} v) \quad \text{if} \quad j+i < 0$$

$$(5.11a)$$

with 
$$S_{j+i}^{\mathbf{x}(\alpha)} = \begin{cases} (a, \mathbf{1}_{j+i} - 1^{|x|} + S_0 - b) & \text{if } j+i < 0 \\ (u, S_0^{\mathbf{x}, (A, B_\alpha)} v) & \text{if } j+i = 0 \\ (u, S_0^{\mathbf{x}, (A, A)} T^{j+i-1} \mathbb{1}_{j+i}^{(A, B_\alpha)} v) & \text{if } j+i > 0 \end{cases}$$
 (5.11b)

Figures 5.4 through 5.6 depict the spectral weight  $S_{1p}^{xx}$  for three different values of  $\lambda$ . The lower part of each plot again shows the deviation from the exact result from Ref [33]. The plot interval  $[0,\pi/3]$  is chosen to emphasize the deviation close to q = 0. Neither the spectral weight itself, nor the deviation reveal anything new beyond this point. All curves follow the tendency observed in the right part of the plots. The calculations were done with F = 25 except for D = 3, where F = 17. The actual choice of the  $\vec{B}_{\alpha}$  and the value of f were manually adjusted to obtain the best possible results from the data.

Equation (2.25) shows that  $S_{1p}^{xx}$  diverges as  $\{\lambda \to 1, q \to 0\}$ . As expected, this singular behaviour is not captured very well at low matrix dimensions D. But again, increased D improves the results.

The spectral weight is also related to the spin-spin correlation function whose real space representation is

$$S_j := \left\langle S_0^{\mathbf{x}} S_j^{\mathbf{x}} \right\rangle = \left\langle \psi_0 | S_0^{\mathbf{x}} S_j^{\mathbf{x}} | \psi_0 \right\rangle \,. \tag{5.12}$$

These correlation functions are known to exhibit exponential decay

$$S_j \propto \exp\left(-\frac{|r_j|}{\xi}\right)$$
, (5.13)

where the correlation length  $\xi$  is given by

$$\xi := \frac{v}{\Delta} = \frac{\max_q \omega_q}{\min_q \omega_q} \ . \tag{5.14}$$

In the subspace of one-particle states, the completeness relation

$$\mathbb{1} = \sum_{i} a_{i}^{\dagger} |\psi_{0}\rangle \langle\psi_{0}|a_{i} .$$
(5.15)

holds. Thereby, the correlation function can be written as

$$S_{j} = \sum_{i} \langle \psi_{0} | S_{0}^{x} a_{i}^{\dagger} | \psi_{0} \rangle \langle \psi_{0} | a_{i} S_{j}^{x} | \psi_{0} \rangle = \sum_{i} m_{i}^{*} m_{i-j}$$
(5.16a)

$$=\sum_{k\ell,\alpha\beta} v_k^{(\alpha)*} v_\ell^{(\beta)} \sum_i \langle \psi_0 | S_0^{\mathbf{x}} | \psi_{i+k}^{\alpha} \rangle \langle \psi_{i+\ell}^{\beta} | S_j^{\mathbf{x}} | \psi_0 \rangle .$$
 (5.16b)



**Figure 5.4:** Spectral weight  $S_{1p}^{xx}$  for the parameter value  $\lambda = 0.5$ . The lower part shows the deviation from Hamer's exact result [33]. In the calculations F = 25 except for D = 3 where F = 17.



Figure 5.5: Spectral weight  $S_{1p}^{xx}$  for parameter value  $\lambda = 0.9$ .



Figure 5.6: Spectral weight  $S_{1p}^{xx}$  for parameter value  $\lambda = 0.999$ . The spectral weight shifts towards q = 0 as the gap closes for  $\lambda \to 1$ . As expected, the divergence is not captured very well at the low matrix dimensions available in the current implementation.

Taking Eq. (5.6a), (5.13) and (5.16b) togeher, one can expect a relation between the correlation length  $\xi$  and the localization length  $\zeta$  of the creation operator

$$\zeta \propto \xi \ . \tag{5.17}$$

Figure 5.7 shows both  $\xi$  and  $\zeta$ . Indeed, both quantities range in the same order of magnitude, if the system is not critical. One would not expect that low matrix dimensions yield results which correctly reproduce the divergence for  $\lambda \to 1$ . Note that  $\zeta$  is always smaller than  $\xi$ . This is not very surprising, remembering the construction of the MPS representation in Chap. 2. In an exact description of the system, the matrix dimension grows exponentially with the distance from the chain edges. This is caused by the accumulated information on correlations. Therefore, it is to be expected, that a highly restricted local matrix size will lead to limited capability of describing long-range correlations.

For small  $\lambda$ , however, the agreement should be better. A reason for this and also the irregularities, e.g., at  $D = 4, \lambda = 0.6$ , can probably be found in the choice of the basis states  $\vec{B}_{\alpha}$ . Since they are not orthogonal,  $V_j$  may need to be replaced in the definition of  $\zeta$  by a more suiteable quantity in future studies.



Figure 5.7: The correlation length  $\xi$  and the decline parameter  $\zeta$  of the combined coefficients  $V_j$  of the creation operator  $a^{\dagger}$ .

#### 5.2.1 Notes on implementation

Since the computation of the Fourier coefficients in Eq. (5.4) involves numerical integration, it is a likely source of errors. Note that both  $v_q$  and  $S_{1p}^{xx}$  are q dependent quantities. Therefore, the relevant  $m_q$  can be computed directly in q space

$$m_q = \sum_j e^{iqr_j} m_j = \sum_j e^{iqr_j} \frac{1}{L} \sum_i \langle \psi_0 | a_{j+i} S_i^{x} | \psi_0 \rangle$$
(5.18a)

$$= \langle \psi_0 | \frac{1}{L} \sum_{ji} e^{iqr_j} a_{j+i} S_i^{x} | \psi_0 \rangle = \langle \psi_0 | a_q S_{-q}^{x} | \psi_0 \rangle$$
(5.18b)

with 
$$a_q := \frac{1}{\sqrt{L}} \sum_j e^{iqr_j} a_j, \qquad S_q^{x} := \frac{1}{\sqrt{L}} \sum_j e^{iqr_j} S_j^{x}.$$
 (5.18c)

By taking the hermitian conjugate of Eq. (5.1), the action of the annihilation operator  $a_q$  on  $\langle \psi_0 |$  is defined by

$$\langle \psi_0 | a_q = \sum_{\alpha} \langle \psi_q^{\alpha} | v_q^{(\alpha)*} = \frac{1}{\sqrt{L}} \sum_{\alpha j} v_q^{(\alpha)*} e^{-iqr_j} \langle \psi_j^{\alpha} |$$
(5.19)

which leads to

$$m_q = \langle \psi_0 | a_q \, S_{-q}^{\mathbf{x}} | \psi_0 \rangle = \frac{1}{L} \sum_{ji\alpha} v_q^{(\alpha)*} e^{iqr_j} e^{-iqr_i} \underbrace{\langle \psi_j^{\alpha} | S_i^{\mathbf{x}} | \psi_0 \rangle}_{s_{j-i}^{(\alpha)}} \tag{5.20a}$$

$$= \frac{1}{L} \sum_{ji\alpha} v_q^{(\alpha)*} e^{-iq(r_j - r_i)} s_{j-i}^{(\alpha)} = \sum_{j\alpha} v_q^{(\alpha)*} s_j^{(\alpha)} e^{-iqr_j}$$
(5.20b)

$$=\sum_{\alpha} v_q^{(\alpha)*} s_q^{(\alpha)} .$$
(5.20c)

The relation in Eq. (5.20) only requires a Fourier series' evaluation to compute  $s_q^{(\alpha)}$ , which can easily be done to higher precision than the integral. Therefore, it is used in the calculations in Fig. 5.4 through 5.6.

## Chapter 6

# Conclusions and outlook

## 6.1 Summary of method

In this thesis, a variational method for deriving effective one-dimensional models was introduced. The method is based on the matrix product state formalism, which is strongly connected to the well established DMRG method.

It was demonstrated that, assuming translational invariance, the MPS ansatz allows a very efficient way of working directly in the thermodynamic limit.

It was shown that the DMRG-like approach of optimizing the local matrix sets on one site and keeping the rest of the chain fixed leads to a generalized eigenvalue problem (EVP). In an iterative procedure, this EVP can be solved, and the solution with the lowest local energy is adopted as new ground state approximation on all other sites in the next step. When converged, this yields an MPS approximation of the ground state and its energy.

The eigenvectors of the ground state search problem were shown to describe excitations in the system. By transforming the EVP and the solution into momentum space, an estimate for the one-particle dispersion was found as the lowest eigenvalue of the EVP in momentum space.

Finally, a real space representation of the local creation operator was constructed. This was done by Fourier transforming the eigenvector of this EVP in momentum space that belongs to the dispersion.

These three results constitue the effective one-particle Hamiltonian. As an example of its application, the one-particle contribution to the spectral weight was calculated.

A key problem was found to be the choice of the correct eigenvectors to use in building the momentum space EVP. At the moment, the selection of the correct eigenvectors that contribute to the dispersion partially relies on human review of the data at an intermediate stage. Here, a better understanding of the behaviour of the momentum space EVP has to be achieved in order to automate this process.

The method is still at an early stage. The current GNU octave implementation leaves plenty of opportunities for improvement, both in efficient use of computational resources and in the algorithms used.

## 6.2 Summary of results

The ground state energy could be obtained to good precision  $(\Delta E < 10^{-3}\Gamma)$  even at low matrix dimensions. In the Ising regime, in which the ground state is twofold degenerate,

the method was found to produce one ground state with the same accuracy as the unique ground state in the disordered phase.

The one-particle energy dispersion could on average be calculated with the same precision as the ground state energy over a large part of the disordered phase. Close to the quantum critical point, the deviations in the precision of the results were found to be much larger. This is mainly due to two reasons. First, that close to criticality, the information is widely spread over the eigenvectors of the real space EVP. And Secondly, the results for the ground state energy and the excitation gap consistently indicate that the phase transition is detected at a too low parameter value for small matrix sizes.

As the key goal of the new method, a representation of the one-particle creation operator was found both in momentum and in real space. In real space, this representation is given by the coefficients of a superposition of states in which one local matrix set is changed from the ground state to one of the eigenvectors of the EVP. These coefficients were found to show exponential decrease as functions of the distance from the lattice site where the quasiparticle is created.

As an application of the derived effective model, the one-particle contribution to the spectral weight was calculated, a quantity that can be directly linked to the observables in neutron scattering experiments. The results were found to agree well with the exact result found by Hamer et al. except for parameter values very close to the quantum critical point. This shows, that the representation of the local creation operator is indeed correct but for critical systems more coefficients are required than were used in the computations for this thesis.

## 6.3 Outlook

A still more detailed analysis will be required to stabilize the behaviour of the momentum space EVP. Preferably, a property of the  $\vec{B}_{\alpha}$  themselves should be found, that identifies the matrices which relevant for the dispersion and the creation operator. If that proves to be impossible, an algorithm needs to be designed, that allows us to effciently transform the whole EVP into momentum space. To handle this full problem, the eigenvalues of the norm matrix need to be analyzed over the full Brillouin zone and the number of discarded eigenvalues must be chosen adaptively.

In further testing other problems or limitations not evident from the present ITF results can be found and resolved. Also, a better understanding of the confidence level of the results can be gained.

Also, the current approach assumes that the boundary conditions are more or less irrelevant for infinite systems. However, in the case of a degenerate ground state, this is no longer true. Thefore, the effective model could only be constructed in the disordered phase of the transverse field Ising model.

Thus, a way to implement a concrete set of boundary conditions needs to be found. Control over the boundary conditions will enable us to treat ground state degeneracy and domain wall excitations. Thus the present approach can then be applied to more complex models.

An extension to include the dynamics of two particles, i.e., two-particle interaction is possible and should be feasible in future research. Finally, the concept of variationally deriving effective models could be extended to two spatial dimensions.
## Bibliography

- S. R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69(19):2863-2866, Nov 1992.
- [2] C. Knetter and G.S. Uhrig. Perturbation theory by flow equations: dimerized and frustrated s = 1/2 chain. *EPJ B*, 13:209-225, 2000.
- [3] H. Krull, N. A. Drescher, and G. S. Uhrig. Enhanced perturbative continuous unitary transformations. *Phys. Rev. B*, 86:125113, Sep 2012.
- [4] S. Dusuel, M. Kamfor, K. P. Schmidt, R. Thomale, and J. Vidal. Bound states in two-dimensional spin systems near the ising limit: A quantum finite-lattice study. *Phys. Rev. B*, 81:064412, Feb 2010.
- [5] K. G. Wilson. The renormalization group: Critical phenomena and the Kondo problem. *Rev. Mod. Phys.*, 47(4):773-840, Oct 1975.
- [6] E. Jeckelmann. Dynamical density-matrix renormalization-group method. Phys. Rev. B, 66:045114, Jul 2002.
- [7] M. A. Cazalilla and J. B. Marston. Time-dependent density-matrix renormalization group: A systematic method for the study of quantum many-body out-of-equilibrium systems. *Phys. Rev. Lett.*, 88(25), Jun 2002.
- [8] S. R. White and A. E. Feiguin. Real-time evolution using the density matrix renormalization group. *Phys. Rev. Lett.*, 93:076401, Aug 2004.
- [9] S. Nishimoto, E. Jeckelmann, F. Gebhard, and R. M. Noack. Application of the density matrix renormalization group in momentum space. *Phys. Rev. B*, 65:165114, Apr 2002.
- [10] T. Xiang. Density-matrix renormalization-group method in momentum space. Phys. Rev. B, 53:R10445-R10448, Apr 1996.
- [11] R. J. Baxter. Dimers on a rectangular lattice. J. Math. Phys., 9(4):650-654, 1968.
- [12] M. Fannes, B. Nachtergaele, and R. F. Werner. Exact antiferromagnetic ground states of quantum spin chains. *Europhys. Lett.*, 10(7):633, 1989.
- [13] M. Fannes, B. Nachtergaele, and R. F. Werner. Finitely correlated states on quantum spin chains. Commun. Math. Phys., 144(3):443-490, 1992.
- [14] S. Östlund and S. Rommer. Thermodynamic limit of density matrix renormalization. *Phys. Rev. Lett.*, 75(19):3537-3540, Nov 1995.

- [15] S. Rommer and S. Östlund. Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Phys. Rev.* B, 55:2164–2181, Jan 1997.
- [16] I. P. McCulloch. Infinite size density matrix renormalization group, revisited. eprint arXiv:0804.2509 [cond-mat.str-el], Apr 2008.
- [17] G. M. Crosswhite, A. C. Doherty, and G. Vidal. Applying matrix product operators to model systems with long-range interactions. *Phys. Rev. B*, 78:035116, Jul 2008.
- [18] G. Vidal. Classical simulation of infinite-size quantum lattice systems in one spatial dimension. *Phys. Rev. Lett.*, 98(7):070201, Feb 2007.
- [19] M. C. Bañuls, M. B. Hastings, F. Verstraete, and J. I. Cirac. Matrix product states for dynamical simulation of infinite chains. *Phys. Rev. Lett.*, 102:240603, Jun 2009.
- [20] Hiroshi Ueda, Isao Maruyama, and Kouichi Okunishi. Uniform matrix product state in the thermodynamic limit. J Phys. Soc. Japan, 80(2):023001, 2011.
- [21] F. Verstraete and J. I. Cirac. Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions. eprint arXiv:cond-mat/0407066 [cond-mat.strel], July 2004.
- [22] V. Murg, F. Verstraete, and J. I. Cirac. Variational study of hard-core bosons in a two-dimensional optical lattice using projected entangled pair states. *Phys. Rev. A*, 75:033605, Mar 2007.
- [23] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. I. Cirac. Classical simulation of infinite-size quantum lattice systems in two spatial dimensions. *Phys. Rev. Lett.*, 101:250602, Dec 2008.
- [24] B. Pirvu, J. Haegeman, and F. Verstraete. Matrix product state based algorithm for determining dispersion relations of quantum spin chains with periodic boundary conditions. *Phys. Rev. B*, 85:035130, Jan 2012.
- [25] H. Saberi, A. Weichselbaum, and J. von Delft. Matrix-product-state comparison of the numerical renormalization group and the variational formulation of the density-matrix renormalization group. *Phys. Rev. B*, 78:035124, Jul 2008.
- [26] U. Schollwöck. The density-matrix renormalization group in the age of matrix product states. Ann. Phys., 326(1):96 - 192, 2011.
- [27] P.G. de Gennes. Collective motions of hydrogen bonds. Solid State Commun., 1(6):132 137, 1963.
- [28] P. Pfeuty. The one-dimensional Ising model with a transverse field. Ann. Phys., 57(1):79-90, 1970.
- [29] P. Jordan and E. Wigner. Über das Paulische Äquivalenzverbot. Z. Phys., 47(9-10):631-651, 1928.
- [30] E. Lieb, T. Schultz, and D. Mattis. Two soluble models of an antiferromagnetic chain. Ann. Phys., 16(3):407 - 466, 1961.
- [31] A. J. A. James, W. D. Goetze, and F. H. L. Essler. Finite-temperature dynamical structure factor of the heisenberg-ising chain. *Phys. Rev. B*, 79:214408, Jun 2009.

- [32] W. Mashall and S. W. Lovesey. Theory of Thermal Neutron Scattering: the Use of Neutrons for the Investication of Condensed Matter. Clarendon Press, Oxford, 1971.
- [33] C. J. Hamer, J. Oitmaa, Zheng Weihong, and Ross H. McKenzie. Critical behavior of one-particle spectral weights in the transverse ising model. *Phys. Rev. B*, 74:060402, Aug 2006.
- [34] S. R. White. Density-matrix algorithms for quantum renormalization groups. Phys. Rev. B, 48(14):10345-10356, Oct 1993.
- [35] I. Peschel et al. Density-Matrix Renormalization. Springer, 1 edition, 1998.
- [36] L. Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. Phys. Rev., 65:117–149, Feb 1944.
- [37] D. Porras, F. Verstraete, and J. I. Cirac. Renormalization algorithm for the calculation of spectra of interacting quantum systems. *Phys. Rev. B*, 73(1):014410, Jan 2006.

## Danksagung

An dieser Stelle möchte ich mich bei Prof. Dr. Götz S. Uhrig für die Gelegenheit bedanken an der Entwicklung dieser Methode mitzuarbeiten, sowie für die stets engagierte Betreuung während des gesamten Projekts. Herrn Prof. Dr. Frithjof Anders danke ich für die Übernahme des Zweitgutachtens der Arbeit.

Weiterhin möchte ich der ganzen Arbeitsgruppe danken, die mir stets mit Rat und Tat zur Seite stand und ein ausgezeichnetes Arbeitsklima bot.

Schließlich danke ich Benedikt Fauseweh für eine vielzahl hilfreicher Gespräche und Laura Möller für das Korrekturlesen.

## Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Masterarbeit mit dem Titel "Effective One-Dimensional Models from Matrix Product States" selbständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Ort, Datum

Unterschrift

## Belehrung

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 Euro geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§ 63 Abs. 5 Hochschulgesetz - HG - ).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z.B. die Software "turnitin") zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen.

 ${\rm Ort},\,{\rm Datum}$ 

Unterschrift