

AARHUS UNIVERSITY

MASTERS THESIS

---

# Engineering Optimal Quantum Phase Transitions through Tensor Networks

---

*Author:*  
Frederik S. Møller  
201303717

*Supervisor:*  
Jakob F. Sherson



Quantum Measurement and Manipulation Group  
Department of Physics and Astronomy

June 3, 2018



# Engineering Optimal Quantum Phase Transitions through Tensor Networks

by Frederik S. Møller  
201303717

## *Abstract*

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...



## *Acknowledgements*

The acknowledgments and the people to thank go here, don't forget to include your project advisor...



# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Ultracold Quantum Gases in Optical Lattices</b>	<b>3</b>
2.1 Optical Lattice Potentials . . . . .	3
2.1.1 Trapping of Neutral Atoms . . . . .	3
2.1.2 Optical Lattices . . . . .	5
2.1.3 Band Structure . . . . .	5
2.1.4 Localized States . . . . .	6
2.2 Bose-Einstein Condensates . . . . .	7
2.2.1 Non-Interacting Particles . . . . .	8
2.2.2 Weakly Interacting Particles . . . . .	9
2.3 Bose-Hubbard Model of Interacting Bosons in a Lattice . . . . .	10
2.3.1 The Bose-Hubbard Hamiltonian . . . . .	11
2.3.2 Phases of the Bose-Hubbard Model . . . . .	12
2.3.3 Mean-Field Approximation of Phase Diagram . . . . .	15
2.4 Non-Equilibrium Dynamics near Critical Points . . . . .	18
2.4.1 Breakdown of Adiabaticity . . . . .	18
2.4.2 The Kibble-Zurek Mechanism . . . . .	18
2.4.3 Light-Cone Propagation of Correlations . . . . .	19
<b>3 Matrix Product States</b>	<b>23</b>
3.1 Entanglement in Quantum Systems and Area Laws . . . . .	23
3.2 Construction of an MPS . . . . .	24
3.3 Canonical Forms . . . . .	26
3.3.1 Left-canonical matrix product state . . . . .	26
3.3.2 Right-canonical matrix product state . . . . .	27
3.3.3 Mixed-canonical matrix product state . . . . .	28
3.4 Matrix Product Operators . . . . .	29
3.4.1 Applying an MPO to an MPS . . . . .	29
3.5 Correlation Functions and Measurement of Local Properties . . . . .	30
3.5.1 Correlation length . . . . .	31
<b>4 Tensor Network Algorithms</b>	<b>35</b>
4.1 Variational Ground State Search . . . . .	35
4.1.1 Efficient Application of a Hamiltonian to an MPS . . . . .	35
4.1.2 Iterative Ground State Search and the DMRG Algorithm . . . . .	38
4.2 Time Evolution of Matrix Product States . . . . .	42
4.2.1 The tDMRG Algorithm . . . . .	42
4.2.2 Modified Time Evolution Algorithm for Bose-Hubbard Model . . . . .	44

<b>5 Quantum Optimal Control Theory</b>	<b>49</b>
5.1 The Gradient-Ascent Pulse Engineering Method - GRAPE . . . . .	50
5.1.1 Gradient of Suzuki-Trotter Propagator . . . . .	54
5.2 Interior Point Methods . . . . .	55
5.2.1 Duality . . . . .	56
5.2.2 Karush–Kuhn–Tucker Conditions . . . . .	57
5.2.3 Basic Primal-Dual Interior Point Algorithm . . . . .	58
5.2.4 Implementation with GRAPE . . . . .	59
5.3 Control Parametrization via Chopped Basis . . . . .	60
5.3.1 Gradient-Optimization Using Parametrization - GROUP . . . . .	61
5.4 Quantum Speed Limit . . . . .	61
<b>6 Results</b>	<b>65</b>
6.1 Determining Parameters . . . . .	66
6.1.1 Boundary Conditions and Constraints . . . . .	66
6.1.2 Seed Selection . . . . .	66
6.2 Characterization of Methods on a Small System . . . . .	67
6.2.1 Determining the Quantum Speed Limit . . . . .	67
6.2.2 Optimized Ramp Sequences . . . . .	69
6.2.3 Control Basis Size . . . . .	70
6.3 Application on Larger Systems . . . . .	70
6.3.1 Truncation of Tensors . . . . .	72
<b>Bibliography</b>	<b>75</b>
<b>A Diagrammatic Representation of Matrix Product States</b>	<b>83</b>
<b>B Example: Building an MPO from a Hamiltonian</b>	<b>85</b>
<b>C Calculation of Condensate Fraction using DMRG Algorithm</b>	<b>87</b>
<b>D Example: Quantum Speed Limit in Landau-Zener Model</b>	<b>91</b>
<b>E Hessian</b>	<b>93</b>

## Chapter 1

# Introduction

Over the last two decades great improvements in the experimental cold atoms toolbox have enabled the study of countless quantum systems and phenomena [1, 2]. Ultracold quantum gases are extremely versatile, as a wide range of Hamiltonians can be mapped unto the system allowing various theoretical models to be realized experimentally. Furthermore, a very high degree of control over the systems is achievable through the manipulation of external fields [3]. A particularly interesting technique is loading the cold atoms into optical lattices created by the interference of laser beams [4]. In a lattice, the cold gases can be used for quantum simulations [5, 6], quantum logic gate operations [7, 8, 9], single-atom transistors [10], and much more. Many of these applications require the system initially being in a Mott-insulating state, where the atoms are evenly distributed and well-localized within the sites of the lattice [11]. Due to the fragile nature of quantum effects, extremely low temperature are necessary in order to experimentally observe these phenomena. Furthermore, many applications of lattice systems require the preparation of certain configurations, such as the Mott, with a minimum of defects. Although the manipulation of cold gases is easily achieved, exercising full control over the systems is very difficult due to the complexity of the many-body dynamics. The preparation of high-quality Mott-insulators is particularly challenging, as the experimental procedure requires the crossing of a quantum phase transition, at which various mechanisms introduces disorder to the system [12, 13]. Currently, many experiments rely on sub-optimal control protocols, which often causes unnecessary heating of the system while being needlessly slow thus making the system vulnerable to decoherence. Therefore, the preparation of states and the dynamical control of systems are some of the central challenges in current cold atoms physics.

Quantum optimal control theory is a framework for designing control schemes for quantum systems [14, 15]. The desired dynamics are achieved by formulating the control problem as the minimization of some cost function, whereby well-established methods from mathematical optimization theory can be employed. Additionally, through the optimization landscape much information can be inferred about the controllability of the system [16]. Optimal control problem are often subject to a series of constraints due to physical and technical limitations. Furthermore, the dynamics of complex systems are often highly non-linear. Therefore, the optimal control functions responsible for generating the desired dynamics can be found analytically for only very few quantum systems, whereby one often has to rely on numerical simulations of the system. Hence, algorithms for optimizing the control functions are central in quantum optimal control theory. These algorithms can roughly be divided into those utilizing derivatives of the cost function [17, 18] versus those relying on classical Nelder-Mead methods [19].

The effectiveness of quantum optimal control is often limited by the difficulty of simulating quantum systems [20]. Quantum many-body systems, such as cold gases in optical lattices, often have Hilbert spaces which scale exponentially with the system size. Therefore, simulating even small systems is extremely resource demanding. In one-dimensional lattice systems, quantum states and operators can be decomposed as tensor networks. These

techniques are incredibly powerful, as the networks by construction puts an upper bound on the entanglement between different domains of the system, whereby most states of the Hilbert space can be neglected [21, 22]. Thus, tensor networks provide a low-entanglement effective theory of the system, which remains valid at even longer durations due to the finite spreading velocity of correlations within the lattice systems [23, 24].

This thesis presents a framework for finding optimal control schemes for precise, dynamical manipulation of one-dimensional lattice systems. The framework combines the tensor network description of quantum states and operators with derivative-based optimal control algorithms. While the tensorial computation of the gradient of the cost function has previously been argued to be too resource demanding and thus impossible [19], the results presented here prove that this is not the case. In fact, when compared to tensor-network-based optimizations utilizing the Nelder-Mead method, the framework presented here performs much better. The high performance can partly be attributed to the use of interior point methods, which are gradient-based optimization techniques for non-linear, constrained problems. Interior point methods are especially applicable to control problems, as they converge using only a minimal number of the otherwise computationally expensive gradient evaluations [25].

To demonstrate the framework, a ground state transfer from the superfluid to the Mott-insulating phase of the Bose-Hubbard model is optimized. Simulating Bose-Hubbard systems is considered fairly difficult due to the non-integrability of the model. Furthermore, the state transfer examined is very relevant for experimental purposes, whereby it is an ideal benchmark for the framework. Although originally developed with the Bose-Hubbard model in mind, the framework can easily be applied to any one-dimensional quantum lattice system.

## Chapter 2

# Ultracold Quantum Gases in Optical Lattices

Ultracold atoms present an extremely powerful tool in quantum research. At the scale of nanokelvins thermal effects no longer destroy the otherwise fragile quantum systems. Thus, quantum phenomena, previously known only from theory, can be studied directly. Atomic physics experiments with such ultracold, quantum degenerate gases offer some unique features: (i) a wide range of Hamiltonians can be mapped to the systems making experiments highly customizable. This can in part be attributed to (ii) the very high degree of control achievable through the manipulation of external fields. Many different configurations of magnetic and optical traps achievable in which the cold atoms can be manipulated very precisely. Furthermore, the collisional properties of the atoms can be tuned through magnetic Feshbach resonances, and the properties of the gas can be probed through interactions between internal energy levels of the atoms and laser light [3, 2].

At temperature scales around  $\approx 100\text{nK}$  many-body phenomena such as Bose-Einstein condensation takes place, where the ground state of the system gains a macroscopic populations. Ever since the first realisation of a Bose-Einstein condensate in 1995 [26], the special properties of this macroscopic quantum state have been used in a wide range of experiments. A powerful tool in the cold-atoms toolbox is the optical lattice, which is an array of potentials created through the atoms dipole interaction with laser beams. The cold atoms are loaded into the highly controllable lattice from where one can perform quantum simulations of various systems, such as spin chains [27], Dirac cones [28], and artificial gauge fields [29]. Furthermore, the properties of cold atoms in optical lattices are very favourable for experiments in quantum information, where quantum gates can be realised through controlled collisions [7] or Rydberg atoms [30].

## 2.1 Optical Lattice Potentials

Cold atoms can be trapped in potentials generated by their own dipole interaction with light. Thus, superimposing laser beams allows one to create optical potentials in various shapes and forms. Such dipole traps realised using far-detuned light have important properties, as (i) they are capable of trapping neutral atoms, and (ii) the optical excitation from the trap is very low resulting in only small losses [4]. Optical lattices are a central component of many experiments, as they not only trap the atoms, but also determine important properties of the system.

### 2.1.1 Trapping of Neutral Atoms

Consider a two level atom in the presence of a time-varying electric field

$$\mathbf{E} = \epsilon E_0 \exp(i(\mathbf{k}\mathbf{r} - \omega_L t)) , \quad (2.1)$$

where  $E_0$  is its amplitude,  $\varepsilon$  is its polarization vector,  $\mathbf{k}$  is its wave vector and  $\omega_L$  is its frequency. The interaction between the atom and the electric field causes a perturbation of the atoms energy levels, otherwise known as the AC Stark-shift. The Hamiltonian describing this interaction is

$$\hat{H}_{int} = \hat{d}\mathbf{E}, \quad (2.2)$$

where  $\hat{d} = -e\hat{\mathbf{r}}$  is the electric dipole operator.

Consider eq. (2.2) as a perturbation to the Hamiltonian of the atom. To first order the non-degenerate, time-independent perturbation of the energy of state  $i$  reads

$$E_i^{(1)} = \langle i | \hat{H}_{int} | i \rangle. \quad (2.3)$$

However, all first-order contributions vanish due to the symmetry of the Hamiltonian (2.2). To second order the perturbation of state  $i$  is given by

$$E_i^{(2)} = \sum_{j \neq i} \frac{|\langle j | \hat{H}_{int} | i \rangle|^2}{\varepsilon_i - \varepsilon_j}. \quad (2.4)$$

The states  $|i\rangle$  and  $|j\rangle$  and their corresponding energies,  $\varepsilon_i$  and  $\varepsilon_j$ , are of the dressed state picture, where one considers the combined system of the atom and the light field [31]. The state  $|i\rangle$  represents the atom being in its ground state, while the light field contains  $n$  photons. Thus, the energy of the combined state is  $\varepsilon_i = n\hbar\omega_L$ , when setting the energy of the atomic ground state to zero. Meanwhile, state  $|j\rangle$  is when the atom has been excited by absorbing one of the photons of the field. Hence, the energy of this state is  $\varepsilon_j = \hbar\omega_0 + (n-1)\hbar\omega_L$ . Defining the detuning  $\Delta = \omega_L - \omega_0$  allows writing the perturbation in the form

$$E_{g/e}^{(2)} = \pm \frac{|\langle e | \hat{d} | g \rangle|^2}{\Delta} |E_0|^2, \quad (2.5)$$

where the upper sign is assigned to the atomic ground state  $|g\rangle$ . This can be rewritten in order to reflect properties of the atom and the field, by considering the intensity of the light,  $I = 2\epsilon_0 c |\mathbf{E}|^2$ , and the decay rate of the atom,  $\Gamma = \frac{\omega_0^3}{3\pi\epsilon_0 c^3} |\langle e | \hat{d} | g \rangle|$ . Thus, equation (2.5) can be written as [4]

$$E_{e/g}^{(2)} = \pm \frac{3\pi c^2}{2\omega_0^3} \frac{\Gamma}{\Delta} I \quad (2.6)$$

This is the AC-Stark shift, which constitutes the dipole potential. For red detuning ( $\Delta < 0$ ) the ground state will experience a negative shift leading to an attractive potential with depth depending on the intensity of the laser. Oppositely, a blue-detuned laser ( $\Delta > 0$ ) will repel the atom. An illustration of this can be seen in figure 2.1. Since the sign of the perturbation is reversed for the excited state, the effect of the detuning becomes opposite, whereby it is important that the atom remains in the ground state. Thus, one has to minimize the scattering with the optical potential. The scattering rate is given as [4]

$$\Gamma_{sc} = \frac{3\pi c^2}{2\hbar\omega_0^3} \left( \frac{\omega}{\omega_0} \right)^3 \left( \frac{\Gamma}{\Delta} \right)^2 I. \quad (2.7)$$

As the detuning becomes small, the laser becomes resonant with the atom causing a large increase in scattered photons. Therefore, one has to choose a large detuning in order for the potential to remain conservative. However, this comes at the cost of a weaker potential. To compensate this, a high laser intensity must be used for the potential to reach sufficient depth. In practice there will be a limit to the laser power available, however, for most

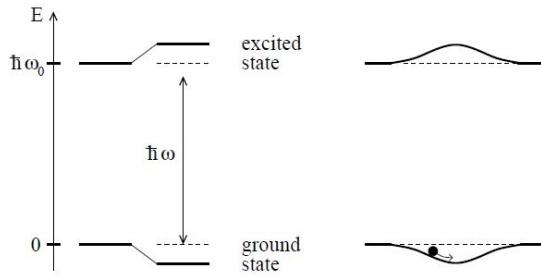


FIGURE 2.1: *Light shifts of a two-level atom. Left-hand side, red-detuned light ( $\Delta < 0$ ) shifts the ground state down and the excited state up by same amounts. Right-hand side, a spatially inhomogeneous field like a Gaussian laser beam produces a ground-state potential well, in which an atom can be trapped. Figure and caption are adopted from [4].*

alkali-metal atoms the detuning is typically chosen to be large compared to the excited-state hyperfine structure splitting, which provides enough depth while being sufficient to suppress scattering events [1].

### 2.1.2 Optical Lattices

The dipole potential in equation (2.6) scales with the intensity of the laser. Thus, superimposing laser beams enables a multitude of different potentials through the interference patterns of the lasers. A standing wave from two counter-propagating light fields will lead to an array of potential wells

$$V(z) = -V_0 \cos^2 kz, \quad (2.8)$$

where  $V_0 = |\frac{3\pi c^2}{2\omega_0^3} \frac{\Gamma}{\Delta} 4I_0|$  is derived from eq. (2.6). In practice, the one-dimensional lattice structure of eq. (2.8) is created by shining a single laser beam at a mirror, whereby it interferes with itself. Adding another standing laser beam in a different direction creates a periodic two dimensional potential. For orthogonal polarization of the two lasers, the resulting potential is purely the sum of the sinusoidal standing wave potential, as no interference term is present [11]. Note, the lattice is only well defined for distances much smaller than the waist of the laser beams, as the lattice is only present within the overlap of the two beams. Various shapes of the lattice can be achieved by adjusting the angle between the beams, however, the most common setup is using two orthogonal beam creating a square lattice of one dimensional tubes, as seen in figure 2.2. In order to create a three dimensional lattice, as seen in figure 2.2, an additional third perpendicular laser beam is needed. In the center of the trap, the lattice potential is then given by

$$V(x, y, z) = -V_0 (\cos^2 kx + \cos^2 ky + \cos^2 kz), \quad (2.9)$$

for distances much smaller than the beam waist. In addition to the lattice, an external harmonic confinement will be present due to the Gaussian profile of the laser beams [1].

### 2.1.3 Band Structure

Consider a periodic potential as described by eq. (2.8). *Bloch's Theorem* states that energy eigenstates of a periodic potential with lattice vector  $\mathbf{R}$  and quasi-momentum  $\mathbf{q}$  can be written as Bloch waves, which take the form

$$\phi_{\mathbf{q}}^{(n)}(\mathbf{r}) = e^{i\mathbf{qr}} u^{(n)}(\mathbf{r}), \quad (2.10)$$



FIGURE 2.2: (a) Two dimensional optical lattice formed by two mutually orthogonal laser beams. These tubes have a characteristic cigar shape, due to the Gaussian profile of the lasers. (b) Upon using three orthogonal laser beams, the result is a three dimensional lattice reminiscent of a cubic crystal. Figure is adopted from [32].

which is a plane wave modulated by a function with the same periodicity as the potential  $u^{(n)}(\mathbf{r}) = u^{(n)}(\mathbf{r} + \mathbf{R})$ . Furthermore, the Bloch waves are periodic in reciprocal space, such that  $\phi_{\mathbf{q}}^{(n)}(\mathbf{r}) = \phi_{\mathbf{q} + \mathbf{G}}^{(n)}(\mathbf{r})$ , where  $\mathbf{G}$  is a reciprocal lattice vector [33]. This leads to an energy spectrum in the shape of bands with the periodicity of the first *Brillouin Zone*. Bands are denoted by the band index  $n$ , and their shape is determined by both the shape and the depth of the potential. The potential depth is often denoted in units of the recoil energy  $E_{\text{rec}} = \frac{\hbar^2 k^2}{2m}$ , where  $m$  is the mass of the atom, and  $k$  is the photon wave number of the light forming the optical lattice. For  $V_0 = 0$ , the particles are free, hence the bands will be parabolic. Meanwhile, for  $V_0 \rightarrow \infty$ , no interaction between different wells of the lattice can take place, as the wavefunctions of the trapped atoms will be confined to their respective well. Thus, the lattice is reduced to an array of independent harmonic oscillators, whereby the bands will appear flat with equal spacing [34].

#### 2.1.4 Localized States

For lattice potential depths within  $5E_{\text{rec}} \leq V_0 \leq 8E_{\text{rec}}$ , the lattice is in the *tight binding limit*. Within this range the wavefunctions of the trapped atoms will only overlap with other wells in their closest proximity. Thus, interactions between wells are almost purely of nearest neighbour nature. Due to how well localized the wavefunctions are, a basis of Wannier functions is ideal for describing the system. Wannier functions are related to Bloch functions through the Fourier transform [35]

$$w^{(n)}(\mathbf{r}) = \frac{1}{\sqrt{N_L}} \sum_{\mathbf{q}} e^{-i\mathbf{q}\mathbf{R}} \phi_{\mathbf{q}}^{(n)}(\mathbf{r}), \quad (2.11)$$

where  $N_L$  is the number of primitive cells of the lattice. The Wannier functions are well localized and centered around the lattice at site  $\mathbf{R}$ . In the case of a separable periodic potential, like that of eq. (2.9), the single-particle problem becomes one-dimensional [36]. Lastly, Wannier functions obey the orthonormality relation

$$\int d^3r \ w^{(n)*}(\mathbf{r} - \mathbf{R})w^{(n')}(\mathbf{r} - \mathbf{R}') = \delta_{n,n'}\delta_{\mathbf{R},\mathbf{R}'} , \quad (2.12)$$

thus forming a complete basis [1]. Figure 2.3.a shows Wannier functions plotted for lattice

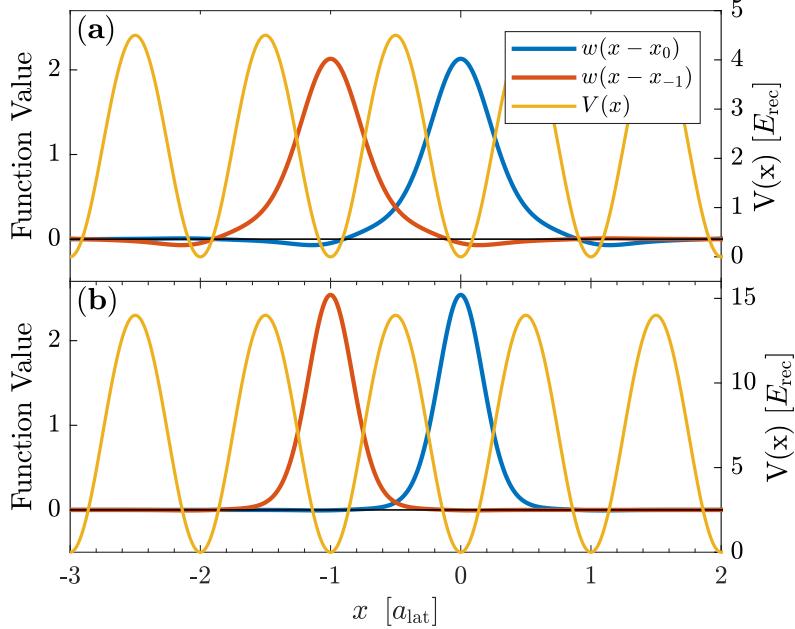


FIGURE 2.3: Two one-dimensional Wannier functions plotted for lattice depths of  $V_0 = 4.5E_{\text{rec}}$  (a) and  $V_0 = 14E_{\text{rec}}$  (b) respectively.

potential depths within the tight binding limit. As evident from the plot, the functions overlap with only their nearest neighbours. In the case of a more shallow lattice, the functions would extend to further wells, while they will tend towards a Gaussian shape as the lattice depth increases [34], which can be seen in figure 2.3.b.

## 2.2 Bose-Einstein Condensates

Bosons are particles of integer spin, whose statistics obey those of a Bose-Einstein distribution

$$n_i = \frac{g_i}{\exp((\varepsilon_i - \mu)/k_B T) - 1} , \quad (2.13)$$

where  $i$  denotes the state,  $n_i$  is the population of the state,  $g_i$  is its degeneracy,  $\varepsilon_i$  is its energy,  $\mu$  is the chemical potential,  $k_B$  is the Boltzmann constant, and  $T$  is the temperature. One important feature of bosons is that, unlike fermions, multiple particles can occupy the same quantum state. At higher temperatures the energy spectrum is practically continuous, whereby this property has little effect, as two particles occupying the same single-particle state is highly unlikely. However, at low temperatures the energy spectrum systems become increasingly discrete, hence the statistics of the particles become very important. As evident

from eq. (2.13), the population of the ground state diverges as  $T \rightarrow 0$ . However, even below a finite temperature,  $T_c$ , one will observe a macroscopic population of the ground state.  $T_c$  is called the critical temperature, and marks the point where multiple particles will start forming a Bose-Einstein Condensate (BEC). [37]

### 2.2.1 Non-Interacting Particles

In the case of non-interacting particles and zero temperature, all particles of a Bose gas can be described by identical single-particles wavefunctions  $\phi(\mathbf{r}_i)$ . Hence, the many-body wavefunction is simply given by the product

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \prod_i^N \phi(\mathbf{r}_i) . \quad (2.14)$$

Such a product state can be described by a single macroscopic wavefunction

$$\psi(\mathbf{r}) = \sqrt{N} \varphi(\mathbf{r}) , \quad (2.15)$$

where  $\phi(\mathbf{r})$  is the wave function of the single-particle state, in which the bosons condensate into [38].

### Second-Quantization

When describing Bose-Einstein condensates it is very convenient to work in second quantization, which describes the number of particle in each state rather than the state of each particle.

First, consider a basis of single particle states  $\{|n\rangle\}$ , namely a Fock basis. In this space particles are created or annihilated through their respective operators

$$\hat{a}_\mu^\dagger |0\rangle_\mu = |1\rangle_\mu . \quad (2.16)$$

For bosons the creation and annihilation operators fulfill the commutation relations

$$[\hat{a}_\nu, \hat{a}_\mu] = [\hat{a}_\nu^\dagger, \hat{a}_\mu^\dagger] = 0 \quad , \quad [\hat{a}_\nu, \hat{a}_\mu^\dagger] = \delta_{\nu,\mu} , \quad (2.17)$$

with the number operator given as

$$\hat{n}_\mu = \hat{a}_\mu^\dagger \hat{a}_\mu . \quad (2.18)$$

In second quantization many-body states are described by the occupation of the individual Fock states. Thus, a creation operator will raise the number of particles in its corresponding state by one, while the annihilation operator will lower it:

$$\hat{a}_\nu^\dagger |N_0, N_1, \dots, N_\nu, \dots\rangle = \sqrt{N_\nu + 1} |N_0, N_1, \dots, N_\nu + 1, \dots\rangle \quad (2.19)$$

$$\hat{a}_\nu |N_0, N_1, \dots, N_\nu, \dots\rangle = \sqrt{N_\nu} |N_0, N_1, \dots, N_\nu - 1, \dots\rangle . \quad (2.20)$$

Likewise, the number operator  $\hat{n}_\mu$  will count the number of particles in its corresponding state. These operators can be combined with an orthonormal basis of spatial wavefunctions  $\{\phi_k\}$  in order to create field operators

$$\hat{\psi}(\mathbf{r}) = \sum_k \phi_k(\mathbf{r}) \hat{a}_k \quad , \quad \hat{\psi}^\dagger(\mathbf{r}) = \sum_k \phi_k^*(\mathbf{r}) \hat{a}_k^\dagger , \quad (2.21)$$

where  $\hat{\psi}^\dagger(\mathbf{r})$  will create a particle at location  $\mathbf{r}$ . For bosons the field operators fulfill the commutation relations [39]

$$[\hat{\psi}(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')] = \delta^3(\mathbf{r} - \mathbf{r}') , \quad [\hat{\psi}(\mathbf{r}), \hat{\psi}(\mathbf{r}')] = 0 . \quad (2.22)$$

Now, consider a gas of non-interacting bosons described by eq. (2.15), where  $\epsilon_k$  is the energy of the  $k$ 'th single-particle state. Due to the completeness of the basis of single-particle wavefunctions,  $\{\phi_k\}$ , the creation operator can be expressed as

$$\hat{a}_k^\dagger = \int d^3r \phi_k^*(\mathbf{r}) \hat{\psi}^\dagger(\mathbf{r}) . \quad (2.23)$$

Thereby the non-interacting Hamiltonian can be written as

$$\begin{aligned} \hat{H}^{(0)} &= \sum_k \epsilon_k \hat{a}_k^\dagger \hat{a}_k \\ &= \sum_k \int d^3r_1 d^3r_2 \epsilon_k \phi_k(\mathbf{r}_2) \phi_k^*(\mathbf{r}_1) \hat{\psi}^\dagger(\mathbf{r}_1) \hat{\psi}(\mathbf{r}_2) \\ &= \int d^3r \hat{\psi}^\dagger(\mathbf{r}) \left( -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right) \hat{\psi}(\mathbf{r}) , \end{aligned} \quad (2.24)$$

where the orthonormality of the wavefunctions

$$\sum_k \phi_k(\mathbf{r}_2) \phi_k^*(\mathbf{r}_1) = \delta^3(\mathbf{r}_1 - \mathbf{r}_2) \quad (2.25)$$

has been utilized.

### 2.2.2 Weakly Interacting Particles

A characteristic of a BEC is its low temperature and density. Thus, is it a valid approximation to only consider two-particle interactions

$$\hat{H}^{(2)} = \frac{1}{2} \sum_{i \neq j} V(\mathbf{r}_i - \mathbf{r}_j) . \quad (2.26)$$

At low energies all interactions can be considered of s-wave nature, because waves of higher angular momentum are reflected by the centrifugal barrier. Furthermore, the thermal de Broglie wavelength for cold gases is much larger than the effective extension of the interaction potential. Therefore, the actual shape of the scattering potential is irrelevant, whereby one can replace it with a pseudo-potential

$$V(\mathbf{r} - \mathbf{r}') = g \delta^3(\mathbf{r} - \mathbf{r}') = \frac{4\pi\hbar^2 a}{m} \delta^3(\mathbf{r} - \mathbf{r}') , \quad (2.27)$$

which results in the same scattering phase as the real, more complicated scattering potential. Thus, the interaction of cold atoms is fully determined by the scattering length,  $a$  [34]. Introducing the density operator

$$\hat{\rho}(\mathbf{r}) = \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}(\mathbf{r}) , \quad (2.28)$$

allows writing  $\hat{H}^{(2)}$  in second quantization

$$\begin{aligned}\hat{H}^{(2)} &= \frac{1}{2} \int d^3r_1 d^3r_2 V(\mathbf{r}_1 - \mathbf{r}_2) \hat{\psi}^\dagger(\mathbf{r}_1) \hat{\psi}(\mathbf{r}_1) (\hat{\psi}^\dagger(\mathbf{r}_2) \hat{\psi}(\mathbf{r}_2) - \delta(\mathbf{r}_1 - \mathbf{r}_2)) \\ &= \frac{1}{2} \int d^3r_1 d^3r_2 \hat{\psi}^\dagger(\mathbf{r}_1) \hat{\psi}^\dagger(\mathbf{r}_2) V(\mathbf{r}_1 - \mathbf{r}_2) \hat{\psi}(\mathbf{r}_1) \hat{\psi}(\mathbf{r}_2).\end{aligned}\quad (2.29)$$

Combining this with the basic Hamiltonian of equation (2.24), gives the full Hamiltonian in second quantization

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(2)}\quad (2.30)$$

$$\begin{aligned}&= \int d^3r \hat{\psi}^\dagger(\mathbf{r}) \left( -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right) \hat{\psi}(\mathbf{r}) \\ &\quad + \frac{1}{2} \int d^3r_1 d^3r_2 \hat{\psi}^\dagger(\mathbf{r}_1) \hat{\psi}^\dagger(\mathbf{r}_2) V(\mathbf{r}_1 - \mathbf{r}_2) \hat{\psi}(\mathbf{r}_1) \hat{\psi}(\mathbf{r}_2)\end{aligned}\quad (2.31)$$

Solving the Heisenberg equations of motion using the field description of the Hamiltonian leads to

$$i\hbar \frac{\partial}{\partial t} \hat{\psi}(\mathbf{r}) = [\hat{\psi}(\mathbf{r}), \hat{H}] = \left( -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) + g \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}(\mathbf{r}) \right) \hat{\psi}(\mathbf{r}),\quad (2.32)$$

which is not analytically solvable. However, if the scattering length  $a$  is much less than the mean inter-particle distance, such that  $na^3 \ll 1$ , where  $n$  is the density of the gas, the mean-field approximation is viable

$$\hat{\psi}(\mathbf{r}) = \psi(\mathbf{r}) + \delta\hat{\psi}(\mathbf{r}),\quad (2.33)$$

where  $\psi(\mathbf{r})$  is the mean-field given by eq. (2.15), and  $\delta\hat{\psi}(\mathbf{r})$  is fluctuations from the mean. If  $\langle \delta\hat{\psi}(\mathbf{r}) \rangle = 0$ , the fluctuations can be neglected, leading to the Gross-Pitaevskii equation [40, 41]

$$i\hbar \frac{\partial}{\partial t} \hat{\psi}(\mathbf{r}) = \left( -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) + g|\psi(\mathbf{r})|^2 \right) \psi(\mathbf{r}).\quad (2.34)$$

The Gross-Pitaevskii equation is very similar to the Schrödinger equation with exception of the non-linear term  $g|\psi(\mathbf{r})|^2$ , which can make the equation hard to solve in regions of low density.

### 2.3 Bose-Hubbard Model of Interacting Bosons in a Lattice

The Bose-Hubbard model describes weakly interacting bosons in a periodic lattice within the tight binding limit. The model is perfectly capable of predicting results with good accuracy, however, two conditions must hold in order for it to be valid: (i) both the thermal energy and the mean interaction energy at a single site must be much smaller than the separation to first excited band,  $\hbar\omega_0$ , and (ii) the Wannier functions decay essentially within the length of the lattice constant [1]. Under these conditions one is assured that only the lowest band is taken into account, and that only nearest-neighbour interactions take place.

The Bose-Hubbard model is interesting, as it supports two distinct phases: the Superfluid phase and the Mott-Insulator phase. Furthermore, the model contains effects such as a quantum phase transition between the two phases mentioned above, which can be crossed without any change in external parameters. While crossing the critical point a state experiences a variety of quantum phenomena, which have great implications for the dynamics of the system.

### 2.3.1 The Bose-Hubbard Hamiltonian

Consider the Hamiltonian for bosonic particles in a trapping potential in one dimension as described by eq. (2.31). For a periodic lattice potential in the tight binding limit it is favourable to work in a basis of localized Wannier functions. Expanding the field operators of the Hamiltonian in eq. (2.31) in the Wannier basis yields [42]

$$\begin{aligned} \hat{H} &= \int dx \sum_{ij} w^*(x - x_i) \hat{a}_i^\dagger \left( -\frac{\hbar^2}{2m} \nabla^2 + V(x) \right) w(x - x_j) \hat{a}_j \\ &\quad + g \int dx \sum_{ijkl} w^*(x - x_i) w^*(x - x_j) w(x - x_k) w(x - x_l) \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k \hat{a}_l \end{aligned} \quad (2.35)$$

$$= - \sum_{ij} J_{ij} \hat{a}_i^\dagger \hat{a}_j + \sum_{ijkl} U_{ijkl} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k \hat{a}_l , \quad (2.36)$$

where

$$J_{ij} = - \int dx w^*(x - x_i) \left( -\frac{\hbar^2}{2m} \nabla^2 + V(x) \right) w(x - x_j) \quad (2.37)$$

$$U_{ijkl} = g \int dx w^*(x - x_i) w^*(x - x_j) w(x - x_k) w(x - x_l) \quad (2.38)$$

Since the system is periodic, one can consider a single site,  $i = 0$ , as representative for the entire lattice. In this way, the different terms of the Hamiltonian can be interpreted as follows:

- $J_{00}$  = constant energy offset
- $J_{01}$  = "overlap matrix element" to neighbouring site
- $J_{02-0\infty}$  = "overlap matrix element" to further sites
- $U_{0000}$  = on-site interaction for two particles
- $U_{0ii0}$  = interaction off-site
- $U_{0001}$  = interaction + tunnelling, off-site

Due to the rapid decay of the Wannier functions, the overlap of wavefunctions is limited to their nearest neighbour. Dropping all exponentially suppressed terms and constant offsets yields the Bose-Hubbard Hamiltonian

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{a}_i^\dagger \hat{a}_j + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) + \sum_i \varepsilon_i \hat{n}_i , \quad (2.39)$$

where  $J = J_{01}$ , the bracket  $\langle i, j \rangle$  denotes only counting neighbouring pairs, and

$$U = U_{0000} = g \int dx |w(x)|^4 . \quad (2.40)$$

The first term of the Bose-Hubbard Hamiltonian describes the kinematics within the model, which takes the form of tunneling between neighbouring sites. This is interpreted by annihilating a particle at site  $j$  while creating a particle at site  $i$ . The second term describes the interaction between particles within a single site, and the last term  $\sum_i \varepsilon_i \hat{n}_i$  takes into account a possible potential offset at different sites. Figure 2.4 illustrates the parameters of the Bose-Hubbard model as function of the lattice depth. The tunneling matrix element

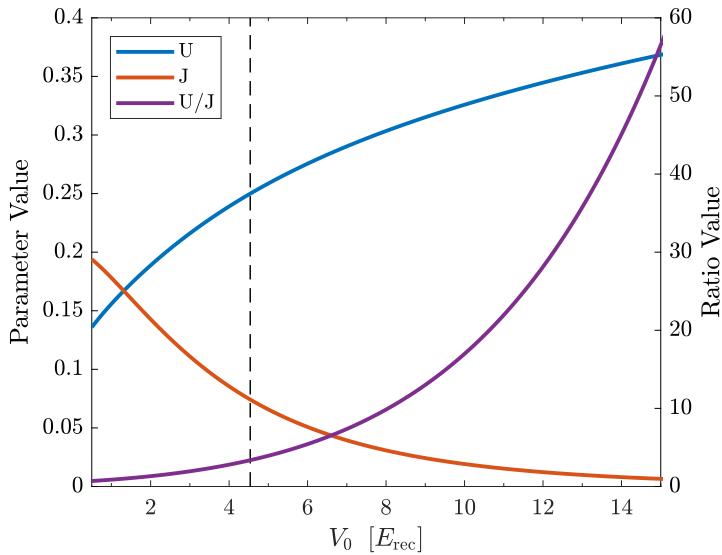


FIGURE 2.4: Tunneling matrix element,  $J$ , of eq. (2.37) and interaction strength,  $U$ , of eq. (2.40) as function of lattice depth. The parameters are calculated for Rubidium-87 atoms in an optical lattice with wavelength  $\lambda = 1064\text{nm}$ . The phase transition between the superfluid and Mott Insulator is marked with a dashed line.

falls off very quickly with increasing lattice depth, as the Wannier functions become localized to their respective wells, thus reducing their overlap. The localization of the Wannier functions also causes the on-site interaction strength to increase.

### 2.3.2 Phases of the Bose-Hubbard Model

The Bose-Hubbard model supports two quantum phases: The *superfluid* phase and the *Mott-Insulator* phase. These phases depend on the ratio  $J/U$ , and can be described separately by examining the ground state of the Bose-Hubbard Hamiltonian in the two extreme limits of the  $J/U$  ratio.

#### Superfluid Phase

For a system where the tunneling matrix element  $J$  is dominant, the lowest energy is obtained by delocalizing the atoms over the entire lattice. Thereby the wavefunction will be a product over single-particle states, and the system will be a superfluid [34].

Consider the case of negligible interactions and a lattice with wells of equal depth. In this scenario the Bose-Hubbard Hamiltonian reduces to

$$\hat{H} = \hat{H}_J = -J \sum_{\langle i,j \rangle} \hat{a}_i^\dagger \hat{a}_j , \quad (2.41)$$

which is completely periodic within the lattice due to the lack of site specific terms. This leads to solutions in the shape of Bloch waves. The Fourier transform of the annihilation and creation operators reads

$$\hat{a}_j = \frac{1}{L} \sum_q e^{iqx_j} \hat{a}_q \quad , \quad \hat{a}_j^\dagger = \frac{1}{L} \sum_q e^{-iqx_j} \hat{a}_q^\dagger , \quad (2.42)$$

where  $L$  is the number of sites in the lattice. Through the Fourier transform the Hamiltonian (2.41) can be expressed in momentum space

$$\hat{H}_J = -J \sum_{q=-\infty}^{\infty} \left( e^{-iqd} + e^{iqd} \right) \hat{n}_q , \quad (2.43)$$

where  $d$  is the lattice distance. In momentum space the Hamiltonian is diagonal and has a continuous energy spectrum

$$E_q = -2J \cos(qd) . \quad (2.44)$$

Thus, the excitation spectrum of the superfluid is said to be gapless. The lowest energy is of the system is obtained for  $q = 0$ , whereby the ground state is

$$|\Psi_{SF}\rangle = \frac{1}{\sqrt{N_p!}} \left( \hat{a}_{q=0}^\dagger \right)^{N_p} |0\rangle = \frac{1}{\sqrt{N_p!}} \left( \frac{1}{L} \sum_{j=1}^L \hat{a}_j^\dagger \right)^{N_p} |0\rangle , \quad (2.45)$$

where  $N_p$  is the number of particles. This state supports a well defined macroscopic phase on each lattice site, since the many-body state is a product over identical single-particle states [34]. With all particle condensed into the  $q = 0$  momentum space state, the particles are completely de-localized in real space, which can be seen by taking the Fourier transform. For large  $N_p, L \rightarrow \infty$  at fixed density  $N_p/L$  the state becomes indistinguishable from having coherent states,  $|\alpha\rangle$ , on all lattice sites [1]. Since bosonic operators at different sites commute, the superfluid state can be factorized into a product of local coherent states

$$|\Psi_{SF}\rangle \approx \prod_j \left( e^{\sqrt{N_p/L} \hat{a}_j^\dagger} \right) |0\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \dots \quad (2.46)$$

Coherent states are eigenstates of the annihilation operator  $\hat{a} |\alpha\rangle = \alpha |\alpha\rangle$ , where  $|\alpha|^2$  can be

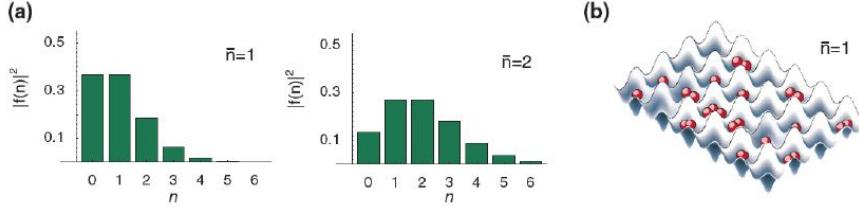


FIGURE 2.5: (a) The statistics for the number of particles per lattice site  $n$  for a filling fraction of  $\bar{n} = 1$  and  $\bar{n} = 2$  in the superfluid phase. (b) Illustration of the particles in the lattice. There will be large fluctuations of the number of particles found at each lattice site at a given time. Figure and caption adopted from [34].

considered the particle density of the system, and  $\alpha = |\alpha|e^{i\phi}$ , with  $\phi$  being a global phase. Utilizing the coherent state form of the wavefunction (2.46) the average filling fraction  $\bar{n}$  can be calculated

$$\bar{n} = \langle \hat{n}_i \rangle = \langle \Psi_{SF} | \hat{a}_j^\dagger \hat{a}_j | \Psi_{SF} \rangle = \frac{N_p}{L} , \quad (2.47)$$

as well as the fluctuations of particle number per site

$$\frac{\sqrt{\Delta \bar{n}^2}}{\bar{n}} \sim \frac{1}{\sqrt{\bar{n}}} . \quad (2.48)$$

Therefore, the probability distribution for the number of atoms at a given site is Poissonian, which can be seen in figure 2.5. Due to the relatively large fluctuation of particle number per site, one would find a somewhat random number of atoms at a given site in a single measurement. However, averaging over many measurements well produce the filling fraction  $\bar{n}$ . In the presence of a finite interaction the resulting distribution would be sub Poissonian due to number squeezing [34].

### Mott-Insulator Phase

In the case of negligible tunneling only the on-site interaction between the atoms has to be taken into account, and the Bose-Hubbard Hamiltonian reduces to

$$\hat{H} = \hat{H}_U = \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) , \quad (2.49)$$

which is quadratic in  $\hat{n}_i$ . This heavily penalizes having multiple particles at the same site. Thus, the ground state of the system will be an equal distribution of all the particles throughout the lattice. Any fluctuations from this average will increase the energy, whereby the phase is said to be incompressible [43]. This incompressibility can be formulated as  $\partial n / \partial \mu = 0$ , where  $\mu$  is the chemical potential, and is the defining property of the Mott-Insulator [1].

Consider the case of average filling of one particle per site,  $\bar{n} = 1$ . The many-body state

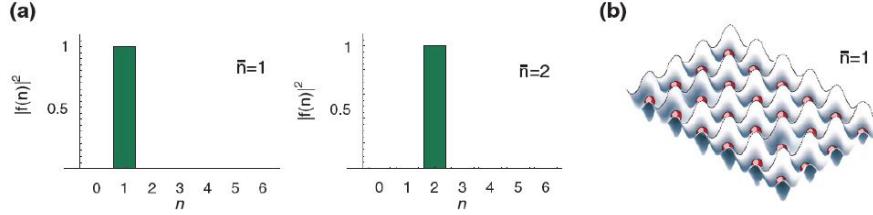


FIGURE 2.6: (a) The statistics for the number of particles per lattice site  $n$  in the Mott-insulator phase, for a filling fraction of  $\bar{n} = 1$  and  $\bar{n} = 2$ . (b) Illustration of the particles in the lattice. The particles does not hop around in the lattice and are equally distributed. Caption and figure are adapted from [34].

can be described by a product of single particles located on each site [1]

$$|\Psi_{Mott}\rangle = \prod_j \hat{a}_j^\dagger |0\rangle . \quad (2.50)$$

Eq. 2.50 is a simple product of local Fock states with precisely one atom per site, which is illustrated in figure 2.6. This configuration of atoms minimizes the energy with regards to the Hamiltonian of eq. (2.49). Any fluctuation from unit occupancy will increase the energy, as merely a single double occupancy will increase the energy by  $U$ . Thereby, unlike the energy spectrum of the Superfluid, the Mott-Insulator spectrum is gapped. As long as the gain in kinetic energy due to hopping,  $J$ , is smaller than the on-site interaction,  $U$ , the atoms remain localized. However, for  $J > 0$  the ground state is no longer the product state described in eq. (2.50) [1]. As  $J$  increases, the gap in the excitation spectrum will gradually decrease until the transition to the superfluid is reached and the spectrum becomes gapless. While the Mott-Insulator phase has complete localization, the phases on the individual sites have obtained maximum uncertainty. Therefore, no phase coherence between different sites is present [34].

### 2.3.3 Mean-Field Approximation of Phase Diagram

Consider once again the full Bose-Hubbard Hamiltonian (2.39). The superfluid and the Mott-Insulator account for the two limits of the ratio  $J/U$ . However, in order to understand the full phase diagram of the Bose-Hubbard model, one has to derive the critical fraction  $(J/U)_{crit}$  for which the phase transition occurs. There are several ways of doing this - one of them is looking at a mean-field solution of the Bose-Hubbard model. For simplicity the zero temperature case is treated here. Even without a change of temperature the phase transition still happens, hence it is referred to as a *quantum phase transition*, since the phase transition can happen without change of external parameters [44].

Applying the mean field approximation to the annihilation operator yields

$$\hat{a}_j = \psi + \delta\hat{a}_j , \quad (2.51)$$

where  $\psi$  is the locally constant mean field, and  $\delta\hat{a}_j$  is the fluctuation term. Inserting this into the Hamiltonian yields

$$\hat{H}_{MF} = -J \sum_{\langle i,j \rangle} \left( \psi^* + \delta\hat{a}_i^\dagger \right) \left( \psi + \delta\hat{a}_j \right) + \text{int.} \quad (2.52)$$

Instead of considering the Hamiltonian as a whole, it can be considered as a sum of local on-site Hamiltonians

$$\hat{H}_{MF} = \sum_i \hat{h}_i . \quad (2.53)$$

For a homogeneous system one can express the kinetic part of the Hamiltonian locally if only small fluctuations from the mean field are assumed

$$\begin{aligned} \hat{a}_i^\dagger \hat{a}_j &= \left( \psi^* + \delta\hat{a}_i^\dagger \right) \left( \psi + \delta\hat{a}_j \right) \\ &= \psi^* \psi + \psi^* \delta\hat{a}_j + \psi \delta\hat{a}_i^\dagger + \delta\hat{a}_i^\dagger \delta\hat{a}_j \\ &\approx \psi^* \psi + \psi^* (\hat{a}_j - \psi) + \psi (\hat{a}_i^\dagger - \psi^*) \\ &= \psi^* \hat{a}_j + \psi \hat{a}_i^\dagger - \psi^* \psi . \end{aligned} \quad (2.54)$$

Since every term only contains a single index, one can let  $j \rightarrow i$  leaving a local description

$$\hat{h}_i = Jz\psi^*\psi - Jz \left( \psi^* \hat{a}_i + \psi \hat{a}_i^\dagger \right) + \frac{U}{2} \hat{n}_i (\hat{n}_i - 1) + \mu_i \hat{n}_i , \quad (2.55)$$

where  $z$  is the number of neighbours of site  $i$  [45]. The last term added is a chemical potential, which is needed due to the lack of a fixed local particle number prompting the use of a Grand Canonical Ensemble description.

Landau theory is a general theory regarding phase transitions, which states that in the vicinity of a critical point, one may expand the free energy in a power series of some order parameter. Thus, in order to find  $(J/U)_{crit}$ , eq. (2.55) must be minimized. here, the order parameter is the mean field, hence

$$E_{MF} = \text{const.} + a|\psi|^2 + b|\psi|^4 + \dots \quad (2.56)$$

Eq. (2.56) constitutes an effective potential, which dictates certain properties of the system. The potential is symmetric in the complex plane allowing the system to have any phase. Furthermore, at some critical value of  $a$ , (here  $a = 0$ ), the potential will shift from having a single central minimum to having minima at some finite  $|\psi|$ . This is a U(1) spontaneous

symmetry breaking, which in Landau theory is associated with a phase transition [46]. Thus, one need to compute the parameter  $a$  of equation (2.56).

One method is using second order perturbation theory, where the  $\hat{\psi} = 0$  case is solved exactly, followed by adding a small  $\hat{\psi}$  as a perturbation. The zeroth order of the local mean-field Hamiltonian reads

$$\hat{h}_i^{(0)} = Jz(\psi^*\psi) + \frac{U}{2}\hat{n}_i(\hat{n}_i - 1) - \mu\hat{n}_i. \quad (2.57)$$

Since the zero-order solution only contains number operators, the ground state of the solution can be expressed in the Fock basis:

$$\begin{aligned} |0\rangle &\text{ for } \mu < 0 \\ |1\rangle &\text{ for } 0 \leq \mu < U \\ |2\rangle &\text{ for } U \leq \mu < 2U \\ &\vdots \end{aligned}$$

The first order perturbation to the energy vanishes, as

$$E_i^{(1)} = \langle n | \delta\hat{h}_i | n \rangle = 0, \quad (2.58)$$

Thus, the use of a second order perturbation is required

$$E_i^{(2)} = \sum_{n \neq m} \frac{|\langle n | \delta\hat{h}_i | m \rangle|^2}{E_m^{(0)} - E_n^{(0)}}. \quad (2.59)$$

Here the perturbation Hamiltonian

$$\hat{h}_i = -zJ(\hat{a}_i\psi^* + \psi\hat{a}_i^\dagger) \quad (2.60)$$

contains only single creation/annihilation operators, whereby

$$\langle n | \delta\hat{h}_i | m \rangle = 0 \quad \text{for } |n - m| \neq 1. \quad (2.61)$$

Hence, only off-diagonal matrix elements contribute to the perturbation

$$\begin{aligned} \langle n | \hat{a}_i^\dagger | n - 1 \rangle &= \sqrt{n} \langle n | n \rangle \\ \langle n + 1 | \hat{a}_i^\dagger | n \rangle &= \sqrt{n + 1} \langle n | n \rangle \\ &\vdots \end{aligned}$$

Thereby, the second order perturbation energy reduces to

$$\begin{aligned} E_i^{(2)} &= (Jz)^2 |\psi|^2 \left( \frac{n}{E_g^{(0)} - E_{n-1}^{(0)}} + \frac{n+1}{E_n^{(0)} - E_{n+1}^{(0)}} \right) \\ &= (Jz)^2 |\psi|^2 \left( \frac{n}{U(n-1) - \mu} + \frac{n+1}{\mu - Un} \right). \end{aligned} \quad (2.62)$$

From eq. (2.56)  $a$  is the pre-factor of all terms proportional  $|\psi|^2$ . Collecting those across from all the perturbations yields the approximation

$$a = Jz + (Jz)^2 |\psi|^2 \left( \frac{n}{U(n-1) - \mu} + \frac{n+1}{\mu - Un} \right). \quad (2.63)$$

As stated earlier  $(J/U)_{crit}$  can be found at the point where  $a = 0$ , hence

$$0 \doteq 1 + \frac{n}{\bar{U}(n-1) - \bar{\mu}} + \frac{n+1}{\bar{\mu} - \bar{U}n}, \quad (2.64)$$

with  $\bar{\mu} = \frac{\mu}{Jz}$  and  $\bar{U} = \frac{U}{Jz}$ . Finally, the solution for the chemical potential is [45]

$$\bar{\mu}_{\pm} = \frac{1}{2} \left( \bar{U}(2n-1) \pm \frac{1}{2} \sqrt{\bar{U}^2 - 2\bar{U}(2n+1)} \right). \quad (2.65)$$

A better understanding of the results can be obtained by examining figure ??, which displays a phase diagram of the Bose-Hubbard model. The  $\mu_{\pm}$  curves encloses the region where the

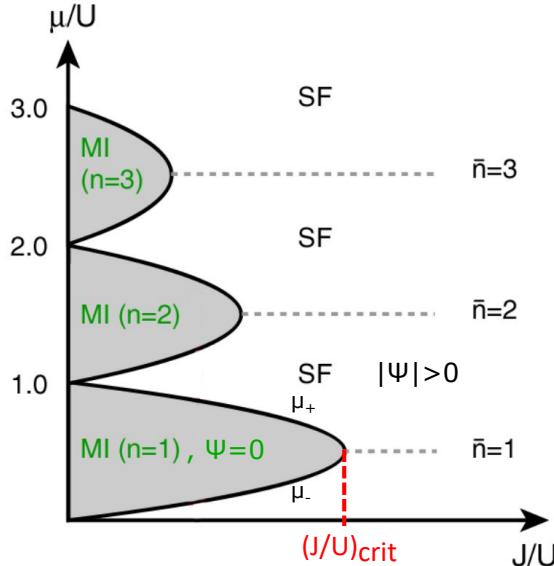


FIGURE 2.7: Phase diagram of Bose-Hubbard model for  $T = 0$ . Grey areas mark the Mott-Insulator phase for different number of particles per site, while white regions mark the superfluid phase. [34]

system is in the Mott-Insulating phase.  $(1/\bar{U})_{crit} = (J/U)_{crit}$  can be read off the graph from the point where the two curves  $\mu_{\pm}$  meet. No fluctuations take place in the Mott Insulator, whereby the particle number per site is well defined. As the chemical potential increases each site can accommodate more particles as long as the increase in chemical potential compensates the increased energy due to interactions between the particles.

The mean-field solution of the Bose-Hubbard model is only an approximation, which proves quite inaccurate for one dimension. This is obvious when comparing the critical ratio for the mean-field approach,  $(U/J)_{crit}^{MF,1D} = 11.66$ , with numerical results computed using the DMRG method,  $(U/J)_{crit}^{DMRG,1D} = 3.37$ , [47]. Nevertheless, the mean-field gives a good intuitive feeling of the physics taking place and how one can describe them without resorting to diagonalizing the Hamiltonian.

## 2.4 Non-Equilibrium Dynamics near Critical Points

The standard procedure for experimentally preparing a Mott-insulator is condensing a gas of cold bosons into a shallow lattice, which effectively forms a superfluid in an equilibrium state. The lattice depth is then slowly ramped up, thereby changing the fraction  $U/J$  and driving the phase transition to the Mott [1, 48]. The constant change of a parameter can be considered a continuous quench of the initial state, whereby the system exhibit out-of-equilibrium dynamics.

This section briefly covers some of the phenomena emerging when crossing a second order phase (continuous) transition like that of the Bose-Hubbard model. The goal is not to provide a full description of these phenomena but rather to illustrate some of the challenges associated with dynamically crossing the critical point.

### 2.4.1 Breakdown of Adiabaticity

Many experimental procedures rely on adiabaticity for state preparation. The adiabatic theorem in quantum mechanics state that under slow enough external perturbations, no transitions between energy levels take place. For systems with an energy gap between the ground and excited states this is particularly useful, as the ground state can only being excited through the Landau-Zener mechanism [49, 50] under small perturbations. Thus, state transfers can be achieved adiabatically in gapped systems. In most many-body systems the distance is exponentially small, whereby the time-scale of an adiabatic parameter variation grows exponentially large [51]. The one-dimensional superfluid phase is gapless, whereby the energy spectrum is continuous. Therefore, any change in the parameters of a superfluid ground state will results in excitations in the form of Bogoliubov quasi-particles [51, 52]. Furthermore, in the vicinity of a critical point the relaxation time-scale of the system diverges, which is known as the *critical slowing down* [12].

Thus, adiabaticity is bound to break down when transitioning from a superfluid to a Mott-insulating state by varying the fraction  $U/J$ . However, changes of parameters are often still referred to as "adiabatic" if they are sufficiently slow compared to the natural time scales of the systems.

### 2.4.2 The Kibble-Zurek Mechanism

The Kibble-Zurek theory describes the emergence of topological defect following a quench through a second-order (continuous) phase transition [12].

Consider a system in equilibrium described by a correlation length  $\xi$ . In the vicinity of a critical point the correlation length diverges as a power law

$$\xi \approx \xi_0 |\epsilon(t)|^{-\nu} , \quad \epsilon > 0 , \quad (2.66)$$

where  $\nu$  is a static critical exponent, and  $\epsilon$  is a reduced coupling, which is zero at the critical point. Thus, near the critical point the equilibrium state is characterized by an infinite correlation length.

Assume a state initially in equilibrium begin evolved towards the critical point [53]. Approaching a second-order phase transition, an equilibrium state can have a relaxation time-scale,  $\tau$ , which diverges according to

$$\tau \approx \tau_0 |\epsilon(t)|^{-\mu} , \quad \epsilon > 0 , \quad (2.67)$$

near the critical point [12]. Here, the parameter  $\mu$  is another critical exponent. This divergence of the response time of the state referred to as the *critical slowing down* and

causes an inevitable breakdown of adiabaticity. As the relaxation time becomes larger than the time before reaching the transition, the systems effectively enters a "freeze-out" period, in which the correlation length no longer can evolve. Thus, the system crosses the critical point without  $\xi$  ever having spanned the entire system, which results in the formation of broken-symmetry phase domains, or defects, with independent values of the underlying order parameter [54]. This phenomenon is known as the Kibble-Zurek mechanism. In the context of the Bose-Hubbard model, the Kibble-Zurek mechanism is dependent on the direction of the phase transition, due the completely different energy spectra of the phases. While the transition from the gapped Mott-insulating state to a superfluid will produce topological defects, the opposite transition will amplify the quantum fluctuations of the initial state. Thus, the number-fluctuations of the initial superfluid state may carry over to the Mott-insulator, however, the magnitude of these fluctuations will depend on the exact form of the parameter variation [55].

### 2.4.3 Light-Cone Propagation of Correlations

An important phenomenon for understanding non-equilibrium dynamics of local Hamiltonian models is the light-cone-like propagation of information within the system [56]. The effective light-cone limits the speed at which a system can respond to sudden changes in the Hamiltonian and is therefore important when considering the dynamical crossing of a phase transition. While the existence of light-cones in relativistic quantum field theories is well known, no strict light-cone behavior is present in general, non-relativistic quantum systems. However, for quantum systems evolving according to a local Hamiltonian, the propagation velocity of information is limited according to the Lieb-Robinson bounds [57]. Consider a general lattice system with the local Hamiltonian

$$\hat{H}(t) = \sum_{\langle i,j \rangle} \hat{h}_{ij}(t) , \quad (2.68)$$

where  $\langle \cdot, \cdot \rangle$  denotes nearest-neighbour sites. For two non-overlapping blocks of the lattice  $A$ ,  $B$  the Lieb-Robinson bound makes a statement about the operator norm of the commutator of any operators  $\hat{O}_A$  and  $\hat{O}_B$  in regions  $A$ ,  $B$  taken at different times

$$\left\| [\hat{O}_A(t), \hat{O}_B(0)] \right\| \leq c \|\hat{O}_A\| \|\hat{O}_B\| \min(|\hat{O}_A|, |\hat{O}_B|) \exp\left(-\frac{L - v|t|}{\gamma}\right) , \quad (2.69)$$

where  $L$  is the distance between  $A$  and  $B$ , and  $|A|$ ,  $|B|$  denote the size (number of sites) in each block [23]. The positive constants  $c, v, \lambda > 0$  are determined by the maximum norm of the local Hamiltonians  $\max_{\langle i,j \rangle} \max_t \|\hat{h}_{ij}(t)\|$  and are therefore system dependent. The time-dependent operator norm of the commutator can be considered a measurement of correlated information separated in space and time, and it is effectively bound by the light-cone emerging from the exponential term of eq. (2.69). A complete understanding of time-dependent Lieb-Robinson bound (2.69) is outside the scope of this thesis, however, two remarkable consequences regarding the spreading of information can be inferred from the bound: (i) information propagating outside the light-cone is exponentially suppressed [23], and (ii) correlations within the system is mediated through entangled quasi-particles [58], whose ballistic propagation velocity is limited by  $v$  [59]. The propagation velocity,  $v$ , different than the sound velocity of the system for curved dispersion relations [60], whereby it is referred to as the *Lieb-Robinson velocity*.

Experimental evidence of an effective light-cone in a one-dimensional Bose-Hubbard system was shown in [58]. In the experiment a Mott-insulating ground state was quenched to near the critical point as shown in figure 2.8(a). The initial ground state was thereby

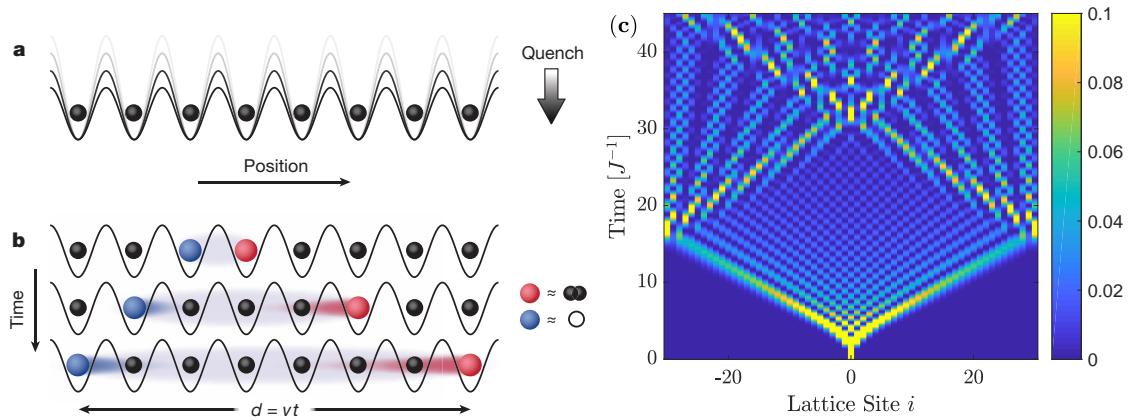


FIGURE 2.8: Illustration of light-cone-like propagation of information. (a) one dimensional ultracold gas of bosonic atoms in a Mott configuration being abruptly quenched to a shallow lattice. (b) Entangled quasi-particle pairs emerge at all sites and propagate ballistically following the quench. (c) Density plot of the evolution of  $|1\rangle_0$  in an  $L = 61$  lattice according to a simple Hubbard Hamiltonian (2.70). Figures (a) and (b) are adopted from [58].

highly excited and acted as source for quasi-particles consisting of *dublons* (excess particle) and *holons* (hole) shown in figure 2.8(b). By measuring the two-point parity correlation function, the spreading of information through the system was shown to propagate at a constant velocity limited by the Lieb-Robinson bound.

To illustrate the propagation of correlation-mediating quasi-particles the evolution of an impurity in a simple Hubbard model was examined. Consider a Hubbard Hamiltonian of the form

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{a}_i^\dagger \hat{a}_j , \quad (2.70)$$

which contains no interactions. A quasi-particle excitation may not have strong interactions with the remaining system, although they may mutually interact. Here a single particle is considered, whereby a negligible interaction is a reasonable approximation. Meanwhile, the quasi-particle will be experiencing the same lattice potential and any underlying harmonic confinement, if it remains in the lowest band. Here, a flat lattice (no underlying potential) is considered. Thus, the model described is equivalent to a quasi-particle excitation on top of a homogeneous Mott-insulator. Figure 2.8(c) shows the density evolution of a single impurity,  $|1\rangle_0$ , initially placed at the central site in a lattice of size  $L = 61$ . The resulting density wave is very reminiscent of a light/cone, as the probability velocity is bounded by a constant. Further examining the emerging cone, one will notice an exponentially suppressed population just outside the cone (more visible near the boundary). Due to the open boundaries of the system, the quasi-particle is reflected at the edge of the lattice, causing it to interfere with itself [59].

It should be stressed that the density wave shown in figure 2.8 is not light-cone but merely the propagation of a single quasi-particle-like impurity. In a physical system quasi-particles will emerge at all sites, whereby the Lieb-Robinson velocity is different from the spreading velocity of non-interacting particles. In fact, the experimental investigation of light-cones in the Bose-Hubbard model found information velocities much higher than single-particle spreading velocities, yet lower than the Lieb-Robinson bound [58]. While the simple Hubbard model illustrates the propagation of a single quasi-particle, the propagation of information within a many-body system is a many-body effect itself.

Due to the breakdown of adiabaticity around the critical point, it is inevitable that quasi-particle excitations will be created when transitioning from a superfluid to a Mott-insulator. Furthermore, a study of the Mott to superfluid phase transition showed that the propagation of correlations was well described for fast ramp sequences by ballistically spreading quasi-particles following the Lieb-Robinson bounds, while slower ramps required the interaction of quasi-particles to be considered [13]. Therefore, engineering a lattice ramping sequence bringing the system from a superfluid to a Mott-Insulating ground state is a highly difficult task. While classical approaches fail to capture the rich dynamics of the phase transition, simulations of the system can be performed through tensor network methods, which provide a low-entanglement effective theory [54]. Although the low-entanglement approximation inevitably fails for long durations, the finite spreading velocity of correlations within the system enables the use of tensor-network methods for simulating dynamics for short durations [56].



## Chapter 3

# Matrix Product States

Numerical analysis of quantum many-body systems using exact diagonalization is highly limited by system size. Due to the large number of possible configurations in a many-body system and entanglement coupling various degrees of freedom, the Hilbert space grows too large for exact methods to search. Consider the Bose-Hubbard model described in eq. (2.39). The dimension of the corresponding Hilbert space is

$$D_{\mathcal{H}} = \frac{(L + N_p - 1)!}{N_p!(L - 1)!} , \quad (3.1)$$

where  $L$  is the number of sites, and  $N_p$  is the number of particles. For unit occupancy,  $N_p/L = 1$ , the Hilbert space grows exponentially with the system size [61]. Therefore, descriptions using exact diagonalization are only possible for small systems. Matrix product states parameterize the quantum states through a product of tensors, which can be depicted as a network. Decomposing the state as an explicitly contracted network of tensors enables efficient application of operators. Furthermore, the construction of matrix product states puts an upper bound on the entanglement entropy within the system, whereby states can be represented efficiently without considering unnecessary degrees of freedom [22].

This chapter will cover the basics of matrix product states including their construction and freedom of gauges called canonical forms [21]. Furthermore, the generalization of MPS to operators called matrix product operators are introduced along with their applications. Algorithms involving matrix products states such as the DMRG method [62, 63] are introduced in the following chapter.

### 3.1 Entanglement in Quantum Systems and Area Laws

Entanglement is a fundamental property of quantum mechanics responsible for correlating different degrees of freedom within a quantum system. Thus, the individual parts of a quantum system can not be described alone, as entanglement with the remainder of the system has to be taken into consideration. This drastically complicates any description of the system, which is why exact descriptions of many-body systems are almost impossible. The measure of entanglement within a quantum many-body system is the entanglement entropy, although it is often more instructive to look at the bipartite entanglement entropy, which measures the entanglement between two partitions of the system. Consider a bipartition of the Hilbert space  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ . A state  $|\psi\rangle \in \mathcal{H}$  can be decomposed using the *Schmidt decomposition* as

$$|\psi\rangle = \sum_{\alpha} \Lambda_{\alpha} |\alpha\rangle_A \otimes |\alpha\rangle_B , \quad (3.2)$$

where the states  $\{|\alpha\rangle_{A(B)}\}$  form an orthonormal basis of  $\mathcal{H}_{A(B)}$ , and  $\Lambda_{\alpha} \geq 0$  are the Schmidt coefficients fulfilling  $\sum_{\alpha} \Lambda_{\alpha}^2 = 1$  [64]. If only one term contributes to the Schmidt decomposition, the state is a product state i.e. the two parts of the Hilbert space are not mutually

entangled. More terms implies that the state is entangled. From the Schmidt decomposition one can determine the reduced density matrix

$$\rho_A = \text{Tr}_B |\psi\rangle\langle\psi| = \sum_{\alpha} \Lambda_{\alpha}^2 |\alpha\rangle_A\langle\alpha|_A , \quad (3.3)$$

by tracing out the subsystem  $B$ . From the Schmidt decomposition one can determine the entanglement entropy of the bipartition, which is defined as the von-Neumann entropy,  $S$ , of the reduced density matrix given by [64]

$$S = - \sum_{\alpha} \Lambda_{\alpha}^2 \log \Lambda_{\alpha}^2 . \quad (3.4)$$

Quite often, one is not concerned with the actual entanglement entropy of the system, but rather how it scales when the region in question grows in size. It is natural to assume that the entanglement entropy scales with the *volume* of the system, since this is the case for thermal systems. However, ground states of one-dimensional quantum many-body systems often follow an *area law*, meaning that the scaling of the entropy is linear in the boundary area of the region. This is especially the case for systems with a gapped and local Hamiltonian [22].

In a one dimensional lattice the boundary size of a bipartition is constant independently of where the cut is made. Hence, the entanglement entropy will be bounded above, if the system follows an area law. This can be understood intuitively, as only degrees of freedom within the correlation length,  $\xi$ , of the boundary will be entangled, no matter where in the system the boundary is present [65]. This observation is the key to the success of numerical computations for one-dimensional systems, as the limited degree of entanglement enables the truncation of unused eigenstates without loss of accuracy in the representation of the state. Thus, many different states can be efficiently described by tensor networks despite existing in an exponentially large Hilbert space.

## 3.2 Construction of an MPS

Matrix products states are parameterizations of one dimensional quantum states through a product of tensors. This parameterization is especially intuitive for lattice systems, as each lattice site can be represented by a single tensor. These tensors have two different kinds of indices; *physical indices*,  $j$ , which corresponds to the local, physical states at a given site, and *bond indices*,  $\alpha$ , which serves to connect neighbouring sites. The dimension of the physical index is equal to that of the underlying Fock space, while the dimension of virtual bond index is related to the entanglement entropy between the two connected sites. Tensors can be merged by contracting the bond connecting them, which is done by summing over their common index. Equations involving matrix product states tend to grow quite long, whereby they often are described through diagrams. In the diagrammatic representation the physical indices are marked by vertical legs, while bond indices are marked by horizontal legs. When two tensors are connected by a leg, it means that the corresponding bond is contracted. Further details and examples regarding tensor diagrams can be found in Appendix A.

Consider a chain of  $L$  sites with each site having a  $d$ -dimensional local Hilbert space  $\{|j_n\rangle\}$ , where  $n = 1, \dots, L$ . An arbitrary quantum state of this system reads

$$|\psi\rangle = \sum_{j_1, \dots, j_L} c_{j_1 \dots j_L} |j_1, \dots, j_L\rangle . \quad (3.5)$$

This description of a general state can be parametrized to the MPS form by applying successive Schmidt decomposition. However, in practice this is done through Singular Value Decompositions (SVD), as these are faster to perform numerically. The two approaches are equivalent, as the singular value decomposition is essentially a restatement of the Schmidt decomposition [64]. Through an SVD, an arbitrary matrix,  $A$ , of dimensions  $(M \times N)$  can be decomposed into three matrices,  $A = USV^\dagger$ , with the following properties [21]:

- $U$  is of dimension  $(M \times \min(M, N))$  and has orthonormal columns, meaning  $U^\dagger U = I$ . If  $M \leq N$ , then it is also unitary  $UU^\dagger = I$ .
- $S$  is a positive, diagonal  $(\min(M, N) \times \min(M, N))$  matrix. The diagonal elements are singular values, and the number of non-zero entries is the Schmidt rank of  $A$ .
- $V^\dagger$  is of dimension  $(\min(M, N) \times N)$  and has orthonormal rows, meaning  $V^\dagger V = I$ . If  $M \geq N$ , then it is also unitary  $VV^\dagger = I$ .

From the general state of eq. (3.5) an MPS can be constructed through the following steps:

1. The  $d^L$ -dimensional vector  $c_{j_1 \dots j_L}$  is reshaped into a  $(d \times d^{L-1})$  matrix  $\Psi_{j_1, (j_2 \dots j_L)}$ . Performing an SVD on  $\Psi$  yields

$$c_{j_1 \dots j_L} = \Psi_{j_1, (j_2 \dots j_L)} = \sum_{\alpha_1}^d U_{j_1, \alpha_1} S_{\alpha_1, \alpha_1} (V^\dagger)_{\alpha_1, (j_2 \dots j_L)} = \sum_{\alpha_1}^d A_{\alpha_1}^{j_1} \Psi_{(\alpha_1 j_2), (j_3 \dots j_L)}, \quad (3.6)$$

where  $U$  has been reshaped into  $d$  row vectors  $A^{j_1}$  with entries  $A_{\alpha_1}^{j_1} = U_{j_1, \alpha_1}$ . In this notation  $A_{\alpha_1}^{j_1}$  is a tensor with physical indices  $j_1$  and bond indices  $\alpha_1$ . Furthermore,  $S$  and  $V^\dagger$  has been multiplied and reshaped into a  $(d^2 \times d^{L-2})$  matrix,  $\Psi_{(\alpha_1 j_2), (j_3 \dots j_L)}$ .

2. The new matrix  $\Psi_{(\alpha_1 j_2), (j_3 \dots j_L)}$  is subjected to an SVD decomposition

$$c_{j_1 \dots j_L} = \sum_{\alpha_1}^d \sum_{\alpha_2}^{d^2} A_{\alpha_1}^{j_1} U_{(\alpha_1 j_2), \alpha_2} S_{\alpha_2, \alpha_2} (V^\dagger)_{\alpha_2, (j_3 \dots j_L)} = \sum_{\alpha_1}^d \sum_{\alpha_2}^{d^2} A_{\alpha_1}^{j_1} A_{\alpha_1, \alpha_2}^{j_2} \Psi_{(\alpha_2 j_3), (j_4 \dots j_L)}, \quad (3.7)$$

where  $U$  has been decomposed into  $d$  matrices  $A^{j_2}$  with entries  $A_{\alpha_1, \alpha_2}^{j_2} = U_{(\alpha_1 j_2), \alpha_2}$ . The matrix  $\Psi_{(\alpha_2 j_3), (j_4 \dots j_L)}$  has dimensions  $(d^3 \times d^{L-3})$  and is yet again a product of the matrices  $S$  and  $V^\dagger$  from the SVD.

3. The procedure detailed in the two previous steps is continued throughout the chain leaving

$$c_{j_1 \dots j_L} = \sum_{\alpha_1, \dots, \alpha_{L-1}} A_{\alpha_1}^{j_1} A_{\alpha_1, \alpha_2}^{j_2} \dots A_{\alpha_{L-2}, \alpha_{L-1}}^{j_{L-1}} A_{\alpha_{L-1}, \alpha_L}^{j_L} \equiv A^{j_1} A^{j_2} \dots A^{j_{L-1}} A^{j_L}, \quad (3.8)$$

where the sum can be recognized as a matrix multiplication allowing a neater notation.

The construction of an MPS is illustrated in figure 3.1. Thereby, the original state (equation 3.5) can be written as a matrix product state in the form [21]

$$|\psi\rangle = \sum_{j_1, \dots, j_L} A^{j_1} A^{j_2} \dots A^{j_{L-1}} A^{j_L} |j_1, \dots, j_L\rangle. \quad (3.9)$$

Following the construction sequence, the dimensions of the matrices  $A^{j_n}$  follow a pyramid-like structure  $(1 \times d), (d \times d^2), \dots, (d^{L/2-1} \times d^{L/2}), (d^{L/2} \times d^{L/2-1}), \dots, (d \times 1)$ , where  $L$  is taken as even for simplicity. However, the state can be represented much more efficiently

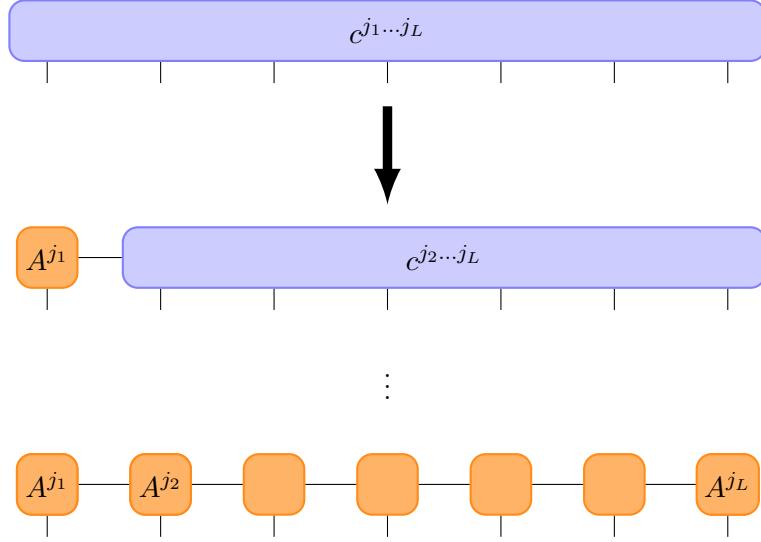


FIGURE 3.1: *Diagrammatic representation of the construction of a left-canonical MPS from an arbitrary quantum state through successive SVD's.*

by truncating the dimensions of the matrices (bonds), due to the area law scaling of entanglement entropy [66]. A bond dimension  $\chi$  infers an entanglement entropy bounded by  $S(\rho_A) = \mathcal{O}(\log \chi)$  [54], whereby one often can discard the contribution of many eigenstates without loss of accuracy. In practice this is done by keeping only the  $D$  largest singular values of the SVD matrix  $S$ , as these contribute the most to the state. Alternatively one can discard all eigenstates contributing with eigenvalues less than some threshold  $\epsilon_t$ .

### 3.3 Canonical Forms

The MPS described in equation 3.9 is not unique, as writing

$$\tilde{A}^{j_n} = X_{n-1} A^{j_n} X_n^{-1} \quad (3.10)$$

describes the same state using different matrices. This gauge freedom allows expressing the MPS in whichever way is most convenient, which can greatly reduce the effort of applying operators and calculating overlaps. By choosing a gauge, the MPS is brought into a *canonical form* [67]. The following examples detail bringing MPSs into canonical forms through construction, however, the same can be achieved for an already built MPS through successive SVDs and regrouping of the resulting matrices.

#### 3.3.1 Left-canonical matrix product state

The process of constructing an matrix product state detailed in Section 3.2 brings the MPS in a *left-canonical* form. This implies that all the matrices are left-normalized, such that

$$\sum_{j_n} A^{j_n \dagger} A^{j_n} = I . \quad (3.11)$$

The left-normalization is a consequence of the matrix  $U$  (from the SVD) fulfilling  $U^\dagger U = I$ . Since the matrices  $A^{j_n}$  are reshaped from  $U$ , these properties persists

$$\begin{aligned}\delta_{\alpha_n, \alpha'_n} &= \sum_{\alpha_{n-1}, j_n} (U^\dagger)_{\alpha_n, (\alpha_{n-1}, j_n)} U_{(\alpha_{n-1}, j_n), \alpha'_n} \\ &= \sum_{\alpha_{n-1}, j_n} (A^{j_n\dagger})_{\alpha_n, \alpha_{n-1}} A_{\alpha_{n-1}, \alpha'_n}^{j_n} \\ &= \sum_{j_n} \left( A^{j_n\dagger} A^{j_n} \right)_{\alpha_n, \alpha'_n}\end{aligned}$$

Figure 3.2 illustrates a contraction of the bonds connecting two left-normalised matrices. As

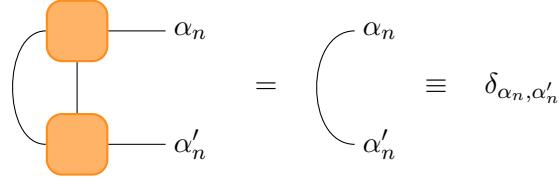


FIGURE 3.2: *Contraction over the left index (shown as the arc) and the physical index of two left-normalised matrices. The result is  $\delta_{\alpha_n, \alpha'_n}$ , adding the  $n$ 'th site to the contraction of all previous sites to the left.*

this contraction results in the identity per definition, one can contract left-normalised matrices without any explicit calculation. The explicit contraction is displayed diagrammatically through an arc, which is equivalent to an identity-tensor with two bond indices.

### 3.3.2 Right-canonical matrix product state

One could also have built a right-canonical MPS from eq. (3.5), had one performed the SVDs from other side of the chain. Constructing the MPS from the right implies multiplying  $U$  and  $S$  into the new  $\Psi$ -matrix, and reshaping  $(V^\dagger)_{(\alpha_{n-1}, j_n), \alpha_n}$  into matrices  $B_{\alpha_{n-1}, \alpha_n}^{j_n}$ . Hence, the right-canonical form of the MPS of eq. (3.9) reads

$$|\psi\rangle = \sum_{j_1, \dots, j_L} B^{j_1} B^{j_2} \dots B^{j_{L-1}} B^{j_L} |j_1, \dots, j_L\rangle . \quad (3.12)$$

In this form all the matrices are right-normalized, whereby

$$\sum_{j_n} B^{j_n} B^{j_n\dagger} = I . \quad (3.13)$$

Note the reverse ordering of the matrix product compared to the left-normalized case. Figure 3.3 illustrates two right-normalised matrices contracted over their physical index.

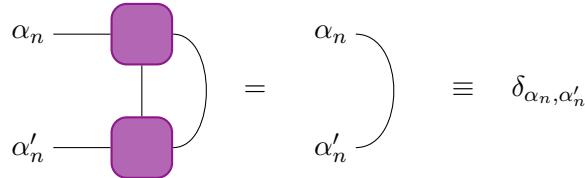


FIGURE 3.3: *Contraction over the right index (shown as the arc) and the physical index of two right-normalised matrices.*

### 3.3.3 Mixed-canonical matrix product state

In practice one rarely finds use for a purely left- or right-canonical MPS. However, combining the two canonical forms listed above yields the mixed-canonical form. The mixed-canonical results in a natural bi-partitioning of the system into a Schmidt decomposition, and it is especially useful for measuring local properties of the state [21].

Consider the procedure of building an MPS in the left-canonical form, where  $c_{j_1 \dots j_L}$  has been decomposed from the left up until site  $n$

$$c_{j_1 \dots j_L} = \sum_{\alpha_n} (A^{j_1} \dots A^{j_n})_{\alpha_n} S_{\alpha_n, \alpha_n} (V^\dagger)_{\alpha_n, (j_{n+1} \dots j_L)} . \quad (3.14)$$

By reshaping  $V^\dagger$  into the matrix  $\Psi_{(\alpha_n j_{n+1} \dots j_{N-1}), j_N}$ , one can initiate a successive decomposition from the right, resulting in a set of right-normalized matrices. The final result reads

$$c_{j_1 \dots j_N} = A^{j_1} \dots A^{j_n} S B^{j_{n+1}} \dots B^{j_L} , \quad (3.15)$$

which is illustrated in figure 3.4a. In the mixed-canonical form the Schmidt decomposition

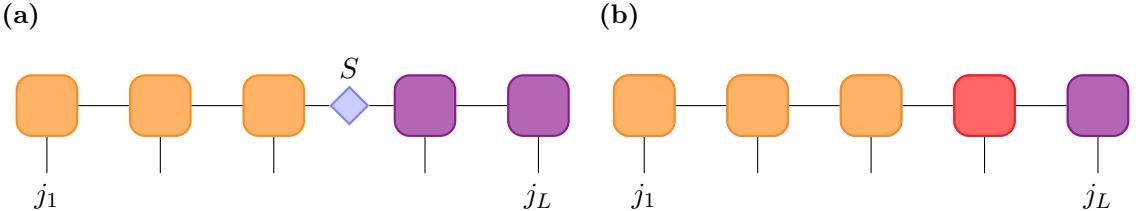


FIGURE 3.4: Two variants of the mixed-canonical form of an MPS. The form (a) retains the matrix  $S$  originating from the SVD, whereby the MPS is directly brought into the form of a Schmidt decomposition. Multiplying  $S$  into either its left or right neighbour creates the form (b), which has a central site containing the normalization of the entire state. The form (b) is well suited for measuring local properties of the state.

can be read directly from the form of the MPS by introducing the vectors

$$|\alpha_n\rangle_A = \sum_{j_1, \dots, j_n} (A^{j_1} \dots A^{j_n})_{1, \alpha_n} |j_1, \dots, j_n\rangle \quad (3.16)$$

$$|\alpha_n\rangle_B = \sum_{j_{n+1}, \dots, j_L} (B^{j_{n+1}} \dots B^{j_L})_{\alpha_n, 1} |j_{n+1}, \dots, j_L\rangle , \quad (3.17)$$

whereby the state can be written in the form

$$|\psi\rangle = \sum_{\alpha_n} S_{\alpha_n, \alpha_n} |\alpha_n\rangle_A |\alpha_n\rangle_B . \quad (3.18)$$

In order for eq. (3.18) to be considered a Schmidt decomposition, the matrix  $S$  must fulfill  $\sum_{\alpha_n} (S_{\alpha_n, \alpha_n})^2 = 1$ , which is fulfilled by default by the SVD. Furthermore, the states  $|\alpha_n\rangle_A$  and  $|\alpha_n\rangle_B$  have to be orthonormal respectively, which they are by construction.

Another useful version of the mixed-canonical form is depicted in figure 3.4b, where the matrix  $S$  of eq. (3.15) has been multiplied unto the matrix to either the left or right of it. The result is a central site, which is neither left- nor right-normalized, but instead contains the normalization of the entire state. Overlaps and expectation values can be computed very efficiently using the mixed-canonical form, as most of the tensor network can be explicitly contracted due to the normalization.

## 3.4 Matrix Product Operators

A Matrix Product Operator (MPO) is an operator expressed in the formalism of an MPS. Thereby, operators can be incorporated as part of the tensor networks, where they are evaluated by the contraction of bonds.

Consider a single coefficient of an MPS of the state general  $|\psi\rangle$

$$\langle j_1, \dots, j_L | \psi \rangle = \langle j | \psi \rangle = M^{j_1} M^{j_2} \dots M^{j_L}. \quad (3.19)$$

Expressing an operator  $\hat{O}$  in the basis of the local states, one can write it in a similar manner

$$\hat{O} = \sum_{j, j'} |j\rangle \langle j| \hat{O} |j'\rangle \langle j'| = \sum_{j, j'} W^{j_1, j'_1} W^{j_2, j'_2} \dots W^{j_L, j'_L} |j\rangle \langle j'|, \quad (3.20)$$

where the coefficients are  $\langle j | \hat{O} | j' \rangle = W^{j_1, j'_1} W^{j_2, j'_2} \dots W^{j_L, j'_L}$ . The operator matrices,  $W^{j_n, j'_n}$ , differ from the state matrices,  $M$ , as the representation of the operators needs both an ingoing and an outgoing physical index. The ingoing physical index is depicted graphically as an extra vertical leg. A pictorial representation of the operator  $\hat{O}$  can be seen in figure 3.5.

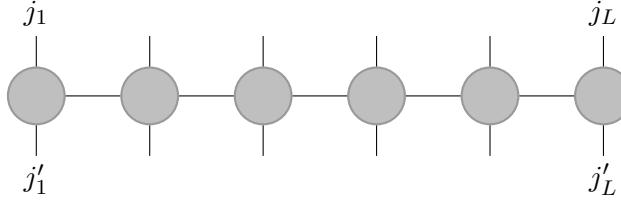


FIGURE 3.5: An operator  $\hat{O}$  expressed in the MPS form (MPO). The resulting matrix product operator has two physical indices marked by vertical lines, which corresponds to an ingoing and outgoing physical state.

### 3.4.1 Applying an MPO to an MPS

Applying a matrix product operator to a matrix product state is simply a matrix multiplication, where the matching physical indices are summed over:

$$\begin{aligned} \hat{O} |\psi\rangle &= \sum_{j, j'} \sum_{\alpha, \beta} \left( M_{1, \alpha_1}^{j'_1} M_{\alpha_1, \alpha_2}^{j'_2} \dots \right) \left( W_{1, \beta_1}^{j'_1, j_1} W_{\beta_1, \beta_2}^{j'_2, j_2} \dots \right) |j\rangle \\ &= \sum_{j, j'} \sum_{\alpha, \beta} \left( M_{1, \alpha_1}^{j'_1} W_{1, \beta_1}^{j'_1, j_1} \right) \left( M_{\alpha_1, \alpha_2}^{j'_2} W_{\beta_1, \beta_2}^{j'_2, j_2} \right) \dots |j\rangle \\ &= \sum_j \sum_{\alpha, \beta} N_{(1, 1), (\alpha_1, \beta_1)}^{j_1} N_{(\alpha_1, \beta_1), (\alpha_2, \beta_2)}^{j_2} \dots |j\rangle \\ &\equiv \sum_j N^{j_1} N^{j_2} \dots |j\rangle = |\phi\rangle \end{aligned} \quad (3.21)$$

The result is a new MPS,  $|\phi\rangle$ , which is described by the matrices  $N^{j_n}$  and is illustrated in figure 3.6. These matrices have the dimensions of the product of the dimensions of the original MPS and MPO. Thus, applying an operator leaves the form of the MPS invariant but increases the matrix dimensions. If the dimension of the MPS is  $D$ , and the dimension of the MPO is  $D_W$ , the total computational cost of the operation is  $\mathcal{O}(Ld^2 D_W^2 D^2)$ .[21, 68]

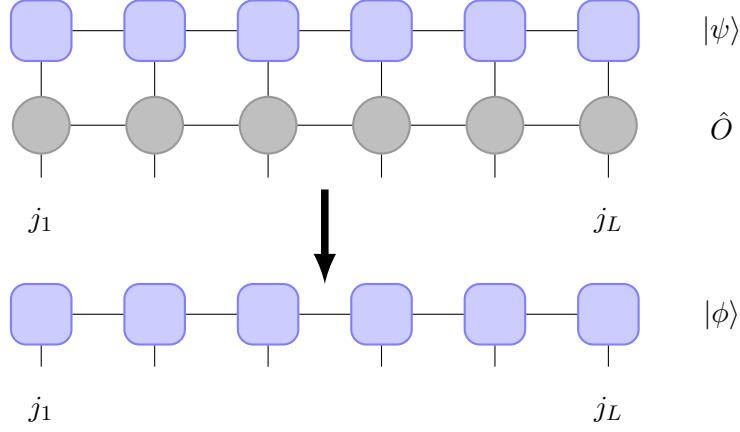


FIGURE 3.6: Application of an MPO,  $\hat{O}$ , to an MPS,  $|\psi\rangle$ . Matching physical indices are contracted resulting in a new MPS,  $|\phi\rangle$ , with increased matrix dimensions.

### 3.5 Correlation Functions and Measurement of Local Properties

Measuring local properties of a state is achieved by calculating the expectation value of local operators. Consider the local operator  $\hat{O}^{[n]}$ , where the square brackets denote the site, on which  $\hat{O}$  operates. As  $\hat{O}^{[n]}$  acts only on site  $n$ , it can be expressed in the local basis of said site

$$\hat{O}^{[n]} = \sum_{j_n, j'_n} O^{j_n, j'_n} |j_n\rangle \langle j'_n|. \quad (3.22)$$

Applying a local operator to a state implies applying the identity operator to all other sites. The full operator,  $\hat{O} = \hat{I}^{[1]} \otimes \hat{I}^{[2]} \otimes \dots \otimes \hat{O}^{[n]} \otimes \dots \otimes \hat{I}^{[L]}$ , can be expressed as an MPO and readily be applied to a matrix product state. However, the calculation is greatly simplified when considering an MPS in a mixed-canonical form, where the central tensor is located on site  $n$ . Thereby the entire tensor network on either side of site  $n$  can be contracted without any calculation. The remaining network needed to be contracted is shown in figure 3.7a, which is simply the sum

$$\langle \psi | \hat{O}^{[n]} | \psi \rangle = \sum_{j_n, j'_n} O^{j_n, j'_n} \text{Tr} \left( M^{j_n \dagger} M^{j'_n} \right) \quad (3.23)$$

with the computational cost  $\mathcal{O}(D^2 d^2)$  [21]. Very similar is the case of multiple-site local

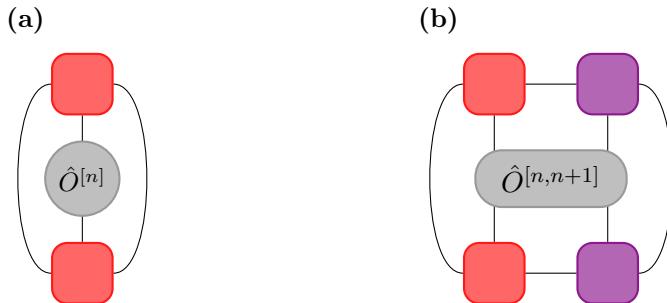


FIGURE 3.7: Measurement of local properties of a matrix product state in the mixed-canonical form.

operator. However, sites on which the operator act can not be contracted explicitly. Thus,

the resulting tensor network will look like the example shown in figure 3.7b, where a two-site operator is measured.

Correlation functions can also be calculated efficiently using matrix product states in the mixed-canonical form. Consider the correlation  $\langle \psi | \hat{O}^{[n]} \hat{Q}^{[n+k]} | \psi \rangle$ , whose corresponding tensor network is shown in figure 3.8. Bringing the MPS into a mixed-canonical form

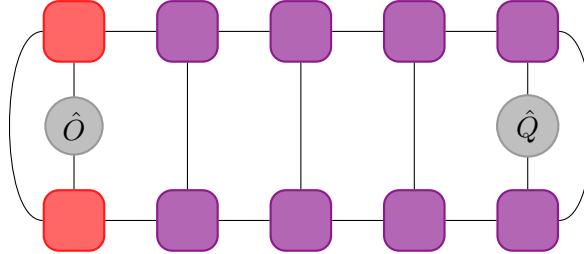


FIGURE 3.8: *Tensor network for the correlation function  $\langle \psi | \hat{O}^{[n]} \hat{Q}^{[n+4]} | \psi \rangle$ . The outer parts of the network are implicitly contracted to identities following the mixed-canonical form of the MPS.*

with the central tensor at site  $m$  results in the outer parts of the network being explicitly contracted as identities. However, the otherwise right-normalized tensors between the two operators can not be contracted as identities, as the right-most operator breaks the canonical form. Thus, one has to contract a network of length  $k + 1$ , which is done most efficiently following the optimal bracketing described in Appendix A. The complexity of the entire contraction is of order  $\mathcal{O}(kD^3d)$ .

### 3.5.1 Correlation length

Although the MPS formalism excels at describing one dimensional systems, it struggles with describing long ranged correlations, due to how it is constructed. Consider a general MPS in no particular canonical form. The *transfer operator* is defined as

$$\hat{E}^{[n]} = \sum_{\alpha_{n-1}, \alpha'_{n-1}} \sum_{\alpha_n, \alpha'_n} E_{(\alpha_{n-1} \alpha'_{n-1}), (\alpha_n \alpha'_n)}^{[n]} (|\alpha_{n-1}\rangle \langle \alpha'_{n-1}|) (|\alpha_n\rangle \langle \alpha'_n|) , \quad (3.24)$$

where  $E^{[n]}$  are the matrix elements of the transfer operator

$$E^{[n]} = \sum_{j_n} M^{[n]j_n*} \otimes M^{[n]j_n} . \quad (3.25)$$

The transfer operator is essentially a complete, positive map from operators defined on a block of the lattice of length  $n - 1$  to a block of length  $n$ , such that

$$\{|\alpha_{n-1}\rangle \langle \alpha'_{n-1}|\} \rightarrow \{|\alpha_n\rangle \langle \alpha'_n|\} . \quad (3.26)$$

The most important property of the transfer matrix is that all its eigenvalues  $|\lambda_k| \leq 1$  if the corresponding state is either left- or right-normalized [21].

Generalizing the transfer operator matrix elements (3.25) to contraction with an operator  $\hat{O}$  at site  $n$  yields

$$E_O^{[n]} = \sum_{j_n, j'_n} O^{j_n, j'_n} M^{[n]j_n*} \otimes M^{[n]j'_n} . \quad (3.27)$$

Thereby, correlation functions can be expressed using only the transfer matrices of eq. (3.25) and (3.27). Assuming a translational invariant transfer matrices for a left-normalized state,

whereby the correlation can be expressed as

$$\begin{aligned}
\langle \psi | \hat{O}^{[i]} \hat{O}^{[j]} | \psi \rangle &= \text{Tr } E^{[1]} \dots E^{[i-1]} E_O^{[i]} E^{[i+1]} \dots E^{[j-1]} E_O^{[j]} E^{[j+1]} \dots E^{[L]} \\
&= \text{Tr } E_O^{[i]} E^{j-i-1} E_O^{[j]} E^{L-j+i-1} \\
&= \sum_{l,k} \langle l | E_O^{[i]} | k \rangle \lambda_k^{j-i-1} \langle k | E_O^{[j]} | l \rangle \lambda_l^{L-j+i-1} \\
&= \sum_k \langle 1 | E_O^{[i]} | k \rangle \lambda_k^{j-i-1} \langle k | E_O^{[j]} | 1 \rangle \quad (\text{for } L \rightarrow \infty)
\end{aligned} \tag{3.28}$$

where  $|k\rangle$  and  $|l\rangle$  are eigenstates of the transfer operator with eigenvalues  $\lambda$ . Since  $|\lambda_k| \leq 1$ , only the leading eigenvalue  $\lambda_1 = 1$  remains as  $L \rightarrow \infty$ . Defining the distance between two sites as  $r = |j - i - 1|$  and the correlation length as  $\xi_k = -1/\ln \lambda_k$ , the correlation function can be written as

$$\frac{\langle \psi | \hat{O}^{[i]} \hat{O}^{[j]} | \psi \rangle}{\langle \psi | \psi \rangle} = c_1 + \sum_{k=2} c_k e^{-r/\xi_k}, \tag{3.29}$$

where  $c_k = \langle 1 | E_O^{[i]} | k \rangle \langle k | E_O^{[j]} | 1 \rangle$ . [21]

According to eq. (3.29), correlation functions are given by a linear combination of ex-

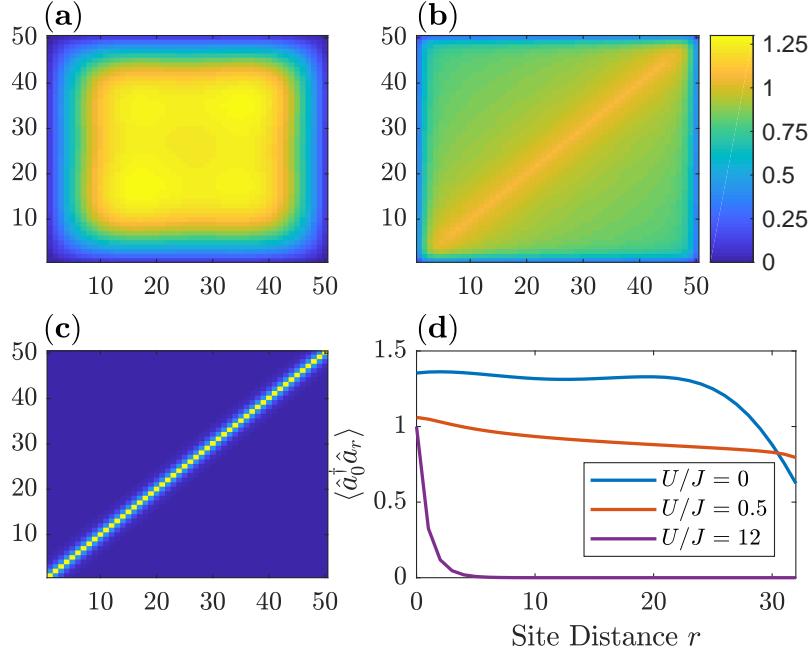


FIGURE 3.9: (a-c) Density matrices of a 50 site system for  $U/J = 0, 0.5, 12$ . (d) entries of the 15th row of the density matrices plotted as a function of distance from the diagonal,  $r$ .

ponential functions in the MPS description. Therefore, matrix product states struggle describing long range correlations, such as the superfluid single-particle correlations. While single-particle correlations decay exponentially for the Mott-Insulator, it decays following a power-law for superfluids

$$\langle \hat{a}_i^\dagger \hat{a}_j \rangle \sim |i - j|^{-K_b/2}, \tag{3.30}$$

where  $K_b$  is the Tomonaga-Luttinger parameter [69]. For short distances eq. (3.29) is able to accurately approximate a power-law, however, only the slowest exponential decay will survive as distances grow larger. Hence, the correlation turns into a pure exponential decay with  $\xi = -1/\ln \lambda$ , where  $\lambda$  is the largest eigenvalue of  $\hat{E}$  contributing to the correlation.

To demonstrate the correlation properties of matrix product states, consider figure 3.9 depicting the density matrix with entries  $\rho_{i,j} = \langle \psi | \hat{a}_i^\dagger \hat{a}_j | \psi \rangle$  for the ground state of a Bose-Hubbard system with  $L = 50$  sites and unit occupancy. The ground state was found at various fractions of  $U/J$  using the DMRG algorithm detailed in Section 4.1.2. Figures 3.9(a) and 3.9(c) show the density matrix plotted for the superfluid and the Mott-Insulator limit respectively. In the superfluid limit, long-range correlations are present, which is seen by large off-diagonal elements. However, since all correlations decay exponentially in the MPS description, the algorithm has difficulty approximating the very long range correlations of the system. This is visualized in figure 3.9(d), where the correlation function is plotted against distance from the diagonal. The superfluid graph has a hump on it, which is an artifact of the MPS description and DMRG algorithm. Due to the very long correlation length, the graph should be almost flat except for the rapid drop in the end, which is due to the closed boundary conditions.

In the Mott-Insulator limit no interaction takes place between sites, and the correlation length is zero. Hence, the correlation matrix contains only diagonal elements of equal magnitude. Figure 3.9(c) shows some off-diagonal elements of non-zero magnitude, since the system is not a pure Mott-Insulator at  $U/J = 12$ . Nevertheless, the correlations are well described by only a single exponential function, as seen by the figure.

Finally, figure 3.9(b) illustrates the density matrix for a system with  $U/J = 0.5$ , which is primarily a superfluid. The particles are not completely de-localized, which is noticeable from the well defined diagonal. Plotting the correlation function reveals the best approximation being a power-law, confirming the system is indeed a superfluid.



## Chapter 4

# Tensor Network Algorithms

Although matrix product states enables analytical treatment of certain classes of quantum states [70, 71], their real power becomes evident when performing numerical computations. These computations require algorithms capable of exploiting the properties of the matrix product states. The Density-Matrix Renormalization Group (DMRG) was developed as the most powerful numerical method to study one-dimensional quantum lattice systems [62, 63]. The DMRG algorithm is a variational ground state search method, which owes its effectiveness to the entanglement area law scaling of most ground states in one-dimensional lattice systems. Therefore, the method only has to search a tiny corner of the otherwise exponentially large Hilbert space [22].

Years later the connection between the DMRG method and matrix product states was made, as these two approaches share many features [72, 73]. Formulating the DMRG method in the MPS language made several extensions to the algorithm possible, which would otherwise have been very difficult to express within the DMRG framework [21]. Among these extensions one finds the tDMRG algorithm [20, 74, 75], which borrows elements from DMRG in order to perform time-evolution of a matrix product state. Due to the light-cone-like propagation of information within the quantum system [60, 23], a state initially following an area law will continue to do so for longer times [23, 24], although the prefactor of the area law scales exponentially with the system size [76]. Therefore, tensor-network methods are capable of efficiently simulating out-of-equilibrium dynamics for short time-scales [56].

### 4.1 Variational Ground State Search

For large systems exact diagonalization of the Hamiltonian is impossible, whereby one must resort to variational methods in order to find ground states. A variational search involves finding the state,  $|\psi\rangle$ , which minimizes

$$E = \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi|\psi\rangle}. \quad (4.1)$$

The Hamiltonian must be expressed as a matrix product operator to be applied to the state. An example is given in Appendix B, where the Bose-Hubbard Hamiltonian is formulated through tensors. Furthermore, in a variational search the state is repeatedly modified until the expectation value of the Hamiltonian is minimized. Therefore, the tensor network corresponding to eq. 4.1 must be contracted efficiently in order for the search to be viable.

#### 4.1.1 Efficient Application of a Hamiltonian to an MPS

Section 3.4 detailed the most effective way of applying a matrix product operator to a state, however, even further reductions to the computational cost are needed, as the overlap with the Hamiltonian is evaluated many times during the ground state search. The DMRG

variational search algorithm updates only a single tensor of the state at a time, whereby most of the tensor network describing the operator overlap remains constant. Hence, most of the network can be stored and used for multiple calculations, which is one of the key properties of the DMRG algorithm.

Consider a matrix product state in a mixed-canonical form

$$\begin{aligned} |\psi\rangle &= \sum_j A^{j_1} \dots A^{j_{n-1}} \Psi^{j_n} B^{j_{n+1}} \dots B^{j_L} |\mathbf{j}\rangle \\ &= \sum_{\alpha_{n-1}, \alpha_n} |\alpha_{n-1}\rangle_A \Psi_{\alpha_{n-1}, \alpha_n}^{j_n} |\alpha_n\rangle_B , \end{aligned} \quad (4.2)$$

where  $\Psi^{j_n}$  is the central site, and  $|\alpha_{n-1}\rangle_A$  and  $|\alpha_n\rangle_B$  are block states introduced in eq. (3.16) and (3.17) respectively. In the basis  $\{|\alpha_{n-1}\rangle, |j_n\rangle, |\alpha_n\rangle\}$  the individual matrix elements of the Hamiltonian can be expressed as

$$\langle \alpha_{n-1} j_n \alpha_n | \hat{H} | \alpha'_{n-1} j'_n \alpha'_n \rangle = \sum_{\mathbf{j}, \mathbf{j}'} W^{j_1, j'_1} \dots W^{j_L, j'_L} \langle \alpha_{n-1} j_n \alpha_n | \mathbf{j} \rangle \langle \mathbf{j}' | \alpha'_{n-1} j'_n \alpha'_n \rangle .$$

Since the basis of local states  $\{|\mathbf{j}\rangle\}$  shares a state with the basis  $\{|\alpha_{n-1}\rangle, |j_n\rangle, |\alpha_n\rangle\}$ , the above expression can be re-written using  $\sum_{\mathbf{j}} \langle j_n | \mathbf{j} \rangle = \sum_{\mathbf{j}^*} \langle \mathbf{j}^* \rangle$ , where "\*" means "excluding  $j_n$ ". Thus,

$$\begin{aligned} &\langle \alpha_{n-1} j_n \alpha_n | \hat{H} | \alpha'_{n-1} j'_n \alpha'_n \rangle = \\ &= \sum_{\mathbf{j}^*, \mathbf{j}'^*} W^{j_1, j'_1} \dots W^{j_n, j'_n} \dots W^{j_L, j'_L} \\ &\quad \times \langle \alpha_{n-1} | j_1 \dots j_{n-1} \rangle \langle \alpha_n | j_{n+1} \dots j_L \rangle \langle j'_1 \dots j'_{n-1} | \alpha'_{n-1} \rangle \langle j'_{n+1} \dots j'_L | \alpha'_n \rangle \\ &= \sum_{\mathbf{j}^*, \mathbf{j}'^*} W^{j_1, j'_1} \dots W^{j_n, j'_n} \dots W^{j_L, j'_L} \\ &\quad \times (A^{j_1} \dots A^{j_{n-1}})_{1, \alpha_{n-1}}^* (B^{j_{n+1}} \dots B^{j_L})_{\alpha_n, 1}^* (A^{j'_1} \dots A^{j'_{n-1}})_{1, \alpha'_{n-1}} (B^{j'_{n+1}} \dots B^{j'_L})_{\alpha'_n, 1} \\ &= \sum_{\alpha_n, \beta_n, \alpha'_n} \left( \sum_{j_1, j'_1} A_{1, \alpha_1}^{j_1*} W_{1, \beta_1}^{j_1, j'_1} A_{1, \alpha'_1}^{j'_1} \right) \left( \sum_{j_2, j'_2} A_{\alpha_1, \alpha_2}^{j_2*} W_{\beta_1, \beta_2}^{j_2, j'_2} A_{\alpha'_1, \alpha'_2}^{j'_2} \right) \dots W_{\beta_{n-1}, \beta_n}^{j_n, j'_n} \\ &\quad \times \left( \sum_{j_{n+1}, j'_{n+1}} B_{\alpha_n, \alpha_{n+1}}^{j_{n+1}*} W_{\beta_n, \beta_{n+1}}^{j_{n+1}, j'_{n+1}} B_{\alpha'_n, \alpha'_{n+1}}^{j'_{n+1}} \right) \left( \sum_{j_L, j'_L} B_{\alpha_{L-1}, 1}^{j_L*} W_{\beta_{L-1}, 1}^{j_L, j'_L} B_{\alpha'_{L-1}, 1}^{j'_L} \right) . \end{aligned} \quad (4.3)$$

While this expression may seem terrible complicated due to all the indices, it is actually rather easy to understand. First, the matrix element is written excluding the local basis state  $|j_n\rangle$ . Next, the Hamilton MPO is projected into the block states of A,  $|\alpha_{n-1}\rangle_A$ , and B,  $|\alpha_n\rangle_B$ . Finally, the matrices are grouped according to their expansion in the local basis. Working with the above expression appears cumbersome, but it is merely a decoupling of the system into three distinct parts, which can be seen in figure 4.1. Since both the left and right side of the network is connected, one can contract these parts into two separate

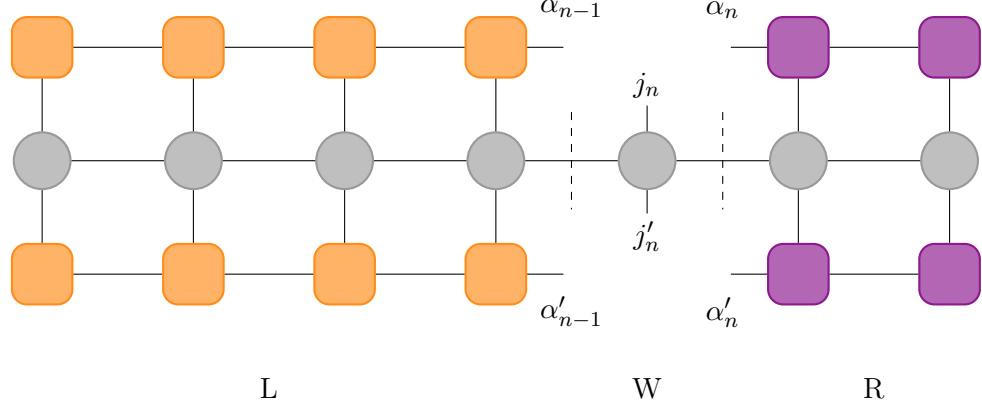


FIGURE 4.1: Representation of the matrix element  $\langle \alpha_{n-1} j_n \alpha_n | \hat{H} | \alpha'_{n-1} j'_n \alpha'_n \rangle$  as a tensor network. The network can be factorize into three distinct parts: The left ( $L$ ) and right ( $R$ ) environments consisting of the contracted network to the left and right of the operator tensor,  $W^{[n]}$ , respectively.

tensors  $L$  and  $R$  called *environments*:

$$L_{\alpha_n, \beta_n, \alpha'_n} = \sum_{\{\alpha_i \beta_i \alpha'_i\}} \left( \sum_{j_1, j'_1} A_{1, \alpha_1}^{j_1*} W_{1, \beta_1}^{j_1, j'_1} A_{1, \alpha'_1}^{j'_1} \right) \dots \left( \sum_{j_n, j'_n} A_{\alpha_{n-1}, \alpha_n}^{j_n*} W_{\beta_{n-1}, \beta_n}^{j_n, j'_n} A_{\alpha'_{n-1}, \alpha'_n}^{j'_n} \right) \quad (4.4)$$

$$R_{\alpha_n, \beta_n, \alpha'_n} = \sum_{\{\alpha_i \beta_i \alpha'_i\}} \left( \sum_{j_{n+1}, j'_{n+1}} B_{\alpha_n, \alpha_{n+1}}^{j_{n+1}*} W_{\beta_n, \beta_{n+1}}^{j_{n+1}, j'_{n+1}} B_{\alpha'_n, \alpha'_{n+1}}^{j'_{n+1}} \right) \dots \times \left( \sum_{j_L, j'_L} B_{\alpha_{L-1}, 1}^{j_L*} W_{\beta_{L-1}, 1}^{j_L, j'_L} B_{\alpha'_{L-1}, 1}^{j'_L} \right) \quad (4.5)$$

From these contractions, the tripartite structure of the Hamiltonian matrix element, as seen in figure 4.1, can be written in a compact way

$$\langle \alpha_{n-1} j_n \alpha_n | \hat{H} | \alpha'_{n-1} j'_n \alpha'_n \rangle = \sum_{\beta_{n-1}, \beta_n} L_{\alpha_{n-1}, \beta_{n-1}, \alpha'_{n-1}} W_{\beta_{n-1}, \beta_n}^{j_n, j'_n} R_{\alpha_n, \beta_n, \alpha'_n} . \quad (4.6)$$

Finally, applying the Hamiltonian in the  $\{|\alpha_{n-1}\rangle, |j_n\rangle, |\alpha_n\rangle\}$  basis to the MPS of eq. (4.2) yields [21]

$$\hat{H} |\psi\rangle = \sum_{\beta_{n-1}, \beta_n} \sum_{\alpha'_{n-1}, j'_n, \alpha'_n} L_{\alpha_{n-1}, \beta_{n-1}, \alpha'_{n-1}} W_{\beta_{n-1}, \beta_n}^{j_n, j'_n} R_{\alpha_n, \beta_n, \alpha'_n} \Psi_{\alpha'_{n-1}, \alpha'_n}^{j'_n} |\alpha_{n-1}\rangle_A |j_n\rangle | \alpha_n\rangle_B . \quad (4.7)$$

Expressing  $\hat{H}$  in the basis of the block states exposes the central site of the MPS such that it can be varied in a ground state search. Evaluating  $\hat{H} |\psi\rangle$  must be done many times during a variational search of the ground state, hence this operation must be executed as fast as possible. Examining eq. (4.7) one will notice, that while the boundaries of  $L$  and  $R$  will change depending on which site is being optimized, the bulk of the two tensors remain constant through a lot of the calculations. Instead of calculating  $L$  and  $R$  from eq. (4.4) and (4.5) for every evaluation of eq. (4.7), one can iteratively build them, since they only change by one column of the network at a time. Thus, a large number of computations can be reused.

Consider the construction of the tensor  $L^{[i]}$ , which can be built iteratively from the left by contracting the previous left-tensor  $L^{[i-1]}$  with the  $i$ 'th column of the network consisting of  $A^{[i]}$ ,  $W^{[i]}$  and  $A^{[i]\dagger}$

$$L_{\alpha_i, \beta_i, \alpha'_i}^{[i]} = \sum_{\substack{j_i, j'_i \\ \alpha_{i-1}, \beta_{i-1}, \alpha'_{i-1}}} W_{\beta_{i-1}, \beta_i}^{[i]j_i, j'_i} \left( A^{[i]j_i\dagger} \right)_{\alpha_i, \alpha_{i-1}} L_{\alpha_{i-1}, \beta_{i-1}, \alpha'_{i-1}}^{[i-1]} A_{\alpha'_{i-1}, \alpha'_i}^{[i]j'_i}. \quad (4.8)$$

The iterative update of  $L^{[i]}$  can be seen illustrated in figure 4.2. The square-bracket notation has been re-introduced to keep track of the tensors relation to the physical sites. In order to remain consistent with notation, the dummy scalars  $L_{\alpha_0, \beta_0, \alpha'_0}^{[0]} = 1 = \alpha_0, \beta_0, \alpha'_0$  have been introduced.

It is important to store every iteration of  $L^{[i]}$ , since  $L$  will grow and shrink constantly

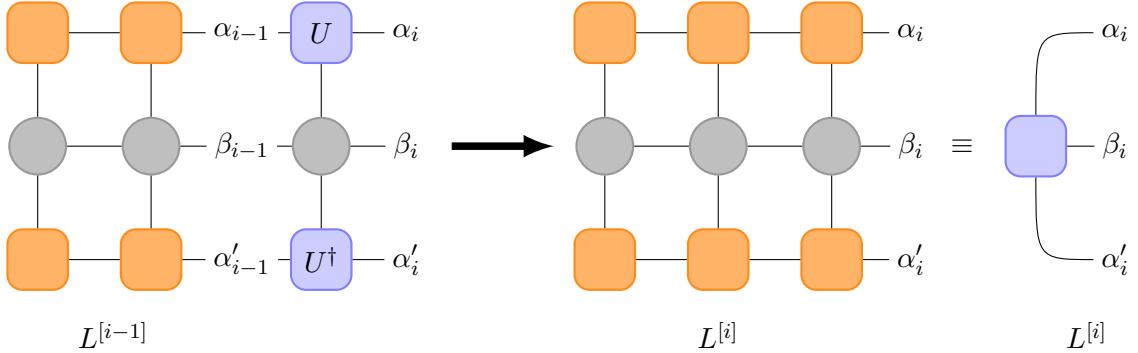


FIGURE 4.2: Iterative update from  $L^{[i-1]}$  to  $L^{[i]}$  through a contraction of  $L^{[i-1]}$  with  $A^{[i]}$ ,  $W^{[i]}$  and  $A^{[i]\dagger}$ . The result is a tensor with three horizontal legs.

throughout the variational search of the ground state, whereby every iteration of  $L^{[i]}$  will be used multiple times. The same applies when building the right environment,  $R$ . Here one starts from the right and moves left when iteratively contracting the tensor. By applying optimal bracketing, the computational cost of updating the environments scales as  $\mathcal{O}(dD^3D_W)$ .

#### 4.1.2 Iterative Ground State Search and the DMRG Algorithm

In order to find the ground state of the system one can introduce a Lagrangian multiplier,  $\lambda$ , and extremize

$$\langle \psi | \hat{H} | \psi \rangle - \lambda \langle \psi | \psi \rangle, \quad (4.9)$$

whereby the desired ground state,  $|\psi\rangle$ , and ground state energy,  $\lambda^0$ , will be reached.

Trying to optimize an entire MPS at once is a highly non-linear problem involving an extremely large number of variables. However, the problem can be linearised by only considering the variables of a single tensor (site) at a time, while keeping the rest of the MPS constant. By varying just a single tensor at a time, one will continuously find states lower in energy, until convergence is reached. However, this procedure is very prone to getting stuck in a local extrema. To circumvent this, one can consider two sites at a time and optimize with regards to a two-site tensor, created by momentarily merging the two sites [63].

Consider the variation of the tensors  $M^{[n]}$  and  $M^{[n+1]}$ . Expressing the minimization problem

of eq. (4.9) in terms of the left and right environments (as done in eq. (4.7)) yields

$$\begin{aligned} \langle \psi | \hat{H} | \psi \rangle = & \sum_{j_n, j'_n} \sum_{\alpha'_{n-1}, \alpha'_n, \alpha'_{n+1}} \sum_{\alpha_{n-1}, \alpha_n, \alpha_{n+1}} \sum_{\beta_{n-1}, \beta_n, \beta_{n+1}} L_{\alpha_{n-1}, \beta_{n-1}, \alpha'_{n-1}}^{[n-1]} W_{\beta_{n_1}, \beta_n}^{j_n, j'_n} W_{\beta_n, \beta_{n+1}}^{j_{n+1}, j'_{n+1}} \\ & \times R_{\alpha_{n+1}, \beta_{n+1}, \alpha'_{n+1}}^{[n+2]} M_{\alpha_{n-1}, \alpha_n}^{j_n} M_{\alpha'_{n-1}, \alpha'_n}^{j'_n *} M_{\alpha_n, \alpha_{n+1}}^{j_{n+1}} M_{\alpha'_n, \alpha'_{n+1}}^{j'_{n+1} *} \end{aligned} \quad (4.10)$$

with the overlap

$$\begin{aligned} \langle \psi | \psi \rangle = & \sum_{j_n, j_{n+1}} \sum_{\alpha'_{n-1}} \sum_{\alpha_{n-1}, \alpha_n, \alpha_{n+1}} \Psi_{\alpha_{n-1}, \alpha'_{n-1}}^A M_{\alpha_{n-1}, \alpha_n}^{j_n} M_{\alpha'_{n-1}, \alpha'_n}^{j'_n *} \\ & \times M_{\alpha_n, \alpha_{n+1}}^{j_{n+1}} M_{\alpha'_n, \alpha'_{n+1}}^{j'_{n+1} *} \Psi_{\alpha_{n+1}, \alpha'_{n+1}}^B, \end{aligned} \quad (4.11)$$

where the Hamiltonian from eq. (4.7) has been re-ordered to accommodate examining two sites,  $n$  and  $n + 1$ , at a time, and

$$\Psi_{\alpha_{n-1}, \alpha'_{n-1}}^A = \sum_{j_1, \dots, j_{n-1}} \left( M^{j_{n-1}\dagger} \dots M^{j_1\dagger} M^{j_1} \dots M^{j_{n-1}} \right)_{\alpha_{n-1}, \alpha'_{n-1}} \quad (4.12)$$

$$\Psi_{\alpha_{n+1}, \alpha'_{n+1}}^B = \sum_{j_{n+2}, \dots, j_L} \left( M^{j_{n+2}} \dots M^{j_L} M^{j_L\dagger} \dots M^{j_{n+2}\dagger} \right)_{\alpha'_{n+1}, \alpha_{n+1}}. \quad (4.13)$$

Further simplifications can be made for mixed-canonical forms, if sites 1 through  $n - 1$  are left-normalized, and sites  $n + 2$  through  $L$  are right-normalized, whereby

$$\Psi_{\alpha_{n-1}, \alpha'_{n-1}}^A = \delta_{\alpha_{n-1}, \alpha'_{n-1}}, \quad \Psi_{\alpha_{n+1}, \alpha'_{n+1}}^B = \delta_{\alpha_{n+1}, \alpha'_{n+1}}. \quad (4.14)$$

Finding the extremum of eq. (4.9) with respect to  $M_{\alpha'_{n-1}, \alpha'_n}^{[n]j'_n *} M_{\alpha'_n, \alpha'_{n+1}}^{[n+1]j'_{n+1} *}$  is done through the following sequence:

### Two-site update for iterative ground state search

1. **Merge:** Contract the two matrices  $M^{[n]}$  and  $M^{[n+1]}$  over the bond  $\alpha_n$  creating a two-site tensor

$$\Theta_{\alpha_{n-1}, \alpha_{n+1}}^{j_n, j_{n+1}} = \sum_{\alpha_n} M_{\alpha_{n-1}, \alpha_n}^{[n]j_n} M_{\alpha_n, \alpha_{n+1}}^{[n+1]j_{n+1}} \quad (4.15)$$

2. **Solve eigenproblem:** This yields an eigenvalue problem, which can be seen by reshaping

$$\begin{aligned} H_{(\alpha_{n-1} j_n j_{n+1}, \alpha_{n+1}), (\alpha'_{n-1} j'_n j'_{n+1}, \alpha'_{n+1})} = & \\ = & \sum_{\beta_{n-1}, \beta_n} L_{\alpha_{n-1}, \beta_{n-1}, \alpha'_{n-1}}^{[n-1]} W_{\beta_{n_1}, \beta_n}^{[n]j_n, j'_n} W_{\beta_n, \beta_{n+1}}^{[n+1]j_{n+1}, j'_{n+1}} R_{\alpha_{n+1}, \beta_{n+1}, \alpha'_{n+1}}^{[n+2]} \end{aligned} \quad (4.16)$$

and

$$v_{\alpha_{n-1} j_n j_{n+1} \alpha_{n+1}} = \Theta_{\alpha_{n-1}, \alpha_{n+1}}^{j_n, j_{n+1}} \quad (4.17)$$

such that

$$Hv - \lambda v = 0. \quad (4.18)$$

Solving eq. (4.18) for the lowest eigenvalue  $\lambda_0$  yields  $v_{\alpha_{n-1}j_n j_{n+1} \alpha_{n+1}}^0$ , which can be reshaped back to the now optimized two-site tensor,  $\tilde{\Theta}_{\alpha_{n-1}, \alpha_{n+1}}^{j_n, j_{n+1}}$ .

3. **Unmerge:** Reshape the updated  $\tilde{\Theta}_{\alpha_{n-1}, \alpha_{n+1}}^{j_n, j_{n+1}}$  to a matrix and perform an SVD yielding

$$\tilde{\Theta}_{(j_n \alpha_{n-1}), (j_{n+1} \alpha_{n+1})} = \sum_{\alpha_n} U_{\alpha_{n-1}, \alpha_n}^{j_n} S_{\alpha_n, \alpha_n} (V^\dagger)_{\alpha_n, \alpha_{n+1}}^{j_{n+1}}. \quad (4.19)$$

This causes the bond dimension to increase  $D \rightarrow dD$ , which must be truncated by keeping only the  $D$  largest singular values of  $S$ .

4. **Update environments:** The last step depends on which direction, one is iterating through the chain. Here, the left- and right-normalization of  $U$  and  $V^\dagger$  is used to update the environments.

*Going right:* Update the left environment

$$\tilde{L}_{\alpha_n, \beta_n, \alpha'_n}^{[n]} = \sum_{\substack{j_n, j'_n \\ \alpha_{n-1}, \beta_{n-1}, \alpha'_{n-1}}} L_{\alpha_{n-1}, \beta_{n-1}, \alpha'_{n-1}}^{[n-1]} U_{\alpha_{n-1}, \alpha_n}^{j_n} W_{\beta_{n-1}, \beta_n}^{[n]j_n, j'_n} U_{\alpha'_{n-1}, \alpha'_n}^{j'_n*}, \quad (4.20)$$

and build the matrix of the right site

$$\tilde{M}_{\alpha_n, \alpha_{n+1}}^{[n+1]j_{n+1}} = \sum_{\alpha_n} S_{\alpha_n, \alpha_n} (V^\dagger)_{\alpha_n, \alpha_{n+1}}^{j_{n+1}}. \quad (4.21)$$

*Going left:* Update the right environment

$$\tilde{R}_{\alpha_n, \beta_n, \alpha'_n}^{[n+1]} = \sum_{\substack{j_{n+1}, j'_{n+1} \\ \alpha_{n+1}, \beta_{n+1}, \alpha'_{n+1}}} R_{\alpha_{n+1}, \beta_{n+1}, \alpha'_{n+1}}^{[n+2]} \left(V^\dagger\right)_{\alpha_n, \alpha_{n+1}}^{j_{n+1}} W_{\beta_n, \beta_{n+1}}^{[n+1]j_{n+1}, j'_{n+1}} \left(V^\dagger\right)_{\alpha_n, \alpha_{n+1}}^{j_{n+1}*}, \quad (4.22)$$

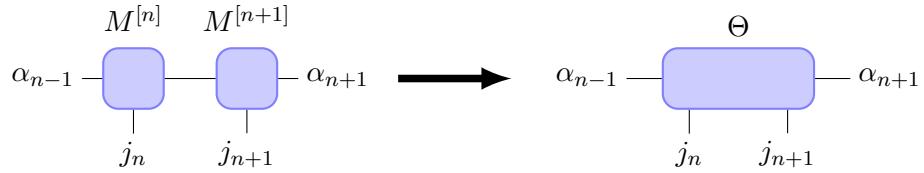
and build the matrix of the left site

$$\tilde{M}_{\alpha_{n-1}, \alpha_n}^{[n]j_n} = \sum_{\alpha_n} U_{\alpha_{n-1}, \alpha_n}^{j_n} S_{\alpha_n, \alpha_n}. \quad (4.23)$$

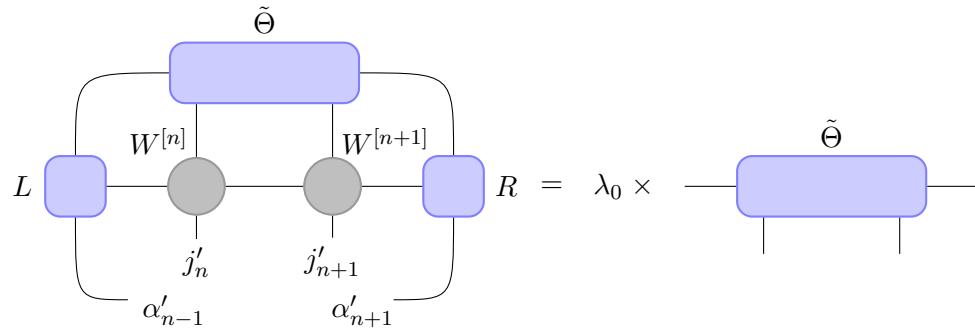
This concludes the two-site update sequence, which is illustrated in figure 4.3. After performing the sequence, one can move *one site* to either left or right, depending on which direction one is iterating.

Some comments regarding the sequence are in order: The matrices of the eigenvalue problem have dimensions  $(d^2 D^2 \times d^2 D^2)$ , which is generally too much for exact diagonalization, however, since only the lowest eigenvalue is of interest, one can use an iterative eigensolver [77]. Furthermore, if the MPS is not in the proper mixed-canonical form, the eigenvalue problem turns into a generalized eigenvalue problem, which can be numerically quite demanding. Thus, updating the left and right environments in step (4) is necessary, since it leads to great simplifications in step (2). Lastly, one could also consider just a single site when updating, however, this method is very prone to getting stuck [78]. By updating two sites at once, one actually optimizes the bond between them. Hence, after updating the two sites, one must only iterate a single site. Optimization is done with regards to the current configuration, whereby it depends on previous updates. To compensate for this, one must sweep through the entire system multiple times, which leads to the following algorithm for iterative ground state search, which follows the structure of the DMRG algorithm:

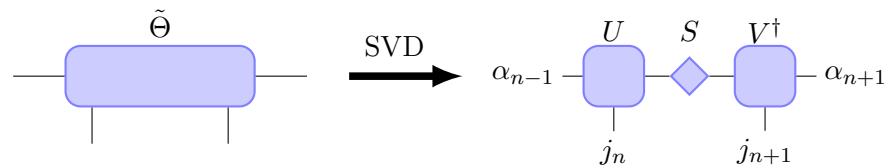
(1) Merge:



(2) Solve eigenproblem:



(3) Unmerge:



(4) Update environments:

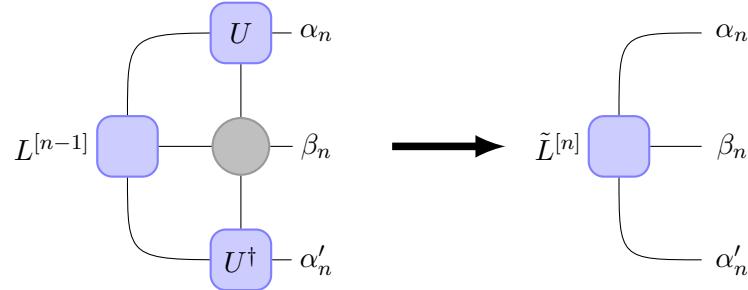


FIGURE 4.3: Diagrammatic representation of the two-site update sequence for iterative ground state search. Step (2) optimizes with regard to two sites, but only a single site is updated to avoid getting stuck. In step (4), when iterating left-to-right in the lattice, the left environment,  $L^{[n]}$ , is updated. The right environment,  $R^{[n]}$ , is updated when moving right-to-left.

---

**Algorithm 1** Iterative ground state search (DMRG)

---

```

Choose  $|\psi\rangle$  right-normalized.
Calculate tensor  $R^{[i]}$  iteratively for  $i = L \dots 1$ .
while Stopping criteria not met do
    for  $n = 1 \dots L - 1$  do                                 $\triangleright$  Left sweep
        Perform two-site update on  $M^{[n]}$  and  $M^{[n+1]}$ .
        Update  $L^{[n]}$  according to eq. (4.20)
    end for
    for  $n = L - 1 \dots 1$  do                             $\triangleright$  Right sweep
        Perform two-site update on  $M^{[n]}$  and  $M^{[n+1]}$ .
        Update  $R^{[n]}$  according to eq. (4.22)
    end for
end while

```

---

Examples of the DMRG algorithm can be found in section 3.5.1, where the ground state of the Bose-Hubbard model is calculated for a lattice with 50 sites. The ground state is used to compute correlation functions of the system, and the example illustrates how matrix product states and the DMRG algorithm struggle when the correlation lengths grow very large. Another example can be found in appendix C, where the condensate fraction for various system sizes is computed in an attempt to determine the point of the superfluid to Mott-Insulator phase transition. Furthermore, the results obtained through the DMRG are compared with those of exact diagonalization.

## 4.2 Time Evolution of Matrix Product States

Several algorithms for time evolving matrix product states exist, however, they all origin from the same ideas proposed in [20, 74]. The most widely used of these algorithms is the tDMRG algorithm [75], which gets its name from its similarity with the ground state search algorithm described in section 4.1.2. The tDMRG algorithm has been utilized in several instances to simulate the dynamics of one-dimensional systems [79, 80, 81], and it has even been previously used in conjunction with the CRAB algorithm to perform optimal control of the superfluid to Mott-Insulator phase transition [82, 19].

Although the standard algorithms for time evolution are quite efficient, further improvements can be made to the algorithms by tailoring them to the problem at hand. Therefore, a modified version of the tDMRG algorithm is proposed in Section 4.2.2, which directly utilizes the properties of the Bose-Hubbard Hamiltonian.

### 4.2.1 The tDMRG Algorithm

Consider the time evolution of a quantum state

$$|\psi(t)\rangle = \hat{\mathcal{U}}(t) |\psi(0)\rangle , \quad (4.24)$$

where  $\hat{\mathcal{U}}(t) = e^{-i\hat{H}t}$  is the time evolution operator. Time evolution of a matrix product states is done in a manner similar to that of ground state search, as the bonds between the tensors are evolved rather than the tensors themselves [20, 74]. Thus, the time evolution operator must be decomposed into two-site tensors. The simplest realization of this is achieved by considering a Hamiltonian containing only nearest-neighbour interactions. Assume the Hamiltonian is a sum of two-site operators of the form  $\hat{H} = \sum_n \hat{h}^{[n,n+1]}$ . One

can decompose this into a sum over even and odd bonds [74]

$$\hat{H} = \hat{H}_{\text{odd}} + \hat{H}_{\text{even}} = \sum_{n \text{ odd}} \hat{h}^{[n,n+1]} + \sum_{n \text{ even}} \hat{h}^{[n,n+1]}. \quad (4.25)$$

Exponentiating the Hamiltonian is non-trivial due to the non-commutativity of the operators

$$[\hat{h}_{\text{odd}}^{[n,n+1]}, \hat{h}_{\text{even}}^{[n,n+1]}] \neq 0. \quad (4.26)$$

Considering a small time slice,  $\Delta t$ , the exponentiation can be achieved through the Trotter-Suzuki expansion [83]. To first order this reads

$$e^{-i\hat{H}\Delta t} = e^{-i\hat{H}_{\text{odd}}\Delta t} e^{-i\hat{H}_{\text{even}}\Delta t} + O(\Delta t^2), \quad (4.27)$$

where the error is due to the non-commutativity of the bond Hamiltonians. Thus, the time evolution operator can be expressed as the product

$$\hat{\mathcal{U}}(\Delta t) \approx \left( \prod_{n \text{ odd}} \hat{\mathcal{U}}^{[n,n+1]}(\Delta t) \right) \left( \prod_{n \text{ even}} \hat{\mathcal{U}}^{[n,n+1]}(\Delta t) \right), \quad (4.28)$$

where

$$\hat{\mathcal{U}}^{[n,n+1]}(\Delta t) = e^{-i\hat{h}^{[n,n+1]}\Delta t}. \quad (4.29)$$

The result is an MPO performing an infinitesimal time step on the odd bonds, and another MPO evolving the even bonds. An illustration of eq. (4.28) is shown in Figure 4.4. The

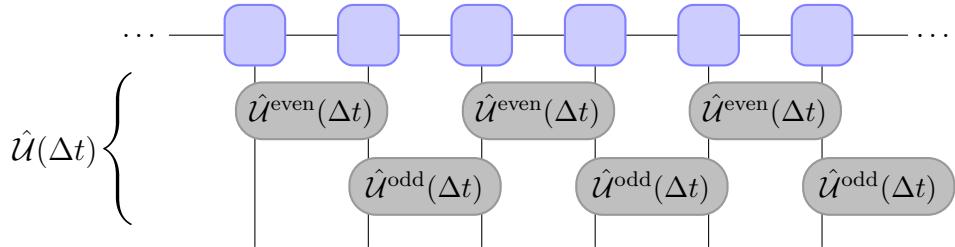


FIGURE 4.4: *Approximation of each time step  $\delta t$  using a Trotter-Suzuki decomposition, such that the time evolution operator is expressed as a product of unitary two-site operators.*

tDMRG algorithm describes the most efficient and accurate way of contracting the tensor network detailed in figure 4.4. The algorithm gets its name from its similarity with the DMRG algorithm detailed in section 4.1.2. In fact, the merge and unmerge procedure of the two algorithms is completely identical, whereby they only differ in the steps of applying the operator and proceeding to the next site. The following procedure details a time evolution step of the  $n$ 'th bond [21].

### Infinitesimal time-step update for tDMRG

1. **Merge:** Contract tensors  $M^{[n]}$  and  $M^{[n+1]}$  over the bond  $\alpha_n$  creating a two-site tensor  $\Theta^{j_n, j_{n+1}}$ .
2. **Apply unitary:** The two-site time evolution operator,  $\hat{\mathcal{U}}^{[n,n+1]}$ , is applied to  $\Theta^{j_n, j_{n+1}}$

$$\tilde{\Theta}_{\alpha_{n-1}, \alpha_{n+1}}^{j_n, j_{n+1}} = \sum_{j'_n, j'_{n+1}} \mathcal{U}^{j_n j_{n+1}, j'_n j'_{n+1}} \Phi_{\alpha_{n-1}, \alpha_{n+1}}^{j'_n, j'_{n+1}}. \quad (4.30)$$

3. **Unmerge:** Reshape  $\tilde{\Phi}_{\alpha_{n-1}, \alpha_{n+1}}^{j'_n, j'_{n+1}}$  to a matrix and decompose it through an SVD. Applying  $\hat{\mathcal{U}}^{[n, n+1]}$  causes an increase in bond dimension,  $D \rightarrow d^2 D$ , which must be truncated by keeping only the  $D$  largest singular values from the SVD.
4. **Progress:** Next, the center cite of the MPS must be shifted by two, in order to update the next even (odd) bond. This is achieved by merging tensors  $M^{[n+1]}$  and  $M^{[n+2]}$  and performing a second SVD, while reshaping the resulting  $U$ -matrices to left-normalised tensors to retain the canonical form. The product of the second SVD must be truncated as well, however, no loss of information will occur, as the Schmidt rank of the matrix  $S$  will be at most  $D$  following the first SVD.

Following the procedure above will leave the MPS in position for application of the unitary  $\hat{\mathcal{U}}^{[n+2, n+3]}$ . The efficiency of the tDMRG algorithm depends on the sequence in which bonds are evolved. A simple, yet effective way is iterating from left to right when evolving even bonds, while iterating right to left when evolving odd bonds. Thereby, the centered cite of the mixed-canonical form is moved continuously through the MPS, rather than having to be reset when reaching the end of the system.

It has been shown that the area law scaling of entanglement of an initial state will remain true even for longer times [23, 24]. Simultaneously, the pre-factor of the area-law is expected to grow exponentially [76]. Therefore, the maximum bond dimension,  $D$ , kept in each iteration of the tDMRG algorithm must be chosen sufficiently high in order to properly capture the relevant dynamics of the system. Alternatively, one can truncate the bond dimensions of the tensors according to some eigenstate-contribution threshold,  $\epsilon_t$ . While this will ensure a consistent description of the state throughout the simulation, the bond dimension may grow very large thereby causing a significant decrease in the efficiency of the algorithm. Thus, analyzing the evolution of the bond dimension is imperative for the success of the simulation when examining phenomena of high entanglement such as phase transitions.

Although the tDMRG algorithm is a very powerful algorithm, an even higher efficiency can be achieved when tailoring the algorithm to the system - in this case the Bose-Hubbard model. Thus, the Hamiltonian contains both nearest-neighbour and on-site terms. Furthermore, the Hamiltonian is time dependent, since the lattice depth is varied when driving a phase transition. The following algorithm is a modification of the tDMRG algorithm, which is tailored for conducting optimal control using the Bose-Hubbard Hamiltonian.

#### 4.2.2 Modified Time Evolution Algorithm for Bose-Hubbard Model

In the control problem treated in this thesis, a Bose-Hubbard system is dynamically manipulated in order to cross the phase transition between the superfluid and the Mott-Insulator. Thus, the Hamiltonian is time-dependent, whereby it must be exponentiated for every time-step to create the propagators required for the time-evolution. Furthermore, any optimization of the control sequence will the corresponding propagators to be re-calculated. Thus, the computational time of the propagators must be taken into account. Typically, operators are exponentiated through their series expansion, however, this is in general a rather expensive operation. Therefore, choosing control-Hamiltonians which are easy to exponentiate is often an advantage.

The phase of the Bose-Hubbard model depends on the ratio  $U/J$ , where  $U$  is the interaction matrix element (2.40), and  $J$  is the tunneling matrix element (2.37). Both parameters are dependent of the lattice depth, which therefore has been used as control function in previous studies [82, 19]. Exponentiating the tunneling Hamiltonian is non-trivial, however, the interaction Hamiltonian is diagonal, whereby the propagator can be computed by simply exponentiating the elements of the matrix. Thus, choosing the ratio  $U/J$  as the control

function and working in units of  $J$  significantly simplifies the calculation of propagators, as the tunneling propagator remains constant. Employing a second-order Suzuki-Trotter expansion

$$\exp(-i(\hat{H}_J + \hat{H}_U)\Delta t) = \exp(-i\hat{H}_U\Delta t/2) \exp(-i\hat{H}_J\Delta t) \exp(-i\hat{H}_U\Delta t/2) + O(\Delta t^3), \quad (4.31)$$

allows separate exponentiation of the two parts of the Bose-Hubbard Hamiltonian, whereby only the interaction needs to be updated. Again, the error is due to the non-commutativity of operators. Thereby, the expanded propagator reads

$$\hat{\mathcal{U}}(\Delta t) = \left( \prod_{i=1}^L \hat{\mathcal{U}}_U^{[i]} \right) \left( \prod_{i \text{ odd}} \hat{\mathcal{U}}_J^{[i,i+1]} \right) \left( \prod_{i \text{ even}} \hat{\mathcal{U}}_J^{[i,i+1]} \right) \left( \prod_{i=1}^L \hat{\mathcal{U}}_U^{[i]} \right), \quad (4.32)$$

where the site-specific propagators, or gates, are given by

$$\hat{\mathcal{U}}_J^{[i,i+1]} = \exp(-iJ(\hat{a}_i^\dagger \hat{a}_{i+1} + \hat{a}_{i+1}^\dagger \hat{a}_i)\Delta t) \quad (4.33)$$

$$\hat{\mathcal{U}}_U^{[i]} = \exp\left(-i\frac{U}{2}\hat{n}_i(\hat{n}_i - 1)\Delta t/2\right). \quad (4.34)$$

A single time step,  $\Delta t$ , using the expanded operator is represented diagrammatically in

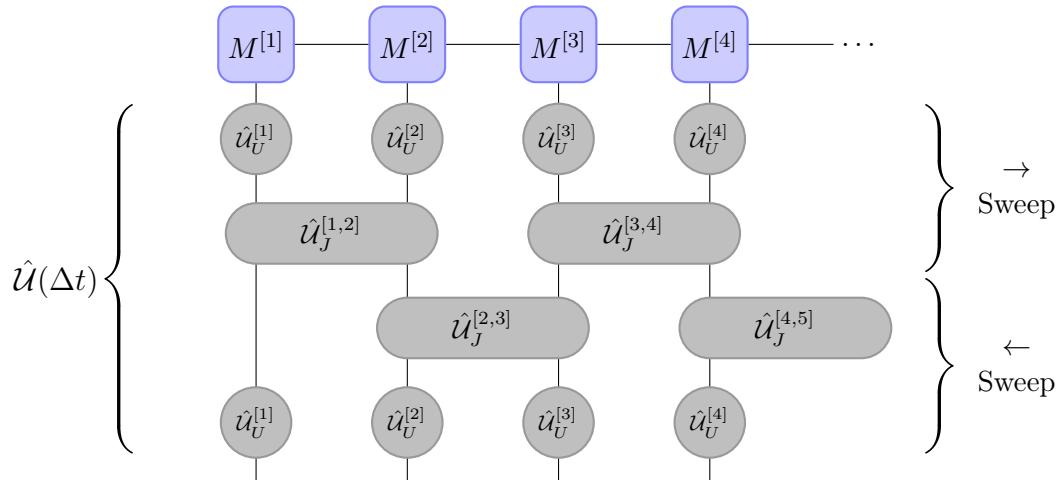


FIGURE 4.5: Tensor diagram depicting a single time step of the modified rDMRG algorithm. The tensors of the upper part of the network are contracted with the MPS while sweeping from left to right, whereas the lower part is applied with a right-to-left sweep.

figure 4.5. At first glance, the tensor network resulting from the Suzuki-Trotter expansion may seem rather extensive, however, it can be contracted in a very efficient manner. The upper part of the network is contracted in a left-to-right sweeping manner, where the position of the center cite, and thereby the normalisation of the MPS, is pushed to the right following each step. Likewise, the lower part of the network is contracted though a right-to-left sweep such that the MPS returns to its original form centered on the first site after applying the final operator. Thereby, the MPS is immediately ready for the subsequent time-propagation.

The sequence of contractions of the left-to-right sweep is shown in figure 4.6. The MPS is initially centered on the first tensor, while its remaining tensors are right-normalised. By contracting the bonds marked with a bold line shown in step 1 and 2, the operators are efficiently applied to the MPS. In the third step, the two-site tensor,  $\Theta$ , is split using an SVD, where the bond dimension of the tensors is truncated. This is crucial in order to

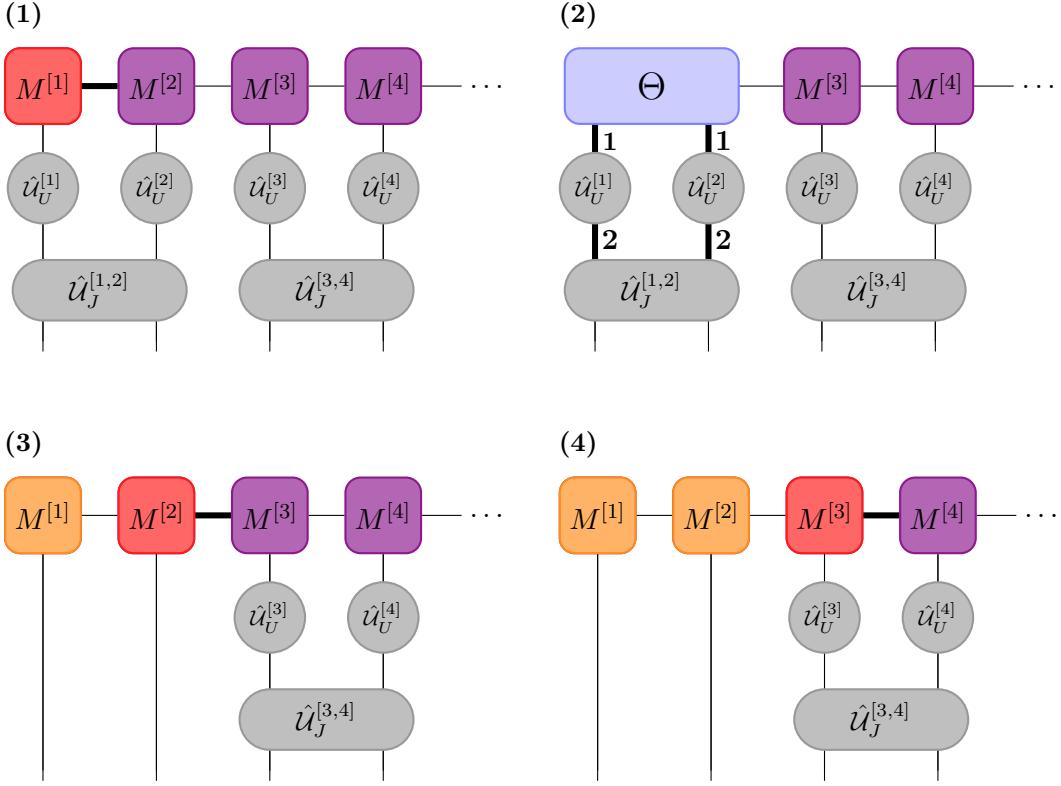


FIGURE 4.6: *Sequence of contractions for modified tDMRG algorithm during left to right sweep.* Step (1): MPS is centered on site 1. Tensors  $M^{[1]}$  and  $M^{[2]}$  are contracted. Step (2): Two-site tensor,  $\Theta$ , is contracted with operators in numbered sequence. The propagated two-site tensor is split using an SVD in step (3), followed by a contraction to the right. Lastly, in step (4), the two-site tensor of  $M^{[2]}$  and  $M^{[3]}$  is split using another SVD, whereby the center (and thereby the normalization) is pushed to site 3.

maintain a reduced dimensionality, which would otherwise result in a significant increase in contraction time. In the final and fourth step, the central cite of the MPS is moved to the start of the next two-site operator through another site merge and subsequent SVD. This step is exactly as in the original tDMRG algorithm. Thereby the normalization of the MPS is "pushed" to the right and contained in a single site, which makes it easy to deal with in the end of the time evolution step.

As the center reaches the end of the MPS, the direction of the sweep is reversed. The right-to-left sweep is very similar to the sequence described above. The main difference is in the order of contractions, as the  $\hat{U}_J^{[i,i+1]}$ -propagator is applied before  $\hat{U}_U^{[i]}$ . As the sweep, and thereby the central cite, reaches the first site of the MPS, the central site is divided by its norm. Thereby the MPS is normalized and in the same configuration as before the time step. Thus, further propagations can be performed readily.

Additional precision is achieved when evaluating the control at the beginning and end points of the time interval [84]. Thus, the left-to-right sweep applies the operator  $\hat{U}_{U(t)}^{[i]}$ , while the right-to-left sweep applies  $\hat{U}_{U(t+\Delta t)}^{[i]}$ .

The time-evolution algorithm described here shares many traits with one proposed in [75], which was tested on the Bose-Hubbard model. The algorithm displayed weak convergence with the maximum bond dimension,  $D$ , whereby precise results can be obtained using only small bond dimensions. Further improvements to the method were suggested in [75], which included utilizing even higher-order expansions of the propagator. By evaluating the

parameters of the propagator at specific points in the time interval,  $\Delta t$ , many algorithmic sweeps can be combined, greatly reducing the computational cost [85, 86].

### Time complexity of time-evolution algorithms

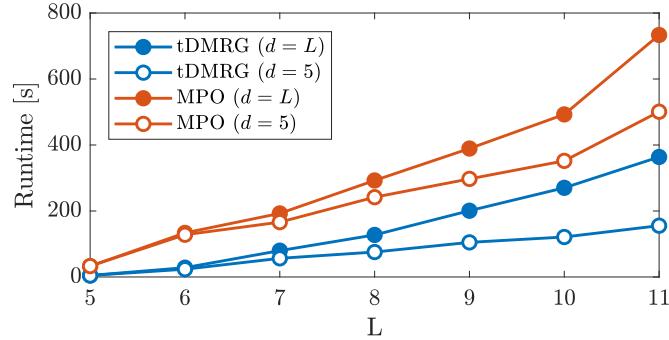


FIGURE 4.7: *Runtime of performing 100 time steps with two different algorithms in Bose Hubbard system with unit occupancy. The methods are tested on systems with both full and truncated Fock spaces.*

Figure 4.7 compares runtimes between the modified tDMRG algorithm and the default time evolution method of the ITensor library [87] for various lattice sizes with unit occupancy. The latter method builds the entire propagator as a single layer MPO following [88], and the resulting MPO is applied to the MPS according to eq. (3.21). The two methods are tested using local Fock space dimensions,  $d$ , equal to the length of the lattice,  $L$ , and set at a constant value of  $d = 5$ .

In Section 3.4 the cost of applying an MPO to an MPS was given by  $\mathcal{O}(Ld^2D_W^2D^2)$ . Since the tDMRG propagator consists of multiple gates rather than a connected layer, the operator bond dimension,  $D_W$ , is unit. Thus, the tDMRG algorithm has a cubic scaling with system size. However, a much better scaling can be achieved by truncating the local Hilbert space such that  $d$  is kept constant, resulting in a runtime scaling linearly with the system size. In the case of the Bose Hubbard model, eq. (2.39), the interaction term scales quadratically with the number of particles at a given site, which causes a huge energy penalty. Thus, for large systems with unit occupation, neglecting contributions from states with a majority of the particles at a single site is a reasonable approximation [75]. Therefore, the modified tDMRG algorithm can be applied efficiently to very large systems, if the dimension of the physical index of the MPS is restricted to a reasonable fraction of the number of particles. A similar practice was reported in [13], analyzing time-evolution of the Bose-Hubbard model across its phase transition. The study utilized a maximum occupation of 6 was used for systems of sizes  $L, N_p = 20$  to 44.



## Chapter 5

# Quantum Optimal Control Theory

The fundamental problem of Quantum Optimal Control Theory is to steer the dynamics of a quantum system in a desired way through external fields [89, 90]. Often, the goal is the transfer from an initial state,  $|\psi_0\rangle$ , to a desired target-state,  $|\psi_{\text{target}}\rangle$ . The fields responsible for controlling the dynamics of the system are parametrized by a set of control parameters or functions. Optimal control theory determines the parameters, which lead to the desired dynamics of the system [15].

In control problems, the Hamiltonian of the system is given as

$$\hat{H} = \hat{H}_0 + \sum_{n=1}^m \hat{H}_n(u_n(t)) , \quad (5.1)$$

where  $\hat{H}_0$  is an uncontrollable drift,  $\hat{H}_n$  are the controllable fields, and  $u_n(t)$  are the control functions.

The control problem presented in this thesis is steering the system from an initial state in the superfluid phase to a target-state in the Mott-Insulator phase. This is achieved by varying the lattice depth, which therefore can be considered the control parameter. The optimal control problem can be stated as follows:

Suppose the system is initially described by the state  $|\psi_0\rangle = |\psi(0)\rangle$ , and the potential is varied in the time interval  $[0, T]$ . The goal is finding the set of control parameters,  $\mathbf{u}(t)$ , which brings the initial state as close as possible to the target-state,  $|\psi_{\text{target}}\rangle$ . This is expressed in terms of a cost function

$$\mathcal{J}_T = \frac{1}{2} (1 - |\langle \psi_{\text{target}} | \psi(T) \rangle|^2) , \quad (5.2)$$

which is given as half the infidelity between the target and the state at  $t = T$ . The cost function becomes zero, when the terminal state matches the target state up to an arbitrary phase. Hence, the optimal control problem can be formulated as a minimization problem of eq. (5.2), whereby one can employ methods from mathematical optimization theory [91]. Large variations in the control parameter is often hard to achieve experimentally. Therefore, an extra term is added to the cost function, which penalizes strong variations in the control. The new cost function reads

$$\mathcal{J} = \mathcal{J}_T + \mathcal{J}_R = \mathcal{J}_T + \frac{\gamma}{2} \sum_{n=1}^m \int_0^T \left( \frac{\partial u_n}{\partial t} \right)^2 dt , \quad (5.3)$$

where  $\gamma$  weighs the relative importance between matching states and smoothness of the control [91]. As the state transfer is considered the highest priority,  $\gamma \ll 1$  such that  $\mathcal{J}_T$  dominates the cost function of eq. (5.3).

## 5.1 The Gradient-Ascent Pulse Engineering Method - GRAPE

A powerful way of performing optimal control is the Gradient-Ascent Pulse Engineering (GRAPE) method. Through GRAPE, the gradient of the cost function (5.3) can be evaluated and used to update the existing set of controls [17]. Thereby, one achieves an optimization of the cost function (5.3).

The gradient of the cost function can be derived in multiple ways. A common method is introducing a Lagrange multiplier [92, 93, 94], which forces the dynamics to obey the Schrödinger equations. This method considers the states and the control as continuous functions, which must be discretized after the derivation of the gradient for numerical purposes. However, postponing the discretization until the very last step causes a loss of accuracy, as a series of higher order corrections due to the discretization are lost. Here, an alternative derivation following [17, 95] is presented in which the discretization is introduced immediately.

Assume the transfer time,  $T$ , is discretized in steps of  $\Delta t = T/N$ . Subjecting the control to a similar discretized yields

$$u_n = (u_n(t_1), \dots, u_n(t_N)) . \quad (5.4)$$

The time-evolution of the system during the time step  $j$  is given by the propagator

$$\hat{\mathcal{U}}_j \equiv \hat{\mathcal{U}}(u(t_j)) = \exp \left\{ -i \left( \hat{H}_0 \sum_{n=1}^m \hat{H}_n(u_n(t_j)) \right) \Delta t \right\} . \quad (5.5)$$

Thereby, the cost function (5.3) becomes

$$\mathcal{J} = \frac{1}{2} \left( 1 - |\langle \psi_{\text{target}} | \prod_{j=1}^N \hat{\mathcal{U}}_j | \psi(0) \rangle|^2 \right) + \frac{\gamma}{2} \sum_{n=1}^m \sum_{j=1}^{N-1} \left( \frac{\Delta u_n(t_j)}{\Delta t} \right)^2 \Delta t , \quad (5.6)$$

where  $\Delta u_n(t_j) = u_n(t_{j+1}) - u_n(t_j)$ . The full gradient of the cost,  $\nabla \mathcal{J}(u)$ , is a vector of partial derivatives  $\partial \mathcal{J}(u)/\partial u_n(t_j)$ , which can be derived analytically.

First, consider the derivative of the regularization

$$\begin{aligned} \frac{\partial \mathcal{J}_R}{\partial u_n(t_j)} &= \frac{\gamma}{2} \left( 2 \frac{u_n(t_j) - u_n(t_{j-1})}{\Delta t^2} - 2 \frac{u_n(t_{j+1}) - u_n(t_j)}{\Delta t^2} \right) \Delta t \\ &= \frac{\gamma}{\Delta t} (2u_n(t_j) - u_n(t_{j+1}) - u_n(t_{j-1})) . \end{aligned} \quad (5.7)$$

The part of the gradient related to the regularization depends only on the control, whereby it can be calculated without considering the state of the system.

Next, consider the derivative of  $\mathcal{J}_T$ . Defining the transfer probability amplitude  $\mathcal{T} \equiv \langle \psi_{\text{target}} | \psi(T) \rangle$ , the derivative of the final infidelity with respect to the control can be written as

$$\frac{\partial \mathcal{J}_T}{\partial u_n(t_j)} = -\frac{1}{2} \frac{\partial}{\partial u_n(t_j)} \mathcal{T}^* \mathcal{T} = -\text{Re} \left( \mathcal{T}^* \frac{\partial \mathcal{T}}{\partial u_n(t_j)} \right) . \quad (5.8)$$

From eq. (5.8) it is evident that the derivative of the infidelity depends only on the derivative of the transfer probability amplitude,  $\partial\mathcal{T}/\partial u_n(t_j)$ . This derivative can be rewritten as

$$\begin{aligned}\frac{\partial\mathcal{T}}{\partial u_n(t_j)} &= \frac{\partial}{\partial u_n(t_j)} \left\langle \psi_{\text{target}} \left| \prod_{j=1}^N \hat{\mathcal{U}}_j \right| \psi(0) \right\rangle \\ &= \left\langle \psi_{\text{target}} \left| \hat{\mathcal{U}}_N \dots \hat{\mathcal{U}}_{j+1} \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \hat{\mathcal{U}}_{j-1} \dots \hat{\mathcal{U}}_1 \right| \psi(0) \right\rangle\end{aligned}\quad (5.9)$$

Multiplying eq. (5.9) with  $\mathcal{T}^*$  yields

$$\mathcal{T}^* \frac{\partial\mathcal{T}}{\partial u_n(t_j)} = \langle \psi(T) | \psi_{\text{target}} \rangle \left\langle \psi_{\text{target}} \left| \prod_{j'=j+1}^N \hat{\mathcal{U}}_{j'} \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \prod_{j'=1}^{j-1} \hat{\mathcal{U}}_{j'} \right| \psi(0) \right\rangle \quad (5.10)$$

$$= i \left\langle \chi(T) \left| \prod_{j'=j+1}^N \hat{\mathcal{U}}_{j'} \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \prod_{j'=1}^{j-1} \hat{\mathcal{U}}_{j'} \right| \psi(0) \right\rangle \quad (5.11)$$

$$= i \left\langle \chi(t_j) \left| \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \right| \psi(t_{j-1}) \right\rangle, \quad (5.12)$$

where  $|\chi(T)\rangle \equiv i |\psi_{\text{target}}\rangle \langle \psi_{\text{target}}| \psi(T)\rangle$  is the projection of the final state onto the target state. Notice how in eq. (5.13) the state  $|\chi(T)\rangle$  has been propagated backwards in time. Finally, inserting eq. (5.12) into the derivative of the transfer cost function (5.8) produces

$$\frac{\partial \mathcal{J}_T}{\partial u_n(t_j)} = -\text{Re} \left\langle \chi(t_j) \left| i \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \right| \psi(t_{j-1}) \right\rangle. \quad (5.13)$$

The derivative of the propagator,  $\hat{\mathcal{U}}_j$ , is non-trivial, due to possible non-commutativity between the Hamiltonian and its derivative. This results in a series of higher order corrections to the derivative of propagator. Expanding the propagator as a Taylor series before taking the derivative gives

$$\begin{aligned}\frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} &= \frac{\partial}{\partial u_n(t_j)} \exp(-i\hat{H}\Delta t) \\ &= \sum_{p=0}^{\infty} \frac{(-i\Delta t)^p}{p!} \frac{\partial \hat{H}^p}{\partial u_n(t_j)}.\end{aligned}\quad (5.14)$$

As mentioned before the Hamiltonian may not commute with its derivative. Therefore, one must be careful when taking the derivative of  $\hat{H}^p$ . Retaining the ordering of the operators while taking the derivative yields

$$\begin{aligned}\frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} &= \sum_{p=1}^{\infty} \frac{(-i\Delta t)^p}{p!} \sum_{q=0}^{p-1} \hat{H}^q \frac{\partial \hat{H}}{\partial u_n(j)} \hat{H}^{p-q-1} \\ &= \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{A^p B A^q}{(p+q+1)!},\end{aligned}\quad (5.15)$$

where  $A \equiv -i\hat{H}\Delta t$  and  $B \equiv -i\partial\hat{H}/\partial u_n(t_j)\Delta t$  have been defined for notational convenience. Through the standard relations of the gamma function,  $\Gamma(z)$ , one can derive the identity

$$\frac{1}{(p+q+1)!} = \frac{1}{p!q!} \int_0^1 (1-\alpha)^p \alpha^q d\alpha . \quad (5.16)$$

Thereby eq. (5.15) can be expressed as

$$\begin{aligned} \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} &= \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{A^p B A^q}{p! q!} \int_0^1 (1-\alpha)^p \alpha^q d\alpha \\ &= \int_0^1 \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(A(1-\alpha))^p}{p!} B \frac{(A\alpha)^q}{q!} d\alpha \\ &= \int_0^1 e^{(1-\alpha)A} B e^{\alpha A} d\alpha \\ &= e^A \int_0^1 e^{-\alpha A} B e^{\alpha A} d\alpha , \end{aligned} \quad (5.17)$$

where the expansion of the exponential function has been used to eliminate the sums. Although eq. (5.17) looks rather simple, evaluating the integral in its current form is a fairly hard task. Instead, the integral can be explicitly solved by applying the Baker–Campbell–Hausdorff expansion

$$e^X Y e^{-X} = \sum_{k=0}^{\infty} \frac{[X, Y]_k}{k!} = Y + [X, Y] + \frac{1}{2!} [X, [X, Y]] + \frac{1}{3!} [X, [X, [X, Y]]] + \dots \quad (5.18)$$

where  $[X, Y]_k = [X, [X, Y]]_{k-1}$  and  $[X, Y]_0 = Y$  is the definition of the recursive commutator [96]. Thus, eq. (5.17) reads

$$\begin{aligned} \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} &= e^A \int_0^1 \sum_{k=0}^{\infty} \alpha^k \frac{(-1)^k}{k!} [A, B]_k d\alpha \\ &= e^A \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} [A, B]_k . \end{aligned} \quad (5.19)$$

Inserting this final expression for the derivative of the propagator into eq. (5.8), one finds the exact derivative of the infidelity

$$\begin{aligned} \frac{\partial \mathcal{J}_T}{\partial u_n(t_j)} &= -\text{Re} \left\langle \chi(t_j) \left| i e^{-i\hat{H}\Delta t} \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} \left[ -i\hat{H}\Delta t, -i \frac{\partial \hat{H}}{\partial u_n(t_j)} \Delta t \right]_k \right| \psi(t_{j-1}) \right\rangle \\ &= -\text{Re} \left\langle \chi(t_{j-1}) \left| \sum_{k=0}^{\infty} \frac{i^k \Delta t^{k+1}}{(k+1)!} \left[ \hat{H}, \frac{\partial \hat{H}}{\partial u_n(t_j)} \right]_k \right| \psi(t_{j-1}) \right\rangle . \end{aligned} \quad (5.20)$$

For small time-steps the higher order corrections can be neglected, however, choosing a larger time-step reduces the run-time of the time-evolution, which is critical when describing many-body systems. Therefore, when using large time-steps, higher-order correlations should be included to preserve accuracy. Computing the higher order corrections can be done efficiently by analytically deriving the commutators beforehand. In section 4.2.2 an alternative propagator is constructed using the Suzuki-Trotter expansion, which causes the gradient to be exact to zeroth order.

Finally, combining the derivatives of eq. (5.7) and (5.20) produces the entries of the gradient vector for the cost function

$$\begin{aligned} \frac{\partial \mathcal{J}(\mathbf{u})}{\partial u_n(t_j)} = & -\text{Re} \left\langle \chi(t_{j-1}) \left| \left( i \frac{\partial \hat{H}}{\partial u_n(t_j)} \Delta t + \text{h.o.} \right) \right| \psi(t_{j-1}) \right\rangle \\ & + \frac{\gamma}{\Delta t} (2u_n(t_j) - u_n(t_{j+1}) - u_n(t_{j-1})) . \end{aligned} \quad (5.21)$$

Here h.o. denotes all higher order terms ( $k > 0$ ).

Through the analytically derived gradient, the cost can be iteratively updated using gradient-based optimization methods until a desired threshold is reached. This forms the framework of the Gradient-Ascent Pulse Engineering (GRAPE) algorithm [17]. The cost is minimized

---

**Algorithm 2** GRAPE Algorithm

---

Choose initial control  $\mathbf{u}^{(1)}$ .

**while**  $\mathcal{J} > \mathcal{J}_{\text{threshold}}$  **do**

Calculate  $|\psi(t_k)\rangle = \prod_{j=1}^k \hat{\mathcal{U}}_j |\psi(0)\rangle$  for  $k = 1 \dots N$ .

Calculate  $|\chi(t_k)\rangle = \prod_{j=N}^k \hat{\mathcal{U}}_j^\dagger |\chi(T)\rangle$  for  $k = N \dots 1$ .

Evaluate  $\frac{\partial \mathcal{J}}{\partial u_n(t_k)}$  for  $k = 1 \dots N$  and  $n = 1 \dots m$  according to eq. (5.21).

Update controls using gradient such that  $J^{(i+1)} < J^{(i)}$ .

**end while**

---

under the initial and terminal condition

$$\mathbf{u}(0) = \text{fixed value} \quad (5.22)$$

$$\mathbf{u}(T) = \text{fixed value} \quad (5.23)$$

$$|\psi(0)\rangle = |\psi_0\rangle \quad (5.24)$$

$$|\chi(T)\rangle = i |\psi_{\text{target}}\rangle \langle \psi_{\text{target}} | \psi(T) \rangle . \quad (5.25)$$

Although the starting guess of the control,  $\mathbf{u}^{(1)}$ , can be completely random, both faster convergence and lower convergence value is achieved by choosing a good starting seed. Clearly, there is no guarantee that the algorithm will converge to the global optimum, as it is based on a gradient ascent procedure [17]. Nevertheless, the algorithm can be made to search a large portion of the parameter space by executing it multiple time for various seeds.

In practice, the GRAPE algorithm requires a lot of computer memory, as the gradient of the cost requires the inner product of the states  $|\psi\rangle$  and  $|\chi\rangle$  at all times,  $t_k$ . The two states cannot be time-evolved simultaneously, as  $|\chi\rangle$  is derived from the final state. To efficiently evaluate the gradient, the state  $|\psi\rangle$  must be stored at each time-step, while entries of the gradient (5.21) are calculated at each step of the back-propagation of  $|\chi\rangle$ . Thus, a total of  $N$  states must be kept in each iteration of the GRAPE algorithm. For complex many-body systems this may become an issue, as tensor-network descriptions of states (described in chapter 3) require a lot of storage. A possible workaround is proposed in [97], which involves calculating the final state,  $|\psi(T)\rangle$ , without storing the state at every step. Subsequently, the two states  $|\psi(T)\rangle$  and  $|\chi(T)\rangle$  are both propagated backwards simultaneously, and the elements of the gradient (5.21) are calculated at each step. This eliminates the need for storing a large amount of information, as a maximum of two states will be kept in the memory at a time. However, this memory-preserving version of the GRAPE algorithm requires an additional time-evolution of  $|\psi\rangle$  in each iteration resulting in an increased runtime.

### 5.1.1 Gradient of Suzuki-Trotter Propagator

In the previous section, the derivative of the cost function with regards to the control was derived for a general propagator (5.21). However, simplifications to the otherwise rather complicated gradient can be achieved when utilizing features of the propagator used for the specific problem. In this instance, the state is time-evolved using a variation of the tDMRG algorithm described in section 4.2.2, where the propagator has been Suzuki-Trotter expanded. Remarkably, the combination of the expanded propagator and a diagonal control Hamiltonian causes all higher order contributions to the cost gradient to vanish. Instead, the precision of the gradient is solely determined by the order of the expansion.

Consider the elements of the transfer cost gradient for the GRAPE algorithm (5.13), where the derivative of a general propagator with respect to the control is given by

$$\frac{\partial \hat{U}_j}{\partial u_n(t_j)} = e^{-i\hat{H}(u_n(t_j))\Delta t} \sum_{k=0}^{\infty} \frac{i^{k-1}\Delta t^{k+1}}{(k+1)!} \left[ \hat{H}(u_n(t_j)), \frac{\partial \hat{H}(u_n(t_j))}{\partial u_n(t_j)} \right]_k . \quad (5.26)$$

In the problem treated here, the system is described by the Bose-Hubbard model (2.39), where the interaction matrix element acts as the control function,  $u_n(t_j) \equiv U(t_j)$ . Thus, the Hamiltonian can be written as

$$\hat{H}(U(t_j)) \equiv \hat{H}_J + U(t_j)\hat{H}_U , \quad (5.27)$$

where the tunneling Hamiltonian,  $\hat{H}_J$ , is regarded as an uncontrollable drift. The time-evolution algorithm used here employs a Suzuki-Trotter expansion of the propagator while considering the control at both start and end of the step. Thereby, the expanded propagator derived in section 4.2.2 reads

$$\hat{U}_j^{\text{ST}} = e^{-iU(t_j)\hat{H}_U\Delta t/2} e^{-i\hat{H}_J\Delta t} e^{-iU(t_{j-1})\hat{H}_U\Delta t/2} \equiv \hat{U}_j^U \hat{U}_j^J \hat{U}_{j-1}^U . \quad (5.28)$$

Since the control at times  $t_j$  and  $t_{j-1}$  contribute to  $\hat{U}_j^{\text{ST}}$ , the elements of the cost gradient are

$$\frac{\partial \mathcal{J}_T}{\partial U(t_j)} = -\text{Re} \left\langle \chi(t_j) \left| i \frac{\partial \hat{U}_j^{\text{ST}}}{\partial U(t_j)} \right| \psi(t_{j-1}) \right\rangle - \text{Re} \left\langle \chi(t_{j+1}) \left| i \frac{\partial \hat{U}_{j+1}^{\text{ST}}}{\partial U(t_j)} \right| \psi(t_j) \right\rangle . \quad (5.29)$$

Further examining the derivative of the first propagator reveals

$$\begin{aligned} \frac{\partial \hat{U}_j^{\text{ST}}}{\partial U(t_j)} &= \frac{\partial \hat{U}_j^U}{\partial U(t_j)} \hat{U}_j^J \hat{U}_{j-1}^U \\ &= \exp(-iU(t_j)\hat{H}_U\Delta t/2) \sum_{k=0}^{\infty} \frac{i^{k-1}\Delta t^{k+1}}{(k+1)!} \left[ U(t_j)\hat{H}_U, \hat{H}_U \right]_k \hat{U}_j^J \hat{U}_{j-1}^U \\ &= (-i\hat{H}_U\Delta t/2) \exp(-iU(t_j)\hat{H}_U\Delta t/2) \hat{U}_j^J \hat{U}_{j-1}^U \\ &= (-i\hat{H}_U\Delta t/2) \hat{U}_j^{\text{ST}} . \end{aligned} \quad (5.30)$$

Since  $\hat{H}_U$  is diagonal, the following relation apply for the recursive commutator

$$\left[ U(t_j)\hat{H}_U, \hat{H}_U \right]_k = \begin{cases} \hat{H}_U, & \text{if } k = 0. \\ 0, & \text{otherwise.} \end{cases} , \quad (5.31)$$

which causes all higher-order contributions to the derivative of the propagator to drop out. Likewise, the derivative of the second propagator is

$$\frac{\partial \hat{U}_{j+1}^{\text{ST}}}{\partial u_n(j)} = \hat{U}_{j+1}^{\text{ST}} \left( -i \hat{H}_U \Delta t / 2 \right). \quad (5.32)$$

Inserting the derivatives of eq. (5.30) and (5.32) into the derivative of the cost (eq. (5.29)) yields

$$\begin{aligned} \frac{\partial \mathcal{J}_T}{\partial U(t_j)} &= -\text{Re} \left\langle \chi(t_j) \left| \left( \hat{H}_U \Delta t / 2 \right) \hat{U}_j^{\text{ST}} \right| \psi(t_{j-1}) \right\rangle - \text{Re} \left\langle \chi(t_{j+1}) \left| \hat{U}_{j+1}^{\text{ST}} \left( \hat{H}_U \Delta t / 2 \right) \right| \psi(t_j) \right\rangle \\ &= -\text{Re} \left\langle \chi(t_j) \left| \hat{H}_U \Delta t / 2 \right| \psi(t_j) \right\rangle - \text{Re} \left\langle \chi(t_j) \left| \hat{H}_U \Delta t / 2 \right| \psi(t_j) \right\rangle \\ &= -\text{Re} \left\langle \chi(t_j) \left| \hat{H}_U \Delta t \right| \psi(t_j) \right\rangle. \end{aligned} \quad (5.33)$$

Thus, the combination of the Suzuki-Trotter expansion and a diagonal control Hamiltonian eliminates all higher order contributions to the gradient. Thereby, the gradient of the cost is exact up to the order of the expansion. A comparison between a numerically calculated

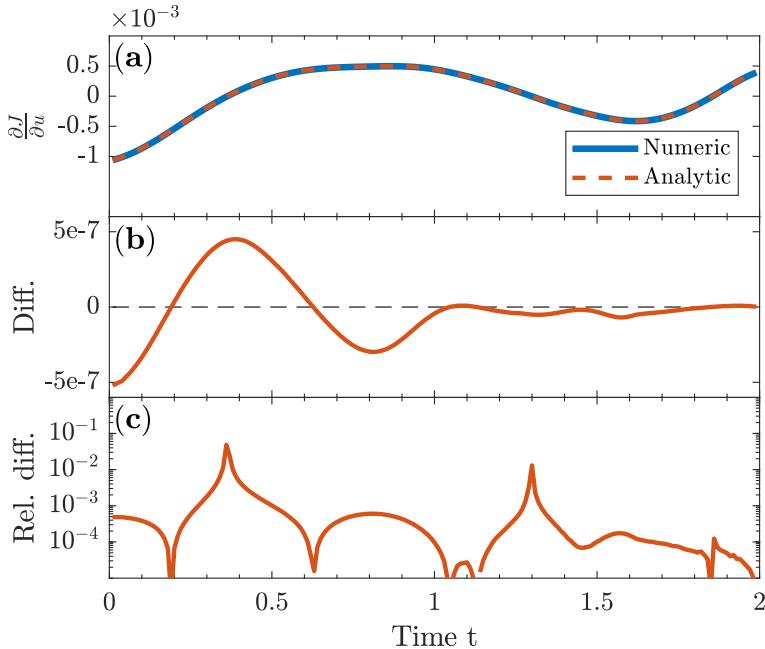


FIGURE 5.1: (a): Numerical gradient along with gradient calculated via eq. (5.33). (b): Difference between the two gradients. (c): Absolute relative difference between the two gradients.

gradient and the analytically derived gradient of eq. (5.33) can be seen in figure 5.1. As evident from the figure, the derived cost gradient is very accurate, which is necessary for an efficient optimization procedure. The difference between the two gradients is largest at the beginning of the sequence, due to the accumulated error from the time evolution of  $|\chi(0)\rangle$ , which is derived from two time evolutions over the entire duration.

## 5.2 Interior Point Methods

The GRAPE method formulates the control problem as the minimization of the cost function (5.3) while providing the derivative of the cost with respect to the control functions (5.21).

Thus, the optimal control can be found using well established methods from mathematical optimization theory [95]. Generally, control problems are very hard to solve, as they are highly non-linear, have many control parameters (even when parametrized), and are often subjected to a series of constraints. An example of the latter can be seen when considering optimizations within the Bose-Hubbard model, as the model is valid only within the tight-binding limit [1]. Thus, the control problem will be subject to the non-linear constraint

$$V_0^{\min} \leq c(u(t)) \leq V_0^{\max}, \quad (5.34)$$

where  $c(u)$  is some constraint function dependent on the parameterization of the problem. Interior point methods are currently considered some of the most powerful algorithms for large-scale non-linear programming [25]. The methods update the control parameters using the derivatives of the cost function via Newton's method, however, they differ from the standard Newton approach in many ways to accommodate constraints [98, 99]. Interior point methods approach the solution from within the feasible region as opposed to other methods like Nelder-Mead, hence the name *interior*. In addition, they provide efficient performance while having better theoretical properties than the standard simplex method [25].

In this section the simplest version of non-linear interior point method is derived, however, many variants of the method exist with respect to update strategies [100, 101], line searches [102, 103, 104], handling non-convexity [105] and more.

Consider the general minimization problem of some objective  $f(x)$

$$\min_x f(x) \quad (5.35a)$$

$$\text{subject to } g(x) = 0 \quad (5.35b)$$

$$h(x) \geq 0, \quad (5.35c)$$

where  $g(x)$  and  $h(x)$  are a set of equality and inequality constraints. The Lagrangian for the general constrained optimization problem (5.35) is defined as

$$\mathcal{L}(x, \lambda, \zeta) = f(x) - \lambda^T g(x) - \zeta^T h(x), \quad (5.36)$$

where  $\lambda$  and  $\zeta$  are vectors of Lagrange multipliers of the equality and inequality constraints respectively. Thus, one must find a point  $(x^*, \lambda^*, \zeta^*)$  minimizing the objective function while satisfying the constraints.

### 5.2.1 Duality

Duality theory describes how an alternative *dual* problem can be derived from the original *primal* problem and how the two problems are related. Thus, the dual problem can be regarded as a different perspective of the original problem. Much information regarding the original problem can be inferred from the solution of its dual, hence many non-linear optimization algorithms solve primal-dual problems [25].

The dual function is defined as

$$q(\lambda, \zeta) = \inf_x \mathcal{L}(x, \lambda, \zeta), \quad (5.37)$$

where the infimum is required to exist and be finite. Note how the Lagrange multipliers of the primal problem are the optimization parameters of its dual. The dual function always satisfies the condition of *weak duality*

$$f(x^*) \geq q(\lambda^*, \zeta^*), \quad (5.38)$$

whereby it provides a lower bound on the solution of (5.35). In order for the direction of the inequality constraint to remain the same, it is necessary that all the inequality Lagrange multipliers fulfill  $\zeta_i \geq 0$ . This requirement can be considered a constraint of the dual, since the Lagrange multipliers are the variables of the dual function [25]. Thus, the dual problem is defined as

$$\max_{\lambda, \zeta} q(\lambda, \zeta) \quad (5.39a)$$

$$\text{subject to } \zeta \geq 0. \quad (5.39b)$$

A very nice property of the dual problem is that it is always convex regardless of  $f(x)$  [106]. Thus, the dual problem is often easier to solve than the primal. Furthermore, the solutions of the dual problem are the optimal Lagrange multipliers of the original problem, which would otherwise be very hard to find. Another useful feature in duality theory is the *duality gap*, which is given by the difference between the primal and dual function,  $f(x) - q(\lambda, \zeta)$ . Considering the condition of weak duality (5.38), the duality gap implies that

$$f(x) - f(x^*) \leq f(x) - g(\lambda, \zeta), \quad (5.40)$$

whereby  $x$  is primal optimal and  $\lambda, \zeta$  are dual optimal, if the duality gap is zero. Thus, eq. (5.40) can be used as a stopping criteria for optimization algorithms, as the duality gap will tend towards zero for  $x \rightarrow x^*$  [25].

### 5.2.2 Karush–Kuhn–Tucker Conditions

The Karush–Kuhn–Tucker (KKT) conditions are a set of first-order necessary conditions for a point  $(x^*, \lambda^*, \zeta^*)$  being an optimum [107]. They only consider properties of the gradient of the objective and constraint function, hence the *first-order* label. The KKT conditions are very important in optimization theory, as many algorithms (including interior point methods) can be interpreted as solving a set of equations directly derived from the conditions [25].

**Theorem 1** *Let the active constraints be all the equality constraints along with the set of inequality constraints fulfilling  $h(x) = 0$  for some feasible  $x$ . Suppose that  $x^*$  is a local solution of (5.35) and the gradients of the active constraints are linearly independent at  $x^*$ . Then Lagrange multiplier vectors  $\lambda$  and  $\zeta$  exists, such that the following conditions are satisfied at  $(x^*, \lambda^*, \zeta^*)$*

$$\nabla_x \mathcal{L}(x^*, \lambda^*, \zeta^*) = 0, \quad (5.41a)$$

$$g(x^*) = 0, \quad (5.41b)$$

$$h(x^*) \geq 0, \quad (5.41c)$$

$$\zeta^* \geq 0, \quad (5.41d)$$

$$\zeta^* h_i(x^*) = 0, \quad \text{for } i = 1, \dots, m \quad (5.41e)$$

The first condition is called the *stationary condition* and simply states that the point  $(x^*, \lambda^*, \zeta^*)$  must be stationary point, which is a general requirement for a point to be considered an optimum. Next, the conditions (5.41b) and (5.41c) are the *primal feasibility*, which coincide with the constraints originally stated in the problem (5.35). The following condition (5.41d) is the *duality feasibility*, which is a constraint of the dual problem (5.39) discussed earlier. The last condition is the *complementary slackness*, which arises from the

requirement  $\zeta^*{}^T h(x^*) = 0$ . The two previous conditions state that  $\zeta_i^* h_i(x^*) \geq 0$ , whereby the condition (5.41e) is needed for each inequality constraint.

If the point  $(x^*, \lambda^*, \zeta^*)$  satisfies all the KKT conditions, then the point is the solution to the primal-dual problem.

### 5.2.3 Basic Primal-Dual Interior Point Algorithm

Interior point methods solve a set of equations derived directly from the KKT conditions using Newtons method. However, Newtons method is in general not compatible with inequality constraints, hence the original problem (5.35) must be reformulated. Instead, the interior point methods solve barrier problems of the form

$$\min_{x,s} f(x) - \mu \sum_{i=1}^m \log s_i \quad (5.42a)$$

$$\text{subject to } g(x) = 0 \quad (5.42b)$$

$$h(x) - s = 0 , \quad (5.42c)$$

where  $\mu$  is the positive barrier parameter, and  $s$  is a vector of slack variables, which transforms the inequality constraint into an equality at the cost of adding additional optimization variables. In order to ensure the direction of the inequality constraint, it is a requirement that  $s \geq 0$ . While this is an inequality constraint itself, it is enforced via the logarithmic barrier term, which diverges as any component of  $s$  approaches zero. Furthermore, the barrier term does not add much complexity to the problem, as it convex and twice differentiable [99].

The step direction towards the KKT-point can be found through Newtons method, however, often one can only take a small step before violating the constraints, hence the pure Newton direction is dubbed the *affine scaling direction* [25]. The barrier approach consists of finding solutions of the barrier problem (5.42) for a sequence of barrier parameters  $\{\mu_k\}$  converging to zero. Several strategies exist for updating the barrier parameters [100, 101], however, their common goal is to bias the step direction of the algorithm, such that the longest possible steps can be taken. The solutions of the barrier subproblems can be denoted  $(x(\mu), s(\mu), \lambda(\mu), \zeta(\mu))$ , and the trajectory of these points is known as the *primal-dual central path*, which converges to  $(x^*, s^*, \lambda^*, \zeta^*)$  as  $\mu \rightarrow 0$  [25].

The KKT conditions of the barrier problem (5.42) can be expressed in a single mapping

$$F \equiv \begin{bmatrix} \nabla_x f(x) - A_g^T(x)\lambda - A_h^T(x)\zeta \\ S\zeta - \mu e \\ g(x) \\ h(x) - s \end{bmatrix} = 0 , \quad (5.43)$$

where  $A_g(x)$  and  $A_h(x)$  are the Jacobian matrices of the constraint functions  $g(x)$  and  $h(x)$ . Additionally,  $S$  and  $Z$  are defined as diagonal matrices with entries given by the vectors  $s$  and  $\zeta$ , while  $e = (1, 1, \dots, 1)^T$ . Note how the introduction of the slack variables and the barrier term has modified the original KKT conditions (5.41). Transforming the constraints has removed the need for the duality feasibility (5.41d), while barrier term "relaxes" the complementary slackness (5.41e), whose original form is recovered as  $\mu \rightarrow 0$  [25].

Applying Newtons method to the set of non-linear equations (5.43) yields

$$\begin{bmatrix} \nabla_{xx}\mathcal{L} & 0 & -A_g^T(x) & -A_h^T(x) \\ 0 & Z & 0 & S \\ A_g(x) & 0 & 0 & 0 \\ A_h(x) & -I & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_s \\ p_\lambda \\ p_\zeta \end{bmatrix} = -F \quad (5.44)$$

where  $p = (p_x, p_s, p_\lambda, p_\zeta)$  is the constrained Newton step direction, and  $\mathcal{L}$  denotes the Lagrangian of the barrier problem (5.42)

$$\mathcal{L}(x, s, \lambda, \zeta) = f(x) - \lambda^T g(x) - \zeta^T (h(x) - s) . \quad (5.45)$$

After computing the step, the parameters are iterated according to

$$\begin{bmatrix} x^{k+1} \\ s^{k+1} \\ \lambda^{k+1} \\ \zeta^{k+1} \end{bmatrix} = \begin{bmatrix} x^k \\ s^k \\ \lambda^k \\ \zeta^k \end{bmatrix} + \alpha \circ p , \quad (5.46)$$

where  $\alpha$  is a vector of step-sizes for each of the parameters. The step-size is commonly determined via a line-search [103, 104], although alternative methods exists [108].

Interior point methods are capable of finding optimal solutions of non-linear constrained point using the methods described above, which can be summarized in the following algorithm [25]:

---

**Algorithm 3** Basic Primal-Dual Interior Point Algorithm

---

```

Choose initial  $\mu_0 > 0$ .
Choose initial point  $(x_0, s_0, \lambda_0, \zeta_0)$ .
while overall stopping criteria not met do
    while subproblem stopping criteria not met do
        Solve eq. (5.44) to obtain search direction  $p$ .
        Compute step-size vector  $\alpha$ .
        Update  $(x_k, s_k, \lambda_k, \zeta_k)$  according to eq. (5.46).
    end while
    Update barrier parameter  $\mu_k$ .
end while

```

---

#### 5.2.4 Implementation with GRAPE

The interior point method is used here for minimizing the cost function (5.3), which is a function of a set of control parameters  $u_n = (u_n(t_1), \dots, u_n(t_N))$ . Examining the set of equations solved by the interior point method (5.44) reveals that in order to compute the search direction,  $p$ , one needs the gradient of the objective function,  $\nabla f(x)$ , the Jacobian matrices of the constraints,  $A_g(x)$  and  $A_h(x)$ , and finally the Hessian of both the objective and constraints contained in  $\nabla_{xx}\mathcal{L}$ . The GRAPE method provides the gradient of the cost function through eq. (5.21). Meanwhile, the constraints are dependent on the specific problem. In this case, the problem is crossing the phase transition between the superfluid and Mott-Insulator of the Bose-Hubbard model. As discussed in section 4.2.2, the control is parametrized as the interaction matrix element,  $U$ , which is directly dependent of the lattice depth as illustrated in figure 2.4. Thus, the control problem is constrained by

$$U_{\min} \leq U(t_j) , \quad \text{for } j = 1, \dots, N , \quad (5.47)$$

for some  $U_{\min}$  chosen such that the lattice depth is not below the tight binding limit, whereby the Bose-Hubbard model remains valid throughout the optimization [1]. In the GRAPE parametrization the constraint function (5.47) is the control itself, whereby the constraint Jacobian is simply the identity matrix. Furthermore, the second derivatives of the constraints completely vanish. Nevertheless, the Hessian of the cost function is still required in order to solve the equations (5.44). Although the second derivative of the cost function (5.3) is analytically derivable (see Appendix E), the numerical cost of computing the Hessian is rather high. Instead, Quasi-Newton methods can be used, where the Hessian can be approximated through various formulas [25]. The implementation of the interior point algorithm used in this thesis is from the IPOPT library [99], which approximates the Hessian through the L-BFGS algorithm, which builds the matrix using the gradients of previous iterations [109]. The combination of interior point methods and the GRAPE algorithm works really well in practice, as interior point methods attempt to solve the problem using as few steps as possible [25], which is ideal, as the GRAPE gradient is very costly to compute.

### 5.3 Control Parametrization via Chopped Basis

The convergence rate of gradient-based optimal control methods such as GRAPE have been shown to be much better than gradient-free methods for multiple problems [91, 110]. One of the main disadvantages of GRAPE is the dimension for the optimization. As the duration of the control pulse is discretized into  $N$  time-steps, the control value at each time slice,  $u_n(t_j)$ , is considered an optimization parameter. For long durations or small time-steps the resulting dimension of the optimization ( $N$ ) will be very large. Thus, this type of control parametrization can be considered having too many degrees of freedom for the optimization. Employing a proper parametrization of the control can drastically reduce the optimization dimension [93]. A common choice is employing a chopped basis, which parametrizes the control as

$$u(t) = u_0(t) + S(t) \sum_{n=1}^M c_n f_n(t) , \quad (5.48)$$

where  $f_n$  are the basis functions, and  $c_n$  are the optimization coefficients. In addition,  $u_0(t)$  is the initial control function, and  $S(t)$  is a shape function enforcing the boundary conditions of the control, whereby  $S(0) = S(T) = 0$  [110]. The shape function used throughout this thesis consists of two steep sigmoids oriented such that  $S$  is unit for most time steps. The basis functions,  $f_n$ , must be chosen based on physical insight of the system [111]. For the control problem discussed in this thesis, employing a basis of sine functions has proven itself useful, as the sine functions are excellent at describing the smoothly varying control pulses desired [110, 112]. Thus, the basis functions read  $f_n = \sin(\omega_n t/T)$ , where  $\omega_n = n\pi$  can be considered the Fourier components of the description. An extension of this type of chopped basis was done in [19, 113], introducing random shifts to the frequencies resulting in the Chopped RAndom Basis or CRAB. In recent years the CRAB parametrization has been used together with the Nelder-Mead hill climbing algorithm to solve various control problems [19, 114, 82, 115].

The great advantage of employing a reduced basis representation is the drastic reduction in the optimization space. However, artificial minima may be introduced, if the chosen basis is incapable of spanning the part of the optimization space containing the optimal solutions [116]. Hence, the chopped basis dimension,  $M$ , must be chosen with consideration.

### 5.3.1 Gradient-Optimization Using Parametrization - GROUP

The Gradient-Optimization Using Parametrization or GROUP algorithm introduced in [110] combines the chopped basis representation with the GRAPE algorithm. A comparison between different optimization algorithms for optimal control demonstrated that GROUP outperformed both GRAPE and Nelder-Mead with CRAB in term of fidelity reached and number of function evaluations required for convergence.

Parameterizing the control alters the gradient of the cost function (5.21). The new gradient can be derived using the chain rule

$$\begin{aligned} \frac{\partial \mathcal{J}(c)}{\partial c_n} &= \sum_{j=1}^N \frac{\partial \mathcal{J}(u)}{\partial u(t_j)} \frac{\partial u(t_j)}{\partial c_n} \\ &= \sum_{j=1}^N \frac{\partial \mathcal{J}}{\partial u(t_j)} S(t_j) f_n(t_j), \end{aligned} \quad (5.49)$$

where only a single control has been chosen for cleaner notation. The gradient in GROUP is more complicated than in GRAPE, however, the cost of computing the gradient is still dominated by the evaluation of the original derivative  $\partial \mathcal{J}/\partial u(t_j)$ . Therefore, the increased computational cost of evaluating the gradient can be considered negligible [110].

In addition, the chopped basis parameterization changes the derivatives of the constraints as well. Thus, the new constraint Jacobian can be derived using the chain rule as well

$$\begin{aligned} (A_h(c))_{i,n} &= \sum_{j=1}^N \frac{\partial h_i(u)}{\partial u(t_j)} \frac{\partial u(t_j)}{\partial c_n} \\ &= \sum_{j=1}^N (A_h(u))_{i,j} S(t_j) f_n(t_j), \end{aligned} \quad (5.50)$$

where  $h(u)$  are the constraint functions, and  $(A_h(u))_{i,j} = \partial h_i(u)/\partial u(t_j)$  are recognized as the elements of the GRAPE parametrized constraint Jacobian. As discussed in section 5.2.4, the constraint Jacobian is simply the identity matrix for the control problem treated here, whereby

$$(A_h(c))_{i,n} = \sum_{j=1}^N \delta_{i,j} S(t_j) f_n(t_j) = S(t_i) f_n(t_i). \quad (5.51)$$

Thus, the elements of the parametrized  $N \times M$  Jacobian matrix remain constant throughout the entire optimization process for this problem.

Lastly, the parameterization of the control also alters the Hessians of the control problem. However, the implementation of the interior point method used here [99] is capable of approximating both the objective and constraint Hessians from previous derivatives using the L-BFGS algorithm [109]. Therefore, calculating the transformed derivatives is sufficient, and the general Hessians will not be derived here.

## 5.4 Quantum Speed Limit

A subtlety in the formulation of optimal control problems is that one is only searching for the control,  $\mathbf{u}(t)$ , which steers the initial state into the target-state at exactly the duration  $t = T$ . However, it is often desirable to obtain the desired state in the shortest timespan possible. If a solution exists at  $t = T_1$ , it might not exist at  $t = T_2 < T_1$ . The shortest duration for which a solution can be found is called the *quantum speed limit* (QSL). The

speed at which the system can evolve is determined by the relevant energy scales. If the system only has access to finite energies, a lower bound for how fast the system can evolve exists, which in turn leads to the QSL [117].

The quantum speed limit can also be interpreted geometrically by examining how the state of the system moves through the Hilbert space. Considering the transfer  $|\psi\rangle \rightarrow |\chi\rangle$ , the fidelity  $F = |\langle\psi|\chi\rangle|$  is a measure of the overlap of the states. The fidelity is zero if the two states are orthogonal, while its maximum value of one is obtained if the states are identical up to a global phase. The distance between the two states in the Hilbert space can be interpreted as an angle [118]

$$\theta = \arccos(\sqrt{F}) , \quad (5.52)$$

where an angle of  $\theta = \pi/2$  corresponds to orthogonal states, while states of unit fidelity are equivalent to  $\theta = 0$ . The velocity of the state evolution can be defined as the rate of change of the angle, which is given by [119]

$$\frac{d\theta}{dt} = \Delta E , \quad (5.53)$$

where  $\Delta E$  is the energy spread given by the variance of the Hamiltonian

$$\Delta E = \sqrt{\langle\psi|\hat{H}^2|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle^2} . \quad (5.54)$$

The energy spread must obey the Heisenberg time-energy uncertainty principle providing a fundamental limit to the velocity in the Hilbert space. The quantum speed limit for the transfer between two arbitrary states  $|\psi\rangle$  and  $|\chi\rangle$  separated by the angle  $\theta$  was derived in [120] as

$$T_{\text{QSL}} = \frac{\theta}{\Delta E} . \quad (5.55)$$

However, the derivation of (5.55) assumes a time-independent Hamiltonian, which is very rarely the case in optimal control problems. The quantum speed limit for a time-dependent energy spread can be derived using arguments from differential geometry [119, 121]

$$\int_0^T \Delta E(t) dt \geq \theta . \quad (5.56)$$

Eq. (5.56) states that the path length transversed in the Hilbert space is bounded by the initial angle  $\theta$ . The quantum speed limit is reached when the time-dependent Hamiltonian realizes the most direct path in the Hilbert space between the states  $|\psi\rangle$  and  $|\chi\rangle$ . Note that eq. (5.56) does not directly provide the quantum speed limit, as the control functions corresponding to the shortest path are not known in general. The presence of a phase transition further complicated the problem, as the rate at which information is spread within the system is bounded by the effective light cone resulting from the local Hamiltonian. Hence, determining the quantum speed limit is often considered an optimization problem in itself.

The quantum speed limit can be extracted from an optimal solution by examining the relative motion of  $|\psi(t)\rangle$  in the Hilbert space. Assuming the state is evolved according the optimal Hamiltonian realizing the equality of eq. (5.56). Thus, the direct velocity in the Hilbert space is given by  $Q_{\text{opt}}(t) = \Delta E_{\text{opt}}(t)$ , and the obtained fidelity can be expressed as

a function of duration [121]

$$F(T) = \sin^2 \left( \int_0^T Q_{\text{opt}}(t) dt \right), \quad (5.57)$$

if the states  $|\psi\rangle$  and  $|\chi\rangle$  are orthogonal. Interpreting the state as moving along a geodesic in the Hilbert space between the two orthogonal states, the direct velocity is given by  $Q_{\text{opt}}(t) = -\frac{d\theta(t)}{dt}$ , whereby

$$F(T) = \sin^2 \left( \int_0^T -\frac{d\theta(t)}{dt} dt \right) = \sin^2 \left( -\theta(t) \Big|_0^T \right) = \sin^2 \left( \frac{\pi}{2} \frac{T}{T_{\text{QSL}}} \right). \quad (5.58)$$

The relation  $\theta(T) = \frac{\pi}{2} \left( 1 - \frac{T}{T_{\text{QSL}}} \right)$  is only valid if the state is moving on a geodesic, whereas eq. (5.57) is more general. A similar behavior to eq. (5.58) was reported in [114], where a number of systems were shown to behave according to the relation. However, it should be stressed that the relation (5.58) requires the initial and final states to be orthogonal and the system to be recast as an effective two-level system.

For an example of optimal control see Appendix D, which illustrates some of the consequences of the quantum speed limit.



## Chapter 6

# Results

Performing optimal control of quantum many-body systems is typically extremely resource demanding [97]. Hence, previous studies of optimization of Bose-Hubbard dynamics have abstained from using gradient-based techniques and instead opted for the gradient-free Nelder-Mead method combined with control parameterizations [19, 82]. The framework presented in this thesis combines tensor description of lattice systems with parameterized optimal control methods and highly-developed gradient-based algorithms from mathematical optimization theory. Here, the framework is used for optimizing the ramp sequence for transitioning from a superfluid to a Mott-Insulator in a Bose-Hubbard system, however, the method can easily be extended to any system in a one-dimensional lattice. Although the individual components of the framework are described in previous chapters, a summary of their unification is brought here:

The physical quantum system is represented using Matrix Product States described in Chapter 3. Through the many Schmidt decompositions used for creating and modifying the tensor networks weakly contributing eigenstates are discarded, whereby one avoids the otherwise exponential scaling of the numerical representation with physical size of the system. For the purposes of this thesis, the ITensor library [87] has proven itself very useful for tensor-operations. To simulate the dynamics of the system, a modified version of the tDMRG algorithm described in Section 4.2.2 is employed. The algorithm efficiently time-evolves the tensor network states while maintaining an accurate description of the system. In order to address the crossing of the phase transition in the Bose-Hubbard model, the state transfer is formulated as the minimization of a cost function (5.3) through the optimal control framework of Chapter 5. The GRAPE method supplies derivatives of the cost function with respect to the control parameters (5.33), whereby the optimization problem can be solved using gradient-based techniques. Further extensions are added through GROUP, which parametrizes the control through smooth functions (5.48) thereby reducing the dimension of the optimization space. Due to limits of the physical model, the control problem is subjected to a series of constraints, which makes the optimization very difficult. Interior Point methods are considered some of the most powerful algorithms for solving non-linear, constrained problems. Using the gradient of the cost function, the algorithm approaches the optimal control values from within the feasible region. Thus, interior point methods require fewer, but more expensive, steps to converge, which is a beneficial feature when using the computationally expensive GRAPE gradient. The version of the Interior Point method used here is implemented in IPOPT library [99].

In this chapter the performance and accuracy of the framework is analyzed. First, suitable optimization settings and boundary conditions are discussed. Next, the framework is applied to a small Bose-Hubbard system of five sites and five particles, and its performance is compared to optimizations performed using the Nelder-Mead method. Finally, the application of the framework on larger systems is discussed, and an analysis of the entanglement entropy during quenches past the critical point is made.

## 6.1 Determining Parameters

As previously discussed in section 4.2.2 the phase of Bose-Hubbard model is solely determined by the fraction  $U/J$ . Expressing the parameters in units of  $J$ , causes the tunneling part of the Hamiltonian to act as a drift, while the interaction matrix element,  $U$ , serves as the control function.

The time evolution is carried out using Suzuki-Trotter decomposed propagator, where the Trotter step has been chosen as  $\Delta t = 10^{-2}$ . The Trotter step is larger than in other studies [19, 82, 13], however, any smaller step-size would result in numerical costs outside the scope of this thesis.

### 6.1.1 Boundary Conditions and Constraints

The Bose-Hubbard model is only well defined within the tight-binding limit. Below this limit the Wannier functions may start overlapping with the next-nearest neighbouring sites thereby facilitating two-site hopping, which is not accounted for in the model. Thus, the control function must be subjected to a lower boundary. In [82, 19] initial lattice depths at  $V_0(0) = 3E_{\text{rec}}$  and  $2E_{\text{rec}}$  were chosen, where it was argued that the model is still a decent approximation at this depth. Taking a conservative approach, the control here is subjected to the constraint  $(U/J)(t) \geq 2$ , which corresponds to a lattice depth around three recoil energies for a gas of Rb-87 atoms and an optical lattice with wavelength  $\lambda = 1064$  nm. The initial control value was set slightly above the minimum value at  $(U/J)(0) = 2.5$ . Initially, the first control value was chosen to be equal to the lower bound, however, initiating the interior point algorithm at the edge of the feasible region caused the method to exhibit a poor convergence rate during its initial cycles. The final control value was set to  $(U/J)(T) = 50$ , which roughly corresponds to the final lattice depth of  $V_0(T) = 14E_{\text{rec}}$  used in [82].

The initial and target state were calculated from the control parameter at the start and end of the duration respectively using the DMRG algorithm. Thus, the states are insured to be the ground state at each end of the duration, whereby the optimized state transfer will bring between two ground states.

### 6.1.2 Seed Selection

The success of an optimization process is often dependent on the quality of the initial starting point or seed. Poor seeding strategies can lead to failure in finding optimal solutions when conducting local searches in complex optimization landscapes [122]. This has been demonstrated to occur in constrained quantum control problems [123]. Hence, the type of seed used for the optimization must be chosen carefully.

Different adiabatic lattice ramps from the superfluid to Mott-Insulator phase were examined in [124]. The study concluded that ramping the lattice slowly around the point of the phase transition results in an improved final fidelity, which is very similar to the process of sweeping over an avoided crossing [1].

Therefore, a ramp sequence was proposed in [124], which has an initial sigmoid shape followed by a slow increase in the lattice depth around the phase transition point. Following this, the lattice follows an exponential ramp to its final depth. Examining other attempts of optimizing the ramp sequence of the Bose-Hubbard model [19, 82] shows similar traits in their results. Therefore, choosing seeds with a slow ramp across the point of the phase transition followed by a rapid increase in lattice depth should yield good optimization results. Figure 6.1 shows the distribution of seeds used for the optimizations. Due to control parameter being the interaction strength of the Bose-Hubbard model, the phase transition occurs at quite a low control value. Hence, the initial part of the seed is a slowly increasing linear ramp, which crosses the point of the phase transition with a small slope. A shape

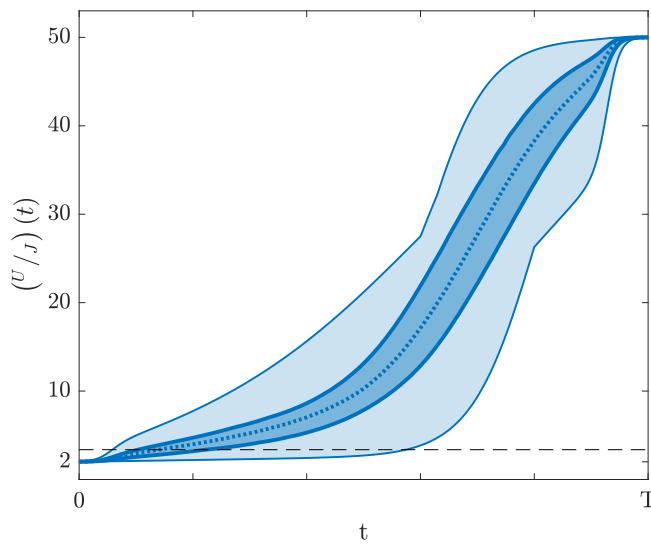


FIGURE 6.1: *Distribution of seeds used for optimization. All seeds lie within the lightly shaded region, while the darker region contains the 25-75 percentile of seeds. Lastly, the dotted line is the median value of the seed. The dashed line signifies the point of the Bose-Hubbard phase transition.*

function has been multiplied to the seeds enforcing a horizontal slope at the start and end of the duration, which helps avoiding any kinks in the control curve.

## 6.2 Characterization of Methods on a Small System

To determine suitable settings for the optimization, an initial analysis of a small system consisting of 5 sites with unit occupancy was made. Systems of different sizes are expected to evolve differently, especially due to the closed boundary conditions of the lattice, which have a greater impact for smaller systems. Although the dynamics are expected to vary with the lattice length, certain parameters can be optimized in a smaller system. The regularization factor,  $\gamma$ , and the dimension of the chopped basis,  $M$ , only affect the shape of the control pulse, which is subjected to constraints independent of the system size. Due to the small size of the system, no maximum bond dimension was required. Instead, a truncation threshold of  $\epsilon_t = 10^{-8}$  was used. Furthermore, the relative tolerance used by the IPOPT library for determining convergence was set at a very low  $1e - 8$ .

### 6.2.1 Determining the Quantum Speed Limit

To illustrate the achieved fidelity's dependence on duration of the control sequence, a series of optimizations were made while scanning over the duration  $T$ . For this a regularization factor of  $\gamma = 10^{-6}$  and an optimization space dimension of  $M = 20$  were used. Since the control is parametrized using a linear combination of smooth functions, the effect of the regularization term is limited. Nevertheless, it does help the optimization algorithm avoid greatly varying controls during its initial iterations.

Figure 6.2 shows the final fidelities obtained for various durations from 100 different initial controls. In part 6.2(a) the 25%- and 75%-quartiles along with the median of the solutions are displayed. At short durations the system can not evolve fast enough from the initial state to reach the target state, hence only low fidelities are obtained. However, as more and

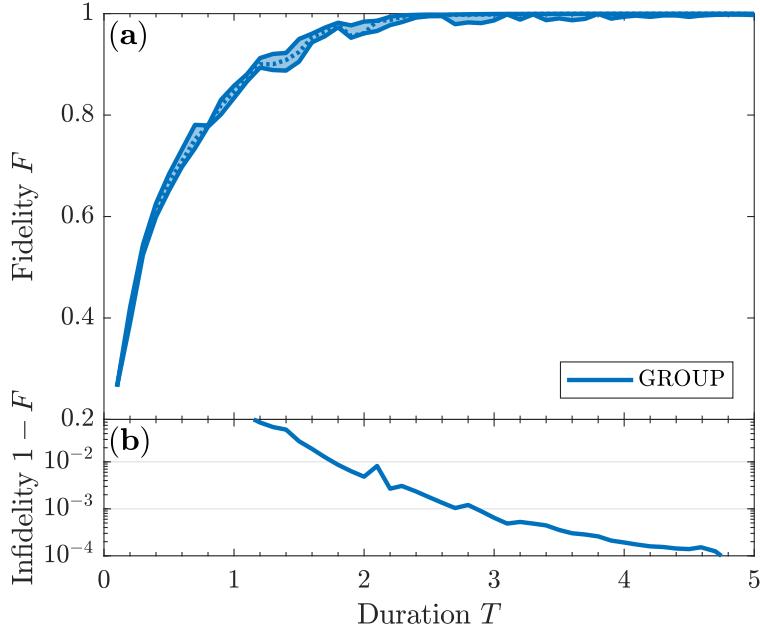


FIGURE 6.2: *Final fidelity obtained for optimal control at various durations. (a)* the dotted line marks the median fidelity achieved, while the shaded area displays the 25%- and 75%-quartiles of the solutions. *(b)* the lowest infidelity achieved for each duration.

more time becomes available, fidelities very close to unit can be reached. Although the final fidelity does not appear to improve beyond  $T = 3$ , figure 6.2(b) reveals the best solutions still improving at longer durations albeit at a reduced rate.

Compared to other figures of merit, the fidelity is a very strong measure, as it only considers the overlap with the target state. Due to the large number of quantum states in lattice systems, achieving unit fidelity is practically impossible, and even reaching high fidelities become increasingly difficult with system size. An alternative figure of merit was used for the optimal control of a Bose-Hubbard system presented in [82]. Instead of determining the convergence through the infidelity, a rescaled variance of the particle numbers in each of the central sites was used. This figure of merit is much less sensitive to the system size, however, it does have some less favorable properties: First, multiple particle configurations different from unit occupancy of each site can achieve a vanishing particle number variance. Furthermore, the pure state of a single particle on each lattice site only becomes the ground state as the lattice depth tends towards infinity. Therefore, the figure of merit used in [82] does not steer the state towards the ground state at the final control value, which could result in residual oscillations of the system after the control sequence. It should be noted that at the final control value used here, the state with a single particle at each site is a very good approximation of the ground state. Nevertheless, the fidelity is a more direct figure of merit, as it steers the system towards a single, well defined state.

Obtaining unit fidelity in a complex quantum system is very difficult, whereby a threshold for accepted final fidelity must be agreed on in order to determine the quantum speed limit. The threshold must be variable with the systems size, due to the sensitivity of the fidelity. In [125] a similar optimization was made, and an infidelity threshold of  $I_{QSL} = 10^{-3}$  was set for a 3-site lattice. Adopting the same threshold for the results shown in figure 6.2 yields a quantum speed limit of  $T_{QSL}^{(5)} = 2.7$  for the 5-site lattice.

Further information regarding the control landscape and the consistency of the algorithm

can be inferred from the spread in fidelities shown in figure 6.2.a. A large variety of final fidelities may be due to a control landscape with many local minima. On the other hand, the same phenomenon could be caused by the inability of an algorithm to accurately converge to the minimum. The interior point method is a very robust algorithm capable of handling non-linear constraints, whereby it is reasonable to assume the variations in solutions being caused by the underlying control landscape. Considering the fairly large variation in seeds shown in figure 6.1, the control landscape of the 5-site problem must be relatively simple, as the variation in obtained fidelity is rather small.

### 6.2.2 Optimized Ramp Sequences

A central question is how the optimized ramp sequence differs from its initial seed in both shape and fidelity reached. From figure 6.2 one observes that durations around  $T = 2.5$  consistently produce solutions with very high fidelity. The lowest infidelities reached at this duration are around  $I \approx 10^{-3}$ . Figure 6.3 displays an optimized ramp sequence at

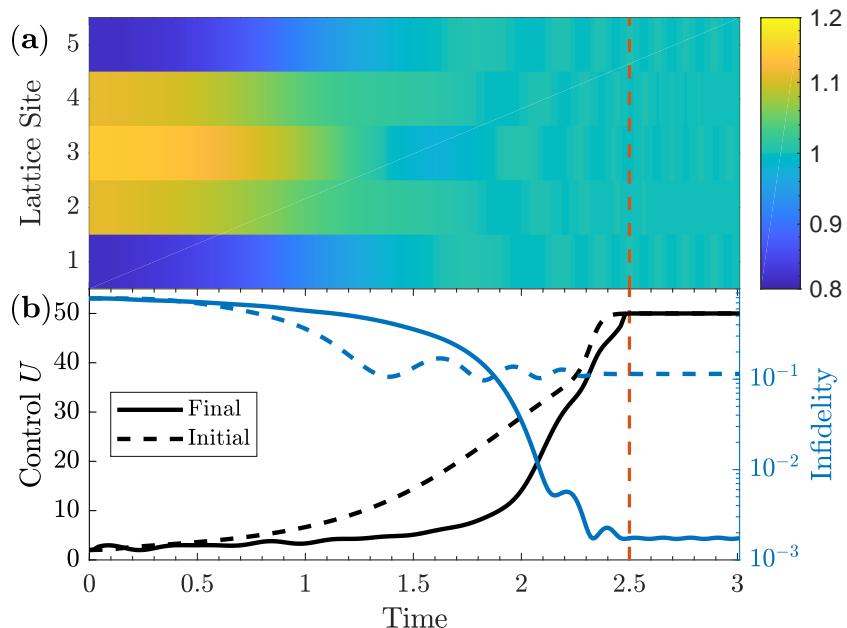


FIGURE 6.3: *Solution with highest fidelity achieved for duration  $T = 2.5$ . The dashed, red line marks the end of the duration, from which the system is further evolved using the final control value. (a) logarithmically scaled expectation value of the number operator,  $\langle \hat{n}_i \rangle$ , for each site as the system is evolved according to the optimized ramp. (b) the initial and optimized ramp sequences along with the corresponding evolution of infidelities.*

duration  $T = 2.5$  along with an illustration of the evolution of the system. First, consider part 6.3(b) in which the initial and optimized control is plotted together with the resulting infidelity. Optimizing the initial seed utilizing the GROUP algorithm has resulted in an increase of two magnitudes in the final fidelity achieved. While the initial ramp is fairly smooth, the optimized ramp wiggles a bit during the first half of the duration followed by a sharp upswing towards the final control value. The steep increase in the control towards the end is very similar to the adiabatic ramp sequence proposed in [124], however, the initial oscillating behavior is non-adiabatic. In fact, examining the optimized ramp sequences for durations shorter than  $T = 2.5$  shows an increase in the amplitude of the ramp oscillations, as solutions are required to be increasingly non-adiabatic in order to reach a good fidelity

within the given duration. Meanwhile, solutions for durations longer than  $T = 2.5$  become increasingly adiabatic-like.

After the given optimization duration,  $T$ , the final control value is kept constant for an additional time span, where the system is further evolved. Any residual evolution of the system at this point is due to the system not being an eigenstate of the Hamiltonian corresponding to final control value. When examining the initial and final ramp sequence after the optimization duration, the initial control appears to reach a more stable state. However, this is an illusion caused by the logarithmic axis, as the optimized ramp produces a state much closer to the target state, which is set as the ground state of  $\hat{H}(U(T))$  through the DMRG algorithm.

Figure 6.3(a) illustrates the expectation value of the number operator,  $\langle \hat{n}_i \rangle$ , for each site during the ramp. Initially the system is in the superfluid state causing a low energetic cost for having multiple particles on each site. Thus, one observes a reduced number of particles at the outer sites of the lattice due to the closed boundary conditions. As the system is evolved towards the Mott-Insulating state, having multiple particles at the same sites becomes very energetically unfavorable. Hence, the target state has a single particle at each site, as it is deep in the Mott regime. The optimized ramp brings the system very close to this configuration of particles, which can be observed near the dashed red line marking the end of the duration,  $T$ . Some imperfections are present, although their magnitude is exaggerated by the logarithmic color axis. The residual oscillations of the infidelity after the optimization durations can be seen in figure 6.3(a) as fluctuations of the particle number.

### 6.2.3 Control Basis Size

Unlike GRAPE, which takes the full optimization space into account, parameterizations like GROUP utilize a chopped basis resulting in a much smaller dimension of the optimization space. When reducing the dimension of the optimization space through a parametrization as described in eq. (5.48), the new control basis must be able to span the solution space. Otherwise artificial minima can be introduced to the control landscape due to the incompleteness of the basis [116]. In the implementation of GROUP used for these calculations, the basis functions of the chopped basis are of the form  $f_n = \sin(\omega_n t/T)$ , where  $\omega_n = n\pi$ . To investigate the basis size required to describe the optimal solutions, a series of optimizations were made for various basis sizes at the duration  $T = 2.5$ . The results of the scan over the basis size can be seen in figure 6.4. From figure 6.4(a) it is evident that a basis size of around 20 and above consistently produces high-fidelity solutions, while the quality of solutions using a smaller basis size fluctuated greatly. Thus, high-fidelity ramps for  $T = 2.5$  have at least 20 frequency components. Although results obtained with a smaller basis are generally worse, high fidelity solutions are still possible to achieve, as the optimal solution may still be realizable using only few frequencies. This can be observed in figure 6.4(b), where the infidelity of the best solution remains constant, while the median fidelity of the solutions decreases. The large spread in final fidelity is a consequence of a more complicated optimization landscape with artificial minima.

## 6.3 Application on Larger Systems

The framework already contains all the components required for applications on larger systems, however, in larger the systems the effects of the open boundaries of the lattice is much smaller, whereby a much higher entanglement entropy is expected. Therefore, optimization of larger systems is infeasible employing the very low degree of truncation used for the smaller system.

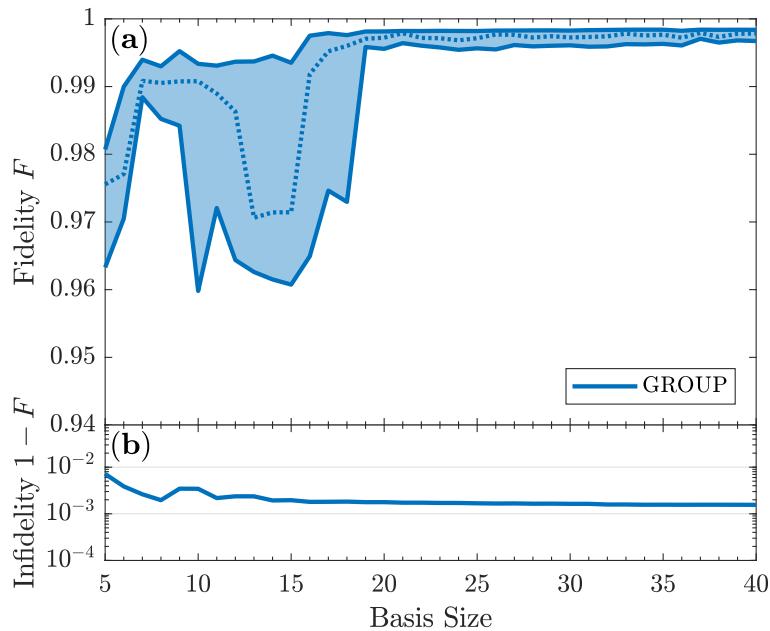


FIGURE 6.4: *Final fidelity obtained for optimal control at various optimization space dimensions for duration  $T = 2.5$ . (a)* the dotted line marks the median fidelity achieved, while the shaded area displays the 25%- and 75%-quartiles of the solutions. *(b)* the lowest infidelities achieved for each basis size.

When applying the same settings to a system of 15 sites the optimization became increasingly slow for longer durations. Since no maximum bond dimension was specified, the matrices of the tensor network continued to grow in size, as entanglement built up in the system following the phase transition. As the cost of applying the tDMRG propagator scales quadratically with the bond dimension of the state, the time-evolution became increasingly slow. Furthermore, the GRAPE algorithm stores the state at each time step to efficiently compute the gradient, which causes the program to run out of available memory in certain cases. Thus, an analysis of the growth of entanglement within the system is necessary, if the framework is to be applied to larger systems.

Further improvements to the convergence rate may be achieved by calculating the Hessian of the cost function. For the purpose of this thesis, the Hessian was approximate using the L-BFGS algorithm, however, this approximation only becomes sufficiently accurate after multiple iterations. Since interior point methods solve a series of subproblems for decreasing barrier heights, the L-BFGS approximation must be restarted multiple times thus reducing its effectiveness [99]. On the other hand, supplying interior point methods with analytically derived gradient should result in superlinear convergence [25].

A full derivation of the general GRAPE Hessian is presented in Appendix E along with the altered version for the Suzuki-Trotter propagator. If both the forward propagated initial state and the backwards propagated target state is kept at each time step, the computation of the Hessian requires an additional  $N^2/2 - N$  step propagations, where  $N$  is the total number of time steps. However, these propagations can be separated into  $N$  parallel calculations, which should significantly reduce the computational time of the Hessian. Note, that storing the backwards propagated target state would result in a doubling of the required memory occupied by the algorithm. Therefore, a suitable truncation of the tensor bond dimensions

is required for the implementation of the Hessian.

In addition to increasing the convergence rate, valuable information regarding the optimization landscape can be extracted from the Hessian [126]. Therefore, the addition of the Hessian to the framework may be a necessary step towards application on large systems, however, its implementation is outside the scope of this thesis.

### 6.3.1 Truncation of Tensors

A central operation in tensor network calculations is the singular value decomposition or Schmidt decomposition. From the entries of the resulting diagonal eigenmatrix, eigenstates with only little contribution to the current state can be inferred. These states can be discarded with minimal loss of accuracy of the tensorial representation. Discarding eigenstates effectively constitutes a truncation of the tensor bond dimension, whereby one avoids their otherwise exponential scaling. The two central parameters for the truncation are the maximum bond dimension,  $D$ , and the threshold (or conversely error) for when a state is discarded,  $\epsilon_t$ . As discussed in ..... the entanglement entropy increases rapidly around the point of the phase transition. Therefore, the bond dimension required to describe a ground state near the critical point is much higher than otherwise needed for the system. As the system is quenched, the initial ground state decays into a superposition of many eigenstates, whereby the truncation error will rise if a maximum bond dimension is kept [75]. Therefore, the system is not in the ground state during the phase transition.

To analyze the dynamical scaling of the bond dimension, a system with  $L, N_p = 15$  and truncation threshold  $\epsilon_t = 10^{-8}$  was time-evolved using an adiabatic-like ramp shown in figure 6.5(a). Infidelities calculated for various maximum bond dimensions are illustrated

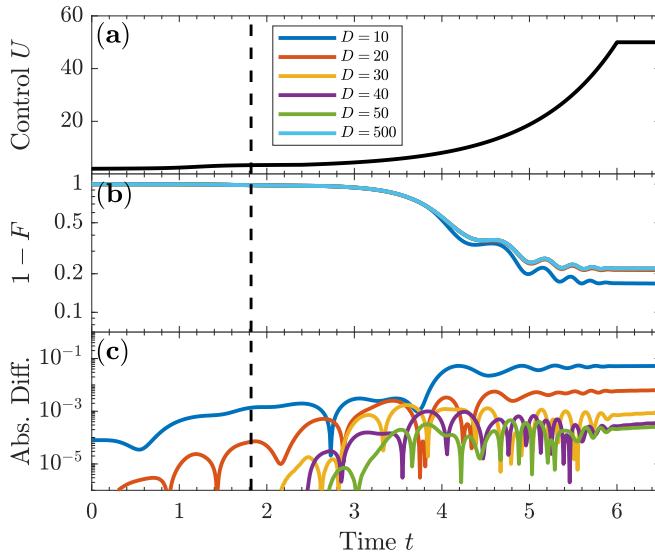


FIGURE 6.5: Ramp sequence for a series of maximum bond dimensions,  $D$ , and constant truncation threshold  $\epsilon_t = 10^{-8}$ . The Bose-Hubbard phase transition is marked by a dashed line. (a) the ramp used. (b) infidelities obtained for various values of  $D$ . (c) absolute difference in infidelity between non-bounded bond dimension ( $D = 500$ ) and remaining bond dimensions.

in 6.5(b-c). Here, the value  $D = 500$  has no influence on the bond dimensions, as the truncation is dominated by the threshold. The results obtained for different values of  $D$  are very similar initially but starts diverging as time passes. This behavior is expected, as

the initial differences between the tensorial representations are amplified during the time-evolution. Nevertheless, the variance in obtained infidelity is relatively small, which confirms the weak scaling with maximum bond dimension reported in [75].

To understand the magnitude of the truncations, all the bond dimension of the non-bounded ( $D = 500$ ) state are plotted for the ramp duration, which can be seen in figure 6.6(a). Initially the bond dimension remains constant, as the ramp is not very steep,

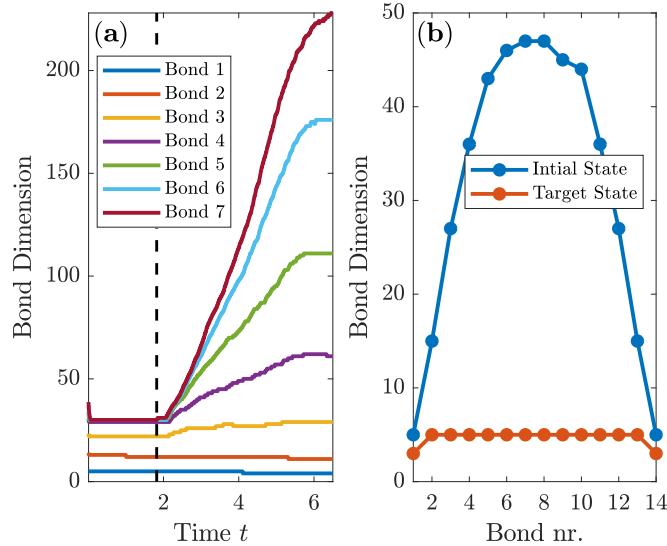


FIGURE 6.6: *Bond dimensions for MPS with  $L, N_p = 15$  and  $\epsilon_t = 10^{-8}$ .* (a) *evolution of bond dimensions as system is subjected to ramp shown in fig. 6.5(a).* The dashed line marks the Bose-Hubbard phase transition. (b) *bond dimensions of boundary states.*

whereby the threshold-based truncation keeps the bond dimension stable. At the point of the phase transition the ramp is still almost flat, while the bond dimension suddenly starts increasing rapidly. Although the bond dimension is expected to increase over time [75], its sudden increase must be attributed to the large entanglement entropy at the critical point. Unexpectedly, the bond dimensions continue to increase at the same rate for a long time after the phase transition. While this is not completely understood, it can be attributed to either physical dynamics of quasiparticles excited at the critical point [13], or mathematical dynamics, as information is propagated at limited speed through tensor network via the Schmidt decompositions. The bond dimensions of the ground state Mott-Insulator at  $U/J = 50$  shown in figure 6.6(b) are much smaller than those of the evolved state at the end of the ramp. Examining the final fidelity displayed in figure 6.5(b) reveals that the final state is indeed not in the ground state. Nevertheless, the final bond dimension is remarkably high, but truncating it has only little effect on the fidelity as discussed earlier. Thus, much of the information held in the bond indices at the end of the duration must be redundant, whereby it could be remnants of the phase transition not yet truncated away.

Although the convergence may only scale weakly with the maximum bond dimension, truncating the tensors of the MPS may affect the gradients, which are required to be accurate in order for the L-BFGS approximation to be valid [95]. Figure 6.7 shows the results of calculating the gradient elements 5.33 for the ramp shown in figure 6.5(a). Compared with the infidelity, the gradient scales more strongly with the maximum bond dimension. Therefore, the truncation is limited by the derivative of the cost function rather than the function itself. Unlike the infidelity, the gradient becomes gradually better during the ramp sequence. The same phenomenon is discussed in section 5.1.1 and is caused by the

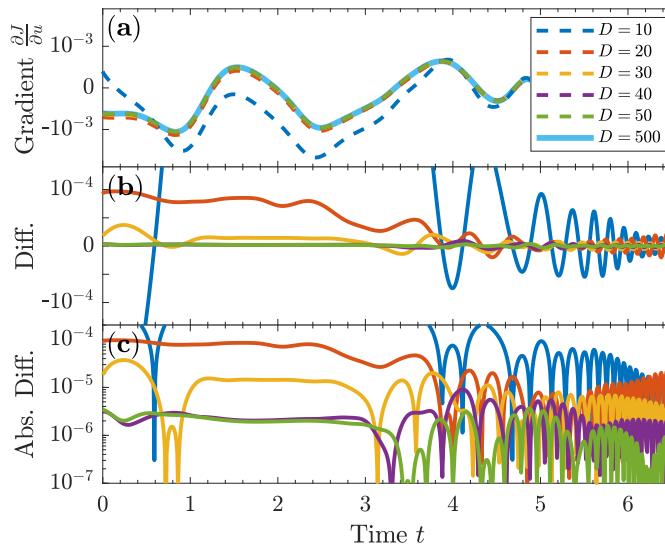


FIGURE 6.7: (a) GRAPE gradient elements calculated for various maximum bond dimension. (b) difference of derivatives between calculations for  $D = 500$  and remaining bond dimensions. (c) absolute differences of derivatives.

accumulated error when calculating  $|\chi(0)\rangle$ . Lastly, the gradient shows no signs of being influenced by the phase transition. Thus, gradient-based methods appear to be very well suited for optimal control problems, which involves crossing a critical point.

# Bibliography

- [1] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger. “Many-body physics with ultracold gases”. In: *Rev. Mod. Phys.* 80 (3 2008), pp. 885–964. DOI: 10.1103/RevModPhys.80.885. URL: <https://link.aps.org/doi/10.1103/RevModPhys.80.885>.
- [2] Immanuel Bloch, Jean Dalibard, and Sylvain Nascimbene. “Quantum simulations with ultracold quantum gases”. In: *Nature Physics* 8.4 (2012), pp. 267–276.
- [3] D. Jaksch and P. Zoller. “The cold atom Hubbard toolbox”. In: *Annals of Physics* 315.1 (2005). Special Issue, pp. 52 –79. ISSN: 0003-4916. DOI: <https://doi.org/10.1016/j.aop.2004.09.010>. URL: <http://www.sciencedirect.com/science/article/pii/S0003491604001782>.
- [4] Rudolf Grimm, Matthias Weidmüller, and Yurii B. Ovchinnikov. “Optical Dipole Traps for Neutral Atoms”. In: ed. by Benjamin Bederson and Herbert Walther. Vol. 42. Advances In Atomic, Molecular, and Optical Physics Supplement C. Academic Press, 2000, pp. 95 –170. DOI: [https://doi.org/10.1016/S1049-250X\(08\)60186-X](https://doi.org/10.1016/S1049-250X(08)60186-X). URL: <http://www.sciencedirect.com/science/article/pii/S1049250X0860186X>.
- [5] Enric Jané et al. “Simulation of quantum dynamics with quantum optical systems”. In: *Quantum Information & Computation* 3.1 (2003), pp. 15–37.
- [6] Dieter Jaksch and Peter Zoller. “Creation of effective magnetic fields in optical lattices: the Hofstadter butterfly for cold neutral atoms”. In: *New Journal of Physics* 5.1 (2003), p. 56.
- [7] D. Jaksch et al. “Entanglement of Atoms via Cold Controlled Collisions”. In: *Phys. Rev. Lett.* 82 (9 1999), pp. 1975–1978. DOI: 10.1103/PhysRevLett.82.1975. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.82.1975>.
- [8] Olaf Mandel et al. “Controlled collisions for multi-particle entanglement of optically trapped atoms”. In: *Nature* 425.6961 (2003), p. 937.
- [9] D Jaksch et al. “Fast quantum gates for neutral atoms”. In: *Physical Review Letters* 85.10 (2000), p. 2208.
- [10] A Micheli et al. “Single atom transistor in a 1d optical lattice”. In: *Physical review letters* 93.14 (2004), p. 140408.
- [11] Maciej Lewenstein et al. “Ultracold atomic gases in optical lattices: mimicking condensed matter physics and beyond”. In: *Advances in Physics* 56.2 (2007), pp. 243–379. DOI: 10.1080/00018730701223200. eprint: <https://doi.org/10.1080/00018730701223200>. URL: <https://doi.org/10.1080/00018730701223200>.
- [12] Wojciech H. Zurek, Uwe Dorner, and Peter Zoller. “Dynamics of a Quantum Phase Transition”. In: *Phys. Rev. Lett.* 95 (10 2005), p. 105701. DOI: 10.1103/PhysRevLett.95.105701. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.95.105701>.
- [13] Simon Braun et al. “Emergence of coherence and the dynamics of quantum phase transitions”. In: *Proceedings of the National Academy of Sciences* 112.12 (2015), pp. 3641–3646.

- [14] Anthony P Peirce, Mohammed A Dahleh, and Herschel Rabitz. “Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications”. In: *Physical Review A* 37.12 (1988), p. 4950.
- [15] J Werschnik and EKU Gross. “Quantum optimal control theory”. In: *Journal of Physics B: Atomic, Molecular and Optical Physics* 40.18 (2007), R175.
- [16] Herschel A Rabitz, Michael M Hsieh, and Carey M Rosenthal. “Quantum optimally controlled transition landscapes”. In: *Science* 303.5666 (2004), pp. 1998–2001.
- [17] Navin Khaneja et al. “Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms”. In: *Journal of magnetic resonance* 172.2 (2005), pp. 296–305.
- [18] Vadim Krotov. *Global methods in optimal control theory*. Vol. 195. CRC Press, 1995.
- [19] Patrick Doria, Tommaso Calarco, and Simone Montangero. “Optimal Control Technique for Many-Body Quantum Dynamics”. In: *Phys. Rev. Lett.* 106 (19 2011), p. 190501. DOI: 10.1103/PhysRevLett.106.190501. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.106.190501>.
- [20] Guifré Vidal. “Efficient classical simulation of slightly entangled quantum computations”. In: *Physical review letters* 91.14 (2003), p. 147902.
- [21] Ulrich Schollwöck. “The density-matrix renormalization group in the age of matrix product states”. In: *Annals of Physics* 326.1 (2011). January 2011 Special Issue, pp. 96 –192. ISSN: 0003-4916. DOI: <https://doi.org/10.1016/j.aop.2010.09.012>. URL: <http://www.sciencedirect.com/science/article/pii/S0003491610001752>.
- [22] J. Eisert, M. Cramer, and M. B. Plenio. “Colloquium: Area laws for the entanglement entropy”. In: *Rev. Mod. Phys.* 82 (1 2010), pp. 277–306. DOI: 10.1103/RevModPhys.82.277. URL: <https://link.aps.org/doi/10.1103/RevModPhys.82.277>.
- [23] S Bravyi, MB Hastings, and F Verstraete. “Lieb-Robinson bounds and the generation of correlations and topological quantum order”. In: *Physical review letters* 97.5 (2006), p. 050401.
- [24] Jens Eisert and Tobias J Osborne. “General entanglement scaling laws from time evolution”. In: *Physical review letters* 97.15 (2006), p. 150404.
- [25] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. second. New York, NY, USA: Springer, 2006.
- [26] Mike H Anderson et al. “Observation of Bose-Einstein condensation in a dilute atomic vapor”. In: *science* 269.5221 (1995), pp. 198–201.
- [27] Jonathan Simon et al. “Quantum simulation of antiferromagnetic spin chains in an optical lattice”. In: *Nature* 472.7343 (2011), pp. 307–312.
- [28] Leticia Tarruell et al. “Creating, moving and merging Dirac points with a Fermi gas in a tunable honeycomb lattice”. In: *Nature* 483.7389 (2012), pp. 302–305.
- [29] Jean Dalibard et al. “Colloquium: Artificial gauge potentials for neutral atoms”. In: *Reviews of Modern Physics* 83.4 (2011), p. 1523.
- [30] M. Saffman, T. G. Walker, and K. Mølmer. “Quantum information with Rydberg atoms”. In: *Rev. Mod. Phys.* 82 (3 2010), pp. 2313–2363. DOI: 10.1103/RevModPhys.82.2313. URL: <https://link.aps.org/doi/10.1103/RevModPhys.82.2313>.
- [31] Claude Cohen-Tannoudji et al. *Atom-photon interactions: basic processes and applications*. Wiley Online Library, 1992.
- [32] Artur Widera. “Constructing correlated spin states with neutral atoms in optical lattices”. PhD thesis. Ph. D. thesis, Johannes Gutenberg-Universität, Mainz, 2008.

- [33] C. Kittel. *Introduction to Solid State Physics*. Wiley, 2004. ISBN: 9780471415268. URL: <https://books.google.dk/books?id=kym4QgAACAAJ>.
- [34] Markus Greiner. “Ultracold quantum gases in three-dimensional optical lattice potentials”. PhD thesis. 2003. URL: <http://nbn-resolving.de/urn:nbn:de:bvb:19-9683>.
- [35] Charles Kittel and Ching-yao Fong. *Quantum theory of solids*. Vol. 3. Wiley New York, 1963.
- [36] Walter Kohn. “Analytic properties of Bloch waves and Wannier functions”. In: *Physical Review* 115.4 (1959), p. 809.
- [37] Christopher J Pethick and Henrik Smith. *Bose-Einstein condensation in dilute gases*. Cambridge university press, 2002.
- [38] Oliver Penrose and Lars Onsager. “Bose-Einstein Condensation and Liquid Helium”. In: *Phys. Rev.* 104 (3 1956), pp. 576–584. DOI: 10.1103/PhysRev.104.576. URL: <https://link.aps.org/doi/10.1103/PhysRev.104.576>.
- [39] H. Bruus and K. Flensberg. *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction*. Oxford Graduate Texts. OUP Oxford, 2004. ISBN: 9780198566335. URL: <https://books.google.dk/books?id=v5vhg1tYLC8C>.
- [40] E. P. Gross. “Structure of a quantized vortex in boson systems”. In: *Il Nuovo Cimento (1955-1965)* 20.3 (1961), pp. 454–477. ISSN: 1827-6121. DOI: 10.1007/BF02731494. URL: <https://doi.org/10.1007/BF02731494>.
- [41] L.P. Pitaevskii. “Vortex Lines in an Imperfect Bose Gas”. In: *Soviet Physics JETP* 13.2 (1961), pp. 451–454.
- [42] D. Jaksch et al. “Cold Bosonic Atoms in Optical Lattices”. In: *Phys. Rev. Lett.* 81 (15 1998), pp. 3108–3111. DOI: 10.1103/PhysRevLett.81.3108. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.81.3108>.
- [43] Nathan Gemelke et al. “In situ observation of incompressible mott-insulating domains in ultracold atomic gases”. In: *Nature* 460.7258 (2009), p. 995.
- [44] Subir Sachdev. *Quantum phase transitions*. Wiley Online Library, 2007.
- [45] D. van Oosten, P. van der Straten, and H. T. C. Stoof. “Quantum phases in an optical lattice”. In: *Phys. Rev. A* 63 (5 2001), p. 053601. DOI: 10.1103/PhysRevA.63.053601. URL: <https://link.aps.org/doi/10.1103/PhysRevA.63.053601>.
- [46] M. Plischke and B. Bergersen. *Equilibrium Statistical Physics*. World Scientific, 2006. ISBN: 9789812560483. URL: <https://books.google.dk/books?id=KYu7igYEkhwC>.
- [47] Till D. Kühner, Steven R. White, and H. Monien. “One-dimensional Bose-Hubbard model with nearest-neighbor interaction”. In: *Phys. Rev. B* 61 (18 2000), pp. 12474–12489. DOI: 10.1103/PhysRevB.61.12474. URL: <https://link.aps.org/doi/10.1103/PhysRevB.61.12474>.
- [48] Markus Greiner et al. “Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms”. In: *nature* 415.6867 (2002), p. 39.
- [49] Lev D Landau. “Zur theorie der energieübertragung. ii”. In: *Phys. Z. Sowjetunion* 2.46 (1932), pp. 1–13.
- [50] Clarence Zener. “Non-adiabatic crossing of energy levels”. In: *Proc. R. Soc. Lond. A* 137.833 (1932), pp. 696–702.
- [51] Anatoli Polkovnikov and Vladimir Gritsev. “Breakdown of the adiabatic limit in low-dimensional gapless systems”. In: *Nature Physics* 4.6 (2008), p. 477.

- [52] R. Ozeri et al. “Colloquium: Bulk Bogoliubov excitations in a Bose-Einstein condensate”. In: *Rev. Mod. Phys.* 77 (1 2005), pp. 187–205. DOI: 10.1103/RevModPhys.77.187. URL: <https://link.aps.org/doi/10.1103/RevModPhys.77.187>.
- [53] Thomas WB Kibble. “Topology of cosmic domains and strings”. In: *Journal of Physics A: Mathematical and General* 9.8 (1976), p. 1387.
- [54] Edward Gillman and Arttu Rajantie. “Kibble Zurek mechanism of topological defect formation in quantum field theory with matrix product states”. In: *Phys. Rev. D* 97 (9 2018), p. 094505. DOI: 10.1103/PhysRevD.97.094505. URL: <https://link.aps.org/doi/10.1103/PhysRevD.97.094505>.
- [55] Ralf Schützhold et al. “Sweeping from the superfluid to the Mott phase in the Bose-Hubbard model”. In: *Physical review letters* 97.20 (2006), p. 200601.
- [56] Jens Eisert, Mathis Friesdorf, and Christian Gogolin. “Quantum many-body systems out of equilibrium”. In: *Nature Physics* 11.2 (2015), p. 124.
- [57] Elliott H Lieb and Derek W Robinson. “The finite group velocity of quantum spin systems”. In: *Statistical Mechanics*. Springer, 1972, pp. 425–431.
- [58] Marc Cheneau et al. “Light-cone-like spreading of correlations in a quantum many-body system”. In: *Nature* 481.7382 (2012), p. 484.
- [59] Pasquale Calabrese and John Cardy. “Time dependence of correlation functions following a quantum quench”. In: *Physical review letters* 96.13 (2006), p. 136801.
- [60] Andreas M Läuchli and Corinna Kollath. “Spreading of correlations and entanglement after a quench in the one-dimensional Bose-Hubbard model”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.05 (2008), P05018.
- [61] J. M. Zhang and R. X. Dong. “Exact diagonalization: the Bose-Hubbard model as an example”. In: *European Journal of Physics* 31 (May 2010), pp. 591–602. DOI: 10.1088/0143-0807/31/3/016. arXiv: 1102.4006 [cond-mat.stat-mech].
- [62] Steven R. White. “Density matrix formulation for quantum renormalization groups”. In: *Phys. Rev. Lett.* 69 (19 1992), pp. 2863–2866. DOI: 10.1103/PhysRevLett.69.2863. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.69.2863>.
- [63] Steven R White. “Density-matrix algorithms for quantum renormalization groups”. In: *Physical Review B* 48.14 (1993), p. 10345.
- [64] Anirban Pathak. *Elements of quantum computation and quantum communication*. Taylor & Francis, 2013.
- [65] M. B. Hastings. “An area law for one-dimensional quantum systems”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.08 (2007), P08024.
- [66] Norbert Schuch et al. “Entropy Scaling and Simulability by Matrix Product States”. In: *Phys. Rev. Lett.* 100 (3 2008), p. 030504. DOI: 10.1103/PhysRevLett.100.030504. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.100.030504>.
- [67] G. Vidal. “Classical Simulation of Infinite-Size Quantum Lattice Systems in One Spatial Dimension”. In: *Phys. Rev. Lett.* 98 (7 2007), p. 070201. DOI: 10.1103/PhysRevLett.98.070201. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.98.070201>.
- [68] Ian P McCulloch. “From density-matrix renormalization group to matrix product states”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.10 (2007), P10014.

- [69] Satoshi Ejima et al. “Characterization of Mott-insulating and superfluid phases in the one-dimensional Bose-Hubbard model”. In: *Phys. Rev. A* 85 (5 2012), p. 053644. DOI: 10.1103/PhysRevA.85.053644. URL: <https://link.aps.org/doi/10.1103/PhysRevA.85.053644>.
- [70] RJ Baxter. “Dimers on a rectangular lattice”. In: *Journal of Mathematical Physics* 9.4 (1968), pp. 650–654.
- [71] Ian Affleck et al. “Rigorous results on valence-bond ground states in antiferromagnets”. In: *Physical review letters* 59.7 (1987), p. 799.
- [72] Stellan Östlund and Stefan Rommer. “Thermodynamic limit of density matrix renormalization”. In: *Physical review letters* 75.19 (1995), p. 3537.
- [73] J Dukelsky. “J. Dukelsky, MA Martín-Delgado, T. Nishino, and G. Sierra, Europhys. Lett. 43, 457 (1998).” In: *Europhys. Lett.* 43 (1998), p. 457.
- [74] Guifré Vidal. “Efficient simulation of one-dimensional quantum many-body systems”. In: *Physical review letters* 93.4 (2004), p. 040502.
- [75] Andrew John Daley et al. “Time-dependent density-matrix renormalization-group using adaptive effective Hilbert spaces”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2004.04 (2004), P04005.
- [76] Norbert Schuch et al. “On entropy growth and the hardness of simulating time evolution”. In: *New Journal of Physics* 10.3 (2008), p. 033032.
- [77] Cornelius Lanczos. “An iteration method for the solution of the eigenvalue problem of linear differential and integral operators”. In: *J. Res. Natl. Bur. Stand. B* 45 (1950), pp. 255–282. DOI: 10.6028/jres.045.026.
- [78] Steven R. White. “Density matrix renormalization group algorithms with a single center site”. In: *Phys. Rev. B* 72 (18 2005), p. 180403. DOI: 10.1103/PhysRevB.72.180403. URL: <https://link.aps.org/doi/10.1103/PhysRevB.72.180403>.
- [79] Frank Verstraete, Juan J Garcia-Ripoll, and Juan Ignacio Cirac. “Matrix product density operators: Simulation of finite-temperature and dissipative systems”. In: *Physical review letters* 93.20 (2004), p. 207204.
- [80] Marko Žnidarič, Tomaž Prosen, and Peter Prelovšek. “Many-body localization in the Heisenberg X X Z magnet in a random field”. In: *Physical Review B* 77.6 (2008), p. 064426.
- [81] MA Cazalilla and JB Marston. “Time-dependent density-matrix renormalization group: A systematic method for the study of quantum many-body out-of-equilibrium systems”. In: *Physical review letters* 88.25 (2002), p. 256403.
- [82] S Van Frank et al. “Optimal control of complex atomic quantum systems”. In: *Scientific reports* 6 (2016).
- [83] Masuo Suzuki. “General theory of fractal path integrals with applications to many-body theories and statistical physics”. In: *Journal of Mathematical Physics* (1991).
- [84] Daniel A Steck. “Quantum and atom optics”. In: *Oregon Center for Optics and Department of Physics, University of Oregon* 47 (2007).
- [85] Haruo Yoshida. “Construction of higher order symplectic integrators”. In: *Physics letters A* 150.5-7 (1990), pp. 262–268.
- [86] M Krech, Alex Bunker, and DP Landau. “Fast spin dynamics algorithms for classical spin systems”. In: *Computer physics communications* 111.1-3 (1998), pp. 1–13.
- [87] E. Miles Stoudenmire and Steven R. White. *ITensor Library*. URL: <http://itensor.org/index.html>.

- [88] Michael P. Zaletel et al. “Time-evolving a matrix product state with long-ranged interactions”. In: *Phys. Rev. B* 91 (16 2015), p. 165112. DOI: 10.1103/PhysRevB.91.165112. URL: <https://link.aps.org/doi/10.1103/PhysRevB.91.165112>.
- [89] Stuart Alan Rice, Meishan Zhao, et al. *Optical control of molecular dynamics*. John Wiley, 2000.
- [90] Moshe Shapiro and Paul Brumer. “Principles of the quantum control of molecular processes”. In: *Principles of the Quantum Control of Molecular Processes, by Moshe Shapiro, Paul Brumer, pp. 250. ISBN 0-471-24184-9. Wiley-VCH, February 2003.* (2003), p. 250.
- [91] Georg Jäger et al. “Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Comparison of gradient-ascent-pulse-engineering and Krotov optimization schemes”. In: *Physical Review A* 90.3 (2014), p. 033628.
- [92] Ulrich Hohenester et al. “Optimal quantum control of Bose-Einstein condensates in magnetic microtraps”. In: *Physical Review A* 75.2 (2007), p. 023602.
- [93] Gregory Von Winckel and A Borzì. “Computational techniques for a quantum control problem with H1-cost”. In: *Inverse Problems* 24.3 (2008), p. 034007.
- [94] A. Borzì and U. Hohenester. “Multigrid Optimization Schemes for Solving Bose–Einstein Condensate Control Problems”. In: *SIAM Journal on Scientific Computing* 30.1 (2008), pp. 441–462. DOI: 10.1137/070686135. eprint: <https://doi.org/10.1137/070686135>. URL: <https://doi.org/10.1137/070686135>.
- [95] P De Fouquieres et al. “Second order gradient ascent pulse engineering”. In: *Journal of Magnetic Resonance* 212.2 (2011), pp. 412–417.
- [96] R. M. Wilcox. “Exponential Operators and Parameter Differentiation in Quantum Physics”. In: *Journal of Mathematical Physics* 8.4 (1967), pp. 962–982. DOI: 10.1063/1.1705306. eprint: <https://doi.org/10.1063/1.1705306>. URL: <https://doi.org/10.1063/1.1705306>.
- [97] Jan-Frederik Mennemann et al. “Optimal control of Bose–Einstein condensates in three dimensions”. In: *New Journal of Physics* 17.11 (2015), p. 113027.
- [98] Narendra Karmarkar. “A new polynomial-time algorithm for linear programming”. In: *Proceedings of the sixteenth annual ACM symposium on Theory of computing*. ACM, 1984, pp. 302–311.
- [99] Andreas Wächter and Lorenz T Biegler. “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming”. In: *Mathematical programming* 106.1 (2006), pp. 25–57.
- [100] Richard H Byrd, Guanghui Liu, and Jorge Nocedal. “On the local behavior of an interior point method for nonlinear programming”. In: *Numerical analysis 1997* (1997), pp. 37–56.
- [101] Anthony V Fiacco and Garth P McCormick. *Nonlinear programming: sequential unconstrained minimization techniques*. Vol. 4. Siam, 1990.
- [102] Richard A Waltz et al. “An interior algorithm for nonlinear optimization that combines line search and trust region steps”. In: *Mathematical programming* 107.3 (2006), pp. 391–408.
- [103] Andreas Wächter and Lorenz T Biegler. “Line search filter methods for nonlinear programming: Motivation and global convergence”. In: *SIAM Journal on Optimization* 16.1 (2005), pp. 1–31.

- [104] Roger Fletcher and Sven Leyffer. “Nonlinear programming without a penalty function”. In: *Mathematical programming* 91.2 (2002), pp. 239–269.
- [105] Robert J Vanderbei and David F Shanno. “An interior-point algorithm for nonconvex nonlinear programming”. In: *Computational Optimization and Applications* 13.1-3 (1999), pp. 231–252.
- [106] R Tyrrell Rockafellar. *Conjugate duality and optimization*. Vol. 16. Siam, 1974.
- [107] Harold W Kuhn and Albert W Tucker. “Nonlinear programming”. In: *Traces and emergence of nonlinear programming*. Springer, 2014, pp. 247–258.
- [108] Thomas F Coleman and Yuying Li. “An interior trust region approach for nonlinear minimization subject to bounds”. In: *SIAM Journal on optimization* 6.2 (1996), pp. 418–445.
- [109] Dong C Liu and Jorge Nocedal. “On the limited memory BFGS method for large scale optimization”. In: *Mathematical programming* 45.1-3 (1989), pp. 503–528.
- [110] Jens Jakob Sørensen et al. “Gradient Based Quantum Optimal Control in a Chopped Basis”. In: *arXiv preprint arXiv:1802.07509* (2018).
- [111] Justin Ruths and Jr-Shin Li. “Optimal control of inhomogeneous ensembles”. In: *IEEE Transactions on Automatic Control* 57.8 (2012), pp. 2021–2032.
- [112] Liang Jiang et al. “Preparation of decoherence-free cluster states with optical superlattices”. In: *Phys. Rev. A* 79 (2 2009), p. 022309. DOI: 10.1103/PhysRevA.79.022309. URL: <https://link.aps.org/doi/10.1103/PhysRevA.79.022309>.
- [113] Tommaso Caneva, Tommaso Calarco, and Simone Montangero. “Chopped random-basis quantum optimization”. In: *Physical Review A* 84.2 (2011), p. 022326.
- [114] Tommaso Caneva et al. “Speeding up critical system dynamics through optimized evolution”. In: *Phys. Rev. A* 84 (1 2011), p. 012312. DOI: 10.1103/PhysRevA.84.012312. URL: <https://link.aps.org/doi/10.1103/PhysRevA.84.012312>.
- [115] S Lloyd and S Montangero. “Information theoretical analysis of quantum optimal control”. In: *Physical review letters* 113.1 (2014), p. 010502.
- [116] Niklas Rach et al. “Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape”. In: *Physical Review A* 92.6 (2015), p. 062343.
- [117] T. Caneva et al. “Optimal Control at the Quantum Speed Limit”. In: *Phys. Rev. Lett.* 103 (24 2009), p. 240501. DOI: 10.1103/PhysRevLett.103.240501. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.103.240501>.
- [118] William K Wootters. “Statistical distance and Hilbert space”. In: *Physical Review D* 23.2 (1981), p. 357.
- [119] J. Anandan and Y. Aharonov. “Geometry of quantum evolution”. In: *Phys. Rev. Lett.* 65 (14 1990), pp. 1697–1700. DOI: 10.1103/PhysRevLett.65.1697. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.65.1697>.
- [120] L. Mandelstam and Ig. Tamm. “The Uncertainty Relation Between Energy and Time in Non-relativistic Quantum Mechanics”. In: *Selected Papers*. Ed. by Boris M. Bolognoscii, Victor Ya. Frenkel, and Rudolf Peierls. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991, pp. 115–123. ISBN: 978-3-642-74626-0. DOI: 10.1007/978-3-642-74626-0\_8. URL: [https://doi.org/10.1007/978-3-642-74626-0\\_8](https://doi.org/10.1007/978-3-642-74626-0_8).
- [121] Miroslav Gajdacz et al. “Time-limited optimal dynamics beyond the quantum speed limit”. In: *Phys. Rev. A* 92 (6 2015), p. 062106. DOI: 10.1103/PhysRevA.92.062106. URL: <https://link.aps.org/doi/10.1103/PhysRevA.92.062106>.

- [122] Jens Jakob WH Sørensen et al. “Exploring the quantum speed limit with computer games”. In: *Nature* 532.7598 (2016), p. 210.
- [123] Dmitry V Zhdanov and Tamar Seideman. “Role of control constraints in quantum optimal control”. In: *Physical Review A* 92.5 (2015), p. 052109.
- [124] Jakub Zakrzewski and Dominique Delande. “Breakdown of adiabaticity when loading ultracold atoms in optical lattices”. In: *Physical Review A* 80.1 (2009), p. 013602.
- [125] Maja Juhl and Julie Kryger. “Optimisation Method Development and Testing for Highly Complex Quantum Dynamics”. MA thesis. Department of Physics and Astronomy, Aarhus University, 2018.
- [126] Zhenwen Shen, Michael Hsieh, and Herschel Rabitz. “Quantum optimal control: Hessian analysis of the control landscape”. In: *The Journal of chemical physics* 124.20 (2006), p. 204106.
- [127] I. B. Spielman, W. D. Phillips, and J. V. Porto. “Condensate Fraction in a 2D Bose Gas Measured across the Mott-Insulator Transition”. In: *Phys. Rev. Lett.* 100 (12 2008), p. 120402. DOI: 10.1103/PhysRevLett.100.120402. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.100.120402>.
- [128] Gerhard C. Hegerfeldt. “Driving at the Quantum Speed Limit: Optimal Control of a Two-Level System”. In: *Phys. Rev. Lett.* 111 (26 2013), p. 260501. DOI: 10.1103/PhysRevLett.111.260501. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.111.260501>.

## Appendix A

# Diagrammatic Representation of Matrix Product States

Due to the many indices, equations describing contractions of matrix product states are often quite hard to read. Therefore, the equations are often represented graphically through diagrams. Many variations of diagrams exists, however, they all follow some general rules:

- Tensors are represented by nodes.
- Indices are represented by lines. Horizontal lines are *bond indices*, while vertical lines are *physical indices*.
- Tensors connected by a line are contracted over said bond.

Through these basic rules, most equations involving matrix product states can be expressed through diagrams. In this instance, the color or shape of a node provides information regarding the properties of the tensor. Therefore, the tensors shown in this thesis follow these guidelines:

- Tensors part of an MPS are square-like. The color of the tensor denotes its normalisation, as shown in figure A.1. These different types of normalisation are part of of MPS canonical forms (see Section 3.3).
- Tensors part of MPO's are grey and circular/elliptical depending on the number of sites they span over. These tensors have two vertical legs compared two the one leg of the MPS tensors.
- Matrices (tensors without any physical index) are depicted as a diamond. These often occur after a singular value decomposition.

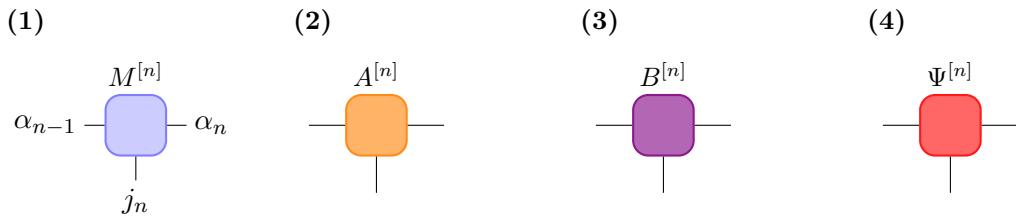


FIGURE A.1: The four different types of MPS tensors used in diagrams. (a) General tensor of un-specified normalisation. The indices corresponding to the tensor are labeled. The remaining tensors are: Left-normalised (b), right-normalised (c), central cite (d).

## Contraction of Tensor Networks

Similarly to matrix multiplication, the number of computations need for contracting a string of tensors can be greatly reduced when performing the contractions in the right order. Consider the general case of two general states  $|\psi\rangle$  and  $|\phi\rangle$  described by the matrices  $M$  and  $\tilde{M}$ . The overlap between these states reads

$$\langle\phi|\psi\rangle = \sum_{j_1, \dots, j_L} \tilde{M}^{j_L\dagger} \dots \tilde{M}^{j_1\dagger} M^{j_1} \dots M^{j_L}, \quad (\text{A.1})$$

which is represented diagrammatically in figure A.2. The evaluation of the overlap can be

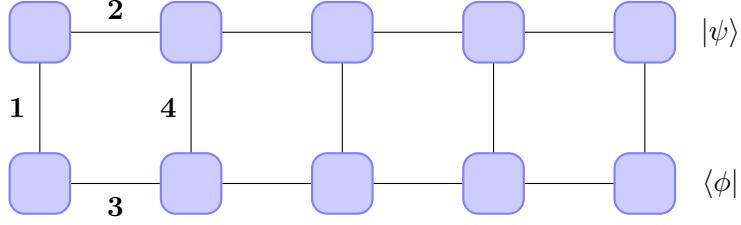


FIGURE A.2: Efficient contraction of the overlap between two general states  $|\psi\rangle$  and  $|\phi\rangle$ . By contracting the bonds in the specified order the number of multiplications is greatly reduced.

drastically sped up by considering the optimal order of contractions, which corresponds to an optimal bracketing of eq. (A.1):

$$\langle\phi|\psi\rangle = \sum_{j_L} \tilde{M}^{j_L\dagger} \left( \dots \left( \sum_{j_2} \tilde{M}^{j_2\dagger} \left( \sum_{j_1} \tilde{M}^{j_1\dagger} M^{j_1} \right) M^{j_2} \right) \dots \right) M^{j_L}. \quad (\text{A.2})$$

Within the innermost bracket a matrix is formed by multiplying a column and a row vector followed by the summation over the first physical index,  $j_1$ . In the remaining set of brackets three matrices are multiplied and the corresponding physical index is summed over. However, after the first contraction the complexity of the operation does not increase. Consider the worst case scenario of all the matrices being of dimension  $(D \times D)$  with a local Hilbert space of dimension  $d$ . The total operational cost of the optimal contraction is  $\mathcal{O}(LD^3d)$  compared to otherwise exponential complexity of a random order of contractions [21]. The order of contractions described in eq. (A.2) is illustrated in figure A.2.

When calculating a norm,  $\langle\psi|\psi\rangle$ , the canonical form of the MPS greatly reduces to computational time, as having an either left- or right-canonical MPS implies a norm of 1 without any contractions needed.

## Appendix B

# Example: Building an MPO from a Hamiltonian

Consider the Bose-Hubbard Hamiltonian , which consists of both nearest-neighbour tunneling terms and on-site interactions terms

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{a}_i^\dagger \hat{a}_j + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) . \quad (\text{B.1})$$

The Hamiltonian can be expressed as a sum of strings of operators - most of these being identities just like in eq. 3.22. Moving through such a string from the right, one will at some point encounter one of 3 non-trivial operators, which can be described by 4 different possible states of the string of operators on a given bond:

1. Only identities to the right of the bond.
2. An  $\hat{a}^\dagger$  operator just to the right of the bond.
3. An  $\hat{a}$  operator just to the right of the bond.
4. A completed tunneling or the interaction term,  $\frac{U}{2} \hat{n} (\hat{n} - 1)$ , somewhere to the right.

When moving through the string of operators from the right, only certain transitions between these state are possible. For instance **1** → **1**, where one starts in state 1, and the next operator is an identity. Likewise, **1** → **2, 3, 4** are all possible, since these transitions represent the next operator being non-trivial. Next, **2** → **4** completes the hopping,  $-J\hat{a}^\dagger \hat{a}$ . A similar transition exists for the other hopping term **3** → **4**. Finally, the transition **4** → **4** is needed to continue iterating through the operator chain after having passed the non-trivial operators. The possible state transfers can be encoded in the operator-valued matrix [21]

$$W^{[i]} = \begin{pmatrix} \hat{I} & 0 & 0 & 0 \\ \hat{a}^\dagger & 0 & 0 & 0 \\ \hat{a} & 0 & 0 & 0 \\ \frac{U}{2} \hat{n} (\hat{n} - 1) & -J\hat{a} & -J\hat{a}^\dagger & \hat{I} \end{pmatrix} . \quad (\text{B.2})$$

When one starts moving through the operator chain from the right side, one obviously begins in state 1 and ends in state 4. This can be encoded in two boundary vectors

$$\vec{v}_{left} = (0, 0, 0, 1) , \quad \vec{v}_{right} = (1, 0, 0, 0)^T .$$

Thus, the Bose-Hubbard Hamiltonian can be written as an MPO

$$\hat{H} = \sum_{\mathbf{j}, \mathbf{j}'} \vec{v}_{left} W^{[1]j_1, j'_1} W^{[2]j_2, j'_2} \dots W^{[N]j_N, j'_N} \vec{v}_{right} |\mathbf{j}\rangle \langle \mathbf{j}'| . \quad (\text{B.3})$$

Often, the two closing vectors are implicitly multiplied unto the outer matrices for a cleaner notation. Observe how all of this was done without having to do a single numerical computation. The resulting MPO can now readily be applied to an MPS.

## Appendix C

# Calculation of Condensate Fraction using DMRG Algorithm

To illustrate the power of the DMRG algorithm, the following numerical analysis was conducted. As the system of interest is the Bose-Hubbard model described in eq. (2.39), a suitable benchmark for the algorithm is attempting to calculate the critical point of the phase transition between the superfluid and Mott-Insulator. The critical point was determined using the DMRG algorithm in [47] by studying the correlations of the system. Here, an alternative approach is presented, which examines the condensate fraction. According to the Penrose-Onsager criterion, a system is in the superfluid phase if and only if the largest eigenvalue,  $\lambda_1$ , of the single-particle density matrix,  $\rho^{(1)}$ , is macroscopic

$$f_c = \frac{\lambda_1}{N_p} > 0 , \quad (\text{C.1})$$

where  $f_c$  is the condensate fraction, and  $N_p$  is the total number of particles [38]. The condensate fraction can be used to determine which phase dominates the system, as

$$\lim_{N_p \rightarrow \infty} f_c^{\text{SF}} \rightarrow 1 \quad (\text{C.2})$$

$$\lim_{N_p \rightarrow \infty} f_c^{\text{MI}} \rightarrow 0 , \quad (\text{C.3})$$

when the filling-fraction,  $n = N_p/L$ , is held constant.

To calculate the condensate fraction, the ground state of the system was found using a version of the DMRG algorithm implemented in the ITensor library [87]. Using the computed ground state,  $|\psi\rangle$ , the entries of the density matrix were calculated

$$\rho_{i,j}^{(1)} = \langle \psi | \hat{a}_i^\dagger \hat{a}_j | \psi \rangle . \quad (\text{C.4})$$

Lastly, the condensate fraction was determined through eq. (C.1). The calculation was performed with varying  $U/J$  for various system sizes of unit occupancy. The DMRG algorithm was set to perform 5 sweeps with a maximum bond dimension of 200. In order to gauge the accuracy of the algorithm, the results for system sizes 4-10 were compared to a similar calculation using exact diagonalisation. The upper part of figure C.1 shows the condensate fraction for various  $U/J$  calculated using the DMRG algorithm. In the limit  $U/J = 0$ , the condensate fraction is unit for the smaller systems, confirming the system is indeed in the superfluid phase. The condensate fraction never reaches zero as  $U/J$  increases, since this is only achieved in the thermodynamic limit. However, the condensate fraction does decrease with increasing particle number, as would be expected. The lower half of figure C.1 displays the results of the DMRG calculation compared with exact diagonalisation. For small systems the two approaches obtain very similar results.

Attempting to use exact diagonalisation for large systems is futile, due to the exponential

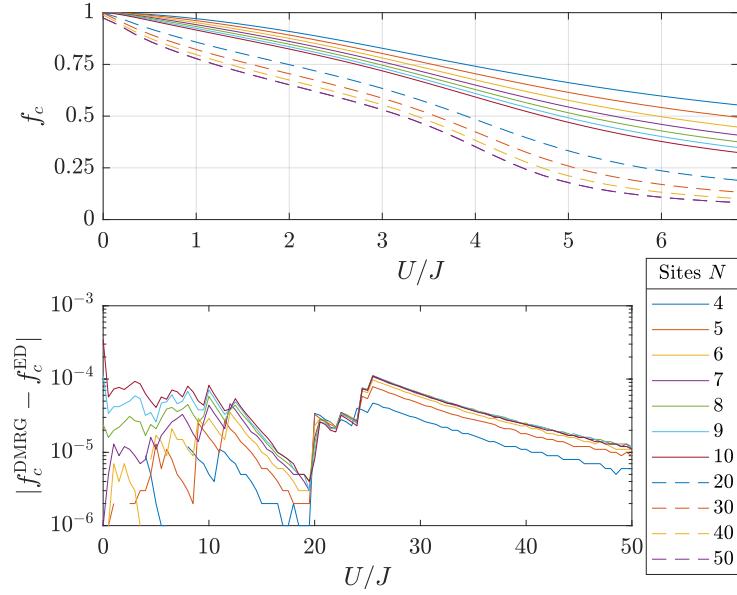


FIGURE C.1: Condensate fraction calculated using the DMRG algorithm with 20 sweeps. The condensate fractions of the smaller systems ( $L = 4 \dots 10$ ) are compared with result obtained through exact diagonalisation.

scaling of the Hilbert space [20]. However, this is not an issue using matrix product states, as the formalism only considers a tiny corner of the Hilbert space by following an area law. The dashed lines of figure C.1 shows the results of the DMRG calculations for up to 50 particles. The condensate fraction behaves as expected in the  $U/J \gg 1$  limit, as it tends towards zero for larger particle numbers. Note, in the superfluid limit the condensate fraction does not quite reach 1, which is due to difficulty of describing long-range correlations when using matrix product states as described in section ???. This is not an issue for smaller systems, as the correlation length is limited by the system size.

Some measures can be taken to minimize errors, when using the DMRG algorithm. Accurately describing long range correlations requires multiple sweeps, as the algorithm has to iteratively approximate an almost constant function with a series of exponentials. Figure C.2 displays the condensate fraction in the Superfluid limit as a function of number of sweeps. Clearly, performing more sweeps yields a better result. Furthermore, increasing

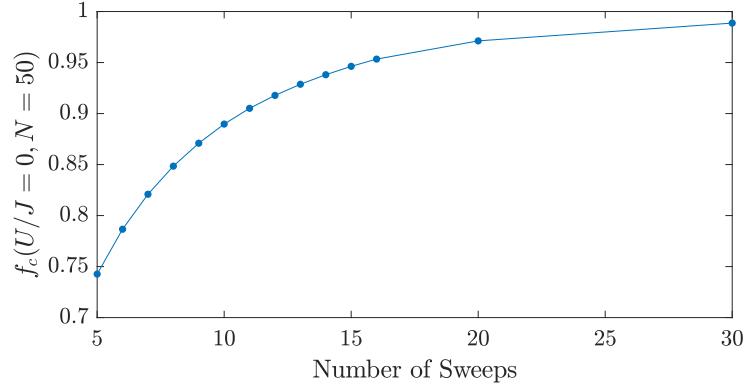


FIGURE C.2: Condensate fraction as a function of number of sweeps of the DMRG algorithm in the superfluid limit. A max bond dimension of  $D = 250$  was used.

the maximal bond dimension,  $D$ , results in a more accurate long-range representation of

correlations. Thus, calculating correlation functions for various values of  $D$  is a great way of estimating the convergence of the correlations for a given length scale [21].

In [47] the critical point of phase transition between the Superfluid and Mott-Insulator is determined as  $(\frac{U}{J})_{crit} = 3.37$ . The result is achieved through examining the correlations of the systems. While correlations in Mott-Insulators decay exponentially, superfluids have a decay of correlations following the power-law given in eq. (3.5). The interface between the two phases has correlations following a power law determined by the Luttinger liquid parameter. The critical point, which is located at the tip of the Mott-lobes, is computed by determining the point at which the Luttinger liquid parameter is  $K = \frac{1}{2}$  [47].

Examining figure C.1, one notices a hump on the graph in the vicinity of the critical ratio, but no clear indication of a phase transition is present. In the thermodynamic limit one would expect the condensate fraction to drop to zero as the critical ratio is reached (as observed in 2D by [127]). However, at 50 particles the condensate fraction is only around  $f_c = 0.5$ . One could extrapolate data from computations using different particle numbers in order to determine the location of the critical point. However, this would require computations using larger systems in order to minimize the boundary effects.



## Appendix D

# Example: Quantum Speed Limit in Landau-Zener Model

To illustrate concepts from quantum optimal control theory, a state transfer within the Landau-Zener model is optimized using the GRAPE algorithm.

The Landau-Zener model is a two-level system with the general Hamiltonian

$$\hat{H}_{\text{LZ}} = \begin{pmatrix} \Delta(t) & \Omega_R \\ \Omega_R & -\Delta(t) \end{pmatrix} = \Omega_R \hat{\sigma}_x + \Delta(t) \hat{\sigma}_z , \quad (\text{D.1})$$

where  $\Omega_R$  is the positive Rabi-frequency,  $\Delta(t)$  is the detuning, and  $\hat{\sigma}_i$  are the Pauli spin matrices. In such a system the detuning will be the control parameter, as it is easily to manipulate in an experimental setup. An advantage of the Landau-Zener system is that an analytical solution to the control problem exists for arbitrary initial and final states. Furthermore, two level systems can be illustrated on the Bloch sphere, which depicts the population of the two levels along with the relative phases. Consider the transfer between the initial state

$$|\psi_0\rangle = \cos\left(\frac{\theta_0}{2}\right)|0\rangle + e^{i\phi_0} \sin\left(\frac{\theta_0}{2}\right)|1\rangle \quad (\text{D.2})$$

and final state

$$|\psi_T\rangle = \cos\left(\frac{\theta_T}{2}\right)|0\rangle + e^{i\phi_T} \sin\left(\frac{\theta_T}{2}\right)|1\rangle , \quad (\text{D.3})$$

which are both expressed in terms of the angles of the Bloch sphere. In [128] the quantum speed limit of such a state transfer was derived to be

$$T_{\text{QSL}} = \left| \frac{\theta_T - \theta_0}{2\Omega_R} \right| . \quad (\text{D.4})$$

Consider the case of  $|\psi_0\rangle = |0\rangle$  and  $|\psi_T\rangle = |1\rangle$ , whereby the quantum speed limit is  $T_{\text{QSL}} = \frac{\pi}{2\Omega_R}$ .

### Optimal control using GRAPE

In this example  $\Omega_R = 1$  for simplicity. Examining the Hamiltonian of eq. (D.1), it is clear that the state transfer  $|0\rangle \rightarrow |1\rangle$  can be achieved efficiently by letting  $u(t) \equiv \Delta(t) = 0$  for the entire duration. To create an interesting example, the control is subjected to the boundary conditions  $u(0) = 0$  and  $u(T) = 2T$ .

Figure D.1 displays the results of optimizing the state transfer using the GRAPE algorithm for duration  $T_1 = \pi/2$ . The algorithm takes an initial linear seed, which clearly does not reach the target state. Meanwhile, the optimized control achieves perfect transfer, which is as expected, as the quantum speed limit for the transfer is  $T_{\text{QSL}} = T_1 = \pi/2$ , whereby a solution should exist for this duration. However, any duration shorter should not be

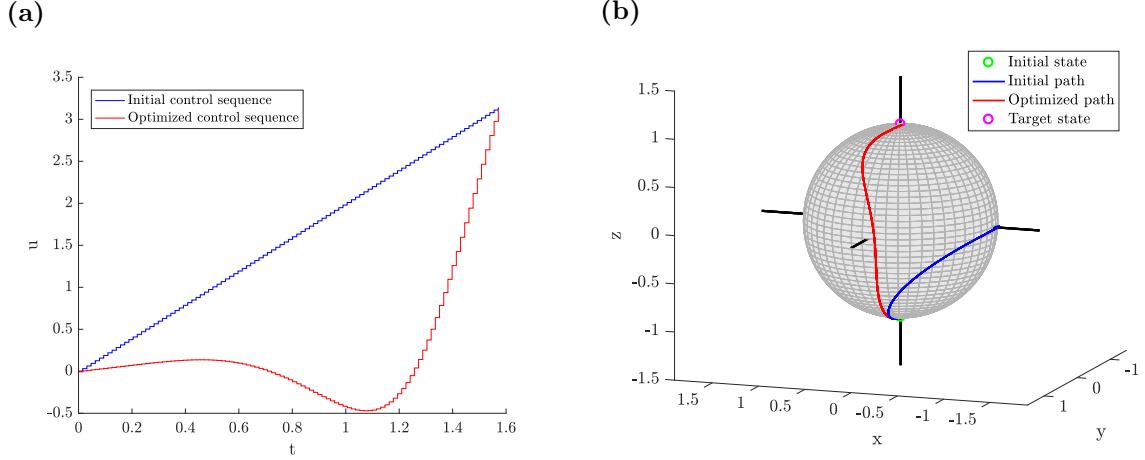


FIGURE D.1: *Optimal control of LZ-system using GRAPE for  $T_1 = \pi/2$ .* (a): Initial and optimized control sequence. (b): Path traced out on the Bloch sphere by the quantum state, as it is evolved according to the control.

sufficient to reach the target state.

Next, consider figure D.2, which shows the optimization results for duration  $T_2 = 2$ . As

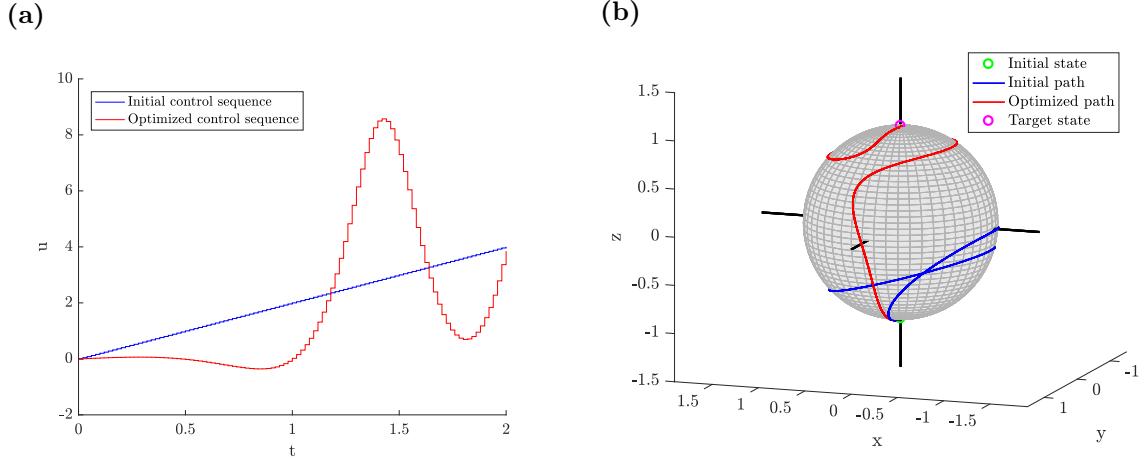


FIGURE D.2: *Optimal control of LZ-system using GRAPE for  $T_2 = 2$ .* (a): Initial and optimized control sequence. (b): Path traced out on the Bloch sphere by the quantum state, as it is evolved according to the control.

this duration is larger than the quantum speed limit, a solution exist for  $T_2$ . However, the path of the state on the Bloch sphere is much less direct than for the previous case. This is a consequence of how the optimal control problem is formulated, as one is searching for a control resulting in  $|\psi(T)\rangle = |\psi_{\text{Target}}\rangle$ . Thus, the path of the state has no impact on the cost. Figures D.1 and D.2 illustrate this point, as many different paths from  $|\psi_0\rangle \rightarrow |\psi_{\text{Target}}\rangle$  are possible for large durations, whereas only few possible solutions are available at durations close to the quantum speed limit.

## Appendix E

# Hessian

Consider the cost

$$\mathcal{J}_T = \frac{1}{2} (1 - |\langle \psi_t | \psi(T) \rangle|^2) = \frac{1}{2} \left( 1 - \left| \left\langle \psi_t \left| \prod_{j=1}^N \hat{\mathcal{U}}_j \right| \psi(0) \right\rangle \right|^2 \right), \quad (\text{E.1})$$

where the general propagator is  $\hat{\mathcal{U}}_j \equiv \hat{\mathcal{U}}(u(t_j)) = \exp\{-i\hat{H}(t_j)\Delta t\}$ , and the Hamiltonian is of the form

$$\hat{H}(t_j) = \hat{H}_0(t_j) + \sum_{n=1}^m \hat{H}_n(u_n(t_j)). \quad (\text{E.2})$$

Defining the transfer probability amplitude  $\mathcal{T} \equiv \langle \psi_t | \psi(T) \rangle$ , the derivative of the cost can be formulated as

$$\frac{\partial \mathcal{J}_T}{\partial u_n(t_j)} = -\frac{1}{2} \frac{\partial}{\partial u_n(t_j)} \mathcal{T}^* \mathcal{T} = -\text{Re} \left( \mathcal{T}^* \frac{\partial \mathcal{T}}{\partial u_n(t_j)} \right), \quad (\text{E.3})$$

where

$$\begin{aligned} \frac{\partial \mathcal{T}}{\partial u_n(t_j)} &= \frac{\partial}{\partial u_n(t_j)} \left\langle \psi_t \left| \prod_{j=1}^N \hat{\mathcal{U}}_j \right| \psi(0) \right\rangle \\ &= \left\langle \psi_t \left| \hat{\mathcal{U}}_N \dots \hat{\mathcal{U}}_{j+1} \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \hat{\mathcal{U}}_{j-1} \dots \hat{\mathcal{U}}_1 \right| \psi(0) \right\rangle \\ &= \left\langle \psi_t(t_j) \left| \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \right| \psi(t_{j-1}) \right\rangle. \end{aligned} \quad (\text{E.4})$$

The matrix elements of the cost Hessian have the form

$$\begin{aligned} \frac{\partial^2 \mathcal{J}}{\partial u_n(t_j) \partial u_n(t_i)} &= -\frac{1}{2} \left( \frac{\partial^2 \mathcal{T}}{\partial u_n(t_j) \partial u_n(t_i)} \mathcal{T}^* + \frac{\partial \mathcal{T}}{\partial u_n(t_j)} \frac{\partial \mathcal{T}^*}{\partial u_n(t_i)} \right. \\ &\quad \left. + \frac{\partial \mathcal{T}^*}{\partial u_n(t_j)} \frac{\partial \mathcal{T}}{\partial u_n(t_i)} + \mathcal{T} \frac{\partial^2 \mathcal{T}^*}{\partial u_n(t_j) \partial u_n(t_i)} \right) \\ &= -\text{Re} \left( \frac{\partial \mathcal{T}}{\partial u_n(t_j)} \frac{\partial \mathcal{T}^*}{\partial u_n(t_i)} \right) - \text{Re} \left( \frac{\partial^2 \mathcal{T}}{\partial u_n(t_j) \partial u_n(t_i)} \mathcal{T}^* \right). \end{aligned} \quad (\text{E.5})$$

The first term of the Hessian is easily computed from the elements of the gradient

$$\frac{\partial \mathcal{T}}{\partial u_n(t_j)} \frac{\partial \mathcal{T}^*}{\partial u_n(t_i)} = \left\langle \psi_t(t_j) \left| \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \right| \psi(t_{j-1}) \right\rangle \left\langle \psi(t_{i-1}) \left| \frac{\partial \hat{\mathcal{U}}_i^\dagger}{\partial u_n(t_i)} \right| \psi_t(t_i) \right\rangle. \quad (\text{E.6})$$

Meanwhile, the second term of the Hessian is more complicated, although it can easily be derived from eq. (E.4)

$$\frac{\partial^2 \mathcal{T}}{\partial u_n(t_j) \partial u_n(t_i)} = \begin{cases} \left\langle \psi_t(t_j) \left| \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \left( \prod_{k=i+1}^{j-1} \hat{\mathcal{U}}_k \right) \frac{\partial \hat{\mathcal{U}}_i}{\partial u_n(t_i)} \right| \psi(t_{i-1}) \right\rangle, & \text{for } i \neq j. \\ \left\langle \psi_t(t_i) \left| \frac{\partial^2 \hat{\mathcal{U}}_j}{\partial u_n^2(t_j)} \right| \psi(t_{i-1}) \right\rangle, & \text{for } i = j. \end{cases} . \quad (\text{E.7})$$

Thus, defining  $|\chi(T)\rangle \equiv i\mathcal{T}|\psi_t\rangle = i|\psi_t\rangle\langle\psi_t|\psi(T)\rangle$ , the matrix elements of the Hessian read

$$\begin{aligned} \frac{\partial^2 \mathcal{J}}{\partial u_n(t_j) \partial u_n(t_i)} = & -\text{Re} \left( \left\langle \psi_t(t_j) \left| \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \right| \psi(t_{j-1}) \right\rangle \left\langle \psi(t_{i-1}) \left| \frac{\partial \hat{\mathcal{U}}_i^\dagger}{\partial u_n(t_i)} \right| \psi_t(t_i) \right\rangle \right) \\ & + \text{Im} \left\langle \chi(t_j) \left| \frac{\partial \hat{\mathcal{U}}_j}{\partial u_n(t_j)} \left( \prod_{k=i+1}^{j-1} \hat{\mathcal{U}}_k \right) \frac{\partial \hat{\mathcal{U}}_i}{\partial u_n(t_i)} \right| \psi(t_{i-1}) \right\rangle (1 - \delta_{i,j}) \\ & + \text{Im} \left\langle \chi(t_i) \left| \frac{\partial^2 \hat{\mathcal{U}}_j}{\partial u_n^2(t_j)} \right| \psi(t_{i-1}) \right\rangle \delta_{i,j} . \end{aligned} \quad (\text{E.8})$$

### Hessian for Suzuki-Trotter Propagator

Consider the case where the control Hamiltonians,  $\hat{H}_n$ , are diagonal, whereby they and their derivatives mutually commute. Expanding the propagator through the Suzuki-Trotter expansion and evaluating the control function at each end of the time-step interval produces the propagator

$$\hat{\mathcal{U}}_j^{\text{ST}} = \left( \prod_{n=1}^m e^{-i\hat{H}_n(t_j)\Delta t/2} \right) e^{-i\hat{H}_0\Delta t} \left( \prod_{n=1}^m e^{-i\hat{H}_n(t_{j-1})\Delta t/2} \right) . \quad (\text{E.9})$$

The derivative of the Suzuki-Trotter propagator is

$$\frac{\partial \hat{\mathcal{U}}_k^{\text{ST}}}{\partial u_n(t_j)} = \left( -i \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \frac{\Delta t}{2} \right) \hat{\mathcal{U}}_j^{\text{ST}} \delta_{j,k} + \hat{\mathcal{U}}_j^{\text{ST}} \left( -i \frac{\partial \hat{H}_n(t_{j-1})}{\partial u_n(t_{j-1})} \frac{\Delta t}{2} \right) \delta_{j,k+1} , \quad (\text{E.10})$$

where the two contributions origin from the end-point evaluation of the control. Thus, the derivatives with respect to  $u_n(t_j)$  stated above will have contributions from both  $\hat{\mathcal{U}}_j^{\text{ST}}$  and  $\hat{\mathcal{U}}_{j+1}^{\text{ST}}$ . Thereby the first order derivative of the transfer probability amplitude reads

$$\frac{\partial \mathcal{T}^{\text{ST}}}{\partial u_n(t_j)} = -i\Delta t \left\langle \psi_t(t_j) \left| \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle , \quad (\text{E.11})$$

from which the second order derivative can be calculated. Starting with the case  $i \neq j$ , the contributions to the derivative from two propagators results in

$$\begin{aligned}
\frac{\partial^2 \mathcal{T}^{\text{ST}}}{\partial u_n(t_j) \partial u_n(t_i)} &= -i\Delta t \left\langle \psi_t(t_{i+1}) \left| \frac{\partial \hat{\mathcal{U}}_{i+1}^{\text{ST}}}{\partial u_n(t_i)} \left( \prod_{k=j+1}^i \hat{\mathcal{U}}_k^{\text{ST}} \right) \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle \\
&\quad - i\Delta t \left\langle \psi_t(t_i) \left| \frac{\partial \hat{\mathcal{U}}_i^{\text{ST}}}{\partial u_n(t_i)} \left( \prod_{k=j+1}^{i-1} \hat{\mathcal{U}}_k^{\text{ST}} \right) \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle \\
&= -i\Delta t \left\langle \psi_t(t_{i+1}) \left| \hat{\mathcal{U}}_{i+1}^{\text{ST}} \left( -i \frac{\partial \hat{H}_n(t_i)}{\partial u_n(t_i)} \frac{\Delta t}{2} \right) \left( \prod_{k=j+1}^i \hat{\mathcal{U}}_k^{\text{ST}} \right) \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle \\
&\quad - i\Delta t \left\langle \psi_t(t_i) \left| \left( -i \frac{\partial \hat{H}_n(t_i)}{\partial u_n(t_i)} \frac{\Delta t}{2} \right) \hat{\mathcal{U}}_i^{\text{ST}} \left( \prod_{k=j+1}^{i-1} \hat{\mathcal{U}}_k^{\text{ST}} \right) \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle \\
&= -\Delta t^2 \left\langle \psi_t(t_i) \left| \frac{\partial \hat{H}_n(t_i)}{\partial u_n(t_i)} \left( \prod_{k=j+1}^i \hat{\mathcal{U}}_k^{\text{ST}} \right) \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle. \tag{E.12}
\end{aligned}$$

The case of  $i = j$  included an extra term, as one must remember the derivative of  $\partial \hat{H}_n(t_j) / \partial u_n(t_j)$ . Thereby the second order derivative reads

$$\begin{aligned}
\frac{\partial^2 \mathcal{T}^{\text{ST}}}{\partial u_n^2(t_j)} &= -i\Delta t \left\langle \psi_t(t_{j+1}) \left| \frac{\partial \hat{\mathcal{U}}_{j+1}^{\text{ST}}}{\partial u_n(t_j)} \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle \\
&\quad - i\Delta t \left\langle \psi_t(t_j) \left| \frac{\partial^2 \hat{H}_n(t_j)}{\partial u_n^2(t_j)} \right| \psi(t_j) \right\rangle \\
&\quad - i\Delta t \left\langle \psi_t(t_j) \left| \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \frac{\partial \hat{\mathcal{U}}_j^{\text{ST}}}{\partial u_n(t_j)} \right| \psi(t_{j-1}) \right\rangle \\
&= -i\Delta t \left\langle \psi_t(t_i) \left| \frac{\partial^2 \hat{H}_n(t_j)}{\partial u_n^2(t_j)} - i\Delta t \left( \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right)^2 \right| \psi(t_j) \right\rangle. \tag{E.13}
\end{aligned}$$

Thus, the second order derivatives of the transfer probability amplitude can be summarized as

$$\frac{\partial^2 \mathcal{T}^{\text{ST}}}{\partial u_n(t_j) \partial u_n(t_i)} = \begin{cases} -\Delta t^2 \left\langle \psi_t(t_i) \left| \frac{\partial \hat{H}_n(t_i)}{\partial u_n(t_i)} \left( \prod_{k=j+1}^i \hat{\mathcal{U}}_k^{\text{ST}} \right) \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle, & \text{for } i \neq j. \\ -i\Delta t \left\langle \psi_t(t_i) \left| \frac{\partial^2 \hat{H}_n(t_j)}{\partial u_n^2(t_j)} - i\Delta t \left( \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right)^2 \right| \psi(t_j) \right\rangle, & \text{for } i = j. \end{cases}. \tag{E.14}$$

Inserting the propagator into the elements of the Hessian (E.8) produces

$$\begin{aligned}
\frac{\partial^2 \mathcal{J}}{\partial u_n(t_j) \partial u_n(t_i)} = & -\operatorname{Re} \left( \left\langle \psi_t(t_j) \left| \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right| \psi(t_j) \right\rangle \left\langle \psi(t_i) \left| \frac{\partial \hat{H}_n(t_i)}{\partial u_n(t_i)} \right| \psi_t(t_i) \right\rangle \right) \Delta t^2 \\
& - \operatorname{Im} \left\langle \chi(t_j) \left| \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \left( \prod_{k=i+1}^j \hat{\mathcal{U}}_k^{\text{ST}} \right) \frac{\partial \hat{H}_n(t_i)}{\partial u_n(t_i)} \right| \psi(t_i) \right\rangle \Delta t^2 (1 - \delta_{i,j}) \\
& - \operatorname{Re} \left\langle \chi(t_i) \left| \frac{\partial^2 \hat{H}_n(t_j)}{\partial u_n^2(t_j)} - i\Delta t \left( \frac{\partial \hat{H}_n(t_j)}{\partial u_n(t_j)} \right)^2 \right| \psi(t_j) \right\rangle \Delta t \delta_{i,j}. \quad (\text{E.15})
\end{aligned}$$

Having already computed the elements of the gradient, the number of step-propagations needed for the computation of the full Hessian is  $N^2/2 - N$  due to the symmetry of the matrix and the lack of new propagations for computing the diagonal elements.