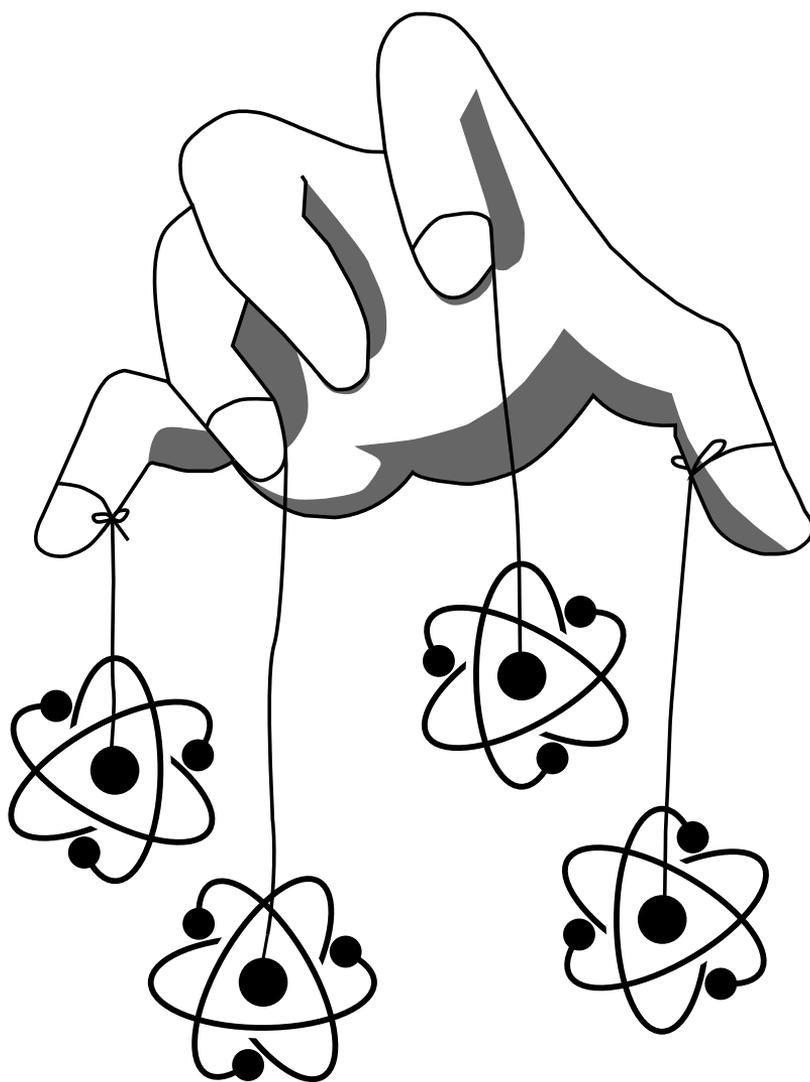


Quantum Control of Ultracold Quantum Systems



Jens Jakob Sørensen
PhD Thesis

Colophon

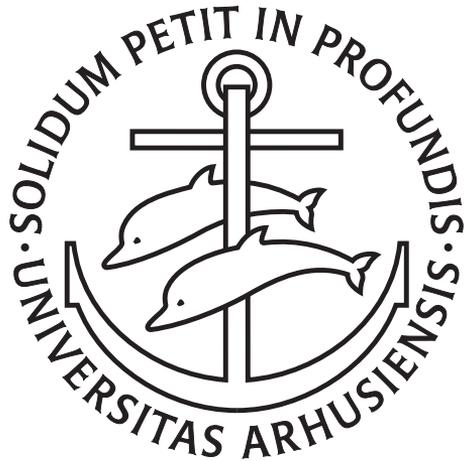
Quantum Optimal Control of Ultracold Systems

PhD thesis by Jens Jakob Sørensen.

The manuscript was typeset using L^AT_EX (memoir class). All figures were made using MATLAB or Inkscape.

PDF version

QUANTUM OPTIMAL CONTROL
OF ULTRACOLD SYSTEMS



JENS JAKOB SØRENSEN
PHD DISSERTATION
DECEMBER 2018

SUPERVISOR: JACOB SHERSON
CO-SUPERVISOR: KLAUS MØLMER

DEPARTMENT OF PHYSICS AND ASTRONOMY
AARHUS UNIVERSITY

*To Vera,
I have known you for a short time
but I will love you for a life time.*

English Summary

Control and engineering of physical phenomena is what transforms the scientific knowledge of today into the technology of tomorrow. Our understanding of quantum mechanics has reached a point where this transformation is possible. Quantum optimal control engineers control policies in a bottom up manner using optimization algorithms.

In this PhD thesis different quantum optimal control algorithms are applied for the manipulation of ultracold quantum systems. Several of the standard methods are compared and extended with global evolutionary methods. This is shown to give improved estimates of the quantum speed limit for certain ultracold optimal control problems.

Optimal control of single atoms with citizen science is also explored where the non-experts in form of players are shown to find useful seeds for the numerical optimization. Clustering analysis of the players' solutions is extended to a citizen science project with remote optimization of experimental Bose-Einstein condensates.

This thesis also explores alternative approaches to quantum optimal control such as numerically optimized adiabatic trajectories and combinations of unitary dynamics and quantum measurements.

Dansk Resume

Kontrol af fysiske fænomener transformerer nutidens videnskabelige viden til fremtidens teknologi. Vores forståelse af kvantemekanik har nået et punkt, hvor denne transformation er mulig. Kvanteeoptimal kontrol bygger ved hjælp af optimeringsalgoritmer kontrolpolitikker fra bunden.

I denne phd-afhandling anvendes forskellige algoritmer inden for kvanteoptimal kontrol til at manipulere ultrakolde kvantesystemer. Mange af standardmetoderne sammenlignes og udvides med globale evolutionære metoder. Dette giver forbedrede estimater af kvantefartgrænsen for udvalgte ultrakolde kvantekontrolproblemer.

Derudover udforskes optimal kontrol af enkelte atomer med borgervidenskab. Her vises det, at lægmænd som spillere er nyttige til at finde gode startpunkter for den numeriske optimering. Klyngeanalysen af spillernes løsninger udvides til et borgervidenskabsprojekt med fjernoptimering af eksperimentelle Bose-Einstein-kondensater.

Denne afhandling udforsker desuden alternative tilgange til kvanteoptimal kontrol såsom numerisk optimerede adiabatisk banekurver og kombinationer af unitær dynamik og kvantemålinger.

Preface

This thesis covers the main scientific results that I have obtained during my PhD studies. My studies were funded by the European Research Council and the Templeton Foundation. The work was conducted from August 2014 to December 2018 under the supervision of Professor Jacob Sherson in the Quantum Measurement and Manipulation Group at the Department of Physics and Astronomy at Aarhus University, Denmark. During my PhD studies the research group also grew to encompass the citizen science project ScienceAtHome that gamifies quantum physics and beyond.

The main focus of my PhD studies has been to establish the quantum control field within my research group. My thesis has a strong methodical focus as this is my primary research interests but also due to the fact that our group was newcomers to the quantum control field when I joined as an undergraduate. The secondary focus of my studies has been to contribute to the formation of a new research group. The development of my research and the group is evident not only in my internal knowledge but also externally from the size of the group. When I joined the group we were a mere handful of people. At the time of writing, the group has grown to more than 30 people including the ScienceAtHome project. I remember how a few weeks after my PhD admission we spent hours discussing the differences between the fundamental Krotov and GRAPE methods. Now at the end of my PhD, we can answer such questions with ease. Zooming in at my own corner in the now large research group, we are three theoretical PhD students and one post doc, in contrast to the beginning of my PhD where my supervisor and I hatched and executed all of our ideas.

A list of my publications follows on the next pages. Most of this thesis is directly reproduced from these publications and each chapter starts with an explicit account of any material duplicated from these references. Articles [3, 5] discuss quantum optimal control theory and they form the

bulk of chapters 4-5. Most of the material in these chapters have been rearranged in order to fit into a cohesive structure and avoid repetitions.

Papers [3] and [7] discuss quantum optimal control of single atoms combined with citizen science, which is presented in chapter 6. Ref. [7] also discusses optimized experimental production of Bose-Einstein condensates and social dynamics in a collaborative citizen science project. These subjects are at the periphery of this thesis and therefore only discussed briefly.

The paper [4] stems from the master thesis by Henrik Mortensen, where I acted in large parts as his co-supervisor. This work considers an analytic approach to quantum control based on fast adiabatic trajectories. The results from this paper are reproduced with only a few paragraphs changed in chapter 7.

Papers [1] and [9] discuss an alternative approach to quantum control that utilizes non-unitary backaction from quantum measurements. This is presented on its own in chapter 8.

Ref. [8] is a technical programming paper that introduces our high performance open source simulation software library QEngine. Quantum optimal control is essentially nonlinear programming applied on quantum problems. Proper implementation of the optimization algorithms requires many important tricks that are typically not documented in papers or PhD theses – including my own. This library is made with the intention that future group members and other researchers can easily access and apply the different methods discussed in this thesis. I have spent a large part of my PhD developing this library in collaboration with Jesper Mohr Jensen and Till Heinzl.

Ref. [2] is an outreach account of our citizen science work and it is not discussed in this thesis.

Acknowledgments

During my PhD studies at Aarhus University, I have had the opportunity to interact, collaborate and simply enjoy the company of many people. I am very grateful for the positive influence you have all had on my PhD studies. As I express my appreciation below, I will undoubtedly fail to mention all.

Initially, I would like to thank my supervisor Jacob Sherson who has always been there to help and guide me through the ups and downs of the PhD project. I have been fortunate to have a supervisor who shows great interest in my work and who despite of his chronically busy schedule always finds time to discuss both physics and life with me.

I would also like to thank Professor Klaus Mølmer, who has also provided indispensable help on parts of this project.

Next, I would like to thank my colleagues at the institute and the Quantum Manipulation and Measurement Group for their countless discussions that both directly or indirectly have aided this thesis. I am grateful to Mads Petersen, Jesper Hasseriis Mohr Jensen, Mogens Dalgaard, Felix Motzoi, Robert Heck, Ottó Elíasson, Pinja Haikka, Till Heinzl, Henrik Mortensen, Mikel Aramburu, Alexander Kiilerich, Jacob Søgaard, Frederik Skovbo, Maja Lassen, Julie Stephansen, Morten Andreasen, Ebbe Stubbe, Kenneth Hansen and Tilo Planke. Besides the scientific discussions you have all contributed towards a joyful atmosphere at the institute. A special thank also goes to the ScienceAtHome development team, who always manages to deliver despite at times impossible deadlines.

Finally, I cannot begin to express my gratitude for having Rikke and Vera in my life. Without your love and support, my PhD studies would not have been possible.

Jens Jakob Sørensen
Aarhus, December 2018.

List of Publications

- [1] M. K. Pedersen, **J. J. Sørensen**, M. C. Tichy, and J. F. Sherson, *Many-body state engineering using measurements and fixed unitary dynamics*, New Journal of Physics 16, 113038 (2014).
- [2] **J. J. Sørensen**, M. K. Pedersen, and J. Sherson, *Spillere hjælper kvantecomputeren på vej*, Kvant 25, 31 (2014).
- [3] **J. J. Sørensen**, M. K. Pedersen, M. Munch, P. Haikka, J. H. Jensen, T. Planke, M. G. Andreasen, M. Gajdacz, K. Mølmer, A. Lieberoth, *et al.*, *Exploring the quantum speed limit with computer games*, Nature 532, 210 (2016).
- [4] H. L. Mortensen, **J. J. Sørensen**, K. Mølmer, and J. F. Sherson, *Fast state transfer in a Λ -system: a shortcut-to-adiabaticity approach to robust and resource optimized control*, New Journal of Physics 20, 025009 (2018).
- [5] **J. J. Sørensen**, M. Aramburu, T. Heinzl, and J. Sherson, *Quantum optimal control in a chopped basis: Applications in control of Bose-Einstein condensates*, Physical Review A 98, 022119 (2018).

- [6] R. Heck, O. Vuculescu, **J. J. Sørensen**, J. Zoller, M. G. Andreassen, M. H. Bason, P. Ejlertsen, O. Eliasson, P. Haikka, J. S. Laustsen, *et. al.*, *Remote optimization of an ultracold atoms experiment by experts and citizen scientists*, Proceedings of the National Academy of Sciences, 201716869 (2018).
- [7] **J. J. Sørensen**, M. Aramburu, T. Heinzl, and J. Sherson, *Quantum Optimal Control of Bose-Einstein Condensates with Combined Global and Local Optimization*, arXiv preprint arXiv:1802.07521 (2018).
- [8] **J. J. Sørensen**, J. H. M. Jensen, T. Heinzl, and J. Sherson, *QEngine: An open-source C++ Library for Quantum Optimal Control of Ultracold Atoms*, arXiv preprint arXiv:1807.11731 (2018).
- [9] **J. J. Sørensen**, M. Dalgaard, A. H. Kiilerich, K. Mølmer, and J. Sherson, *Quantum Control with Measurements and Quantum Zeno Dynamics*, arXiv preprint arXiv:1806.07793 (2018).

Units

Unless otherwise state $\hbar = 1$ throughout this thesis.

List of Abbreviations

GPE	Gross-Pitaevskii equation
BFGS	Broyden-Fletcher-Goldfarb-Shanno
HILO	Heuristically Initialized Local Optimizer
CDT	Crossed Dipole Trap
WCDDT	Wide Crossed Dipole Trap
NCDDT	Narrow Crossed Dipole Trap
HT	Hybrid Trap
WHT	Wide Hybrid Trap
STA	Shortcuts-To-Adiabaticity
COSTACT	Combining STA and control theory
STIRAP	Stimulated Raman Adiabatic Passage
MEDO	Multiple Evenly Distributed Observables

MUM Mutually Unbiased Measurements

FUMES Fixed Unitary Evolution and Measurements

Contents

English Summary	i
Dansk Resume	ii
Preface	iii
Acknowledgments	iv
List of Publications	v
Units	vi
List of Abbreviations	vi
1 Introduction	1
2 Ultracold Atoms	5
2.1 Optical Dipole Potentials	6
2.2 Optical Lattice	7
2.3 Atom Chips	12
2.4 Multiparticle systems	15
2.5 Bose-Hubbard Model	18
3 Quantum Optimal Control	23
3.1 The Mathematical Structure of Quantum Control Problems	24
3.2 Necessary Conditions for Optimality	27
3.3 GRAPE Gradient Calculation	33
3.4 Advanced Methods for Gradient Based Quantum Control	34
3.5 Krotov's Method	40
3.6 Chopped Basis and Nelder-Mead	43
3.7 Gradient Optimization using Parameterization - GROUP	46
3.8 Controllability	47
3.9 Conclusion	49
4 Optimal Control of Bose-Einstein Condensates	51

4.1	The Control Problem	52
4.2	Results	55
4.3	Conclusion	65
5	Time-Optimal Control	66
5.1	The Quantum Speed Limit	67
5.2	Pontryagin's Maximum Principle	71
5.3	Time-Optimal Control of Bose-Einstein Condensates with Global-Local Optimization	77
5.4	Conclusion	88
6	Optimal Control of Single Atoms	90
6.1	Basic Architecture	92
6.2	Bring Home Water	95
6.3	Optimization Landscape Structure	101
6.4	Experimental Landscape Structure	107
6.5	Discussion of D. Sels' Results	111
6.6	Conclusion	114
7	Combining Shortcuts-To-Adiabaticity and Control The- ory	116
7.1	Shortcuts-To-Adiabaticity	117
7.2	The Λ -system	121
7.3	Combining Shortcuts and Control Theory	125
7.4	Conclusion	131
8	Measurement Based Control	132
8.1	Control-Free Control	133
8.2	Fixed Unitary Dynamics and Measurements	136
8.3	Measurement Based Control With Zeno dynamics	144
8.4	Conclusion	158
9	Conclusion and Outlook	160
9.1	Outlook	162
	Bibliography	166

1 Introduction

Technology building on the laws of quantum mechanics may usher in a wide range of new technologies. During the last century, we have experienced major advances in our understanding of quantum mechanics. Advances that today indicate new opportunities for developing quantum enhanced communication networks [10–12], high precision sensors [13–16], and fundamentally new paradigms for computation and simulation [17–19].

Initially, our understanding of quantum mechanics has resulted in groundbreaking technologies such as lasers, light-emitting diodes (LEDs) and transistors. These technologies are a result of the quantum description of energy levels in solid-state systems. However, these technologies do not utilize the quirky or counter-intuitive aspects of quantum theory such as superpositions (an object being in multiple states at the same time) or entanglement (objects being connected without direct physical interaction). A long term goal is the creation of quantum computers that may tackle problems unsolvable by current supercomputers [20–23]. Solving these problems would principally enable new technologies such as tailored chemical processes [24, 25], new materials such as high temperature superconductors [26] and advancing machine learning [27–29].

Despite major advances by several research groups, most of these ideas are several years if not decades away from commercial viability [30, 31]. A key stumbling block is that these technologies require careful manipulation of microscopic phenomena at unprecedented levels. Our ability to *control* quantum systems will determine the success in turning these scientific ideas into technology. Quantum control theory studies how such manipulations may be achieved at the atomic and molecular scale, typically using external electromagnetic fields [32]. Quantum control studies how to design external controls that realize a given task for a quantum system in the most optimal manner. All new quantum



technologies require composing sequences of such pulses to control the quantum system. This applies whether the system of interest is quantum gate operations in a quantum computer [33–35] or state preparation for quantum sensing at or beyond the Heisenberg limit [36, 37]. Quantum control combines mathematical control theory and numerical mathematics in order to design the most suitable control pulses for a given problem. Although the experimental implementations vary greatly, quantum control is a common mathematical framework [32].

Quantum optimal control attempts to engineer quantum control pulses in a bottom up manner by optimizing a cost function [38, 39]. In recent years, this has yielded many examples of surprising and counter-intuitive quantum control techniques that are still highly efficient and robust [40–46]. Important examples of current applications are magnetic field sensing using vacancy centers in diamond [45], state preparation of Bose-Einstein condensates [47] and high fidelity quantum gates in superconducting systems [48, 49].

The success of quantum optimal control theory stems from the fact that once the quantum control problem is encoded in the cost function, it effectively becomes a nonlinear optimization problem whereby all the machinery from mathematical programming may be directly applied [39]. Analytical methods typically rely on Pontryagin’s Maximum Principle that generalizes the classical Hamilton’s equations to a larger class of problems [32]. The typical process is first to find the extremal controls by solving the associate equations and secondly to select the extremal controls that also minimize the cost [50–53]. Although this process is straightforward in theory, it is far from trivial and in most cases intractable. In many cases, the default choice is to use a numerical rather than an analytical approach. The quantum control community has developed a plethora of different numerical optimization algorithms for solving numerical optimal control problems such as gradient ascent methods or Krotov-type methods [54–56]. These methods have also been extended to account for experimental imperfections by ensuring that the controls are robust against these imperfections. A key challenge within this field is the continued refinement of these methods allowing them to tackle larger and more experimentally relevant problems in terms of Hilbert space size, robustness and experimental constraints.

The focus of this thesis is the control of ultracold physics. In this field, neutral atoms are routinely trapped and cooled with laser light into the microkelvin regime whereby the wavelike behavior of the atoms emerge [57–60]. If the atoms have a bosonic or fermionic character this allows for the creation of a Bose-Einstein condensate [61–63] or quantum degenerate Fermi gas [64, 65]. Using optical lattices, radio-frequency dressing and



all the tricks from atomic physics, ultracold systems enable the study of a wide-range of quantum phenomena [66]. Much research is focused on creating strongly correlated systems mimicking well known systems from solid-state physics such as high temperature superconductors. This approach is known as quantum simulators [26].

During my PhD studies, I have applied several of the quantum optimal control methods on problems from ultracold physics. I have also contributed towards the continued development of these methods. The thesis is organized as follows:

- Chapter 2 introduces ultracold physics and the Hamiltonians used for modeling the control problems in the later chapters. We discuss the physics behind manipulation of single atoms with far detuned off-resonant light in so-called optical tweezers. The concept of trapping and manipulating Bose-Einstein condensates on atom chips is also introduced. Atom chips offer a platform with exceptional control over the spatial potential. These systems have been used to realize matter wave interferometry with a Mach-Zehnder type interferometer or create correlated beams of matter. Both these applications rely on careful state preparation using optimal control theory. This chapter also introduces the Bose-Hubbard model describing interacting bosons in an optical lattice.
- Chapter 3 introduces quantum optimal control theory. I introduce a number of the central methods from the field. The methods are all applied in the later chapters. The chapter ends with discussing how the popular chopped basis methods and gradient based methods may be combined in Gradient Optimization Using Parametrization (GROUP).
- Chapter 4 discusses a comparison of the different optimal control algorithms from chapter 3. The algorithms are compared on an experimentally relevant state preparation problem on an atom chip.
- Chapter 5 introduces the concept of time-optimal control, which is typically relevant in problems where a state must be prepared faster than a system's decoherence rate. The connection between time-optimality and the quantum speed limit is also explored. This chapter presents the analytical solutions to time-optimal control of the two-level system that reveals a surprising connection between the optimal time and the quantum speed limit. Finally, this chapter discusses time-optimal control of Bose-Einstein condensates trapped on atom chips, which is calculated numerically using a combination of global and local optimization.



- Chapter 6 introduces optimal control of single atoms using optical tweezers. In this work, we took the unusual approach of creating a citizen science game allowing non-experts to aid us in identifying time-optimal solutions. We applied machine learning techniques to identify different solution strategies. This work was later extended to another science effort where non-experts helped find optimal experimental settings for creating Bose-Einstein condensates.
- Chapter 7 introduces the shortcuts to adiabaticity formalism, which is an analytical method for finding optimal controls that is more approachable than Pontryagin's Maximum Principle. This formalism actually offers an entire solution space and the chapter discusses how regular control theory may be used to search this space for the most optimal solutions. This method is applied to find improved solutions for an experimentally realized population transfer in an ultracold Λ -system.
- Chapter 8 discusses another approach to quantum control that utilizes the non-unitary backaction from quantum measurements. This backaction allows for so-called control-free control where the Hamiltonian is constant or even zero and the state is only manipulated through the measurements. We discuss methods that simultaneously measure the system and change the Hamiltonian.
- Chapter 9 finishes the thesis with some closing remarks.



2 Ultracold Atoms

Ultracold quantum gases are a diverse platform for studying quantum phenomena. The initial hallmark of this field was the observation of a macroscopic population in the ground state mode of a low temperature bosonic systems known as a Bose-Einstein condensate [63]. These systems are typically created in optical or magnetic traps inside a high quality vacuum chamber.

For neutral atoms as Rubidium-87 there is no Coulomb interaction requiring that the atoms are trapped by perturbing the energy levels using a spatially dependent the atom-light interaction. This interaction has a dissipative and conservative component. The dissipative interaction is due to absorption of photons followed by subsequent reemission whereby the atoms experience a force due to the momentum transfer from the absorbed and emitted photons. This light force is central in neutral atom trapping techniques such as magneto optical traps [67] or optical molasses [68]. These methods have the advantage that they both cool and trap the atoms.

The conservative interaction is due to the interaction between the light field and an induced dipole moment in the atom causing an energy shift known as the AC-stark shift [69]. For sufficiently off-resonant light the dissipative dynamics may be neglected and the interaction becomes purely conservative and proportional to the intensity of the light. In a crossed dipole trap the atoms are trapped at the minimum of a conservative light potential. In a typical ultracold atoms experiment one starts with cooling and trapping the atoms in a dissipative magneto optical trap and then transfers the cooled atoms to a conservative trap.

Counter propagating off-resonant laser beams can be used to trap atoms in periodic potential landscapes or so-called optical lattices [66]. These are the starting place for realizing the ultracold quantum gas analogue of solid state systems, which have a high level of tunability due to the experimental control of the laser light.



In Sec. 2.1 we will briefly outline the physics of optical dipole traps as they are the starting point for controlling and manipulating the atoms in the later chapters. In Sec. 2.3 we will discuss purely magnetic trapping of atoms using atom chips. In the later chapters, it is discussed in detail how to control the dynamics of the atoms trapped on such atom chips. In Sec. 2.4 the Gross-Pitaevskii equation is introduced, which is repeatedly utilized in this thesis. Sec. 2.5 introduces the Bose-Hubbard model and the last chapters in this thesis discusses how to manipulate the dynamics of this model.

2.1 Optical Dipole Potentials

An electric field \mathbf{E} induces an electric dipole \mathbf{p} on an atom. The amplitude of the polarization is proportional the electric field $p(\mathbf{r}) = \alpha\mathbf{E}(\mathbf{r})$ with some complex polarizability $\alpha(\omega)$. The induced potential is

$$V_{\text{dip}} = -\frac{1}{2}\langle\mathbf{p} \cdot \mathbf{E}\rangle_{\text{avg}} = -\frac{1}{2\varepsilon_0 c}\Re(\alpha)I(\mathbf{r}), \quad (2.1)$$

where $I(\mathbf{r}) = 2\varepsilon_0 c|\mathbf{E}(\mathbf{r})|^2$ is the intensity of the light and the bracket denotes a time average over one periode [69]. As usual, the constants ε_0 and c are the permeability and the speed of light. This interaction is conservative since it only perturbs the energy levels. This important relation Eq. (2.1) states that the conservative part of the atom-light is proportional to the laser intensity and the real part of α , which creates the dipole force $\mathbf{F}_{\text{dip}}(\mathbf{r}) = -\nabla V_{\text{dip}}(\mathbf{r})$. The interaction also absorbs and spontaneously re-emits radiation giving rise to the scattering rate [69]

$$\Gamma_{\text{sc}}(\mathbf{r}) = \frac{P_{\text{abs}}}{\hbar\omega} = \frac{\langle\dot{\mathbf{p}} \cdot \mathbf{E}\rangle_{\text{avg}}}{\hbar\omega} = \frac{1}{\hbar\varepsilon_0 c}\Im(\alpha)I(\mathbf{r}). \quad (2.2)$$

In order to use these relations we must calculate the polarizability, which may be estimated from a classical Lorentz model with equation of motion $\ddot{x} + \gamma\dot{x} + \omega_0^2 x = -eE(t)/m_0$ where e and m_0 are the electron charge and mass. The damping rate γ is given by Larmor's formula $\gamma(\omega) = e^2\omega^2/6\pi\varepsilon_0 m_0 c^3$. Integrating the equation of motion gives

$$\alpha(\omega) = 6\pi\varepsilon_0 c^3 \frac{\gamma_0/\omega_0^2}{\omega_0^2 - \omega^2 - i\gamma_0\omega^3/\omega_0^2}, \quad (2.3)$$

where we have defined the on-resonance damping rate $\gamma_0 = \gamma\omega_0^2/\omega^2$ [69].

A more appropriate approach is a semiclassical method where the polarizability is calculated by considering the atom as a two-level system interacting with a classical radiation field. One finds, when saturation



effects are neglected in the dipole approximation, that the semiclassical approach yields exactly the same result as the Lorentz model. The only exception is that the damping rate γ from Larmor's formula is replaced by [69]

$$\Gamma = \frac{\omega_0^3}{3\pi\epsilon_0\hbar c^3} \left| \langle e | \hat{d} | g \rangle \right|^2, \quad (2.4)$$

with $\hat{d} = -e\mathbf{r}$ being the dipole operator. Plugging this expression into Eqs. (2.1) and (2.2) gives the important relations

$$V_{\text{dip}} = -\frac{3\pi c^2}{2\omega_0^3} \left(\frac{\Gamma}{\omega_0 - \omega} + \frac{\Gamma}{\omega_0 + \omega} \right) I(\mathbf{r}) \simeq \frac{3\pi c^2}{2\omega_0^3} \frac{\Gamma}{\Delta} I(\mathbf{r}), \quad (2.5)$$

$$\Gamma_{\text{sc}} = \frac{3\pi c^2}{2\hbar\omega_0^3} \left(\frac{\omega}{\omega_0} \right)^3 \left(\frac{\Gamma}{\omega_0 - \omega} + \frac{\Gamma}{\omega_0 + \omega} \right)^2 I(\mathbf{r}) \simeq \frac{3\pi c^2}{2\hbar\omega_0^3} \left(\frac{\omega}{\omega_0} \right)^3 \left(\frac{\Gamma}{\Delta} \right)^2 I(\mathbf{r}), \quad (2.6)$$

where $\Delta = \omega - \omega_0$ is the detuning. The left hand side is valid for any driving field with two terms being the resonant contributions at $\omega = \omega_0$ and the so-called counter rotating term at $\omega = -\omega_0$ [69]. It is customary to apply the rotating wave approximation where the detuning is assumed so small that the counter rotating terms may be neglected whereby the expressions on the right hand are obtained.

From Eq. (2.5) we see that the sign of the detuning determines whether the optical potential is attractive $\Delta < 0$ (red detuned) or repulsive $\Delta > 0$ (blue detuned). In red (blue) detuned light fields the atoms are attracted towards intensity maxima (minima). As $\Gamma_{\text{sc}} \sim 1/\Delta^2$, most experiments prefer a large detuning in order to decrease the scattering [70, 71]. The cost of this choice is per Eq. (2.5) a weaker trapping potential, which is compensated by a high laser intensity [66].

2.2 Optical Lattice

A conservative potential trap can as described above be created using far red detuned laser light. The dipole potential in Eq. (2.5) is proportional to the intensity enabling a multitude of different potentials through the interference pattern of multiple laser beams. A single laser beam typically has a Gaussian profile,

$$I(r, z) = I_0 \frac{w_0^2}{w^2(z)} \exp\left(-\frac{2r^2}{w^2(z)}\right), \quad (2.7)$$

where r denotes the radial distance from the beam axis, I_0 is the peak intensity and $w(z) = w_0 \sqrt{1 + z^2/z_R^2}$ is the width, and $z_R = \pi w_0^2/\lambda$ is the Rayleigh length [69].



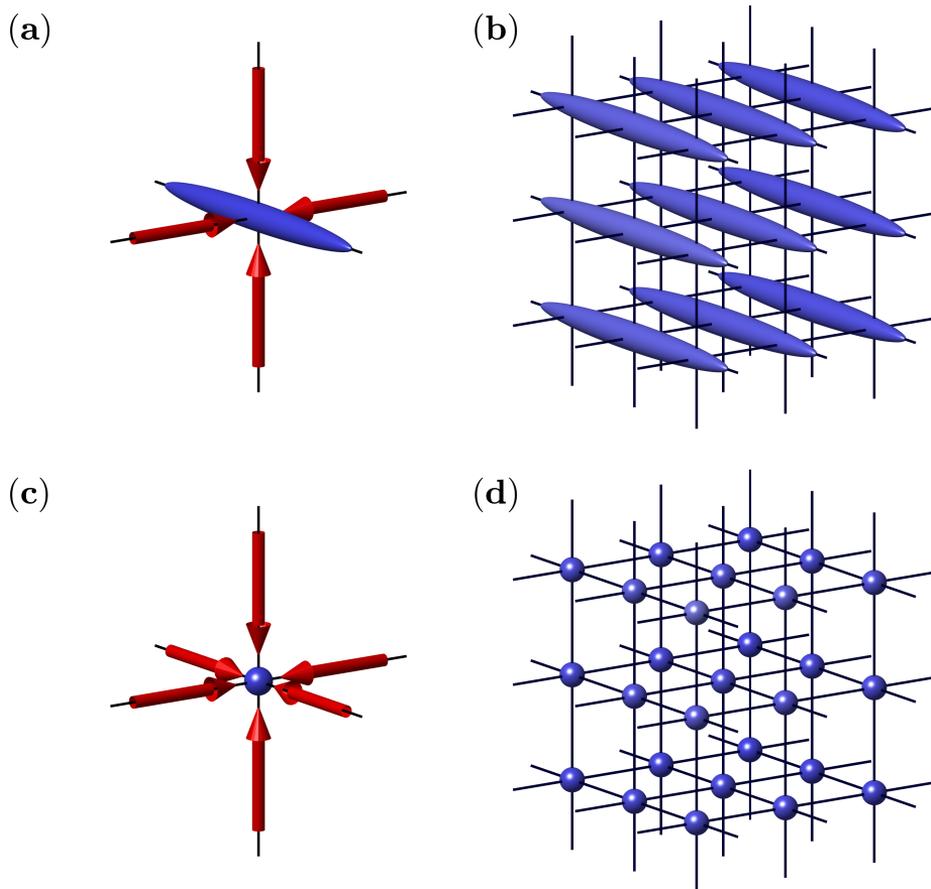


Figure 2.1: (a) and (b) show a single atom and an array of trapped atoms in a two dimensional optical lattice. The lattice is formed from four mutually orthogonal laser beams and the system has a characteristic cigar shape along the remaining axes due to the Gaussian profile of the lasers. (c) and (d) show how six orthogonal laser beams fully confine the atoms giving a result similar to a three dimensional primitive cubic lattice.

Two counter propagating laser beams with the same wavelength λ creates a one-dimensional standing wave with period $a = \lambda/2$. The distance a is known as the lattice spacing. In practice the counter propagating beam is created by shining a single beam on a mirror, whereby it reflects back on itself. From Eqs. (2.7) and (2.5) the associated potential



of the two beams is

$$V(r, z) = -V_0 \exp\left(-\frac{2r^2}{w^2(z)}\right) \cos^2(kz), \quad (2.8)$$

where $V_0 = \left|\frac{3\pi c^2}{2\omega_0^3} \frac{\Gamma}{\Delta} 4I_0\right|$ with $k = 2\pi/\lambda = \pi/a$ where the factor of four is due to constructive interference. The subscript on the potential is suppressed for brevity. Four beams create tube shaped states as seen in Fig. 2.1. In the case of six perpendicular beams the potential close to the trap center is

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

where the Gaussian profile has been neglected as we are close to trap center.

2.2.1 Band Structure

Consider a cubic optical lattice as given in Eq. (2.9) that has terms depending on each coordinate individually. This decouples the time-independent Schrödinger equation into three independent differential equations for each spatial dimension. For instance, consider the x -coordinate

$$\left(-\frac{1}{2m} \frac{\partial^2}{\partial x^2} - V_0 \cos^2(kx) \right) \varphi_q^{(n)}(x) = E \varphi_q^{(n)}(x). \quad (2.10)$$

This is a periodic potential so the eigenstates are described by Bloch's Theorem [72]

$$\varphi_q^{(n)}(x) = e^{iqx} u_q^{(n)}(x), \quad (2.11)$$

which is a plane wave modulated by a function with the same periodicity as the potential $u_q^{(n)}(x) = u_q^{(n)}(x + d)$. The quantum number q is known as the quasi-momentum and it is discretized within the first Brillouin zone due to the periodic boundary conditions as $q = 2\pi m/L$ [72].

If Eq. (2.11) is inserted into Eq. (2.10) we find an equation for $u_q^{(n)}$

$$\left[-\frac{1}{2m} \left(\frac{\partial}{\partial x} + iq \right)^2 + V(x) \right] u_q^{(n)} = E_q^{(n)} u_q^{(n)}. \quad (2.12)$$

It is convenient to expand the potential $u_q^{(n)}$ and $V(x)$ in terms of a Fourier series,

$$u_q^{(n)} = \sum_l c_l^{(n,q)} e^{2iklx}, \quad \text{and} \quad V(x) = \sum_m v_m e^{2ikmx}. \quad (2.13)$$



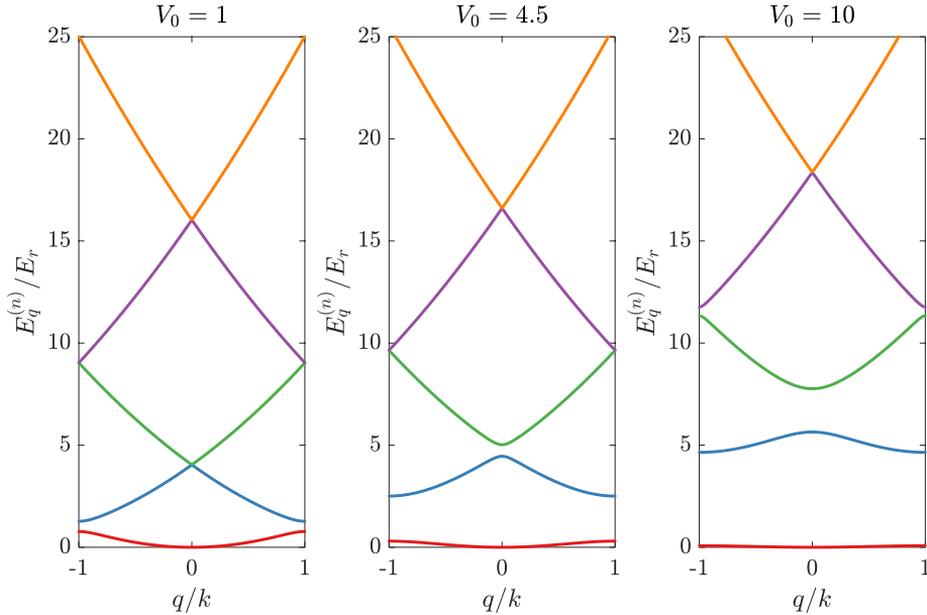


Figure 2.2: The energy spectra (band structure) for the optical lattice Hamiltonian Eq. (2.12) with different lattice depths V_0 . Each energy band is plotted with its own color. Note, how the gaps at the level crossings increase with the lattice depth.

Expanding $V(x) = -V_0 \cos^2(kx)$ shows that its Fourier series only contains three terms $v_{\pm 1} = -V_0/4$ and $v_0 = -V_0/2$. Plugging the expansions into Eq. (2.12) and rearranging gives

$$\sum_l \left[\left(\frac{1}{2m} (q + 2kl)^2 - E_q^n - \frac{V_0}{2} \right) c_l^{(n,q)} - \frac{V_0}{4} (c_{l-1}^{(n,q)} + c_{l+1}^{(n,q)}) \right] e^{2ilkx} = 0. \quad (2.14)$$

The plane waves form an orthogonal basis implying that the square bracket must vanish for all l , which gives a set of algebraic equations for each l ,

$$\left(\frac{1}{2m} (q + 2kl)^2 - \frac{V_0}{2} \right) c_l^{(n,q)} - \frac{V_0}{4} (c_{l-1}^{(n,q)} + c_{l+1}^{(n,q)}) = E_q^n c_l^{(n,q)}. \quad (2.15)$$

This is an eigenvalue problem and it can be expressed as a diagonalization



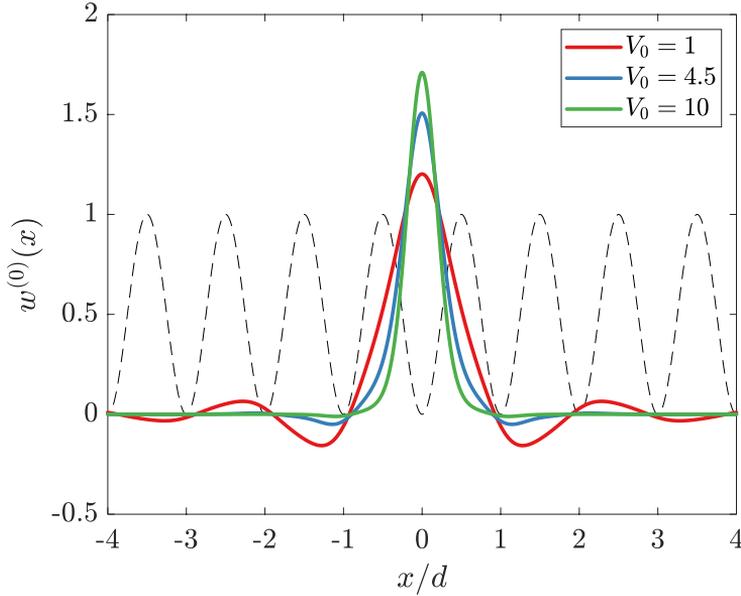


Figure 2.3: The Wannier states plotted for different potential depths (V_0) - see legend. The period of the potential is indicated by a dashed line. The Wannier states have a significant side lobe overlapping with the nearest neighbor site for low values of the lattice depth, which gives rise to tunneling between the sites.

with deeper lattices and for a shallow lattice they have characteristic side lobes at the neighboring sites. As we will discuss later, these side lobes cause tunneling between neighboring sites.

2.3 Atom Chips

Besides optical traps an attractive platform for trapping ultracold atoms are magnetic microtraps formed by currents on microchips, which is known as *atom chips*. These traps offer a very high level tunability of the fields, which is ideal for control applications. However, they only permit trapping of significantly fewer atoms than the optical traps. In this section we give a brief introduction. Assuming a weak magnetic field such that $\hat{\mathbf{F}}$ is a good quantum number then the Zeemann Hamiltonian is

$$\hat{H} = g_F \mu_B \mathbf{B} \cdot \hat{\mathbf{F}}, \quad (2.19)$$

where g_F is the Landé g-factor, μ_B is the Bohr magneton, and $\hat{\mathbf{F}}$ is the collective angular spin operator [74]. The sign of the Landé g-factor



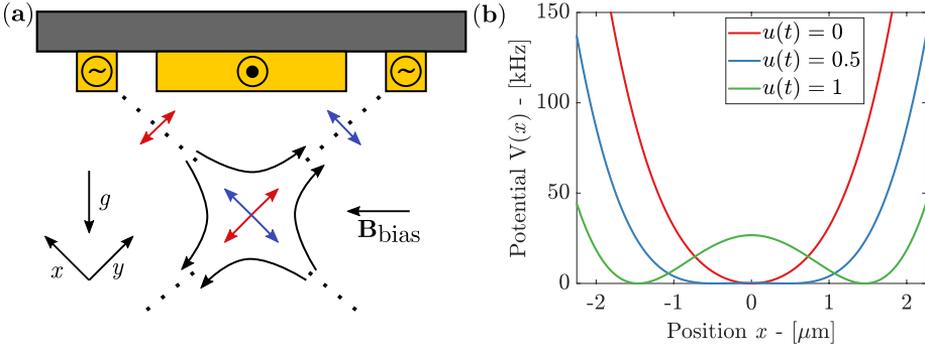


Figure 2.4: **(a)** A graphical illustration of a typical atom chip experiment. The trapping potential is created by the current running in the wires on the top. **(b)** The trapping adiabatic potential plotted for different values of $u(t)$ - see legend. For $u(t) = 0$ the potential is a single anharmonic well and it deforms into a double well as $u(t)$ changes into $u(t) = 1$.

distinguishes states, which are either aligned parallel or anti-parallel to the external magnetic field. These are known as weak and strong field seekers respectively. This effects allows trapping of atoms in extrema of the magnetic fields. At a glance Eq. (2.19) offers a complete freedom to design potential landscapes by designing the magnetic field, however Maxwell's equations constrain the allowed field configurations. For instance, Earnshaw's theorem states that there are no magnetic field maxima inside a source-free region [75]. This restriction may be avoided using time-dependent fields.

Later in this thesis, we will discuss applications of a trap setup consisting of a so-called Ioffe-Pritchard configuration [76] together with external radio-frequency fields. The trap geometry can be realized by current carrying wires mounted on an atom chip [77, 78] that enables tight confinement due to the high magnetic field gradients [79]. At any point in space, we may pick the static field ($\mathbf{B}_S(\mathbf{r})$) as the quantization axis. Assuming a linearly polarized magnetic field with respect to this quantization axis the Hamiltonian becomes

$$\hat{H} = g_F \mu_B \left(|\mathbf{B}_S(\mathbf{r})| \hat{F}_z + B_{\text{RF}} \cos(\omega t) \hat{F}_x \right), \quad (2.20)$$

where B_{RF} and ω are the magnitude and angular frequency of the time-dependent radio frequency field. The second term in Eq. (2.20) may be



rewritten as

$$\cos(\omega t)\hat{F}_x = \frac{1}{4}\left(\hat{F}_+e^{i\omega t} + \hat{F}_+e^{-i\omega t} + \hat{F}_-e^{i\omega t} + \hat{F}_-e^{-i\omega t}\right), \quad (2.21)$$

where \hat{F}_\pm are the raising and lowering operators. It is convenient to transform this Hamiltonian into a frame rotating with the time-dependent field as $|\tilde{\psi}\rangle = \hat{U}|\psi\rangle$ where $\hat{U} = \exp(i\frac{g_F}{|g_F|}F_z\omega t)$. In the transformed frame the Hamiltonian becomes

$$\tilde{H} = \hat{U}\hat{H}\hat{U}^\dagger + i(\partial_t\hat{U})\hat{U}^\dagger. \quad (2.22)$$

Evaluating the two terms in this transformed Hamiltonian for Eq. (2.21) one finds

$$i(\partial_t\hat{U})^\dagger\hat{U} = -\frac{g_F}{|g_F|}\omega\hat{F}_z, \quad (2.23)$$

$$\begin{aligned} \hat{U}\cos(\omega t)\hat{F}_x\hat{U}^\dagger &= \frac{1}{4}\left[\hat{F}_+\left(1 + e^{2i\frac{g_F}{|g_F|}\omega t}\right) + \hat{F}_-\left(1 + e^{-2i\frac{g_F}{|g_F|}\omega t}\right)\right] \\ &\simeq \frac{1}{2}\hat{F}_x, \end{aligned} \quad (2.24)$$

where we used the rotating wave approximation neglecting the terms proportional to 2ω on the last identity. Collecting the results from above Hamiltonian (2.20) in the transformed frame reads

$$\begin{aligned} \tilde{H} &= g_F\mu_B\left(|\mathbf{B}_S(\mathbf{r})| - \frac{\omega}{|g_F|\mu_B}\right)\hat{F}_z + \frac{g_F\mu_B}{2}B_{\text{RF}}\hat{F}_x \\ &= \Delta(\mathbf{r})\hat{F}_z + \Omega_{\text{RF}}(\mathbf{r})\hat{F}_x. \end{aligned} \quad (2.25)$$

The reader may recognize this as the Landau-Zener model if the spin is $F = 1/2$, which is discussed more in Sec. 3.1.1. The adiabatic potentials are the eigenvalues of Eq. (2.25)

$$V(\mathbf{r}) = m'_F g_F \mu_B \sqrt{\Delta^2(\mathbf{r}) + \Omega_{\text{RF}}^2(\mathbf{r})}, \quad (2.26)$$

As with m_F , m'_F has $2F + 1$ values from $-F$ to F defining a range of possible adiabatic potentials. For ^{87}Rb in the $F = 1$ state there is only three adiabatic potentials whereof one has a trapping geometry. In the adiabatic approximation the atoms always reside in the same adiabatic potential due to the slow spatial and temporal variation of the Hamiltonian. These adiabatic potentials are also known as the (magnetically) dressed states, which have been used in a number of experiments beyond the applications discussed in this thesis. A recent review is given by the experts Garraway



and Perrin [80]. Later in Sec. 5.3, we will use the potential in Eq. (2.26) to study optimal coherent splitting of a Bose-Einstein condensate. This potential has been realized experimentally on an atom chip [81, 82]. The intensity of the radio-frequency field (B_{RF}) can experimentally be varied as $B_{\text{RF}} = (0.5 + 0.3u(t))\text{G}$ in units of Gauss [46]. As $u(t)$ changes from zero to one the trap changes from a single to a double well - see Fig. 2.4(b). The deformation from a single to a double well has been used to study coherent splitting of Bose-Einstein condensates, which can be used to realize a Mach-Zehnder type interferometer using matter waves [83].

2.4 Multiparticle systems

In this thesis we discuss optimal control of ultracold systems with bosons, whose statistics obey the Bose-Einstein distribution

$$n_i = \frac{g_i}{\exp((E_i - \mu)/k_B T) - 1}, \quad (2.27)$$

where i denotes a state, n_i is the number of particles in the state, E_i is the energy, g_i is the degeneracy, k_B is the Boltzmann constant, T is the temperature, and μ is the chemical potential. The key difference to fermions is that multiple particles can occupy the same state. At high temperatures the energy spectrum is effectively continuous and the distinction is of little importance. However, at low temperatures the spectrum is discrete and the statistics become very important. As seen from Eq. (2.27) the population of the ground state diverges for $T \rightarrow 0$. Even below some critical finite temperature T_C there is a macroscopic population of the ground state that signifies the point where the particles form a Bose-Einstein condensate [84].

2.4.1 Second Quantization

Whenever discussing a system of multiple particles it is convenient to use the second quantization formalism. If $\{|\varphi_j\rangle\}$ is a basis for the single particle states then we can construct a Fock space as all the correctly symmetrized product states from these single-particle states. Let $|n_1, n_2, \dots, n_j, \dots\rangle$ denote such a multi-particle Fock state where the single-particle state $|\varphi_j\rangle$ enters n_j times. The Fock states form an orthonormal basis and they are interlinked by the creation and annihilation operators \hat{a}_j^\dagger and \hat{a}_j that add and remove a particle in the state $|\varphi_j\rangle$. These operators



satisfy the relations [85],

$$\hat{a}_j^\dagger |n_1, n_2, \dots, n_j, \dots\rangle = \sqrt{n_j + 1} |n_1, n_2, \dots, n_j + 1, \dots\rangle \quad (2.28a)$$

$$\hat{a}_j |n_1, n_2, \dots, n_j, \dots\rangle = \sqrt{n_j} |n_1, n_2, \dots, n_j - 1, \dots\rangle. \quad (2.28b)$$

These operators also satisfy the bosonic commutation relations

$$[\hat{a}_j, \hat{a}_i] = [\hat{a}_j^\dagger, \hat{a}_i^\dagger] = 0, \quad [\hat{a}_j, \hat{a}_j^\dagger] = \delta_{i,j}. \quad (2.29)$$

From Eqs. (2.28a) and (2.28b) it follows that any Fock state can be constructed by consecutive operations on the vacuum state

$$|n_1, n_2, \dots, n_j, \dots\rangle = \frac{(\hat{a}_1^\dagger)^{n_1} (\hat{a}_2^\dagger)^{n_2} \dots (\hat{a}_j^\dagger)^{n_j} \dots}{\sqrt{n_1! n_2! \dots n_j! \dots}} |0\rangle. \quad (2.30)$$

In order for this formalism to be applicable it is also necessary to include operators. If \hat{O} is an N -particle operator in first quantization then the second quantized version is [85]

$$\hat{O} = \frac{1}{N!} \sum_{i_1, \dots, i_N} \sum_{j_1, \dots, j_N} \hat{a}_{i_1}^\dagger \dots \hat{a}_{i_N}^\dagger \langle n_{i_1}, \dots, n_{i_N} | \hat{O} | n_{j_1}, \dots, n_{j_N} \rangle \hat{a}_{j_1} \dots \hat{a}_{j_N}. \quad (2.31)$$

In order to apply \hat{O} we need the matrix elements of the operator in the single particle basis $\{|n_i\rangle\}$.

2.4.2 Gross-Pitaevskii Equation from Weakly Interacting Particles

In this section we express the Hamiltonian in second quantized formalism. The Hamiltonian can be written as $\hat{H} = \hat{H}^{(0)} + \hat{H}^{(2)}$ denoting respectively the non-interacting and interacting parts of the Hamiltonian. The two refers to the fact that we only consider two-particle interactions. It is insightful to express the Hamiltonian using the field operators,

$$\hat{\psi}(\mathbf{r}) = \sum_k \varphi_k(\mathbf{r}) \hat{a}_k, \quad \hat{\psi}^\dagger(\mathbf{r}) = \sum_k \varphi_k^*(\mathbf{r}) \hat{a}_k^\dagger. \quad (2.32)$$

that annihilates and creates a particle at position \mathbf{r} expressed as a linear combination of the single particle energy eigenstates $|\varphi_k\rangle$ with energy E_k . This allows us to express the non-interacting part of the Hamiltonian as

$$\begin{aligned} \hat{H}^{(0)} &= \sum_k E_k \hat{a}_k \hat{a}_k = \sum_k \iint \hat{\psi}^\dagger(\mathbf{r}_1) \varphi_k^*(\mathbf{r}_1) E_k \varphi_k(\mathbf{r}_2) \hat{\psi}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\ &= \int \hat{\psi}^\dagger(\mathbf{r}) \left(-\frac{1}{2m} \nabla^2 + V(r) \right) \hat{\psi}(\mathbf{r}) d\mathbf{r}, \end{aligned} \quad (2.33)$$



where we used that φ_k are the energy eigenstates and the orthonormality of the wave functions $\sum_k \varphi_k(\mathbf{r}_2)\varphi_k^*(\mathbf{r}_2) = \delta^3(\mathbf{r}_1 - \mathbf{r}_2)$.

Due to the low temperature of ultracold experiments it is valid only to consider two-particle interactions

$$\hat{H}^{(2)} = \frac{1}{2} \sum_{i \neq j} V(\mathbf{r}_i - \mathbf{r}_j). \quad (2.34)$$

Neutral atoms interact through the short range Lennard-Jones potential [66]. However, at low energies the atoms can only probe the s -wave scattering state, since the higher momentum states are reflected by the centrifugal barrier [66]. Furthermore the thermal de Broglie wavelength of cold atoms is much larger than the effective interactive range, whereby the details of the scattering potential become irrelevant and it may be replaced by a pseudo-potential

$$V(\mathbf{r}_1 - \mathbf{r}_2) = g\delta^3(\mathbf{r}_1 - \mathbf{r}_2) = \frac{4\pi\hbar^2 a}{m} \delta^3(\mathbf{r}_1 - \mathbf{r}_2). \quad (2.35)$$

This potential gives the same phase for s -wave scattering as the actual potential characterized by the scattering length a [66].

The pseudo-potential allows us to write the interaction in second quantized form

$$\begin{aligned} \hat{H}^{(2)} &= \frac{1}{2} \iint \hat{\psi}^\dagger(\mathbf{r}_1)\hat{\psi}^\dagger(\mathbf{r}_2)V(\mathbf{r}_1 - \mathbf{r}_2)\hat{\psi}(\mathbf{r}_1)\hat{\psi}(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2 \\ &= \frac{g}{2} \int \hat{\psi}(\mathbf{r})^\dagger(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r})\hat{\psi}(\mathbf{r})d\mathbf{r} \end{aligned} \quad (2.36)$$

This together with Eq. (2.33) gives the full second quantized Hamiltonian. Solving Heisenberg equations of motion for the field operators gives

$$i\frac{\partial}{\partial t}\hat{\psi}(\mathbf{r}) = [\hat{\psi}(\mathbf{r}), \hat{H}] = \left(-\frac{1}{2m}\nabla^2 + V(\mathbf{r}) + g\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r})\right)\hat{\psi}(\mathbf{r}). \quad (2.37)$$

This set of equations is not analytically solvable. However, if the scattering length a is much less than the mean inter-particle distance $na^3 \ll 1$, where n is the density of the gas, it is viable to use a mean-field description

$$\hat{\psi}(\mathbf{r}) = \psi(\mathbf{r}) + \delta\hat{\psi}(\mathbf{r}), \quad (2.38)$$

where $\psi(\mathbf{r})$ is the c -number mean-field and $\delta\hat{\psi}(\mathbf{r})$ is a fluctuation. Neglecting all such fluctuations leads to the celebrated Gross-Pitaevskii equation [86, 87]

$$i\frac{\partial}{\partial t}\psi(\mathbf{r}) = \left(-\frac{1}{2m}\nabla^2 + V(\mathbf{r}) + g|\psi(\mathbf{r})|^2\right)\psi(\mathbf{r}). \quad (2.39)$$



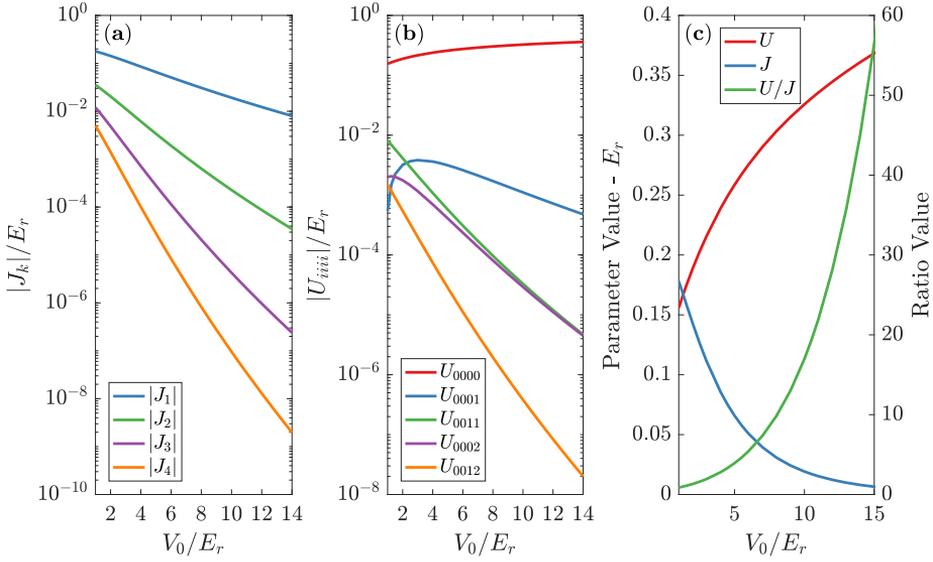


Figure 2.5: **(a)** The absolute value of the tunneling $|J_k|$, see legend. **(b)** The value of the on-site and off-site two particle interactions. **(c)** The tunneling (J) and interaction strength (U) matrix elements left and their ratio plotted right. All calculations are performed for Rubidium-87 atoms in an optical lattice with wavelength $\lambda = 1064\text{nm}$ and transverse lattice depths of $20E_r$.

The Gross-Pitaevskii equation is very similar to the regular Schrödinger equation except for the nonlinear repulsion term $g|\psi(\mathbf{r})|^2$, which makes it difficult to solve in regions with low density. It follows that in the weakly interacting limit all the particles condense into the same state, but it is not the ground state of a single particle problem but rather the ground state of the Gross-Pitaevskii equation. This equation is very successful in describing both qualitative and quantitative phenomena and later in Chap. 4-5 we will discuss quantum optimal control with the Gross-Pitaevskii equation.

2.5 Bose-Hubbard Model

The Bose-Hubbard model describes bosons in an optical lattice and it is the starting point for many studies on ultracold bosons [66, 88]. In Chap. 8 we will discuss different approaches for controlling the dynamics of this model. The Bose-Hubbard model provides an adequate description



of the dynamics whenever the optical lattice is sufficiently deep. The lattice must be so deep that the band gaps in Fig. 2.2 are enough large to restrict the dynamics to the lowest band.

In an optical lattice given by Eq. (2.9) the Wannier states form a natural single particle basis. For this reason the field operators are expanded in these states $\hat{\psi}(\mathbf{r}) = \sum_k w(x - x_j)$, where only the Wannier functions from the lowest band are included. This is known as the lowest band approximation. The Hamiltonian becomes [89],

$$\begin{aligned} \hat{H} &= \int \sum_{i,j} w^*(\mathbf{r} - \mathbf{r}_i) \hat{a}_i^\dagger \left(-\frac{1}{2m} \nabla^2 + V(\mathbf{r}) \right) w(\mathbf{r} - \mathbf{r}_j) \hat{a}_j \mathbf{d}\mathbf{r} \\ &+ \frac{g}{2} \int \sum_{ijkl} w^*(\mathbf{r} - \mathbf{r}_i) w^*(\mathbf{r} - \mathbf{r}_j) w(\mathbf{r} - \mathbf{r}_k) w(\mathbf{r} - \mathbf{r}_l) \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k \mathbf{d}\mathbf{r} \\ &= - \sum_{ij} J_{ij} \hat{a}_i^\dagger \hat{a}_j + \frac{1}{2} \sum_{ijkl} U_{ijkl} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k, \end{aligned} \quad (2.40)$$

where

$$J_{ij} = - \int w^*(\mathbf{r} - \mathbf{r}_i) \left(-\frac{1}{2m} \nabla^2 + V(\mathbf{r}) \right) w(\mathbf{r} - \mathbf{r}_j) \mathbf{d}\mathbf{r}, \quad (2.41)$$

$$U_{ijkl} = \int w^*(\mathbf{r} - \mathbf{r}_i) w^*(\mathbf{r} - \mathbf{r}_j) w(\mathbf{r} - \mathbf{r}_k) w(\mathbf{r} - \mathbf{r}_l) \mathbf{d}\mathbf{r}. \quad (2.42)$$

Due to the translational invariance of the system over a lattice spacing we have that $J_{ij} = J_{|i-j|} = J_k$, which gives an interpretation of the J_k coefficients as the hopping or tunneling strengths between two sites k sites apart. The zero range terms J_0 are removable energy offsets. In Fig. 2.5(a) the coupling strength is calculated for different values of the lattice depth. In this calculation it has been utilized that the Wannier functions are products $w(\mathbf{r} - \mathbf{r}_i) = w_x(x - x_i)w_y(y - y_i)w_z(z - z_i)$ for a cubic lattice. The non-vanishing values of $J_k, k > 0$ are due to the characteristic side lobes of the Wannier functions illustrated in Fig. 2.3. As the lattice depth increases the Wannier functions become increasingly localized causing the matrix elements J_k to drop off rapidly. Already for $V_0/E_r > 2$ the next nearest neighbor coupling strengths $J_k \geq 2$ are two order of magnitude below J_1 - see Fig. 2.5(a).

In a similar manner to J_{ij} the matrix elements of U_{ijkl} are also invariant under index permutations and shifts due to the translational invariance over a lattice site. Taking the site $i = 0$ as a fixed site, U_{0000} can be interpreted as the on-site two particle interaction. The other terms are longer range interactions due to the two-particle interactions, which are heavily suppressed due to the localized nature of the Wannier states as illustrated in Fig. 2.5(b).



These observations motivate the so-called tight binding approximation that neglects all terms except $J = J_1$ and $U = U_{0000}$, which are given by

$$J = \int w^*(x - x_0) \left(-\frac{1}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right) w(x - x_1) dx, \quad (2.43)$$

$$U = \frac{4\pi\hbar^2 a}{m} \left(\int |w(x)|^4 dx \right)^3, \quad (2.44)$$

where the power of three is a geometric factor. It is possible to derive approximate relations for J and U in terms of lattice depth measured in units of the recoil energy $s = V_0/E_r$ [66],

$$J \approx \frac{4}{\sqrt{\pi}} E_r s^{3/4} e^{-s\sqrt{s}}, \quad (2.45)$$

$$U \approx \sqrt{\frac{8}{\pi}} ka E_r s^{3/4}. \quad (2.46)$$

The relations are accurate within 10% for recoil energies above 15. These relations capture the qualitative behavior in Fig. 2.5, namely that J (U) decrease (increase) with the lattice depth. The precise ratio U/J is calculated numerically from Eqs. (2.43)-(2.44) in Fig. 2.5(c)

Within the tight-binding approximation, the Hamiltonian Eq. (2.40) reduces to the Bose-Hubbard Hamiltonian

$$\hat{H} = -J \sum_{\langle i,j \rangle} \hat{a}_i^\dagger \hat{a}_j + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1). \quad (2.47)$$

The notation $\langle i,j \rangle$ refers to a sum only over the nearest neighbours and $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$ is the number operator for site i . The first term describes the kinematics as tunneling between neighbouring sites whereas the second term describes the interaction between particles on the same site.

2.5.1 Phases of the One-Dimensional Bose-Hubbard Model

A particularly intriguing aspect of the Bose-Hubbard model is that it has two distinct quantum phases being the superfluid phase for low values of U/J and Mott insulator for high values of U/J . For simplicity, we only discuss the one-dimensional Bose-Hubbard model in this section.

In the limit $J \gg U$ the second term in Eq. (2.47) can be neglected, which gives

$$\hat{H} = \hat{H}_J = -J \sum_i \hat{a}_i^\dagger \hat{a}_{i+1} + \hat{a}_i \hat{a}_{i+1}^\dagger. \quad (2.48)$$



This can be diagonalized by a transformation into Bloch waves $\hat{a}_i = \sum_q \hat{a}_q e^{-iqr} / \sqrt{L}$ where L is the number of lattice sites. Inserting this into Eq. (2.48) gives

$$\hat{H}_J = -2J \sum_q \cos(qd) \hat{a}_q^\dagger \hat{a}_q, \quad (2.49)$$

where the eigenenergies can be identified as $E_q = -2J \cos(qd)$, which is a gapless spectrum. The lowest energy state is the superfluid state

$$|\text{SF}\rangle = \frac{1}{\sqrt{N}} (\hat{a}_{q=0}^\dagger)^N |0\rangle = \frac{1}{\sqrt{N}} \left(\frac{1}{\sqrt{L}} \sum_{j=1}^L \hat{a}_j^\dagger \right)^N |0\rangle, \quad (2.50)$$

where N is the number of particles. This is a product state where all the individual particles are in the same single particle state similar to the Gross-Pitaevskii equation. This state is completely delocalized over the entire lattice. In the limit of $N, L \rightarrow \infty$ at fixed filling $n = N/L$ the superfluid state becomes indistinguishable from having

$$|\text{SF}\rangle \simeq \prod_j \exp \left(\sqrt{\frac{N}{L}} \hat{a}_j^\dagger \right) |0\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \dots, \quad (2.51)$$

where $|\alpha_i\rangle$ are coherent states defined as eigenstates of the annihilation operator for the i 'th site $\hat{a}_i |\alpha_i\rangle = \alpha_i |\alpha_i\rangle$ [66]. From Eq. (2.51) it can be identified that $n = N/L = \langle \hat{n}_i \rangle = |\alpha_i|^2$ showing that the norm of the coherent states is the on-site density. The fluctuations can also be calculated as $\langle \Delta \hat{n}_i \rangle = n$ whereby the on-site probability distribution can be identified as a Poissonian distribution [66]. This is an important fact that we will utilize later in Chap. 8.

In the opposite limit $J \ll U$ the tunneling term in Eq. (2.47) may be neglected and the Hamiltonian becomes

$$\hat{H}_U = \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1). \quad (2.52)$$

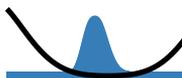
The operator is already diagonal in the Fock states and it penalizes sites with multiple particles causing the ground state to prefer identical occupancy on all sites. Consider the special case of unit filling $n = 1$ then the many-body ground state is a product,

$$|\text{Mott}\rangle = \prod_j \hat{a}_j^\dagger |0\rangle. \quad (2.53)$$

This state is known as the Mott state and it has precisely one particle per site and minimal energy $E = 0$. Unlike the superfluid the Mott insulator



has full localization on each site and no variance in the local atom number. Any fluctuations in the local atom number will increase the energy by U , hence the energy spectrum is gapped unlike the superfluid. This also causes the Mott state to become incompressible $\partial n / \partial \mu = 0$ where μ is the chemical potential [66]. For finite J the system remains in a localized state as long as the kinetic energy gain from J is small compared to the on-site interactions U but the ground state no longer has a simple product form as in Eq. (2.53) [66]. If J becomes on the order of U the kinetic energy gain outweighs the repulsion and the system undergoes a quantum phase transition to a superfluid. This transition is sharp in the thermodynamic limit since the superfluid unlike the Mott exhibits long range order in the single particle density matrix [66].

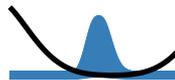


3 Quantum Optimal Control

Most of Sec. 3.2.2 and Sec. 3.3 is copied from the appendix in [5]. The discussion of a chopped basis is copied from Ref. [5]. Fig. 3.2 and caption are taken from Ref. [5]. Parts of the discussion in Sec. 3.5 is copied from the presentation in [5]. The presentation in Sec. 3.7 is copied from [8].

Technological advances in the experimental toolbox in physical chemistry and atomic, molecular, and optical physics currently enable exciting new developments in the manipulation of complex quantum systems. Gradually, the focus is shifting from verifying the validity of theoretical models towards manipulating quantum systems for specific technological applications [66, 90]. The vision is to engineer and *control* instead of just observing it. Some examples of this trend are quantum state preparation [91, 92], atomic clocks [93], quantum based sensors [94–96], quantum simulators [26], and quantum computers [97, 98].

These applications require the ability to steer the quantum dynamics precisely using external electromagnetic fields. The details of how a particular quantum system is manipulated may greatly differ from application to application but problems share a common mathematical framework in quantum control theory. It is beyond the scope of any thesis to treat everything within this field, and this thesis focuses on subfield of quantum *optimal* control. In this subfield the quality of a control is measured by a cost function, which can describe e.g. distance to some target state [46], similarity with a unitary gate operator [99], or the amount of experimental signal [100]. This goal in quantum optimal control is to optimize this cost functional and preferably find high quality solutions efficiently [30, 38, 101]. This theory has been studied in a wide range of physical systems [45, 46, 102–104].



This chapter starts by defining the state-to-state problem in Sec. 3.1. The optimality equations are introduced in Sec. 3.2. The equations describe necessary conditions for any control to be optimal. In Sec. 3.3 the optimality equations are used to calculate the gradient of the state-to-state problem. This is a methods chapter as this chapter aims to give a cohesive introduction to some of the primary optimal control algorithms applied in the chapters 4-6. Gradient based methods (GRAPE) are discussed in Sec. 3.4. In Sec. 3.5 I discuss Krotov's method that is an alternative derivative based method. In Sec. 3.6 the popular gradient-free method Nelder-Mead with CRAB is introduced. In Sec. 3.7 I discuss how to calculate the gradient in a chopped basis whereby more efficient optimization than Nelder-Mead becomes possible. For completeness, this chapter also includes a brief introduction to the concept of controllability in Sec. 3.8. The chapter is concluded in Sec. 3.9

3.1 The Mathematical Structure of Quantum Control Problems

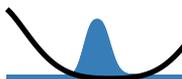
A finite dimensional control system is a system of ordinary differential equations displaying one or more time varying functions of time $u(t)$ known as the controls [32]. Hence, a control system has the form

$$\dot{\mathbf{x}} = f(t, \mathbf{x}, u), \quad (3.1)$$

where \mathbf{x} represents the state completely describing the system and f is the dynamical equation of motion. Translated into the quantum setting, the state is $|\psi(t)\rangle$ and the dynamical equation is the Schrödinger equation,

$$|\dot{\psi}\rangle = -i\hat{H}(u)|\psi\rangle, \quad (3.2)$$

where \hat{H} is the system Hamiltonian. Later, we will also discuss problem where the dynamical equation is the Gross-Pitaevskii equation Eq. (2.39). This thesis concerns so-called state-to-state control problems where some initial state $|\psi_0\rangle$ must be transported into a target state $|\psi_t\rangle$. The quality of such a transfer is quantified by the fidelity $F = |\langle\psi_t|\psi(T)\rangle|^2$ describing the fraction of the final state $|\psi(T)\rangle$ in the target state $|\psi_t\rangle$. In quantum optimal control theory this problem is expressed as a minimization problem,



The State-to-State Control Problem

Given an initial state $|\psi_0\rangle$ and a target state $|\psi_t\rangle$ the state-to-state control problem is,

$$\min_{\psi, u} J(\psi, u) = \frac{1}{2} \left(1 - |\langle \psi_t | \psi(T) \rangle|^2 \right) + \frac{\gamma}{2} \int_0^T \dot{u}^2 dt, \quad (3.3a)$$

$$\text{s.t. } i\dot{\psi} = \hat{H}(u)\psi, \quad \text{in } 0 \leq t \leq T, \quad (3.3b)$$

$$\psi(0) = \psi_0, \quad u(0) = u_i, \quad u(T) = u_T, \quad (3.3c)$$

where u_i and u_T are the fixed initial and final values of the control.

The first term in Eq. (3.3a) is proportional to the infidelity $1 - F$ favoring high quality transfers. The second term is known as the regularization that penalizes rapidly fluctuating controls, which are typically not experimentally feasible. The weight factor γ sets the ratio between the two terms. The final time T is called the duration whereas time refers to the time during the transfer. In most applications, we require very high fidelities $F \geq 0.99$. Typically the Hamiltonian has a bilinear form

$$\hat{H}(u) = \hat{H}_0 + \sum_n \hat{H}_n u_n(t), \quad (3.4)$$

where the fixed Hamiltonian \hat{H}_0 is known as the drift Hamiltonian.

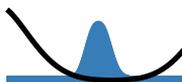
Often the transfer is between eigenstates of the Hamiltonian at $t = 0$ and $t = T$ making it easy to find solutions at high durations due to the adiabatic theorem. However, it is much more challenging to find solutions at short durations.

3.1.1 The Landau-Zener Model

A simple yet educational example of a state-to-state control problem is the Landau-Zener model. Despite the simplicity of this model it is ubiquitous in quantum physics. It is also one of the few models where the state-to-state control problem can be solved analytically. The Hamiltonian is

$$\hat{H}_{\text{LZ}} = \begin{pmatrix} \Delta(t) & \Omega_R \\ \Omega_R & -\Delta(t) \end{pmatrix} = \Omega_R \sigma_x + \Delta(t) \sigma_z, \quad (3.5)$$

where Ω_R is the positive Rabi-frequency, $\Delta(t)$ is the unbounded detuning, and σ_x, σ_z are the usual Pauli matrices. In this example, the detuning is the control $u(t) = \Delta(t)$. This problem is linear in the control and the drift Hamiltonian is $\hat{H}_0 = \Omega_R \sigma_x$. Whenever discussing the two level



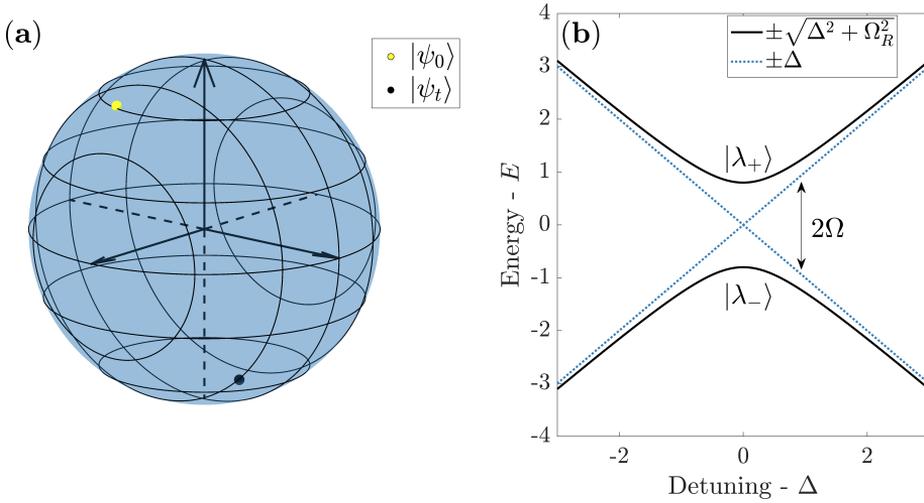


Figure 3.1: (a) An illustration of some initial state $|\psi_0\rangle$ and a target state $|\psi_t\rangle$ on the Bloch sphere. (b) The energy levels in the Landau-Zener model as a function of Δ . The blue dotted lines are the energies with $\Omega_R = 0$ and for $\Omega_R \neq 0$ an avoided crossing appears.

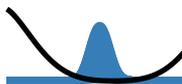
systems it is beneficial visualize states on the Bloch sphere. A state on the Bloch sphere with polar angle θ and azimuthal angle φ is given as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle, \quad (3.6)$$

with the phase convention $\langle\psi|\psi\rangle = 1$. An example state-to-state control problem is shown in Fig. 3.1(a), clearly illustrating that many possible curves exist on the Bloch sphere connecting the initial and target state. Hence, one should in general expect multiple solutions to exist any state-to-state control problem. In many applications there are experimental constraints beyond just reaching the target state such as robustness or power consumption, implying that all solutions may not be equally good. Chap. 7 explores this observation in more detail when designing optimal adiabatic shortcuts.

A straightforward way to diagonalize the Landau-Zener Hamiltonian is to rewrite the Hamiltonian in the form,

$$\hat{H}_{\text{LZ}} = \hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \sqrt{\Delta^2(t) + \Omega_R^2}, \quad \hat{\mathbf{n}} = \frac{1}{\sqrt{\Delta^2(t) + \Omega_R^2}} \begin{pmatrix} \Omega_R \\ 0 \\ \Delta(t) \end{pmatrix}. \quad (3.7)$$



On the Bloch sphere, the associated time evolution operator $\hat{U} = \exp(-i\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \Omega t)$ rotates a state around the xz -plane confined vector $\hat{\mathbf{n}}$ with $\Omega = \sqrt{\Delta^2(t) + \Omega_R^2}$. The eigenstates of the Hamiltonian must be stationary with respect to this rotation implying they are respectively parallel and antiparallel with $\hat{\mathbf{n}}$,

$$|\lambda_+\rangle = \cos \frac{\theta(t)}{2} |0\rangle + \sin \frac{\theta(t)}{2} |1\rangle, \quad |\lambda_-\rangle = \sin \frac{\theta(t)}{2} |0\rangle - \cos \frac{\theta(t)}{2} |1\rangle, \quad (3.8)$$

with corresponding eigenvalues $\pm\Omega$ and $\theta(t)$ is the angle between $\hat{\mathbf{n}}$ and the z -axis,

$$\theta(t) = \arctan \frac{\Omega_R}{\Delta(t)} \quad \text{for } \Delta \geq 0 \quad \text{and} \quad \theta(t) = \pi + \arctan \frac{\Omega_R}{\Delta(t)} \quad \text{for } \Delta < 0. \quad (3.9)$$

The eigenvalues are plotted in Fig. 3.1(b) as a function of Δ and for a non-zero Ω_R an avoided crossing appears. The avoided crossing allows the Landau-Zener model to approximate the spectrum of complex systems in the vicinity of a single avoided crossing.

3.2 Necessary Conditions for Optimality

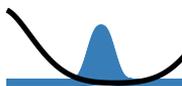
Recall the principle of least action, stating that a classical particle always follow the state trajectory minimizing the action. The trajectories satisfying $\delta S = 0$ give rise to the Euler-Lagrange equations. In a similar spirit we can ask, what equations characterize the trajectories with least cost $\delta J = 0$. This will allow us to identify Euler-Lagrange type equations for the state-to-state control problem. Later these equations will allow us to develop powerful numerical algorithms for minimizing the cost function (3.3a).

A proper definition of a variational derivative is the Gâteaux derivative: If $f : X \rightarrow Y$ is a function, where X and Y are normed vector spaces then the Gâteaux derivative at $g \in X$ in the direction of $\delta h \in X$ is [105]

$$D_{\delta g} f(g) = \lim_{\varepsilon \rightarrow 0} \frac{f(g + \varepsilon \delta g) - f(g)}{\varepsilon} = \left. \frac{d}{d\varepsilon} f(g + \varepsilon \delta g) \right|_{\varepsilon=0}. \quad (3.10)$$

The first equality in the definition Eq. (3.10) is good for numerical implementation whereas the second is most useful for analytic calculations. This derivative expresses that we are calculating the derivative of a function f at g along the direction of another function δg , i.e. it is a direct generalization of a regular directional derivative.

A constrained optimization problem such as the state-to-state control problem may solved within the framework of Lagrange multipliers, which



introduces the Lagrangian function

$$\mathcal{L}(\psi, u, \chi) = J(\psi, u) + \Re \int_0^T \langle \chi | Z(\psi, u) \rangle dt = J(\psi, u) + J_Z(\psi, u, \chi). \quad (3.11)$$

where χ is the Lagrange multiplier and $Z(u, \psi)$ is the constraint given by [101, 106]

$$Z(u, \psi) = i\dot{\psi} - \hat{H}\psi - \beta|\psi|^2\psi. \quad (3.12)$$

Without loss of generality, the Gross-Pitaevskii equation is used as the equation of motion since the regular linear Schrödinger equation is a special case $\beta = 0$. When minimizing the Lagrangian the equation of motion is not a constraint as in Eq. (3.3) but a penalty. A necessary condition for an optimal point (ψ^*, u^*, χ^*) is that all the derivatives along all directions vanish [39],

$$D_{\delta\psi}\mathcal{L}(\psi, u, \chi) = 0 \quad \text{for all } \delta\psi, \quad (3.13a)$$

$$D_{\delta u}\mathcal{L}(\psi, u, \chi) = 0 \quad \text{for all } \delta u, \quad (3.13b)$$

$$D_{\delta\chi}\mathcal{L}(\psi, u, \chi) = 0 \quad \text{for all } \delta\chi, \quad (3.13c)$$

under the constraint that $\delta\psi(0) = 0$ and $\delta u(0) = \delta u(T) = 0$. This set of equations are known as the optimality system [39]. Any point satisfying the optimality system is a stationary point. In the remainder of this section each of these derivatives is evaluated one-by-one.

Variation of χ : Setting the definition of the Gâteaux variation to zero gives

$$D_{\delta\chi}\mathcal{L}(\psi, u, \chi) = \Re \int_0^T \frac{d}{d\varepsilon} \langle \chi + \varepsilon\delta\chi | i\dot{\psi} - \hat{H}\psi - \beta|\psi|^2\psi \rangle dt \Big|_{\varepsilon=0} \quad (3.14)$$

$$= \Re \int_0^T \langle \delta\chi | i\dot{\psi} - \hat{H}\psi - \beta|\psi|^2\psi \rangle dt. \quad (3.15)$$

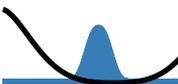
Invoking that the variation $\delta\chi$ is arbitrary leads to the conclusion $i\dot{\psi} = \hat{H}\psi + \beta|\psi|^2\psi$, which is just the Gross-Pitaevskii equation. This calculation shows that the Lagrange multiplier ensures that optimal points always satisfy the Gross-Pitaevskii equation.

Variation of ψ : We start with

$$0 = D_{\delta\psi}\mathcal{L}(\psi, u, \chi) = \frac{d}{d\varepsilon} \left(J(\psi + \varepsilon\delta\psi, u) + J_Z(\psi + \varepsilon\delta\psi, u, \chi) \right) \Big|_{\varepsilon=0} \quad (3.16)$$

The derivative of the first term is,

$$D_{\delta\psi}J = -\Re \langle \psi(T) | \psi_t \rangle \langle \psi_t | \delta\psi(T) \rangle \quad (3.17)$$



The next calculation is the derivative of the second term with the constraint Z ,

$$\begin{aligned} D_{\delta\psi} J_Z(\psi, u, \chi) &= \frac{d}{d\varepsilon} \int_0^T \Re \langle \chi | (i\partial_t - \hat{H} - \beta|\psi + \varepsilon\delta\psi|^2)(\psi + \varepsilon\delta\psi) \rangle dt \Big|_{\varepsilon=0} \\ &= \Re \int_0^T \left(\langle \chi | (i\partial_t - \hat{H} - 2\beta|\psi|^2) - \langle \chi^* | \beta\psi^{*2} \right) | \delta\psi \rangle dt \end{aligned} \quad (3.18)$$

The last step is to move the time derivative from $|\delta\psi\rangle$ to $\langle\chi|$ using integration by parts,

$$\int_0^T i \langle \chi | \delta\dot{\psi} \rangle dt = i \langle \chi(T) | \delta\psi(T) \rangle - \int_0^T i \langle \dot{\chi} | \delta\psi \rangle dt, \quad (3.19)$$

where it was used that $\delta\psi(0) = 0$. Combining the two Gâteaux derivatives gives

$$\begin{aligned} D_{\delta\psi} \mathcal{L}(\psi, u, \chi) &= \Re \left(- \langle \psi(T) | \psi_t \rangle \langle \psi_t | \delta\psi(T) \rangle + (i \langle \chi(T) | \delta\psi(T) \rangle) \right) \\ &\quad + \Re \int_0^T \langle i\dot{\chi} - \hat{H}\chi - 2\beta|\psi|^2\chi - \beta\psi^2\chi^* | \delta\psi \rangle dt \end{aligned} \quad (3.20)$$

Now we are ready to invoke that $\delta\psi$ is an arbitrary variation. For $t = T$ the first line in Eq. (3.20) gives

$$- \langle \psi(T) | \psi_t \rangle \langle \psi_t | + i \langle \chi(T) | = 0 \quad (3.21)$$

This can be rearranged into,

$$i\chi(T) = - \langle \psi_t | \psi(T) \rangle \psi_t. \quad (3.22)$$

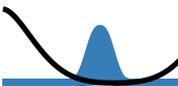
This initial condition for the Lagrange multiplier is a complex number times the target state. For $0 \leq t < T$ we obtain from the second line in Eq. (3.20)

$$i\dot{\chi} = \left(\hat{H} + 2\beta|\psi|^2 \right) \chi + \beta\psi^2\chi^*. \quad (3.23)$$

This equation of motion for the Lagrange multiplier mixes the real and imaginary parts of χ for non-zero β and it reduces to the Schrödinger equation for $\beta = 0$. In the common case $\beta = 0$, the Lagrange multiplier can be interpreted as the target state propagated backwards in time.

Variation of u : This calculation starts with

$$D_{\delta u} J = \frac{d}{d\varepsilon} \left(J(\psi, u) + J_Z(\psi, u + \delta u, \chi) \right) \Big|_{\varepsilon=0} \quad (3.24)$$



Again calculating the derivative of the two terms individually one finds,

$$D_{\delta u} J_Z = \frac{d}{d\varepsilon} \Re \int_0^T \langle \chi | (i\partial_t - \hat{H}(u + \varepsilon\delta u) - \beta|\psi|^2) |\psi\rangle dt \Big|_{\varepsilon=0} \quad (3.25)$$

$$= -\Re \int_0^T \left\langle \chi \left| \frac{\partial \hat{H}(u)}{\partial u} \right| \psi \right\rangle \delta u dt \quad (3.26)$$

Moving on to the J derivative gives

$$D_{\delta u} J = \gamma \int_0^T \dot{u} \delta \dot{u} dt = -\gamma \int_0^T \ddot{u} \delta u dt \quad (3.27)$$

where we used integration by parts and the constraint $\delta u(0) = u\delta(T) = 0$. Putting the the two terms together yields

$$D_{\delta u} J = \int_0^T \left(-\Re \left\langle \chi \left| \frac{\partial \hat{H}(u)}{\partial u} \right| \psi \right\rangle - \gamma \ddot{u} \right) \delta u dt \quad (3.28)$$

Invoking the fact that δu is arbitrary gives the final optimality equation,

$$\gamma \ddot{u} = -\Re \left\langle \chi \left| \frac{\partial \hat{H}(u)}{\partial u} \right| \psi \right\rangle. \quad (3.29)$$

After all these calculations we have found the set of equations equivalent to the optimality system in Eqs. (3.13).

$$i\dot{\psi} = \left(\hat{H} + \beta|\psi|^2 \right) \psi, \quad \psi(0) = \psi_0 \quad (3.30a)$$

$$i\dot{\chi} = \left(\hat{H} + 2\beta|\psi|^2 \right) \chi + \beta\psi^2 \chi^*, \quad i\chi(T) = -\langle \psi_T | \psi(T) \rangle \psi_T \quad (3.30b)$$

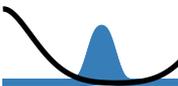
$$\gamma \ddot{u} = -\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial u} \right| \psi \right\rangle, \quad u(0) = u_0, \quad u(T) = u_T \quad (3.30c)$$

On the left are the optimality equations and on the right the boundary conditions. Eq. (3.30b) is known as the adjoint equation [39]. For this reason the calculations performed in this chapter are sometimes referred to as the adjoint method [107].

3.2.1 Optimality System for the Landau-Zener Model

As an instructive example I will solve the optimality system for the Landau-Zener model Eq. (3.5) with initial and target states $|\psi_0\rangle = |0\rangle$ and $|\psi_T\rangle = |1\rangle$. For simplicity, consider a constant control $u(t) = \Delta$. The Schrödinger equation may be solved by direct exponentiation,

$$\begin{aligned} |\psi(t)\rangle &= \exp(-i\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}\Omega t) |0\rangle \\ &= \left(\cos \Omega t - i \frac{\Delta}{\Omega} \sin \Omega t \right) |0\rangle - i \sin \Omega t \frac{\Omega_R}{\Omega} |1\rangle. \end{aligned} \quad (3.31)$$



From this equation, one may directly identify the solution as a $\pi/2$ -pulse, i.e. $\Delta(t) = 0$ for $0 \leq t \leq \pi/2\Omega = T$. In this example, this result is verified using the optimality system Eq. (3.30). The solution to Eq. (3.30b) is

$$|\chi(t)\rangle = \exp(-i\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}\Omega t)|\chi(0)\rangle, \quad (3.32)$$

However, from Eq. (3.30b) we have the boundary condition $\chi(T)$ at the final time $t = T$ not at the beginning $t = 0$. Instead we solve Eq. (3.30b) *backwards* in time. A calculation gives,

$$\begin{aligned} |\chi(t)\rangle &= \exp(i\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}\Omega(T-t))|\chi(T)\rangle = \frac{\Omega_R \sin \Omega T}{\Omega} \exp(i\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}\Omega(T-t))|1\rangle \\ &= \frac{\Omega_R \sin \Omega T \sin \Omega(T-t)}{\Omega^2} \left(i\Omega_R |0\rangle + (\Omega \cot \Omega(T-t) - i\Delta) |1\rangle \right). \end{aligned} \quad (3.33)$$

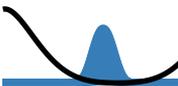
From the second identity, it is explicitly seen how the Lagrange multiplier is the backwards propagated target state. Note that this state is not normalized to unity due to the complex factor from the boundary condition. This factor is sometimes called the overlap factor. Calculating the last optimality equation (3.30c) gives

$$0 = -\Re \left\langle \chi \left| \frac{\partial \hat{H}(u)}{\partial u} \right| \psi \right\rangle = \frac{2\Delta\Omega_R^2 \sin \Omega T \sin \Omega t \sin \Omega(T-t)}{\Omega^3} \quad (3.34)$$

This equation has the simple solution $\Delta = 0$ as expected. However, we do not learn anything from the optimality equations about the length of the duration. Indeed $\Delta = 0$ is only a solution if $T = (2n+1)\pi/2\Omega_R$ with $n = 1, 2, 3, \dots$. The reason is that the optimality system only provides *necessary* and not *sufficient* conditions for optimality. Sufficient conditions also take the second order derivatives into account [39].

This example highlights the two main drawbacks with the optimality system (3.30) i) the coupled differential equations are almost always too complex to be solved analytically and ii) the solutions are not guaranteed to be optimal. Instead the common approach within quantum optimal control theory is to use numerical algorithms that minimize the state-to-state control problem. These algorithms can be formulated based on the optimality system.

The steps taken for solving the optimality system were, i) first solve Eq. (3.30a) forward in time, ii) then solve Eq. (3.30b) backwards in time and finally iii) solve Eq. (3.30c). These are precisely the same steps as used in the numerical algorithms introduced later in Sec. 3.3.



3.2.2 Reduced Cost Functional

In the definition of the state-to-state control problem Eq. (3.3) the cost functional is given as

$$J(\psi, u) = \frac{1}{2} \left(1 - |\langle \psi_t | \psi(T) \rangle|^2 \right) + \frac{\gamma}{2} \int_0^T \dot{u}^2 dt \quad (3.35)$$

This depends on both the states ψ and the control $u(t)$. For a given control $u(t)$ there exist a unique set of states $\psi_u(t)$ given by the solution to the (nonlinear) Schrödinger equation [39]. This fact allows the introduction of the so-called reduced cost functional

$$\hat{J}(u) = J(\psi_u, u). \quad (3.36)$$

In numerical simulations, we solve the Gross-Pitaevskii equation for a given $u(t)$ making it is more natural to consider the reduced cost functional. Note that, $\hat{J}(u) = \mathcal{L}(\psi_u, u, \chi)$ since the constraint term in Lagrangian (3.11) is $Z(u, \psi_u) = 0$. This is correct for any value of the Lagrange multiplier χ , hence χ is a free variable that we can choose conveniently. The derivative of the reduced cost functional is given by the chain rule

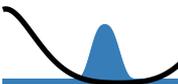
$$D_{\delta u} \hat{J}(u) = D_{\delta u} \mathcal{L}(\psi(u), u, \chi) + D_{\delta \psi_u} \mathcal{L}(\psi(u), u, \chi), \quad (3.37)$$

The derivative $D_{\delta \psi_u} \mathcal{L}$ is the induced variation in ψ_u from the variation of u . This extra term appears since ψ_u depends implicitly on u through the Gross-Pitaevskii equation. This was not the case when discussing the Lagrangian Eq. (3.11), since ψ was taken to be a free variable and Gross-Pitaevskii equation is a constraint. Notice that performing this variation is formally the same as the $D_{\delta \psi}$ just with the induced variation instead. Hence, we would get the same equations as above for $D_{\delta \psi} \mathcal{L}$ just with $\delta \psi$ replaced with $\delta \psi_u$. Specifically, we find Eq. (3.20) again for the induced variation. Recall that χ is a free variable we can select. If we pick χ to satisfy Eq. (3.30b) then the induced variation $D_{\delta \psi_u}$ will vanish, leading to the conclusion

$$D_{\delta u} \hat{J}(u) = D_{\delta u} \mathcal{L}(\psi_u, u, \chi(u, \psi_u)) \quad (3.38)$$

$$= \int_0^T \left(-\Re \left\langle \chi(u, \psi_u) \left| \frac{\partial \hat{H}}{\partial u} \right| \psi_u \right\rangle - \gamma \ddot{u} \right) \delta u dt, \quad (3.39)$$

where we have used the result from Eq. (3.28). This is an important result that is repeatedly used in the next sections. For notational convenience we will omit the implicit dependencies of ψ and χ in the rest of the thesis. Eq. (3.38) is a well-known result from mathematical optimal control theory [39].



3.3 GRAPE Gradient Calculation

Instead of solving the optimality system directly we numerically iterate some initial control u_1 such that

$$\hat{J}(u_1) \geq \hat{J}(u_2) \geq \dots \geq \hat{J}(u_{k-1}) \geq \hat{J}(u_k), \quad (3.40)$$

since the cost functional \hat{J} is bounded from below such a sequence must converge. The simplest such method is steepest descent that updates the control along the negative gradient

$$u_k = u_{k-1} - \alpha_k \nabla J_X(u_k). \quad (3.41)$$

where k is the iteration index and α_k is some suitable step size. This algorithm can be considerably improved and more advanced methods are discussed in Sec. 3.4. In this section we discuss how to compute the gradient $\nabla J_X(u_k)$. The gradient under the norm X is defined as the unique element satisfying

$$D_{\delta u} \hat{J} = \langle \nabla_X \hat{J}, \delta u \rangle_X, \quad (3.42)$$

where δu is an arbitrary variation. The most common choices of norm are $X = L^2$ and $X = H^1$ defined as

$$\langle f, g \rangle_{L^2} = \int_0^T f(t)g(t)dt, \quad \langle f, g \rangle_{H^1} = \int_0^T \dot{f}(t)\dot{g}(t)dt. \quad (3.43)$$

$D_{\delta u} \hat{J}$ was calculated in Eq. (3.39). For the L^2 inner product this gives

$$\langle \nabla_{L^2} \hat{J}, \delta u \rangle_{L^2} = \int_0^T \nabla_{L^2} \hat{J} \delta u dt = \int_0^T \left(-\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial u} \right| \psi \right\rangle - \gamma \ddot{u} \right) \delta u dt. \quad (3.44)$$

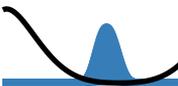
From this equation we can immediately recognize that

$$\nabla_{L^2} \hat{J} = -\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial u} \right| \psi \right\rangle - \gamma \ddot{u}. \quad (3.45)$$

If one uses the H^1 inner product the gradient becomes

$$\langle \nabla_{H^1} \hat{J}, \delta u \rangle_{H^1} = \delta u \frac{d}{dt} \nabla_{H^1} \hat{J} \Big|_0^T - \int_0^T \delta u \frac{d^2}{dt^2} \nabla_{H^1} \hat{J} dt \quad (3.46a)$$

$$= \int_0^T \left(-\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial u} \right| \psi \right\rangle - \gamma \ddot{u} \right) \delta u dt \quad (3.46b)$$



Algorithm 1 Calculation of the GRAPE-gradient.

- 1: Solve the Gross-Pitaevskii equation for a given u *forward* in time.
- 2: Solve the adjoint equation (3.30a) with u and ψ *backwards* in time.
- 3: Compute the L^2 - or H^1 -gradient Eqs. (3.45)-Eq. (3.47).

The first term in Eq. (3.46a) vanishes since $\delta u(0) = \delta u(T) = 0$. From this expression we can directly read off the H^1 -gradient

$$\frac{d^2}{dt^2} [\nabla_{H^1} \hat{J}(u)] = \gamma \ddot{u} + \Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial u} \right| \psi \right\rangle = -\nabla_{L^2} \hat{J} \quad (3.47)$$

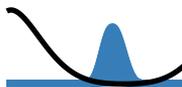
The L^2 -gradient may not vanish at the boundaries of the time interval $t = 0$ and $t = T$ in which case updating the control according to Eq. (3.41) causes the control to change for $t = 0$ and $t = T$. This is not desired as the initial and final values of the control are typically fixed whereby one must artificially enforce $\nabla_{L^2} J(0) = \nabla_{L^2} J(T) = 0$. This can lead to serious decreases in the stability of the method [108]. In solving the Poisson Eq. (3.47) for the H^1 -gradient we may conveniently choose Dirichlet boundary conditions $\nabla_{H^1} J(0) = \nabla_{H^1} J(T) = 0$ such that the boundary condition in Eq. (3.30c) is always fulfilled. This gives a theoretical motivation for using the H^1 -gradient over the L^2 -gradient.

Calculating the gradient using Eq. 3.45 or 3.47 is known as Gradient-Ascent Pulse Engineering (GRAPE) within the optimal control literature [39, 109, 110].

In Algorithm 1 there is a summary of how to compute the gradient of $\hat{J}(u)$ for a given u , which is the central algorithm in many of the later applications. In fact, we have already applied this algorithm analytically in Sec. 3.2.1 where we computed the gradient of a constant control pulse in the Landau-Zener model. Indeed in light of this section, Eq. (3.34) can be understood as requiring that $\nabla \hat{J}(u(t)) = 0$ for all t , which is a necessary condition optimal controls.

3.4 Advanced Methods for Gradient Based Quantum Control

This section introduces established gradient based optimization methods that may be used to minimize $\hat{J}(u)$. As discussed in the previous section the choice of inner product changes the formulation and numerical implementation of the optimization problem. Seminal textbooks on optimization such as Ref. [111] are directly performed in Euclidean spaces and care must be taken when generalizing the methods to function spaces



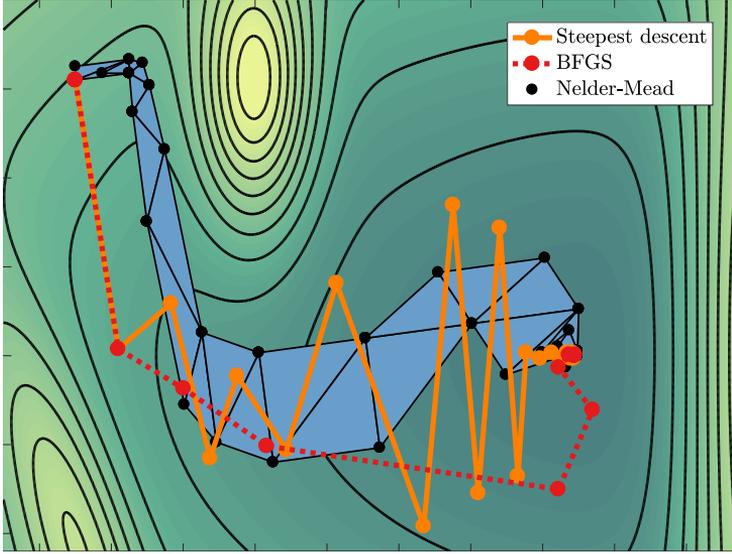
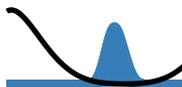


Figure 3.2: Comparison of gradient-based and gradient-free optimization methods in an artificial landscape. The shaded blue triangles show the gradient-free method, Nelder-Mead. The solid orange line and the dotted red line are gradient-based algorithms steepest descent and BFGS respectively. Steepest descent exhibits the characteristic zigzag type behavior, which BFGS avoids due to the inverse Hessian approximation. Nelder-Mead, steepest descent, and BFGS respectively use 45, 34, and 15 iterations for convergence.

with other inner products. The goal of this section is to review these methods with arbitrary inner products.

The goal is to solve the optimization problem $\min_u \hat{J}(u)$ and we can calculate the cost $\hat{J}(u)$ and gradient with $\nabla \hat{J}(u)$ from e.g. Eqs. (3.45) or (3.47) without worrying about especially the temporal discretization of the Schrödinger (Gross-Pitaevskii) equation. In some applications this assumption cannot be justified and the discretization method must be directly induced in the gradient expressions. These complications are not discussed in this thesis.

In the last section the steepest descent method was introduced. As detailed in Ref. [112] steepest descent always take steps orthogonal to the current level set. This leads to a characteristic inefficient zigzag motion when the optimization moves along a narrow valley - see Fig. 3.2. In this section we discuss two possible improvements known as nonlinear conjugate gradient and quasi-Newton methods.



Algorithm 2 General Line search algorithm.

- 1: Initialize p_0 ;
 - 2: Set $k = 0$;
 - 3: **while** $k < k_{max}$ and $\|p_k\|_X > \tau$ **do**
 - 4: Calculate the descent direction p_k ;
 - 5: Compute the step size α_k along p_k using a step size algorithm;
 - 6: Set $u_{k+1} = u_k + \alpha_k p_k$;
 - 7: Set $k \leftarrow k + 1$;
 - 8: **Return** u_k .
-

3.4.1 Nonlinear Conjugate Gradient Methods

The class of algorithms discussed here are so-called line search methods that iteratively minimize the cost functional along a descent direction p_k [111]. This class of algorithms are outlined in Algorithm 2 and they update the control along some descent direction p_k

$$u_{k+1} = u_k + \alpha_k p_k. \quad (3.48)$$

For steepest descent the direction is the negative gradient $p_k = -\nabla \hat{J}(u_k)$. These methods consist of two components an algorithm to choose the descent direction and to find the step size.

First we discuss the step size. Convergence can be guaranteed if the step size satisfies the Armijo condition

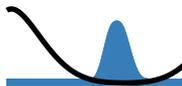
$$\hat{J}(u_k + \alpha_k p_k) \leq \hat{J}(u_k) + c_1 \alpha_k \langle \nabla \hat{J}(u_k), p_k \rangle_X, \quad (3.49)$$

together with the Wolfe condition

$$\langle \nabla \hat{J}(u_k), p_k \rangle_X < c_2 \langle \nabla \hat{J}(u_k + \alpha_k p_k), p_k \rangle_X, \quad (3.50)$$

where $0 < c_1 < c_2 < 1/2$. The Armijo condition expresses that the decrease should be better than a linearization at $\alpha_k = 0$, but this condition is always fulfilled for $\alpha_k \rightarrow 0$. Too small step sizes are prevented by the Wolfe condition that prevents the graph of \hat{J} from changing too fast beyond the minimum. These algorithms are very important as optimization algorithms typically spend more than 60% execution time on finding a proper step size. The reason is that when algorithm tests some step size then it must propagate along the entire control to get the fidelity and thereby the cost. It is the repeated solutions of the Schrödinger equation that consumes the most runtime.

Nonlinear conjugate rate gradient improves upon the convergence rate of steepest descent by choosing a better descent direction. Nonlinear



conjugate gradient methods are an extension of the linear methods [111]. In the standard variant the descent direction is

$$p_{k+1} = -g_{k+1} + \beta_{k+1}p_{k+1}, \quad (3.51)$$

where $g_k = \nabla \hat{J}(u_k)$ and $p_{(0)} = -g_{(0)}$. The β_k parameter is chosen such that Eq. (3.51) reduces to the linear conjugate gradient if i) \hat{J} is a strictly convex quadratic function and ii) α_k is the exact one-dimensional minimizer of \hat{J} along p_k . If these conditions are fulfilled then the algorithm quickly converges to the optimum [39, 111]. There is a number of different proposals for β_k where two prominent options are due to Dai and Yuan [113]

$$\beta_k^{\text{DY}} = \frac{\langle g_{k+1}, g_{k+1} \rangle_X}{\langle p_k, y_k \rangle_X}. \quad (3.52)$$

where $y_k = g_{k+1} - g_k$ and the other is due to Hager and Zhang [114]

$$\beta_k^{\text{HZ}} = \frac{\langle \sigma_k, g_{k+1} \rangle_X}{\langle p_k, y_k \rangle_X}, \quad \sigma_k = y_k - 2p_k \frac{\langle y_k, y_k \rangle_X}{\langle y_k, p_k \rangle_X}. \quad (3.53)$$

These choices of β_k together with the Armijo and Wolfe conditions make it possible to prove that the linesearch is globally convergent [39, 111]. An advantage of the nonlinear conjugate gradient method is that it only uses the last gradient and not all previous gradients resulting in a low requirement for memory.

3.4.2 Quasi-Newton Methods

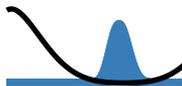
Quasi-Newton methods are an approximation of Newton methods. Assuming for the moment that we are minimizing a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$. A second order approximation of f gives

$$f(u_k + p_k) \simeq f(u_k) + p_k^\top \nabla f(u_k) + \frac{1}{2} p_k^\top \nabla^2 f(u_k) p_k \equiv m_k. \quad (3.54)$$

A minimizer u_k^* of this function satisfies $\nabla m_k(u_k^*) = 0$ and solving this equation gives the Newton direction

$$p_k = -[\nabla^2 f(u_k)]^{-1} \nabla f(u_k), \quad (3.55)$$

where $[\nabla^2 f(u_k)]^{-1}$ is the inverse Hessian of $f(u_k)$. This direction is a descent direction whenever the inverse Hessian is positive definite. Unfortunately a direct calculation of the Hessian is very expensive. An alternative is quasi-Newton methods where the descent direction is $p_k = -B_k^{-1} g_k$



with B_k being some positive definite approximation of the Hessian. The matrix B_k is taken to satisfy the Taylor expansion

$$\nabla f(u_k + s_k) \simeq \nabla f(u_k) + B s_k, \quad (3.56)$$

where $s_k = \alpha_k p_k$ is the difference between successive steps. This is the so-called secant equation, which is underdetermined requiring additional constraint on $H_k = B_k^{-1}$ for a unique solution [111]. In Broyden-Fletcher-Goldfarb-Shanno (BFGS) the constraint is $H_{k+1} = \operatorname{argmin} \|H - H_k\|$ subject to $H = H^T$ using the Frobenius norm. This criteria leads to the celebrated BFGS update rule

$$\begin{aligned} H_{k+1} &= H_k + \frac{s_k^\top y_k + y_k^\top H_k y_k}{(s_k^\top y_k)^2} (s_k s_k^\top) - \frac{H_k y_k s_k^\top + s_k y_k^\top H_k}{s_k^\top y_k} \\ &= H_k + d_k c_k s_k s_k^\top - c_k z_k s_k^\top - c_k s_k z_k^\top, \end{aligned} \quad (3.57)$$

where we have introduced the abbreviations for the scalars $c_j = (s_j^\top y_j)^{-1}$ and $d_j = 1 + c_j y_j^\top z_j$.

This completes our brief introduction of the standard BFGS formalism. It is not straightforward to apply this method when the domain of f is not \mathbb{R}^n . A solution to this was presented in the seminal reference [108]. Assume that $X = L^2$ or $X = H^1$. In this case we can define the action of outer products $(x \otimes y)z = \langle y, z \rangle_X x$ [39, 108]. In general function spaces Eq. (3.57) becomes

$$H_{k+1} = H_k + d_k c_k s_k \otimes s_k - c_k z_k \otimes s_k - c_k s_k \otimes z_k. \quad (3.58)$$

Equipped with this formulation and the definition $(x \otimes y)z = \langle y, z \rangle_X x$ we find,

$$\begin{aligned} H_k x &= H_{k-1} x + d_{k-1} c_{k-1} \langle s_{k-1}, x \rangle_X s_{k-1} - c_{k-1} \langle s_{k-1}, x \rangle_X z_{k-1} \\ &\quad - c_{k-1} \langle z_{k-1}, x \rangle_X s_{k-1} \\ &= H_{k-1} x + c_{k-1} \langle s_{k-1}, x \rangle_X r_{k-1} - c_{k-1} \langle z_{k-1}, x \rangle_X s_{k-1}, \end{aligned} \quad (3.59)$$

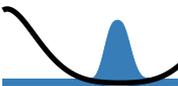
where we have introduced $r_j = d_j s_j - z_j$. This equation can be used to recursively eliminate all H_k giving

$$H_k x = H_0 x + \sum_{j=0}^{k-1} c_j \left(\langle s_j, x \rangle_X r_j - \langle z_j, x \rangle_X s_j \right). \quad (3.60)$$

Using this equation for $x = y_k$ and $x = -g_{k+1}$ gives

$$z_k = H_0 y_k + \sum_{j=0}^{k-1} c_j \left(\langle s_j, y_k \rangle_X r_j - \langle z_j, y_k \rangle_X s_j \right), \quad k \geq 1, \quad (3.61)$$

$$p_{k+1} = -H_0 g_{k+1} - \sum_{j=0}^k c_j \left(\langle s_j, g_{k+1} \rangle_X r_j - \langle z_j, g_{k+1} \rangle_X s_j \right). \quad (3.62)$$

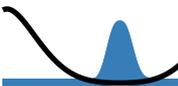


Algorithm 3 BFGS Linesearch

-
- 1: Initialize $g_0 = \nabla \hat{J}(u_0)$ and $p_0 = -g_0$.
 - 2: Set $k = 0$;
 - 3: **while** $k < k_{max}$ and $\|g_{k-1}\|_X > \tau$ **do**
 - 4: Calculate $u_{k+1} = u_k + \alpha_k p_k$ using a step size algorithm.
 - 5: Calculate $g_{k+1} = \nabla \hat{J}(u_{k+1})$, $y_k = g_{k+1} - g_k$, and $s_k = \alpha_k p_k$.
 - 6: Calculate z_k using Eq. (3.61).
 - 7: Calculate $c_k = \langle s_k, y_k \rangle_X^{-1}$, $d_k = 1 + c_k \langle y_k, z_k \rangle_X$, and $r_k = d_k s_k - z_k$.
 - 8: Calculate p_{k+1} using Eq. (3.62).
 - 9: Set $k \leftarrow k + 1$;
 - 10: **Return** u_k .
-

With these equations we can compute a BFGS direction in some function space X . The algorithm starts with some initial guess u_0 and corresponding gradient $g_0 = \nabla \hat{J}(u_0)$. In the first iteration the steepest descent direction is used $p_0 = -g_0$ implying $H_0 = \mathbb{1}$. In every iteration a suitable α_k is found using a step size algorithm. In the subsequent iterations we can find $y_k = g_{k+1} - g_k$ and $s_k = \alpha_k p_k$. The algorithm is outlined in Algorithm 3. In Fig. 3.2 BFGS is compared with steepest descent and BFGS clearly shows superior convergence rate. Similarly to the nonlinear conjugate gradient it is possible to prove this method is globally convergent provided the Wolfe conditions are always satisfied [39].

Note, that in both nonlinear conjugate gradient and BFGS the resulting control u_k is a linear combination of the current and all previous gradients. However the two methods uses the last gradient information in very different ways, the nonlinear conjugate gradient update in Eq. (3.51) only uses the *last* gradient to find the next search direction. In comparison the quasi-Newton methods uses *all* past gradients to build an approximation of p_k since it uses all previous values of $\{s_j, y_j, z_j\}$ in every iteration. As the number of iterations increase it takes progressively longer and requires more memory to calculate the BFGS direction. In conclusion calculating the nonlinear conjugate gradient direction is both faster and more memory efficient than BFGS. However, BFGS typically offer better directions and still converge faster. Due to the high cost of calculating the step size it is almost always beneficial to use BFGS over nonlinear conjugate gradient in quantum control problems. In this thesis I only apply BFGS type descent.



3.4.3 Control Constraints

In a number of experimentally relevant cases there is also a bound on the maximal and minimal values of the control $u_{\min} \leq u(t) \leq u_{\max}$. These bounds can be accommodated by using a non-linear transformation that makes the control unconstrained [8]. Another option is also to project the control back inside the constraints in which case the control update Eq. (3.48) becomes

$$u_{k+1} = \mathcal{P}(u_k + \alpha_k p_k), \quad (3.63)$$

where \mathcal{P} is a projection operator on the interval $u_{\min} \leq u(t) \leq u_{\max}$ [39].

An alternative is to add a term in the cost function Eq. (3.3a) that penalizes controls outside the bounds. The latter approach, is known as soft bounds and a possible choice is a parabolic cost penalty

$$J_b = \frac{\sigma}{2} \int_0^T \Theta(u_{\min} - u)(u - u_{\min})^2 + \Theta(u - u_{\max})(u - u_{\max})^2 dt, \quad (3.64)$$

where Θ is the Heaviside step function. The weight factor σ is typically of the order $\sigma \sim 10^3 - 10^4$ to heavily penalize controls outside the bounds. It is straightforward to calculate the L^2 -gradient of this term, which is

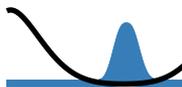
$$\nabla_{L^2} J_b = \sigma \left(\Theta(u_{\min} - u)(u - u_{\min}) + \Theta(u - u_{\max})(u - u_{\max}) \right). \quad (3.65)$$

This gradient is added to Eq. (3.45) steering the optimization towards a region inside the bounds. In most of our applications, we choose this option.

3.5 Krotov's Method

An alternative to the Lagrange multiplier method for deriving optimal control algorithms is Krotov's method [54, 55]. The presentation of Krotov's method in this chapter is based on Refs. [55, 115, 116] but with many modifications in order to have a cohesive notation with the previous chapters. For simplicity, we only discuss the linear Schrödinger equation ($\beta = 0$). In Krotov's method, the cost functional is rewritten such that the Schrödinger equation appears explicitly. The regularization term is not included in this method so the cost functional is simply

$$\begin{aligned} J(\psi, u) &= J(\psi, u) - \int_0^T \frac{d\varphi}{dt} dt + \varphi(T, \psi(T)) - \varphi(0, \psi_0) \\ &= J(\psi, u) - \int_0^T \frac{\partial \varphi}{\partial t} + D_\psi \varphi dt + \varphi(T, \psi(T)) - \varphi(0, \psi_0), \end{aligned} \quad (3.66)$$



where $\varphi(t, \psi)$ is some scalar field we may select freely. In Krotov's method this field is chosen such that the Schrödinger equation appears explicitly, which is different from the Lagrange multiplier method outlined earlier where the Schrödinger equation enters as a constraint. A suitable choice is

$$\varphi(t, \psi) = \frac{1}{2} \left(\langle \xi | \psi \rangle + \langle \psi | \xi \rangle \right), \quad (3.67)$$

where the expansion state $|\xi\rangle$ is a free parameter. Inserting this into Eq. (3.66) and calculating the Gâteaux derivative gives

$$J(\psi, u) = L(\psi, u) = G(\psi(T)) - \int_0^T R(t, \psi, u) dt + \varphi(0, \psi(0)). \quad (3.68)$$

where

$$G(\psi(T)) = J(\psi, u) + \varphi(T, \psi(T)), \quad (3.69)$$

$$R(t, \psi, u) = \frac{\partial \varphi}{\partial t} + D_{\dot{\varphi}} \varphi = \Re \langle \dot{\xi} + i \hat{H} \xi | \psi \rangle. \quad (3.70)$$

This shows that minimizing J is equivalent to minimizing L . A reader familiar with optimal control theory may recognize these steps as the standard way of transforming a Mayer type optimization problem into a Bolza type [32]. In Krotov's method, it is directly required that the cost decreases at each iteration $J_{k+1} \leq J_k$, which is equivalent to $0 \leq L_k - L_{k+1} = \Delta_1 + \Delta_2 + \Delta_3$ where,

$$\Delta_1 = G(\psi_k(T)) - G(\psi_{k+1}(T)), \quad (3.71a)$$

$$\Delta_2 = \int_0^T R(t, \psi_{k+1}, u_{k+1}) - R(t, \psi_{k+1}, u_k) dt, \quad (3.71b)$$

$$\Delta_3 = \int_0^T R(t, \psi_{k+1}, u_k) - R(t, \psi_k, u_k) dt. \quad (3.71c)$$

A sufficient condition for a decrease in the cost is that the Δ 's are positive. Seeking fields and controls that make these three terms positive we require that the first order derivatives of R and G vanish with respect all variations in $\delta\psi$

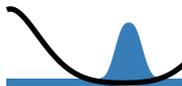
$$D_{\delta\psi} R = \Re \langle \dot{\xi} + i \hat{H} \xi | \delta\psi \rangle = 0. \quad (3.72)$$

We require that this derivative vanish along the current iteration that is

$$|\dot{\xi}_k\rangle = -i \hat{H}(u_k) |\xi_k\rangle, \quad (3.73)$$

which is just the usual Schrödinger equation. The associate boundary condition for ξ can be found from the variation of G ,

$$\begin{aligned} D_{\delta\psi} G(\psi(T)) &= D_{\delta\psi} J + D_{\delta\psi} \varphi(T, \psi(T)) \\ &= \Re \langle \xi(T) - \langle \psi(T) | \psi_t \rangle \psi_t | \delta\psi(T) \rangle = 0, \end{aligned} \quad (3.74)$$



where we reused the result from Eq. (3.17). From this equation we can identify

$$|\xi_k(T)\rangle = |\psi_t\rangle\langle\psi_t|\psi_k(T)\rangle. \quad (3.75)$$

If we compare Eqs. (3.73)-(3.75) with the adjoint equation (3.30b), then it incidentally occurs that $|\xi(t)\rangle$ is the Lagrange multiplier simply with another phase $|\chi(t)\rangle = i|\xi(t)\rangle$.

We now discuss the additional conditions that ensure that each Δ is positive one-by-one. Starting with Δ_1 . Plugging the boundary condition for ξ Eq. (3.75) into the definition of G Eq. (3.69) gives

$$G(\psi(T)) = \frac{1}{2} \left(1 - \langle\psi(T)|\hat{P}_t|\psi(T)\rangle + \langle\psi_k(T)|\hat{P}_t|\psi(T)\rangle + \langle\psi(T)|\hat{P}_t|\psi_k(T)\rangle \right) \quad (3.76)$$

where $\hat{P}_t = |\psi_t\rangle\langle\psi_t|$ is a projector on the target state. Writing out the expression for Δ_1 reveals

$$\Delta_1 = G(\psi_k(T)) - G(\psi_{k+1}(T)) = \frac{1}{2} \langle\Delta\psi|\hat{P}_t|\Delta\psi\rangle \geq 0, \quad (3.77)$$

where $|\Delta\psi\rangle = |\psi_{k+1}\rangle - |\psi_k\rangle$. The expression in Eq. (3.77) is non-negative due to the positivity of the \hat{P}_t operator.

Let us move on to Δ_2 . It is possible to ensure that Δ_2 is positive by taking a step along the gradient of R as

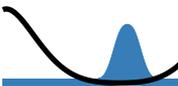
$$u_{k+1} = u_k + \alpha S(t) \left. \frac{\partial R}{\partial u} \right|_{\psi_{k+1}, u_k}, \quad (3.78)$$

where α is some suitable step-size and $S(t)$ is a shape function that ensures the boundary condition $u_{k+1}(0) = u_i$ and $u_{k+1}(T) = u_T$. Consider Δ_2 and Taylor expand the first term around u_k

$$\begin{aligned} \Delta_2 &= \int_0^T R(t, \psi_{k+1}, u_{k+1}) - R(t, \psi_{k+1}, u_k) dt \\ &\simeq \int_0^T R(t, \psi_{k+1}, u_k) + \left. \frac{\partial R}{\partial u} \right|_{\psi_{k+1}, u_k} (u_{k+1} - u_k) - R(t, \psi_{k+1}, u_k) dt. \\ &= \alpha \int_0^T S(t) \left(\left. \frac{\partial R}{\partial u} \right|_{\psi_{k+1}, u_k} \right)^2 dt \geq 0. \end{aligned} \quad (3.79)$$

Finally, let us discuss the last Δ_3 . This term is

$$\begin{aligned} \Delta_3 &= \int_0^T R(t, \psi_{k+1}, u_k) - R(t, \psi_k, u_k) dt \\ &= \int_0^T \Re\langle\dot{\xi}_k + i\hat{H}(u_k)\xi_k|\psi_{k+1}\rangle - \Re\langle\dot{\xi}_k + i\hat{H}(u_k)\xi_k|\psi_k\rangle dt = 0, \end{aligned} \quad (3.80)$$



where we have inserted the definition of the R -function Eq. (3.70). Both the terms in this equation are zero due to the choice of $|\xi_k\rangle$ as the solution of Eq. (3.72).

In conclusion we have shown that we get a guaranteed decrease in the cost if we update the control according to

$$u_{k+1} = u_k + \alpha S(t) \Re \left\langle \chi_k \left| \frac{\partial \hat{H}}{\partial u} \right|_{u_k} \right| \psi_{k+1} \rangle, \quad (3.81)$$

for some non-zero small value of α . This equation is known as the Krotov update rule. α is the step-size parameter that must be selected for proper convergence. Note that both the current iteration index i and the next index $i + 1$ appear in Eq. (3.81). Specifically, the control at the next iteration $u^{(i+1)}$ depends on the states in the next iteration $\psi^{(i+1)}$. This makes Krotov's method a sequential algorithm where the next control $u^{(i+1)}$ is being calculated while the equations of motion are being solved along that control. This is very different from the other methods presented here, where the control is updated concurrently for all values of $0 < t < T$.

At first sight the promise of a guaranteed decrease in the fidelity is very appealing. However it is important to note that all the gradient based methods presented in this section also have this property. If p_k is a descent direction then there exist a non-zero α_k such that $u_{k+1} = u_k + \alpha_k p_k$ has a lower cost [111]. The only exception is if the control is at the boundary of the domain of admissible controls, but such exceptions are also a problem for Krotov's method. In my opinion, monotonic decrease in the cost should be considered a necessary property of any optimal control algorithm and one should only consider applying a method if it has this property.

Note, that if one considers a nonlinear Schrödinger equation then a monotonic decrease in the cost can only be guaranteed if Eq. (3.81) includes an additional term that is proportional to the difference in the states between iterations [55, 110, 115]. The derivative is typically with respect to the next iteration ($u^{(i+1)}$), but for the small values of α used here it is acceptable to use the current iteration ($u^{(i)}$) [55, 110].

3.6 Chopped Basis and Nelder-Mead

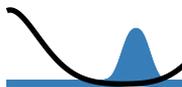
In a chopped basis, the space of admissible controls is parametrized by a set of smooth functions [104]. The high-quality solutions often lie in some lower dimensional subspace inside the space of admissible controls [117]. Hence, a proper low dimensional parametrization of the control space can *a priori* steer the optimization in the correct direction.



Algorithm 4 One Step of the Nelder-Mead Algorithm

-
- 1: Order the $n + 1$ vertices to satisfy $f(\mathbf{x}_1) \leq f(\mathbf{x}_2) \leq \dots \leq f(\mathbf{x}_{n+1})$,
 - 2: Compute reflection point $\mathbf{x}_r = (1 + \rho)\bar{\mathbf{x}} - \rho\mathbf{x}_{n+1}$ and $f_r = f(\mathbf{x}_r)$.
 - 3: **if** $f_r \leq f_1$ **then** (\mathbf{x}_e best point try moving further this way)
 - 4: Compute expansion point $\mathbf{x}_e = (1 + \rho\chi)\bar{\mathbf{x}} - \rho\chi\mathbf{x}_{n+1}$ and $f_e = f(\mathbf{x}_e)$.
 - 5: **if** $f_e < f_r$ **then**
 - 6: Replace \mathbf{x}_{n+1} by \mathbf{x}_e and terminate iteration.
 - 7: **else**
 - 8: Replace \mathbf{x}_{n+1} by \mathbf{x}_r and terminate iteration.
 - 9: **else**
 - 10: **if** $f_r < f_n$ **then** (f_r is neither best or worst)
 - 11: Replace \mathbf{x}_{n+1} by \mathbf{x}_r and terminate iteration.
 - 12: **else**
 - 13: **if** $f_r < f_{n+1}$ **then** (Outside contraction)
 - 14: Compute $\mathbf{x}_c = (1 + \rho\gamma)\bar{\mathbf{x}} - \rho\gamma\mathbf{x}_{n+1}$ and $f_c = f(\mathbf{x}_c)$.
 - 15: **if** $f_c \leq f_r$ **then**
 - 16: Replace \mathbf{x}_{n+1} by \mathbf{x}_c and terminate iteration.
 - 17: **else**
 - 18: Shrink simplex.
 - 19: **else** (Inside contraction)
 - 20: Compute $\mathbf{x}_{cc} = (1 - \gamma)\bar{\mathbf{x}} + \gamma\mathbf{x}_{n+1}$ and $f_{cc} = f(\mathbf{x}_{cc})$.
 - 21: **if** $f_{cc} < f_{n+1}$ **then**
 - 22: Replace \mathbf{x}_{n+1} by \mathbf{x}_{cc} and terminate the iteration.
 - 23: **else**
 - 24: Shrink simplex.
-

When performing the numerical optimization of the cost functional it is necessary to discretize the control in steps of Δt , where Δt is set by the required accuracy when numerically solving the Gross-Pitaevskii equation. This means that the control becomes represented by a vector of length $N = \lfloor T/\Delta t \rfloor + 1$. Typically, in GRAPE and Krotov's method the dimension for the space of admissible controls (M) is the same as for the simulation $N = M$. However, with a proper change of basis the optimal controls could adequately be described by another basis with much smaller dimension [117, 118]. In this sense setting $N = M$ often introduces too many degrees of freedom for the control problem. Often we have $N \simeq 3000$ whereas by a proper choice of basis we have only $M \simeq 50$. These potential large reductions in the dimension of the control problem is the motivation for reduced or chopped basis methods. A chopped basis



(CB) is defined as

$$u(t) = u_0(t) + S(t) \sum_{n=1}^M c_n f_n(t), \quad (3.82)$$

where the f_n 's are the basis functions. $S(t)$ is a shape function that enforces the boundary condition in Eq. (3.30c), i.e. $S(0) = S(T) = 0$. In this chopped basis we optimize the coefficients c_n instead of full control $u(t)$ implying that the cost function is $\hat{J}(\mathbf{c})$ where $\mathbf{c} = (c_1, c_2, \dots, c_M)$. We also refer to a chopped basis as a parameterization. The functions f_n must be chosen sensibly based on physical insight, which would typically be sinusoidal functions around characteristic frequencies. This insight is what can help steer an optimization algorithm *a priori*. A very popular choice in the literature is the so-called Chopped Random Basis (CRAB)

$$u(t) = u_0(t) \left(1 + \frac{1}{\lambda(t)} \sum_{n=1}^M A_n \sin(\omega_n t) + B_n \cos(\omega_n t) \right), \quad (3.83)$$

where $\lambda(t)$ enforces the boundary conditions $\lambda(t) \rightarrow \infty$ for $t \rightarrow 0$ and $t \rightarrow T$ [104, 119]. This choice is a special case of Eq. (3.82). The random part of CRAB is a randomized frequency shift $\omega_k = 2\pi k(1 + r_k)/T$ where $-0.5 \leq r_k \leq 0.5$. The optimization is repeated a number of times with different values of r_n 's. This is a central idea in CRAB, since it allows the algorithm to explore different basis functions with slightly similar frequencies starting from the same u_0 .

The prevalent choice in the literature is to perform optimization in a chopped basis using derivative-free methods such as the Nelder-Mead algorithm that operate by directly evaluating the cost functional in carefully selected points [47, 104, 119, 120]. More recently, it has been proposed to use gradient-free methods in quantum optimal control such as Brent's principal axis [121]. Gradient-free methods have the advantage that there is no need to implement code that calculates the gradient by solving e.g. Eq. (3.47), which also requires solving the adjoint equation for the Lagrange multiplier Eq. (3.30b). This is particularly an advantage whenever calculating the gradient is infeasible or too resource consuming [119].

In the remainder of this section, we briefly introduces the Nelder-Mead method. Nelder-Mead is a popular derivative-free method that maintains $n + 1$ points in \mathbb{R}^n whose convex hull forms a simplex [111, 122]. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be the optimization function. The points are always ordered by their cost $f(\mathbf{x}_i) \leq f(\mathbf{x}_{i+1})$. In each iteration of the Nelder-Mead algorithm the goal is to improve the worst point $f(\mathbf{x}_{n+1})$ by reflecting,



expanding or contracting the simplex along the line joining the worst vertex with the centroid of the remaining vertices [111]. In special cases the simplex undergoes a shrinkage where all points are replaced by $\mathbf{x}_i \rightarrow (\mathbf{x}_1 - \mathbf{x}_i)$. Pictorially, this algorithm may be understood as a multidimensional simplex “rolling” down the objective function. The method is outlined in Algorithm 4 [123]. The conventional choice of parameters is $\rho = 1$, $\chi = 2$, $\gamma = 1/2$, and $\sigma = 1/2$.

In two dimensions the Nelder-Mead algorithm maintains three points forming a triangle. In Fig. 3.2 it is shown how Nelder-Mead “rolls” down a two-dimensional optimization landscape. However, compared with steepest descent and BFGS Nelder-Mead is much slower at reaching the optimal point. This inefficient descent is the main disadvantage of popular methods such as Nelder-Mead with CRAB. In the next section, I introduce Gradient Optimization using Parameterization (GROUP) where the gradient within a chopped basis is calculated analytically. This allows for the best of both worlds, namely the analytic gradients and the guided optimization from the parametrization (chopped basis).

3.7 Gradient Optimization using Parameterization - GROUP

In Gradient Optimization Using Parameterization (GROUP) the best features of chopped basis and GRAPE are combined. This method was introduced in Ref. [5] with me as first author. We parametrize the control in some basis as in Eq. (3.82). However, instead of using a gradient-free method such as Nelder-Mead to search in the chopped basis, we use the gradient-descent methods from GRAPE. Since we are now optimizing the expansion coefficients \mathbf{c} the Gâteaux variations of \hat{J} wrt. $\{\psi(t), \chi(t), \mathbf{c}\}$ are set to zero

$$D_{\delta\psi}\hat{J} = D_{\delta\chi}\hat{J} = D_{\delta\mathbf{c}}\hat{J} = 0. \quad (3.84)$$

This produces the same equations of motion as Eqs. (3.30a)-(3.30c) for ψ and χ . In the chopped basis we have $\hat{J} = \hat{J}(\mathbf{c})$, which has domain \mathbb{R} . Choosing the inner product to be the usual vector dot product for $X = \mathbb{R}^M$, the corresponding gradient of \hat{J} wrt. \mathbf{c} is then defined as the element $\nabla_{\mathbb{R}^M}\hat{J}$ fulfilling the relation Eq. (3.42)

$$D_{\delta\mathbf{c}}\hat{J} = \langle \nabla_{\mathbb{R}^M}\hat{J}, \delta\mathbf{c} \rangle_{\mathbb{R}^M} = \nabla_{\mathbb{R}^M}\hat{J} \cdot \delta\mathbf{c} = \sum_{m=1}^M \frac{\partial\hat{J}}{\partial c_m} \delta c_m, \quad (3.85)$$

where $\delta\mathbf{c}$ is an arbitrary variation. This definition of the gradient coincides with the usual definition of the gradient as a column vector of partial



derivatives. The GROUP gradient entries become

$$\begin{aligned}\frac{\partial \hat{J}}{\partial c_m} &= \int_0^T \left(-\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial u} \right| \psi \right\rangle - \gamma \ddot{u} \right) S(t) f_m(t) dt \\ &= \int_0^T \nabla_{L^2} \hat{J}(t) S(t) f_m(t) dt,\end{aligned}\quad (3.86)$$

where we identified the term in parenthesis to simply be the GRAPE L^2 -gradient from Eq. (3.45). Calculating the GROUP gradient amounts to first calculating the usual GRAPE L^2 -gradient and subsequently performing M inexpensive one-dimensional integrals. The coefficients are then iterated according to the line search prescription

$$\mathbf{c}^{(k+1)} = \mathbf{c}^{(k)} + \alpha^{(k)} \mathbf{p}^{(k)}, \quad (3.87)$$

where $\mathbf{p}^{(k)}$ is either the steepest descent direction or the BFGS direction both utilizing the gradient Eq. (3.85) as discussed in section 3.4. Note that when optimizing $\hat{J}(\mathbf{c})$, the optimization is done in the usual Euclidean space and not H^1 -space, so the standard methods for BFGS in Eq. (3.57) can be directly applied. Later in chapter 6, we compare this method with Nelder-Mead using CRAB, where GROUP is at least an order of magnitude faster.

3.8 Controllability

Until now we have only discussed methods for finding solutions to the state-to-state problem. In this section, we discuss sufficient conditions for solutions with unit fidelity to exist. In this section we assume the quantum system is described by a finite dimensional bilinear Hamiltonian, i.e.

$$|\dot{\psi}\rangle = -i\hat{H}(u(t))|\psi(t)\rangle = \left(\hat{H}_0 + \sum_n u_n \hat{H}_n \right) |\psi(t)\rangle, \quad (3.88)$$

The general solution this equation is $|\psi(t)\rangle = X(t)|\psi(t_0)\rangle$ where $X(t)$ is the solution to

$$\dot{X} = i \left(\hat{H}_0 + \sum_n u_n \hat{H}_n \right) X, \quad X(0) = \mathbb{1}. \quad (3.89)$$

This is a linear operator that evolves in the Lie group of unitary matrices $U(N)$. If the control allows us to realize any unitary operator then we can also reach any target state from any initial state. This observation gives us definition of controllability for a quantum system



Controllability of a Quantum System

A quantum system is controllable if it is possible to reach any target state from any initial state. That is a quantum system is controllable if the set of unitary matrices realizable by Eq. (3.89) is all of $U(N)$.

A sufficient and necessary condition for controllability is given by the so-called Lie-algebra rank criterion. Consider the Lie subalgebra

$$L_0 = \text{Lie}\{i\hat{H}_0, i\hat{H}_1, \dots, i\hat{H}_m\}, \quad (3.90)$$

which is the subspace with $i\hat{H}_0, i\hat{H}_1, \dots, i\hat{H}_m$ and all their repeated iterated commutators,

$$[i\hat{H}_0, i\hat{H}_1], [i\hat{H}_0, i\hat{H}_1], \dots, [i\hat{H}_0, i\hat{H}_m], \quad (3.91a)$$

$$[[i\hat{H}_0, i\hat{H}_1], i\hat{H}_1], [[i\hat{H}_0, i\hat{H}_1], i\hat{H}_2], \dots, [[i\hat{H}_0, i\hat{H}_1], i\hat{H}_m], \quad (3.91b)$$

$$\dots \quad (3.91c)$$

Due to the structure of $U(N)$ it is possible to prove the following theorem [32, 39].

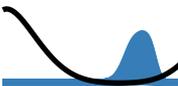
Lie Algebra Rank Criterion

A quantum system is controllable in $U(N)$ if and only if the Lie algebra (3.90) has the dimension N^2 .

As an illustrative example let us consider the two-level system with Hamiltonian

$$\begin{aligned} \hat{H}(u(t)) &= \begin{pmatrix} E & \Omega_R \\ \Omega_R & E \end{pmatrix} + \Delta(t) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= E\mathbb{1} + \Omega_R\sigma_x + \Delta(t)\sigma_z. \end{aligned} \quad (3.92)$$

where we get the previously discussed Landau-Zener Hamiltonian Eq. (3.5) for $E = 0$. This system is controllable if the Lie algebra generated by the repeated commutators has dimension $N = 4$. In order to check this explicitly, we construct a matrix W representing a basis for the Lie algebra. Generally, this matrix can be constructed as follows. First vectorize the Hamiltonian \hat{H}_0 and \hat{H}_k for all k and add them to W as column vectors. Next compute the commutator $[\hat{H}_0, \hat{H}_k]$ for all k and add the vectorized form to W only if it is linearly independent from the existing columns. Repeated commutators between the columns are



calculated until no new linearly independent columns can be added. For the Hamiltonian in Eq. (3.92) we find,

$$W = \begin{pmatrix} E & \Delta & 0 & 0 \\ \Omega_R & 0 & 2\Delta\Omega_R & 4\Delta^2\Omega_R \\ \Omega_R & 0 & -2\Delta\Omega_R & 4\Delta^2\Omega_R \\ E & -\Delta & 0 & 0 \end{pmatrix} \quad (3.93)$$

This matrix has $\text{rank}(W) = 4$ hence the system is completely controllable. If $E = 0$ as in the Landau-Zener model then $\text{rank}(W) = 3$ and the system is only completely controllable in $SU(2)$ and there is no control of the global phase, which is irrelevant in most applications.

There is a number of problems with the Lie algebra rank criterion. Firstly, as demonstrated in the example above it is difficult to check in most cases. Alternative improved criteria do exist, but they are also difficult to apply in many of cases [39, 107]. Secondly, knowing that a Hamiltonian is controllable does not give the optimal controls, implying that the optimization algorithms must still be used. Finally, this criterion does not give any information on the duration T , it simply states that a control exists but not how much time it takes to reach the target state. It can be proven under mild conditions that if a system is controllable at some duration T^* then it also controllable for all later $T \geq T^*$ [39].

However, later we shall see that the concept of controllability is closely intimately connected with the shape of the quantum control landscapes. Especially, one can show that if a system is controllable then all local maxima are also global maxima [124]. This will be discussed more in Sec. 5.3.

3.9 Conclusion

In this chapter I presented the state-to-state problem and the associate equations of optimality. Different gradient based methods such as GRAPE and Krotov's method have also been introduced. An alternative popular gradient free method is Nelder-Mead with CRAB. I introduced how the convergence rate of this method could be improved by computing the gradients in a chopped basis in GROUP. Finally, there was a brief discussion of the controllability concept. This chapter serves as background information for the applications in the later chapters. In particular, I will benchmark all these algorithms on a Bose-Einstein condensate control problem in the next chapter.

A major challenge in the quantum control field is the efficiency of quantum control algorithms both in terms of runtime and final fidelities



depend delicately on the quality of implementation. In order to make the methods presented in this chapter more accessible, I have spent considerable efforts during my PhD on developing the open-source C++ library QEngine [8]. This library offers optimized code for most of the optimization methods discussed in this chapter.



4 Optimal Control of Bose-Einstein Condensates

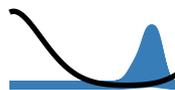
Most of this chapter is copied from Ref. [5] aide from a number of minor changes to the formulations. The introduction has been updated to fit this thesis' presentation. The discussion of dressing in Sec. 4.2.1 is copied from Ref.. [8].

In the last chapter, quantum optimal control theory was introduced along with a plethora of numerical algorithms for solving the state-to-state control problem. These local algorithms can broadly be divided along two axes. The first axis is derivative-based versus derivative-free algorithms and the second is optimization in a full or reduced-basis for admissible controls.

Derivative-free algorithms operate by directly evaluating the cost functional in carefully selected points. A particularly prevalent example in quantum optimal control is the Nelder-Mead algorithm [47, 104, 119, 120]. On the other hand, derivative-based methods use both the functional and derivative information, which can speed up the convergence rate. Important examples from quantum optimal control theory are the Gradient-Ascent Pulse Engineering (GRAPE) [109] and Krotov's method [55, 115].

Reduced-basis or chopped basis methods parametrize the control in some basis [104, 119] whereby the optimization may be steer in the correct direction. These methods also avoid the typically large dimensionan associated with the control problem. Within quantum optimal control, reduced-basis methods are typically only used in combination with derivative-free algorithms such as Nelder-Mead [47, 104, 119, 120].

Ideally, one should combine both of these two approaches and perform derivative-based optimization in a reduced-basis. In the past years, steps have been taken to implement this in a number of different ways



[108, 118, 125–127]. We will collectively refer to such methods as Gradient Optimization Using Parametrization (GROUP). In part due to the frequent success of derivative-free search in a reduced-basis, GROUP-type methods have not been widely adopted in the quantum optimal control community.

In this chapter, we compare GROUP with the many other quantum optimal control algorithms introduced in the last chapter in the context of controlling a Bose-Einstein condensate trapped in a magnetic microtrap. These systems are particularly challenging to control due to the nonlinearity in the equations of motion. We will investigate fast excitation from the ground state to the first-excited state in a single well. This system has attracted much attention in quantum optimal control community and it has been investigated both experimentally and theoretically [46, 47, 81, 110]. Due to the interest in this system, we see it as an ideal test bed for quantum control algorithms.

This chapter is organized as follows in Sec. 4.1 the control problem is introduced. In Sec. 4.2 the different quantum optimal control algorithms from the previous chapter are benchmarked against each other.

The content of this chapter was published in Ref. [5] with me as first author and as such I made major contributions to all parts of the project.

4.1 The Control Problem

In this chapter, we discuss optimal control of Bose-Einstein condensates trapped on atom chips [92]. The fundamentals of atom trapping in these systems is outlined in Sec. 2.3. In this chapter, we focus on the specific experimental configuration described in Refs. [81, 92], where a source for stimulated emission of matter waves in twin beams is created by transferring a condensate into the collective first-excited state. The typical decay rate of the system is 3 ms, so it is very important to find optimal controls that can transfer the condensate into the excited state faster than this decay rate and still allow time for subsequent experiments [47, 92]. We call this control problem *condensate driving* and it has been investigated using a number of different quantum optimal control algorithms in Refs. [46, 47, 110]. For more details on the experimental setup we refer the reader to Refs. [47, 74, 81, 92]. In a mean-field treatment, the dynamics of the condensate are well described by an effective one-dimensional Gross-Pitaevskii equation,

$$\begin{aligned} i\frac{\partial\psi}{\partial t} &= -\frac{1}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x,u)\psi + \beta|\psi|^2\psi \\ &= (\hat{H} + \beta|\psi|^2)\psi, \end{aligned} \tag{4.1}$$



where β is the effective nonlinear self-interaction, and \hat{H} is the Hamiltonian. The atom chip experiment is tightly confined along two transverse directions and weakly confined along the axial direction. One of the transverse directions is so strongly confining that the state is frozen into the ground state. The dynamics along the axial direction is slow compared with the other transverse direction, which allows for a description with an effective one-dimensional Gross-Pitaevskii equation with a nonlinear dependence of β on the atom number [47, 128]. For 700 atoms, one finds $\beta = 1830\hbar \text{ Hz}\mu\text{m}$ [47].

The potential is an anisotropic Ioffe-Pritchard trap dressed by a radio-frequency potential [47, 74, 81] as outlined in Sec. 2.3. If the radio-frequency field is weak the potential is close to a single well as shown in Fig. 2.4(b). In Sec. 2.3 this adiabatically dressed potential [Eq. (2.26)] was found using the rotating wave approximation. However, due to the high sensitivity of this control problem this approximation is not valid. Instead one must numerically diagonalize the Zeemann Hamiltonian Eq. 2.19 for each value of x to find the potentials. The result of this numerical calculation can be well approximation by a sixth order polynomial

$$V(x, u(t)) = p_2(x - u(t))^2 + p_4(x - u(t))^4 + p_6(x - u(t))^6, \quad (4.2)$$

where the control $u(t)$ is the displacement of the trap. The coefficients are given by $p_2 = 2\pi\hbar \cdot 310r_0^{-2} \text{ Hz}$, $p_4 = 2\pi\hbar \cdot 13.6r_0^{-4} \text{ Hz}$, and $p_6 = -2\pi\hbar \cdot 0.0634r_0^{-6} \text{ Hz}$, with $r_0 = 172\text{nm}$ [47]. For an in-depth discussion of the experimental setup we refer the reader to Refs. [47, 92]. The goal is to transfer the initial state ψ_0 into the target state ψ_t after a duration of T . The initial state is the ground state for $u(t=0) = 0$ and the target state is the first-excited state for $u(t=T) = 0$.

This is a state-to-state control problem as described in Eq. (3.3a) thereby allowing all the methods from the last section to be directly applied. Here we investigate the control problem at $T = 1.09\text{ms}$, which was reported in Ref. [47] to the lowest durations where solutions could be found.

4.1.1 Choice of Chopped Basis

In order to apply the chopped basis methods from Secs. 3.6-3.7 it is necessary to introduce an expansion according to Eq. 3.82. For the control problem discussed here, we found the following expansion useful

$$u(t) = u_0(t) + S(t) \sum_{n=1}^M c_n \sin\left(\frac{\omega_n t}{T}\right), \quad (4.3)$$



where $\omega_n = n\pi$ is a set of frequencies. This basis can be changed into a chopped random basis as described in Sec. 3.6 by randomly shifting the frequencies as $\omega_n = (n + r_n)\pi$, where $-0.5 \leq r_n \leq 0.5$ are initially chosen random numbers.

4.1.2 Filter Function

In order to obtain a close match to the experimental conditions in the condensate driving control problem, it is also necessary to include the finite bandwidth of the control electronics. This effect causes the control $u(t)$ to become distorted into $v(t)$ and the atoms experience the potential from $V(x, v)$ [47]. The distortion is large enough to cause the fidelity to drop by a couple of percent, so it must be included in the modeling. It has previously been discussed how to include this type of effect into GRAPE [129] and Krotov-type methods [130, 131]. The distorted control is given by a convolution with the filter $h(\tau)$

$$v(t) = (h * u)(t) = \int_0^t h(\tau)u(t - \tau)d\tau. \quad (4.4)$$

This effect changes the expressions for the gradients. Again, we have to calculate the variation. This can be done with the chain rule for variations,

$$\begin{aligned} D_{\delta u} \hat{J}(v) = D_{D_{\delta u} v} \hat{J}(v) &= \int_0^T \left(-\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial v} \right| \psi \right\rangle - \gamma \dot{v} \right) (h * \delta u)(t) dt. \\ &+ \gamma \dot{v}(T) (h * \delta u)(T) \end{aligned} \quad (4.5)$$

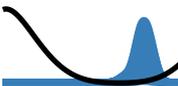
The expression inside the bracket is evaluated along the distorted control $v(t)$. An extra term appears here since it may happen that $v(T) \neq 0$ although $u(T) = 0$. Calculating the two terms individually one finds,

$$\int_0^T \nabla \hat{J}_{L^2}(v) \delta v(t) dt = \int_0^T \int_t^T \delta u(t) \nabla \hat{J}_{L^2}(v) h(t' - t) dt' dt, \quad (4.6)$$

$$\gamma \dot{v}(T) \delta v(T) = \int_0^T \delta u(t) \gamma \dot{v}(T) h(T - t) dt. \quad (4.7)$$

Combining these two results gives

$$\int_0^T \nabla \hat{J} \delta u dt = \int_0^T \delta u(t) \left(\gamma \dot{v}(T) h(T - t) + \int_t^T \nabla \hat{J}_{L^2}(v) h(t' - t) dt' \right) dt. \quad (4.8)$$



From this expression we can read off the L^2 gradient including a filter function

$$\nabla \hat{J} = \gamma \dot{v}(T) h(T-t) + \int_t^T \left(\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial v} \right| \psi \right\rangle + \gamma \ddot{v} \right) h(t'-t) dt'. \quad (4.9)$$

However, the first of these terms is not zero when the filter is included. From this expression of the L^2 -gradient one may directly compute the H^1 . Similarly the GROUP gradient becomes

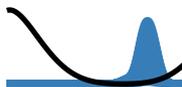
$$\begin{aligned} \frac{\partial \hat{J}(\mathbf{c})}{\partial c_n} = & - \int_0^T \left(\Re \left\langle \chi \left| \frac{\partial \hat{H}}{\partial v} \right| \psi \right\rangle + \gamma \ddot{v} \right) (h * S f_n)(t) dt \\ & + \gamma \dot{v}(T) (h * S f_n)(T) \end{aligned} \quad (4.10)$$

With these expressions for the gradients we can apply the quantum optimal control algorithms to the condensate driving control problem.

4.2 Results

We applied the following quantum optimal control algorithms GRAPE, Krotov, Nelder-Mead with CRAB and GROUP with CB and CRAB to the condensate driving control problem. We applied the algorithms to the same 100 initial controls, which were randomly generated using Eq. (4.3). The convergence behavior of the different methods is illustrated in Fig. 4.1. Here the median and 25% and 75% quartiles are shown for the different algorithms, which gives an impression of the expected behavior for each method on this problem. One function evaluation is a solution of the Gross-Pitaevskii equation or the adjoint equation for the Lagrange multiplier [Eq. (3.30b)]. Note, that the presentation of Krotov's method in Sec. 3.5 uses a linear Schrödinger equation implying that the Krotov update rule in Eq. (3.81) does not strictly guarantee a monotonic decrease in the cost. For the small values of the nonlinear interaction (β) considered here, one typically just the same back propagated state as in GRAPE and forfeit the strict guarantee of a decreased cost [110].

Throughout the optimization GROUP achieves the lowest infidelities. At the end of the optimization GROUP has the best infidelity followed by GRAPE and Krotov. There is no particular difference between GROUP using CB or CRAB so we will refer to them both as GROUP. Nelder-Mead using CRAB has the slowest convergence rate of the four methods, since it does not utilize derivative information. This is in accordance with the picture presented in Fig. 3.2, which shows that derivative-based methods are typically faster than derivative-free.



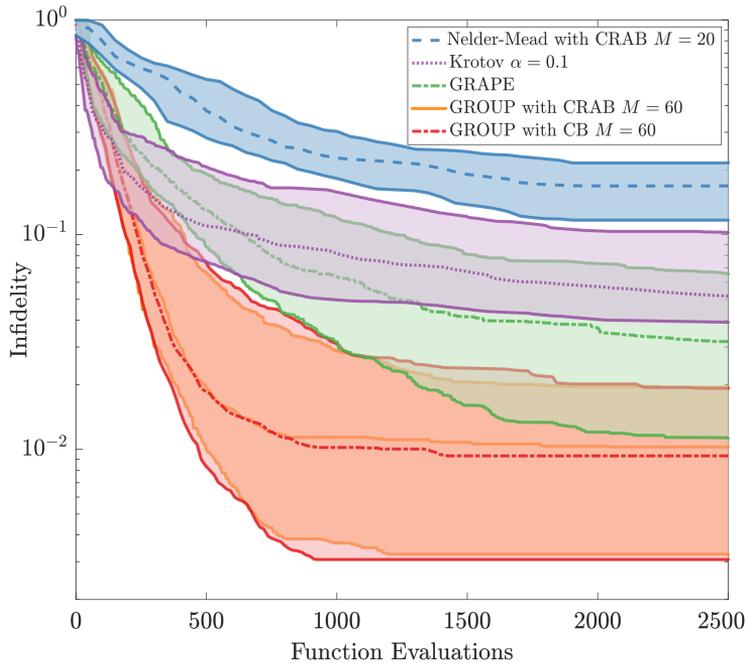
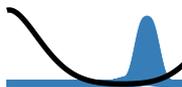


Figure 4.1: The infidelity as a function of the number of evaluations for each algorithm. One function evaluation is a solution of the Gross-Pitaevskii equation or Lagrange multiplier Eq. (3.30b). The different algorithms are shown at the basis size or steps size where they performed the best (see legend). The dotted line shows the median and the shaded area indicates the 25%- and 75%-quartiles found from 100 different random initial controls. The quasi-Newton method BFGS was used together with GRAPE and GROUP.

The optimization curves in Fig. 4.1 can be split into two regimes: one at a high number of function evaluations after 600 and one below. Below 600 evaluations, the three derivative-based methods have similar rates of convergence, but GROUP and Krotov perform better than GRAPE. In the high number of function evaluations regime GROUP performs better than GRAPE and Krotov. We attribute this to the fact that the basis gradually steers GROUP towards a more profitable part of the optimization landscape. This shows the complexity of the optimization landscape, since although the algorithms start at the same point and perform local greedy optimization they converge towards different points with different infidelities. This is different from the situation in Fig. 3.2 where all the algorithms converge to the same point since the two-dimensional



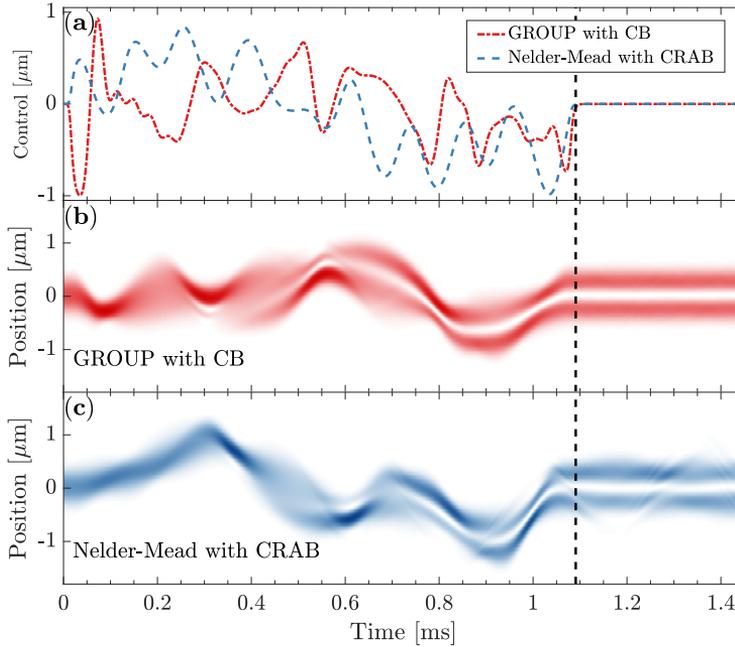
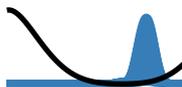


Figure 4.2: **(a)** The best case controls found using GROUP with CB (red) and Nelder-Mead with CRAB (blue) after 2500 iterations. The control is held constant after the vertical dashed line. **(b)** The density for the condensate ($|\psi(x, t)|^2$) found when using the solution from GROUP. After the vertical line, the density is constant since the state has converged with $F = 0.999$ to the first excited state. **(c)** Here the same is shown as **(b)** for Nelder-Mead with CRAB, which has residual oscillations since the final fidelity is $F = 0.95$.

landscape is much simpler than the control problem, which has dimension $M \simeq 3500$ or $M \simeq 50$.

In the high function evaluations regime Krotov and GRAPE switch places and GRAPE finds better infidelities than Krotov. These results show that Krotov achieves fast initial reductions in the infidelity, but it slows considerably down when approaching the optimum. GRAPE does not exhibit this behavior, which we attribute to the Hessian approximation from BFGS, since the cost function can be well described by a second-order expansion close to the optimum [111]. Similar results were also reported in Ref. [110]. In principle, Krotov's method can also be combined with a BFGS type method [116], but it has been reported in Refs. [110, 116] that this does not significantly improve the convergence.



This comparison focuses on the expected behavior. Each of the methods have a few optimization runs that perform significantly better, which reflects that each algorithm has some specific seeds where it just happens to search the optimization landscape in the most favorable manner. GROUP had the individual optimization runs with the lowest infidelity.

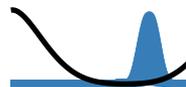
The best controls found for GROUP with CB and Nelder-Mead with CRAB are shown in Fig. 4.2, with respectively, $F = 0.999$ and $F = 0.95$. This figure also shows the density for the condensate when propagated along the optimal controls. After T , the control is held constant at $u = 0$. The Nelder-Mead with CRAB solution has a residual oscillation after T due to residual excited-state components in the solution.

4.2.1 Reduced-Basis Size and Dressed Methods

A limitation of the reduced-basis methods is that the optimization might get caught in an artificial trap introduced by limited bandwidth of the parametrization. As an example, consider the expansion in Eq. (4.3) at some fixed M . If the optimal solution requires frequencies above M , then the optimization can never converge to this solution and it will get caught in an artificial trap. This effect would favor larger values of M . On the other hand, increasing M also increases the dimension of the control problem, which is exactly what the reduced-basis methods attempt to avoid.

In Fig. 4.3(a) we compare the trade-off between the risk of artificial traps and retaining a low dimension for Nelder-Mead with CRAB and GROUP with CB. As expected, a too small basis $M \leq 10$ gives poor results for both algorithms since the optimal controls cannot be adequately described with these small basis sizes. Nelder-Mead with CRAB clearly has an optimal basis size around $M = 20$ and it becomes worse with a larger dimension. This is due to the fact that the Nelder-Mead algorithm cannot effectively search within a large dimension. Surprisingly, the performance in GROUP is very robust with respect to the basis size, and it even seems to prefer large basis sizes. We attribute this result to the fact that the large frequencies have small weights in the gradient, meaning that they contribute little to the search. This type of behavior is illustrated in Fig. 4.4, which shows the median initial gradient for the 100 initial points. Above $n = 50$, the partial derivatives are much smaller than for $n = 5$, and hence it is the low frequencies that dominate the search.

A solution to the artificial trap problem was proposed in Ref. [132] where a so-called dressed CRAB (dCRAB) was introduced. In general



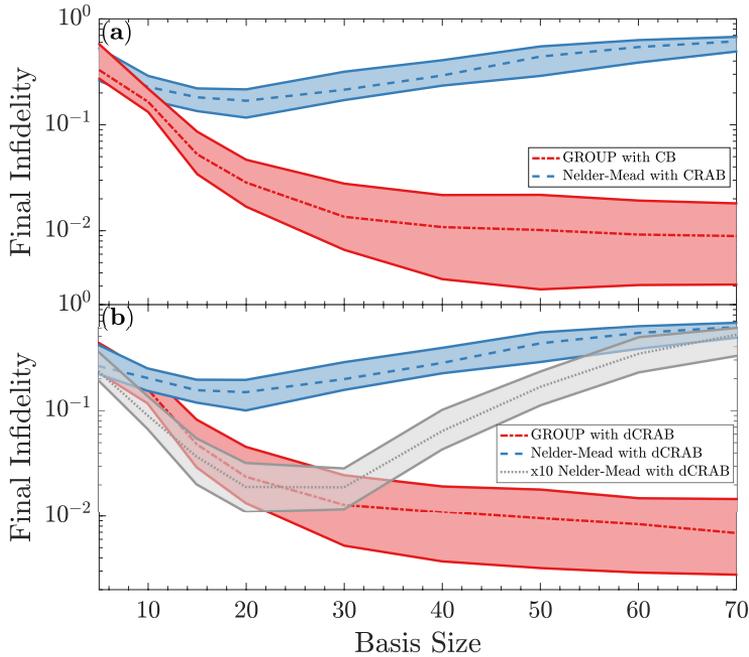


Figure 4.3: (a) The final infidelity after 2500 function evaluations for GROUP with CB (red) and Nelder-Mead with CRAB (blue) for different basis sizes. The dotted line is the median and the shaded area display the 25% and 75% quartiles found from the 100 different random initial controls. (b) The same as (a) for Nelder-Mead and GROUP with dressed CRAB. The gray plot shows the final infidelity for Nelder-Mead with dressed CRAB after 25,000 iterations rather than 2500 iterations.

we may let $f_m = f_m(t; \mathbf{r}_m)$ where \mathbf{r}_m is a vector of values that is usually drawn at random. As describe above we may choose $f_m(t, r_m) = \sin((m + r_m)\pi t/T)$ where $-0.5 \leq r_m \leq 0.5$ is drawn from a uniform distribution. Then, if the algorithm gets trapped at (possibly) an artificial minimum, we set $u_0(t) \leftarrow u(t)$, re-initialize the algorithm with coefficients $\mathbf{c} = \mathbf{0}$, and draw a new set of basis functions f_m defined by a new set of values \mathbf{r}_m ,

$$u_0(t) \leftarrow u(t), \quad \mathbf{c} \leftarrow \mathbf{0}, \quad f_m(t; \mathbf{r}_m) \leftarrow f_m(t; \mathbf{r}_m^*), \quad (4.11)$$

where \mathbf{r}_m^* are new random values. This changes the topology of the optimization landscape and the artificial trap may have been eliminated. Effectively, this corresponds to restarting the GROUP algorithm with a new parametrization basis. These restarts are known as superiterations [132].



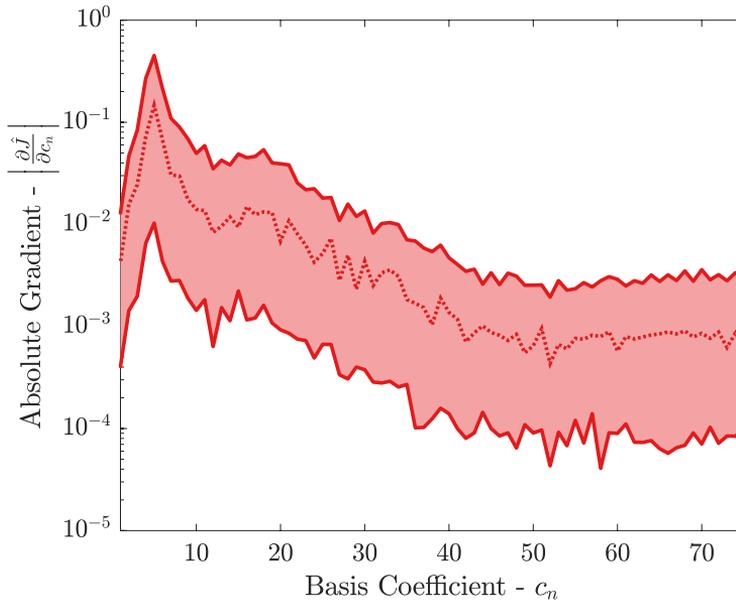
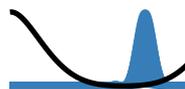


Figure 4.4: The absolute value of the partial derivative with respect to basis coefficients $|\partial\hat{J}(\mathbf{c})/\partial c_n|$ found using Eq. (4.10). The dotted line shows the median of the 100 gradients taken at the start of the optimization. The shaded area indicates the 25% and 75% quartiles.

This modification is referred to as dressed GROUP, or dGROUP for short [5]. Clearly, this formalism cannot be used in a normal chopped basis (CB). We apply the same methodology to GROUP by simply following the same procedure for the superiterations and calculate the gradients using Eq. (4.10).

In Fig. 4.3(b), we compare the effect of dressing for different basis sizes. GROUP in combination with dCRAB (dGROUP) only does slightly better than GROUP with CB, which indicates that the solutions were already highly optimal after 2500 evaluations. We observe the same behavior for Nelder-Mead with dCRAB after 2500 evaluations. However, if Nelder-Mead with dCRAB is allowed to perform 25,000 evaluations, then it finds much better results especially for basis sizes around $M = 20$. This shows that for low basis sizes, Nelder-Mead with dCRAB can escape the artificial traps and greatly improve the infidelity with enough function evaluations. The Nelder-Mead with dCRAB using 25,000 evaluations finds similar infidelities to GROUP with dCRAB using 2500 evaluations, which shows that if Nelder-Mead runs for longer times, it can find results similar to gradient descent. At large basis sizes, dCRAB does not notably improve the infidelity, which



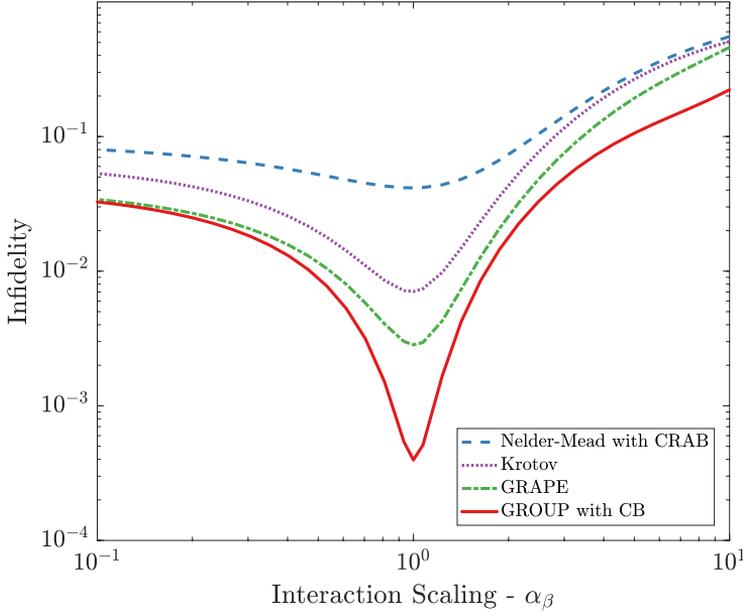


Figure 4.5: The robustness of the optimal solutions from each algorithm to rescaling of the self-interaction in Eq. (4.1): $\tilde{\beta} = \alpha_\beta \beta$, where α_β is the scaling.

we interpret as the optimization being blocked by Nelder-Mead's inability to efficiently search in high-dimensional landscapes rather than artificial traps.

4.2.2 Robustness Analysis

The solutions presented in this chapter can only be directly applied in an experiment if they are stable against unavoidable experimental fluctuations in the system's parameters. As reported in Ref. [47], the primary experimental fluctuations occur in the self-interaction (β) in Eq. (4.1) due to variations in the atom number. This gives variations in β below 14%. For this reason, we have investigated the stability of the solutions against scaling of the self-interaction. In Fig. 4.5, the infidelity is calculated with the best solution from each algorithm using $\tilde{\beta} = \alpha_\beta \beta$, where α_β is the scaling factor. The solutions are generally stable with respect to fluctuations in the self-interaction, since atom interactions have little influence over the relative short duration investigated here. The GROUP solution clearly exhibits superior behavior over the entire range of parameter variation. The GROUP solution remains above $F = 0.99$ in



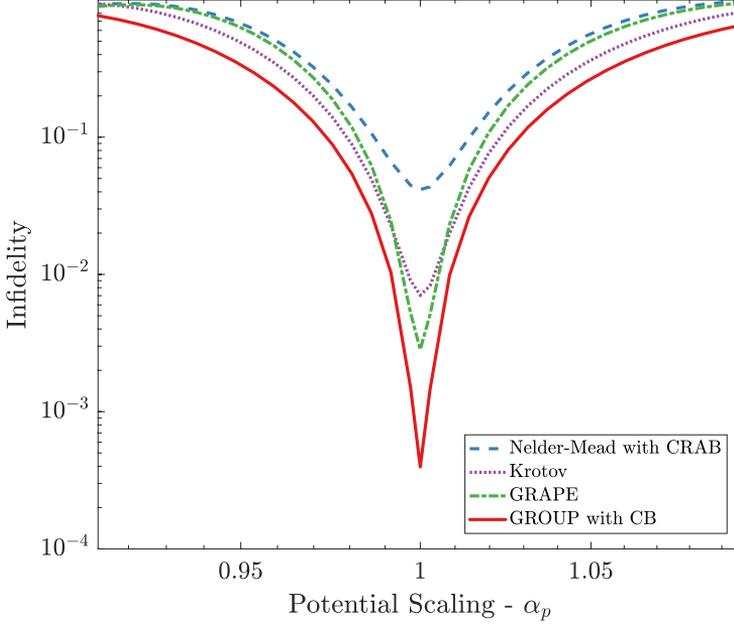


Figure 4.6: The robustness of the optimal solutions from each algorithm to rescaling of the potential coefficients in Eq. (4.2): $\tilde{p}_i = \alpha_p p_i$, $i = 2, 4, 6$, where α_p is the scaling.

the range $0.5 \leq \alpha_\beta \leq 1.5$, which is above the experimental fluctuations reported in Ref. [47].

We have also investigated the robustness of the solutions against perturbations in the potential. In Fig. 4.6, the infidelity is calculated for each algorithm's best solution with rescaled potential coefficients ($\tilde{p}_i = \alpha_p p_i$) in Eq. (4.2). The solutions are much more sensitive to perturbations in the potential compared with the self-interaction.

We note, that it would most likely be possible to find much more stable solutions by including the stability directly in the optimization. This could, for example, be achieved by optimizing over an ensemble of different scaled interactions and potentials and defining the average cost over the ensemble.

4.2.3 Convergence Behavior at Different Parameter Values

The results presented in this section have so far only discussed a single choice of the system parameter values, meaning that the underlying optimization problem has been the same. The performance of an op-



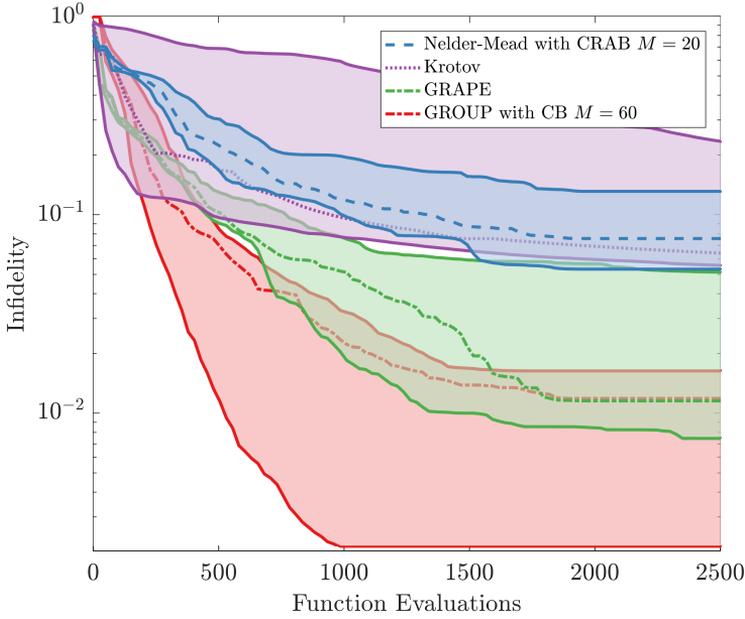


Figure 4.7: Comparison of the convergence behavior for different values of the self-interaction for each algorithm. The optimizations are performed for 35 different scaled values of the self-interaction $\tilde{\beta} = \alpha_{\beta}\beta$, where α_{β} is the scaling. The dotted line shows the median and the shaded area indicates the 25% and 75% quartiles. The quasi-Newton method BFGS was used together with GRAPE and GROUP.

timization algorithm may depend on the choice of problem, so in this section we investigate the performance of each algorithm for different problems by rescaling the system’s parameters. In Fig. 4.1, we compared the convergence behavior for a number of different seeds but fixed system parameters. In Figs. 4.7 and 4.8, we hold the seed fixed and vary the system’s parameters examined in the previous section. The seeds are those that gave the best result for each algorithm in the previous analysis. As in Fig. 4.1, we have performed the analysis for different basis sizes and step sizes and we present the best case behavior in Figs. 4.7 and 4.8.

In Fig. 4.7, we compare the convergence behavior of each algorithm when optimizing with 35 different values of $\tilde{\beta} = \alpha_{\beta}\beta$ ranging from $\alpha_{\beta} = 0.1$ to $\alpha_{\beta} = 10$ as in the previous section. Although GRAPE and GROUP have similar medians for their final infidelities, GROUP has a better overall convergence behavior. Both GROUP and GRAPE have faster rates of convergence than Nelder-Mead with CRAB and Krotov’s method. In



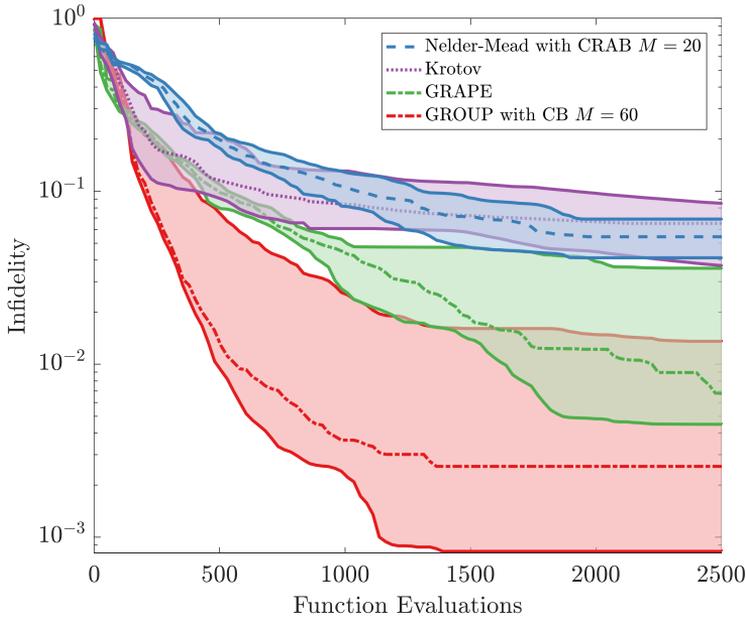


Figure 4.8: Comparison of convergence behavior for different values of the potential coefficients. The optimizations are performed for 35 different scaled values of the potential coefficients $\tilde{p} = \alpha_p p_i$ with $i = 2, 4, 6$, where α_p is the scaling. The dotted line shows the median and the shaded area indicates the 25% and 75% quartiles. The quasi-Newton method BFGS was used together with GRAPE and GROUP.

Fig. 4.1, Krotov’s method has a better final infidelity than Nelder-Mead with CRAB. However in Fig. 4.7, Nelder-Mead with CRAB catches up with Krotov after a high number of function evaluations. In general, Krotov’s method is struggling for the high values of β , which we attribute to the lack of including the term proportional to the difference in Eq. (3.81). We believe that correctly including this term could substantially improve the convergence rate of Krotov’s method. Note that here we present the results with the best basis for Nelder-Mead after optimizing over a range of basis sizes. The performance is significantly worse for nearby basis sizes 15 and 30.

In Fig. 4.8, we perform a similar analysis for each algorithm just optimizing 35 different values of the potential with rescaled coefficients $\tilde{p}_i = \alpha_p p_i$, with $i = 2, 4, 6$ ranging from $\alpha_p = 0.91$ to $\alpha_p = 1.09$ as in the previous section. Again, we observe that GROUP followed by GRAPE finds the solutions with the lowest infidelities. Krotov’s method has fast initial

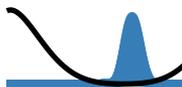


reductions in the infidelity below 600 function evaluations, but it slows down above 1000 function evaluations. In this case, we also observe that Nelder-Mead with CRAB catches up with Krotov's method after a high number of function evaluations. Similarly to before, we observe in our other optimizations that Nelder-Mead has significantly worse performance for the nearby basis sizes 15 and 30.

We have also performed a similar analysis for rescaled values of the regularization factor (γ) in Eq. (3.3), which showed that the regularization factor only has a minor impact on the results for all algorithms.

4.3 Conclusion

In this chapter, I have presented a detailed comparison of the different optimal control algorithms on a particular problem. In particular, I compared gradient-based (GROUP) and gradient-free optimization (Nelder-Mead) in a chopped basis, which shows that gradient-based optimization in GROUP can greatly improve the rate of convergence. Our analysis shows that GROUP is competitive with other quantum control algorithms. It would be very interesting to compare these methods on other control problems in order to better understand the advantages and disadvantages of each method. The formalism for GROUP type methods presented here can also be applied to systems with a linear Schrödinger equation such as control of many-body systems [47, 119]. I believe our method and analysis is a relevant addition to the repository of quantum control algorithms.



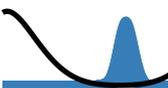
5 Time-Optimal Control

Secs. 5.3 is copied from Ref. [6]. The chapter conclusion is also based on Ref. [6].

We are not satisfied with any solution to the state-to-state control problem in a number of applications. Often, we seek the fastest possible solution, i.e. a control minimizing the duration T subject to the constraint that $F = |\langle \psi_t | \psi(T) \rangle|^2 = 1$. The motivation for fast solutions is typically that the experiment must be completed faster than the decoherence time-scale. We already saw an example of this in the last chapter where the condensate driving control problem that has a decoherence time-scale of 3 ms. There are numerous applications but a few important examples are superconducting qubits [102, 118, 133], spin-networks [134–136], and nitrogen vacancies [45, 48]. The fastest possible solution is often known as the time-optimal solution.

However, Heisenberg’s time-energy uncertainty principle limits the rate of quantum evolution and therefore also sets lower limits on how quickly quantum systems can be manipulated. This is often reformulated in terms of so-called quantum speed limits. In particular, it has been shown how quantum speed limits sets the maximal rate of quantum information communications [137], maximal rate of quantum information processing [138], maximal rate of quantum entropy production [139], and best precision in quantum metrology [14]. In the context of this thesis the most exciting results are how quantum speed limits defines the shortest time-scale for quantum optimal control algorithms to converge [120, 140]. In this chapter, I will discuss some of the connections between time-optimality and quantum speed limits.

Analogues to the state-to-state control problem, I define the time-optimal state-to-state control problem as follows.



Time-Optimal State-to-State Control Problem

Given an initial state $|\psi_0\rangle$ and a target state $|\psi_t\rangle$ the time-optimal state-to-state problem is,

$$\min_u J = \int_0^T 1 dt = T \quad (5.1a)$$

$$\text{s.t. } i\dot{\psi} = \hat{H}(\psi, u)\psi, \quad \text{in } 0 \leq t \leq T, \quad (5.1b)$$

$$\psi(0) = \psi_0, u(0) = u_i, u(T) = u_T, \quad (5.1c)$$

$$|\langle \psi_t | \psi(T) \rangle|^2 = 1, \quad (5.1d)$$

where u_i and u_T are the initial and final values of the control.

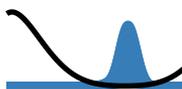
The solutions to this optimization problem are called time-optimal solutions with $T = T_{\min}$ being the minimal time. I will follow the standard terminology of minimal time but in the context of this thesis a more appropriate term would be minimal duration. In most cases, it is not possible to solve the time-optimal state-to-state problem analytically. To my knowledge solutions only exist for two-level systems [52, 141], three-level systems [142, 143], certain coupled two-level systems [135, 144] and a special case of the quantum harmonic oscillator [145].

In Sec. 5.1 I discuss how T_{\min} is non-zero due to quantum speed limits. Later, in Sec. 5.2 I discuss how to analytically solve the time-optimal Landau-Zener control problem using Pontryagin's Maximum principle. These analytic solutions will reveal connections between the quantum speed limit and the minimal time. The solutions will also allow me to comment on the "universal behavior" below the quantum speed limit. In Sec. 5.3, I discuss how to numerically estimate T_{\min} using a combination of global and local optimization in control problems with Bose-Einstein condensates. The chapter is concluded in Sec. 5.3.

The content of Sec. 5.3 is published in Ref. [6] that at the time of writing is under revision. As first author, I made primary contributions to all parts of the project.

5.1 The Quantum Speed Limit

A fundamental question is how quickly can a quantum state be transformed into another, i.e. to what extent does quantum mechanics itself impose a lower bound on T_{\min} . Indeed based on the Heisenberg's time-energy uncertainty relation $\Delta t \Delta E \geq 1/2$ we expect such a lower bound



could exist. This insight can be formalized with the concept of quantum speed limits.

The quantum speed limit is defined as the time it takes a state $|\psi(0)\rangle$ to become orthogonal to itself [146]. Note, that with this definition the quantum speed limit is, contrary to its name, not actually a speed but a time. For this reason early papers on quantum speed limits used the term orthogonalization time [147]. Here we follow the derivation from [137]. Expanding the current state $|\psi_0\rangle$ in the energy eigenbasis gives

$$|\psi(t)\rangle = \sum_n c_n e^{-iE_n t} |\varphi_n\rangle. \quad (5.2)$$

Consider the self-correlation amplitude $S(t) = \langle\psi(0)|\psi(t)\rangle$. The quantum speed limit is the lowest time where $|S(T_{\text{QSL}})|^2 = 0$ that may be found from

$$|S(t)|^2 = \sum_{n,n'} |c_n|^2 |c_{n'}|^2 \cos((E_n - E_{n'})t). \quad (5.3)$$

Using the trigonometric inequality $\cos(x) \geq 1 - 4x \sin(x)/\pi^2 - 2x^2/\pi$ and rearranging gives the result

$$|S(t)|^2 \geq 1 + \frac{4t}{\pi^2} \frac{d}{dt} |S(t)|^2 - \left(\frac{2t\Delta E}{\pi^2}\right)^2, \quad (5.4)$$

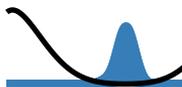
Noting that $|S(t)|^2 \geq 0$ we can inspect this inequality at the orthogonalization time $S(T_{\text{QSL}}) = 0$ where the second term in the above equation vanishes,

$$0 \geq 1 - \left(\frac{2T_{\text{QSL}}\langle\Delta E\rangle}{\pi^2}\right)^2. \quad (5.5)$$

Rearranging this expression gives the so-called Mandelstam-Tamm bound [146, 148],

$$T_{\text{QSL}} = \frac{\pi}{2\Delta E}. \quad (5.6)$$

The reader may recognize this expression as Heisenberg's time-energy uncertainty relation. Unlike the regular uncertainty relation between position and momentum, this relation is not a statement about how precisely two incompatible observables can be measured but it rather sets an intrinsic time scale for any time-independent system to self-orthogonalize. This bound has been rederived in a number of different ways by a number of authors [149–151]. As pointed by in Refs. [146, 152] a conceptual issue with Eq. (5.6) is that the variance ΔE may be arbitrarily large even if the energy is bounded.



It is possible to derive an alternative form of the quantum speed limit without this issue. Consider the real part of $S(t)$,

$$\begin{aligned}\Re S(t) &= \sum_n |c_n|^2 \cos(E_n t) \geq \sum_n |c_n|^2 \left[1 - \frac{2}{\pi} (E_n t + \sin(E_n t)) \right] \\ &= 1 - \frac{2}{\pi} E t + \frac{2}{\pi} \Im S,\end{aligned}\quad (5.7)$$

where the trigonometric inequality $\cos(x) \geq 1 - 2(x + \sin(x))/\pi$ for $x \geq 0$ has been used. For orthogonal initial and final states we have $\Re S = \Im S = 0$ implying that the minimal time between two orthogonal states is

$$T_{\text{QSL}} = \frac{\pi}{2E}.\quad (5.8)$$

In the literature, this is known as the Margolus-Levitin bound [146, 147]. The Margolus-Levitin bound does not have the conceptual issue that it depends on the variance of some observable.

As a consequence the minimal time it takes a state to travel between two orthogonal states for a time-independent Hamiltonian is

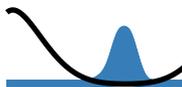
$$T_{\text{QSL}} = \max\left\{ \frac{\pi}{2\Delta E}, \frac{\pi}{2E} \right\}.\quad (5.9)$$

This is the so-called unified bound, which is tight in the sense that it is only saturated for states of the form $|\psi\rangle = (|\varphi_0\rangle + |\varphi_1\rangle)/\sqrt{2}$ where $\hat{H}|\varphi_k\rangle = E_k|\varphi_k\rangle$ [137]. These bounds were observed experimentally in shaken lattice experiments with Bose-Einstein condensates [153].

A fundamental problem with these quantum speed limits is that they are all derived under the assumption of a time-independent Hamiltonian. When designing a control pulse the resulting Hamiltonian function $\hat{H}(u(t))$ is almost surely not time-independent, hence we cannot just apply these bounds to some time-optimal state-to-state problem and get a reasonable estimate of T_{min} .

As an alternative to the quantum speed limits in Eq. (5.9), we discuss the so-called geometric approach [146, 154, 155]. This presentation is based on the seminal Ref. [151]. Let $|\psi(t)\rangle$ be a state trajectory such that $|\psi(t=0)\rangle = |\psi_0\rangle$ and $|\psi(t=T)\rangle = |\psi_t\rangle$ and note that in general several possible state trajectories exist connecting the initial state and target state. Let us consider how quickly the states change along this state trajectory

$$|\psi(t+dt)\rangle = |\psi(t)\rangle + \frac{d}{dt}|\psi(t)\rangle dt + \frac{1}{2} \frac{d^2}{dt^2}|\psi(t)\rangle dt^2 + \mathcal{O}(dt^3).\quad (5.10)$$



The Schrödinger Equation gives

$$\frac{d}{dt}|\psi(t)\rangle = -i\hat{H}|\psi(t)\rangle, \quad \frac{d^2}{dt^2}|\psi(t)\rangle = -i\frac{d\hat{H}}{dt}|\psi(t)\rangle - \hat{H}^2|\psi(t)\rangle. \quad (5.11)$$

Utilizing that both \hat{H} and $d\hat{H}/dt$ are Hermitian gives for the fidelity that

$$\begin{aligned} |\langle\psi(t)|\psi(t+dt)\rangle|^2 &= 1 - \left(\langle\psi(t)|\hat{H}^2|\psi(t)\rangle - \langle\psi(t)|\hat{H}|\psi(t)\rangle^2\right)dt^2 + \mathcal{O}(dt^3) \\ &= 1 - \Delta E(t)^2dt^2 + \mathcal{O}(dt^3), \end{aligned} \quad (5.12)$$

where $\Delta E(t)$ is the instantaneous energy uncertainty. The distance between two states with fidelity $F = |\langle\psi|\varphi\rangle|^2$ may be characterized by the Wotters' distance [156, 157]

$$\theta(\psi, \varphi) = 2 \arccos(\sqrt{F}). \quad (5.13)$$

The rate of change in θ along the state trajectory is using Eq. (5.12)

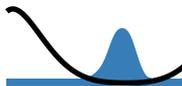
$$d^2\theta = 4(1 - |\langle\psi(t)|\psi(t+dt)\rangle|^2) = 4\Delta E^2(t)dt^2. \quad (5.14)$$

This equation may be understood as expressing the Fubini-Study metric on the projective Hilbert space [151]. Rearranging Eq. (5.14) gives

$$\frac{d\theta}{dt} = 2\Delta E. \quad (5.15)$$

In conclusion the current energy uncertainty bounds the rate of change in a quantum state. The advantage of this equation is that ΔE can be understood as a proper *speed limit* and not an orthogonalization *time*. This result also has a nice geometric interpretation as the magnitude of change along the perpendicular component of the state [158]. Similar to the Mandelstam-Tamm bound (5.6) fast solutions should have a high ΔE as to facilitate fast motion in Hilbert space.

A common situation is that we seek a fast state transfer from an eigenstate of the Hamiltonian at $t = 0$ to another eigenstate at $t = T$. According to Eq. (5.15) we should try to create a high ΔE at all times in order to move as fast as possible. However, we also have the opposing criterion that $\Delta E \rightarrow 0$ at $t \rightarrow 0$ and $t \rightarrow T$ since we are moving between eigenstates. For this reason optimal control pulses often go through an initial phase with strong excitations creating a high ΔE facilitating rapid motion in Hilbert space towards the target state. Later when approaching the target state, ΔE must decrease so the state is rapidly deexcited. This pattern of excitation and deexcitation is very common and in the later chapters with numerical applications, we will observe this pattern several times.



It is also possible to use Eq. (5.15) to find a time-dependent version of the Mandelstam-Tamm bound. We may define the distance along the state trajectory

$$s = 2 \int_0^T \Delta E(t) dt = \int d\theta, \quad (5.16)$$

This distance is independent of the $\hat{H}(t)$ used to realize $|\psi(t)\rangle$ enabling a geometric interpretation. For orthogonal initial and final states the minimal distance is along a geodesic with distance $s \leq \pi$. Inserting Eq. (5.16) into this inequality gives

$$\langle \Delta E \rangle_{\text{avg}} T_{\text{QSL}} \geq \frac{1}{2}, \quad (5.17)$$

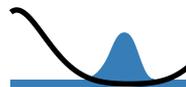
where $\langle \Delta E \rangle_{\text{avg}}$ is the time averaged energy uncertainty. It is tempting to use this relation to estimate T_{min} for a control problem. However, one must remember that this expression only estimates T_{QSL} along a particular $|\psi(t)\rangle$ induced by some control. In other words, this only estimates an orthogonalization time along a particular route from initial to target state and in general there exist several possible routes each with their own quantum speed limit. This makes it difficult to apply Eq. (5.17) to estimate T_{min} on actual control problems [159].

5.2 Pontryagin's Maximum Principle

The general problem of controlling a system described by a differential equation extends beyond quantum mechanics in the mathematical field called control theory. An important result in this field is Pontryagin's Maximum Principle, which can be used to solve general control problems. This principle is outlined below in the general manner and not specialized to quantum mechanics although we shall shortly see how it can be applied in the Landau-Zener control problem. Consider a state $\mathbf{x} \in \mathbb{R}^N$ with the dynamical equation $\dot{\mathbf{x}} = f(t, \mathbf{x}, u)$. Our job is to minimize the cost functional

$$J(u) = \int_0^T r(\mathbf{x}(t), u(t)) dt + g(\mathbf{x}(T)), \quad (5.18)$$

where T is the duration, r is the running cost and g is the terminal cost. Whenever discussing time-optimal control this cost functional has the simple form $r(\mathbf{x}, u) = 1$ and $g(\mathbf{x}(T)) = 0$ as can be seen from Eq. (5.1). An optimal control is a minimizer of this cost functional, i.e. a u^* satisfying $J(u^*) \leq J(u)$ for all u . Now we can formulate Pontryagin's Maximum in the free time and fixed final state formulation.



Pontryagin's Maximum Principle for the Time-Optimal Control

For a control system governed by the ordinary differential equations as $\dot{\mathbf{x}} = f(t, \mathbf{x}, u)$ and cost functional $J(u) = \int_0^T 1 \, dT$ and corresponding optimal trajectory $x^*(t)$ there exists a function $p^*(t)$ for $0 \leq t \leq T^*$ satisfying

$$\dot{\mathbf{x}}^*(t) = \nabla_p \mathcal{H}_c(\mathbf{x}^*, \mathbf{p}^*, u^*) \quad (5.19)$$

$$\dot{\mathbf{p}}^*(t) = -\nabla_x \mathcal{H}_c(\mathbf{x}^*, \mathbf{p}^*, u^*) \quad (5.20)$$

where the control Hamiltonian

$$\mathcal{H}_c = f(\mathbf{x}, u) \cdot \mathbf{p} - 1. \quad (5.21)$$

Furthermore, \mathcal{H}_c is maximal when evaluated in the optimal control function

$$\mathcal{H}_c(\mathbf{x}^*, \mathbf{p}^*, u^*) = \max_{u \in \mathcal{U}} \mathcal{H}_c(\mathbf{x}^*, \mathbf{p}^*, u), \quad (5.22)$$

and we have $\mathcal{H}_c(\mathbf{x}^*, \mathbf{p}^*, u^*) = 0$ for $0 \leq t \leq T_{\min}$.

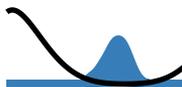
At first sight this formulation sounds very promising. However, the catch is that there is no guarantee that a given control actually connects the initial and final states. This must be insured in another way making this approach challenging and solutions are only known for a limited set of low-dimensional quantum systems such as two-level and three-level systems [52, 141–143].

5.2.1 Time-Optimal Control of the Landau-Zener Model

In this section we apply the time-optimal form of the Pontryagin's Maximum Principle to find the time-optimal solutions for the Landau-Zener model from Sec. 3.1.1. This will give insights into the connection between the optimal-time T_{\min} and the quantum speed limit T_{QSL} . The presentation in this section is based on Refs. [52, 160]. We have the equations of motion

$$|\dot{\psi}\rangle = -i \begin{pmatrix} \Delta(t) & \Omega_R \\ \Omega_R & -\Delta(t) \end{pmatrix} |\psi\rangle. \quad (5.23)$$

A technical point is that Pontryagin's Maximum Principle is formulated in terms of real variables necessitating to transform Eq. (5.23). It is also possible to formulate Pontryagin's Maximum Principle in terms of complex variables but this is not standard in the literature and it will



be presented here. In the next section, we give a brief discussion of how to find time-optimal controls for the Landau-Zener control problem. A convenient choice is to decompose the time-evolution operator in terms of three Eulerian rotation angles [161]. A proper definition of the angles gives

$$\hat{U}(t, 0) = e^{-i\sigma_z\tau_3(t)/2} e^{-i\sigma_1\tau_1(t)/2} e^{-i\sigma_y\tau_2(t)/2}. \quad (5.24)$$

The goal is then to design the τ_i functions such that they solve the state-to-state problem $\hat{U}(T, 0)|\psi_0\rangle = |\psi_t\rangle$ up to some arbitrary phase factor. Inserting this time-evolution operator into the Schrödinger equation and performing a number of manipulations results in a system with three equations for $\dot{\tau}_i(t)$,

$$\dot{\tau}_1(t) = 2\Omega_R \cos \tau_3(t), \quad (5.25a)$$

$$\dot{\tau}_2(t) = -2\Omega_R \frac{\sin \tau_3(t)}{\cos \tau_1(t)}, \quad (5.25b)$$

$$\dot{\tau}_3(t) = 2\Delta(t) + 2\Omega_R \frac{\sin \tau_3(t) \sin \tau_1(t)}{\cos \tau_1(t)}. \quad (5.25c)$$

After these manipulations we have transformed the Schrödinger equation with two time-dependent complex numbers into this set of equations with three time-dependent real numbers. Our quantum problem is now on a form where we directly can apply Pontryagin's maximum principle. The associated control Hamiltonian is

$$\mathcal{H}_c = \dot{\tau}_1 p_1 + \dot{\tau}_2 p_2 + \dot{\tau}_3 p_3 - 1. \quad (5.26)$$

We drop the star on the control and the costate \mathbf{p} . From Pontryagin's maximum principle we can calculate the equations for the costate,

$$\dot{p}_1 = 2\Omega_R p_2(t) \frac{\sin \tau_3 \sin \tau_1}{\cos^2 \tau_1} - 2\Omega_R p_3 \frac{\sin \tau_3}{\cos^2 \tau_1}, \quad (5.27a)$$

$$\dot{p}_2 = 0, \quad (5.27b)$$

$$\dot{p}_3 = 2\Omega_R \left(p_1 \sin \tau_3 + p_2 \frac{\cos \tau_3}{\cos^2 \tau_1} - \frac{\cos \tau_3 \sin \tau_3}{\cos \tau_1} \right). \quad (5.27c)$$

From the second equation we see that $p_2(t) = c_2$ is constant. The maximum principle Eq. (5.22) gives that the control Hamiltonian is zero, $\partial\mathcal{H}_c/\partial\Delta = 2p_3(t) = 0$ implying that $p_3(t) = 0$. This allows us to rewrite Eqs. (5.27a) and (5.27c) as

$$\dot{p}_1 = 2\Omega_R c_2 \frac{\sin \tau_3 \sin \tau_1}{\cos^2 \tau_1}, \quad (5.28)$$

$$0 = 2\Omega_R p_1 \sin \tau_3 + 2\Omega_R c_2 \frac{\cos \tau_3}{\cos \tau_1}. \quad (5.29)$$



We have $\mathcal{H}_c = 0$ for the optimal trajectory giving the equation,

$$2\Omega_R p_1 \cos \tau_3 - 2\Omega_R c_2 \frac{\sin \tau_3}{\cos \tau_1} = 1. \quad (5.30)$$

Multiplying Eq. (5.29) with $\cos \tau_3(t)$ and Eq. (5.30) with $\sin \tau_3$ and subtracting gives

$$\frac{2\Omega_R c_2}{\cos \tau_1} = -\sin \tau_1. \quad (5.31)$$

Solving Eqs. (5.29) and (5.30) for \dot{p}_1 gives

$$\dot{p}_1 = \frac{\Delta}{\Omega_R} \sin \tau_3(t) - \frac{\sin^2 \tau_3(t) \sin \tau_1(t)}{\cos \tau_1(t)}. \quad (5.32)$$

It is also possible to find another expression for \dot{p}_1 by plugging Eq. (5.31) into Eq. (5.28)

$$\dot{p}_1 = -\frac{\sin^2 \tau_3(t) \sin \tau_1(t)}{\cos \tau_1(t)}. \quad (5.33)$$

After all these manipulations we are now in a position to compare Eqs. (5.32) and (5.33) showing that $\Delta(t) \sin \tau_3(t) = 0$. The only options are $\Delta(t) = 0$ or $\sin \tau_3(t) = 0$. Assume $\sin \tau_3(t) = 0$ then $\dot{\tau}_3(t) = 0$ but then we see from Eq. (5.25c) that $\Delta = 0$ so both options are equivalent. This is an important result, it shows that for the Landau-Zener model the optimal control is *always* to have $\Delta(t) = 0$ except possibly at the boundary of the time interval. In Sec. 3.2.1, we found a similar result based on the optimality equations but here we assumed a constant pulse *a priori* unlike this result that does not have any restrictions on the pulse initially.

When $\Delta(t) = 0$ then the Landau-Zener Hamiltonian reduces to $\hat{H} = \Omega_R \sigma_x$ having the associate time evolution operator $\hat{U} = \exp(-i\Omega_R \sigma_x t)$. The action of this operator is illustrated in Fig. 5.1(a) on the Bloch sphere where it can be recognized as a rotation around the x -axis. However, as illustrated in Fig. 5.1(a) $\Delta(t) = 0$ does not transfer any initial state into any target state. This necessitates initial and final δ -like pulses with zero duration

$$\alpha_i = \lim_{\varepsilon \rightarrow 0} \int_0^\varepsilon \Delta_i(t) dt, \quad \alpha_f = \lim_{\varepsilon \rightarrow 0} \int_0^\varepsilon \Delta_f(t) dt. \quad (5.34)$$

Both of these pulses vanish in the limit where $|\Delta_{i,f}| \rightarrow \infty$. This gives us the following expression for the optimal time evolution operator

$$\hat{U}(t, 0) = e^{-i\alpha_f \sigma_z / 2} e^{-i\Omega_R T \sigma_x / 2} e^{-i\alpha_i \sigma_z / 2}. \quad (5.35)$$

Note, that since $\exp(i\pi \sigma_z) = -1$ we are free to add multiples of π to each of the α pulses. For a given initial and target state we now have to



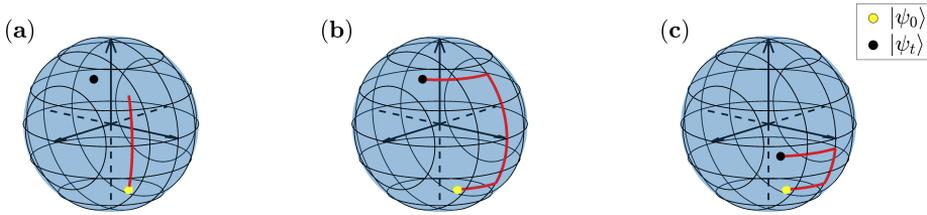


Figure 5.1: An illustration of the time-optimal trajectories on the Bloch sphere without constraints on the control. The initial (target) state is a yellow (black) dot. **(a)** The state transfer without the δ -like pulses that does not transfer from the initial state into the target state. **(b)** The time-optimal transfer from the initial state to the target state. **(c)** This the optimal transfer the different initial and target states. Note, that the optimal trajectory is not the shortest path (geodesic) on the Bloch sphere even for these closely laying states.

find the minimal T among all the possible values of α_i and α_f . For the specifics of this calculation we refer to Ref. [52]. The resulting minimal time is

$$T_{\min} = \frac{|\theta_t - \theta_i|}{2\Omega_R}, \quad (5.36)$$

This result has a nice interpretation on the Bloch sphere. We see from Eq. (5.35) that the optimal time evolution consists of three rotations around the z -axis, then x -axis and finally the z -axis again. This trajectory is shown in Fig. 5.1(b). It is also shown in Fig. 5.1(c) that the time-optimal trajectory between two points is not the geometrically shortest distance between the two points on the Bloch sphere.

5.2.2 Relation to the Quantum Speed Limit

It is particularly illuminating to relate the minimal time T_{\min} to the quantum speed limit T_{QSL} . The optimal control has $\Delta(t) = 0$ so for a general Bloch state with angles θ and φ Eq. (3.6) the energy variance is

$$\Delta E = \Omega_R \sqrt{\langle \sigma_x^2 \rangle - \langle \sigma_x \rangle^2} = \Omega_R \sqrt{1 - \sin^2 \theta \cos^2 \varphi}. \quad (5.37)$$

The optimal control rotates the state into the yz -plane where ΔE is maximal. The geometric quantum speed limit in Eq. (5.15) is proportional to ΔE whereby the optimal control can be understood as creating the largest possible ΔE in order to facilitate quick motion in Hilbert space. This is in accordance with the excitation and deexcitation pattern discussed



previously. In the special case of polar states $\theta_i = 0$ and $\theta_f = \pi$ the optimal time and the Mandelstamm-Tamm bound are the same

$$T_{\min} = \frac{\pi}{2\Omega_R} = \frac{\pi}{2\Delta E} = T_{\text{QSL}}. \quad (5.38)$$

This striking result was first observed in Ref. [140] where the minimal time in the Landau-Zener model was found numerically using Krotov's method. Due to this result the terms time-optimal and quantum speed limit are often used interchangeably perhaps due to the appeal of the term quantum speed limit. This is a bit unfortunate but it is standard terminology and I will follow this convention in the later sections.

It is important to note that $T_{\min} = T_{\text{QSL}}$ is only true for polar states. The reason for the equality in Eq. (5.38) is that the optimal control is a constant i.e. a time-independent Hamiltonian. For non-polar states the optimal Hamiltonian is time-dependent and in this case it is tempting to calculate T_{QSL} using Eq. (5.17) or Eq. (5.9) along a geodesic on the Bloch sphere as proposed in Ref. [140]. In both of these cases the equality in Eq. (5.38) is violated and for certain initial and final states it may even occur that $T_{\min} < T_{\text{QSL}}$. The problem is that the quantum speed limit and time-optimal control have two slightly different perspectives. The geometric quantum speed limit in Eq. (5.17) sets a lower bound on how quickly a system can follow a particular state trajectory $|\psi(t)\rangle$. When discussing the optimal time for a control problem we seek the fastest possible state trajectory $|\varphi(t)\rangle$ such that $|\varphi(0)\rangle = |\psi_0\rangle$ and $|\varphi(T_{\min})\rangle = |\psi_t\rangle$ up to a phase factor. In this sense, each state trajectory has its own quantum speed limit not related to the optimal time. In fact, it is possible to have one state trajectory saturating its own quantum speed limit while still taking longer time to reach the target state than another state trajectory. A simple analogy is to drive to a city along two different routes. Although you race at the speed limit along the first route, you may still get to the city faster along the second simply because it is shorter or it allows you to go even faster.

It is also possible to extend the result in Eq. (5.36) below the optimal time. Let $|\psi_F\rangle$ be a state with fidelity F , which is then a distance $\theta_F = 2 \arccos(\sqrt{F})$ away from the target state by Eq. (5.13). There exist a number of such states. As the detuning is unbounded, azimuthal rotations on the Bloch sphere are infinitely fast. This implies that the closest in time state has the form $|\psi_F\rangle = \cos \theta_t/2|0\rangle + e^{i\varphi_t} \sin \theta_t/2|1\rangle$. From Eq. (5.36) this state may be reached in time

$$T_{\min}^F = T_{\min} - \frac{\theta_F}{2\Omega_R}. \quad (5.39)$$



Rearranging this equation gives the important result

$$F = \cos^2\left(\frac{|\theta_i - \theta_f|}{2} - \Omega_R T_{\min}^F\right). \quad (5.40)$$

In the special case of polar states $\theta_i = 0$ and $\theta_f = \pi$ the fidelity scales as $\sin^2(\Omega_R T_{\min}^F)$. This \sin^2 -behavior has been reported as a “universal behavior” below the quantum speed limit (optimal time) in Refs. [47, 120, 162] based on numerical calculations. These considerations show that any effective two-level system always has a \sin^2 -behavior of the $F(T)$ -curve for polar states. The original Ref. [162] considered time-optimal control between two ground states across a quantum phase transition featuring a narrow gap in the level spectrum close to the critical point. If the dynamics was constrained to the two lowest energy levels then the system could be described by a Landau-Zener model close to the critical point and we should expect a \sin^2 -behavior. This may offer an explanation for their numerical observations. For systems with more than two levels we can neither confirm nor deny a \sin^2 -behavior. However, for some of our numerical examples below we do not observe it.

5.3 Time-Optimal Control of Bose-Einstein Condensates with Global-Local Optimization

In most cases, it is not possible to solve the time-optimal state-to-state control problem analytically using Pontryagin’s Maximum Principle. As an alternative one typically performs multiple numerical solutions of the state-to-state problem for different fixed durations T . This gives some numerical estimate of the quantum speed limit. A common observation is that the control problems are typically easy to solve for long durations but it becomes increasingly difficult to find initial controls or seeds that optimize to $F > 0.99$ for short durations. It has even been shown that given a controllable Hamiltonian and arbitrarily long durations then *any* seed will converge towards $F > 0.99$ [124, 163, 164]. This is also known as a trap-free or benign control problem whereas a control problem where seeds converge towards solutions below $F > 0.99$ is complex. There is a growing number of studies that support the complexity of quantum control problems in the vicinity of the quantum speed limit [165–168]. This demonstrates that local optimization may fail as the quantum speed limit is approached.

This can be understood in terms of the so-called fitness or optimization landscape, which is the graph for the (reduced) cost-functional $\hat{J}(u)$



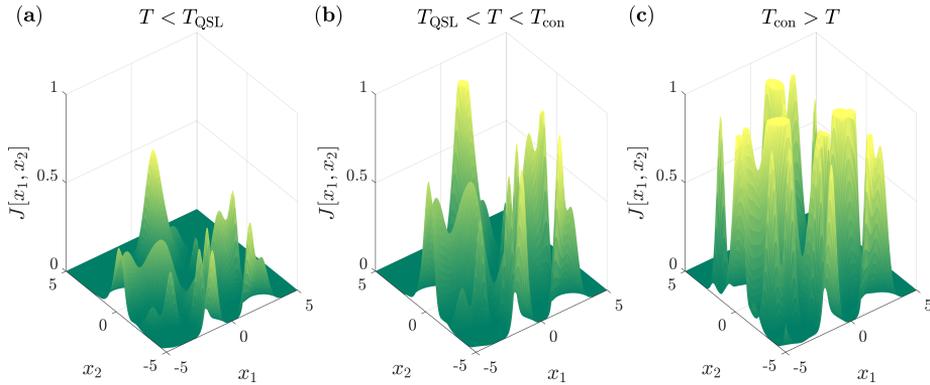
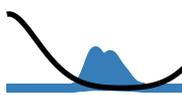


Figure 5.2: A graphical illustration of the Optimization Landscape Hypothesis using an artificial optimization landscape. **(a)** The optimization landscape far below the quantum speed limit where only suboptimal peaks remain. **(b)** The optimization landscape in the vicinity of the quantum speed limit, which is complex due to the low number of peaks with a high fidelity. **(c)** The optimization landscape at long durations that is benign since all maxima are also global maxima.

Eq. (3.36). Local or greedy optimization algorithms such as those presented in Chap. 3 climb to the most nearby optimum or peak in the optimization landscape. When speaking of the optimization landscape one typically refers to maximization rather than minimization. The control problem considerations may be collected in the *Optimization Landscape Hypothesis* that is illustrated in Fig. 5.2. This hypothesis states that above some controllability duration T_{con} the landscape is benign and trap free. As the duration is gradually decreased some controls lose their high fidelity and the landscape becomes complex. At the quantum speed limit only a few or a single optimal control remain. Below the quantum speed limit there are no high fidelity solutions and the landscape remains complex. We will return more to the discussion of the optimization landscape in Secs. 6.2-6.3.

Local optimization in a complex landscape might get trapped unable to reach the global optimum. An alternative is global optimization methods such as genetic algorithms that take stochastic steps in an attempt to avoid the traps, which typically comes at the cost of dramatically increased runtime compared with the local solvers. In computer science [169], machine-learning [170] and even social science [171], one typically refers to the balance between local and global optimization as



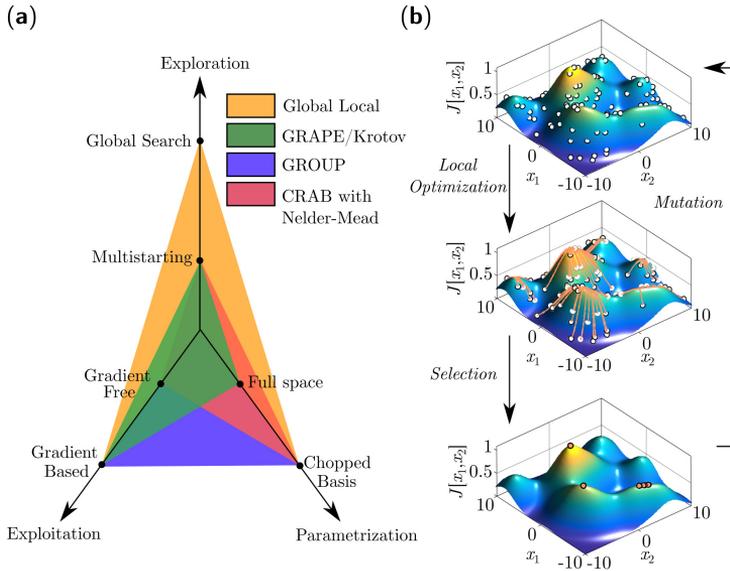
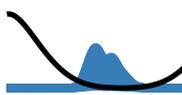


Figure 5.3: (a) The combined search has three main components namely exploration, exploitation and control parametrization. It is shown how other standard methods Gradient Ascent Pulse Engineering (GRAPE) [109], Krotov’s method [54, 55] and Nelder-Mead with Chopped Random Basis (CRAB) [104, 119, 120], are placed within this categorization. (b) The three main steps in Global-Local algorithm being *mutation*, *local optimization*, and *selection*. The landscape is an artificial mathematical problem being maximized rather than minimized for ease of visualization.

the *exploitation-exploration* trade-off. Exploration is a global search for new candidate solutions in sparsely probed parts of the control space, whereas exploitation is intense investigation of a small local portion of the control space enhancing the best solution [169]. As we have seen in the previous chapter it may also be beneficial to parametrize the space of admissible controls in a chopped basis, which eases the search for optimal solutions [5, 104, 117]. Within quantum control the most popular method for exploration is simply to restart the optimization at many different initial controls or points in the optimization landscape and hope to find satisfactory solutions. This method is known as multistarting [172] and the wide spread popularity may be attributed to the ease of finding good solutions at long durations. However, it may generally not be the most efficient or systematic way of adding exploration.

Here I propose a hybrid Global-Local algorithm that combines Dif-



ferential Evolution with local algorithms. This achieves a better balance between *exploration*, *exploitation* and *parameterization* as shown in Fig. 5.3(a). We apply this method to the control of Bose-Einstein condensates trapped on magnetic microtraps and observe improvements in the estimate of the quantum speed limit. We study the shape of the fidelity curve below the quantum speed limit and we do not observe a \sin^2 -behavior. This work therefore strengthens our doubts on universal behavior as already discussed at the end of Sec. 5.2.2. This indicates that further algorithmic advances are necessary when searching for the ultimate limits of time-optimal control.

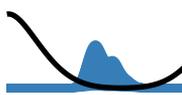
5.3.1 Global-Local Optimization

If a local optimization algorithm e.g. GROUP converges to a local sub-optimum (a control with $F < 1$) then the result may be improved by simply optimizing another initial control. This multistarting procedure has a constant probability of success given by how likely it is to randomly select a control that optimizes to $F = 1$. An alternative to multistarting is using a global optimization algorithm such as Particle Swarm Optimization, Simulated Annealing, Differential Evolution and Covariance Matrix Adaptation Evolutionary Strategy [173–176]. However, these are domain general algorithms that only focus on one of the three aspects in Fig. 5.3(a) lacking the analytic gradients (exploitation) and good problem parametrization. In order to add these aspects we use a combined Global-Local algorithm, where the global algorithm replaces the multistarting strategy. In principle, the global optimization could be done with any global optimization algorithm, but we choose to focus on Differential Evolution due to its good performance in quantum optimal control problems and general optimization contests [168, 175, 177]. Global evolutionary algorithms are also well established tools in closed loop optimal control of shaped laser pulses [178–182].

Differential Evolution updates a population of points (members) $P = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_N\}$ where each point is the vector of expansion coefficients in a chopped basis i.e. $\mathbf{c}_i = (c_1, c_2, \dots, c_M)$. We use the same type of expansion as in the last chapter,

$$u(t) = u_0(t) + S(t) \sum_{n=1}^M c_n f_n(t), \quad (5.41)$$

where M is the size of the basis and $f_n = \sin(n\pi t/T)$ are the basis functions multiplied by shape function $0 \leq S(t) \leq 1$. The main three steps of the algorithm, *mutation*, *local optimization*, and *selection* are



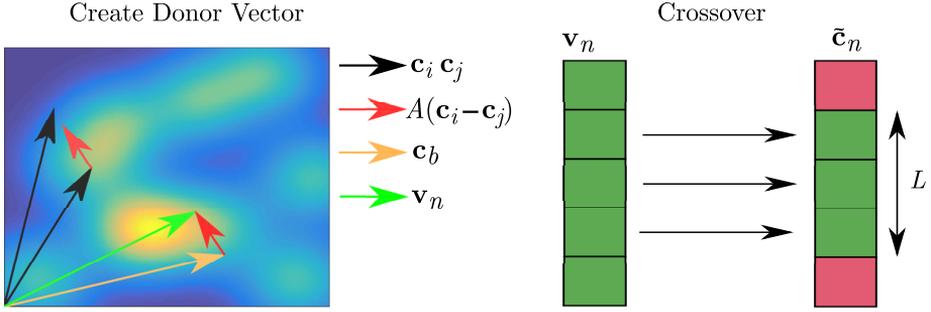
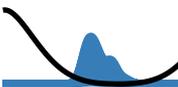


Figure 5.4: The mutation strategy from Differential Evolution. The left figure graphically illustrates the process for generating a single donor vector \mathbf{v}_n and the right figure shows the crossover step generating the trial vector $\tilde{\mathbf{c}}_n$.

illustrated in Fig. 5.3(b), which together constitutes a single generation. First a new trial population \tilde{P} is formed from the original population P using the mutation strategy from Differential Evolution, which is explained below. In the next step we locally optimize all the members in \tilde{P} using GROUP. Finally, in the last selection step the $\tilde{\mathbf{c}}_i$ member in \tilde{P} replaces the \mathbf{c}_i member in P if it has a lower cost. This procedure is iterated in a number of generations. Before discussing the results we give a brief account of Differential Evolution’s mutation strategy [175]. Differential evolution randomly selects two distinct members $\mathbf{c}_i, \mathbf{c}_j, i \neq j$ and the current best member \mathbf{c}_b . A donor vector is created from these three members $\mathbf{v}_n = \mathbf{c}_b + A(\mathbf{c}_i - \mathbf{c}_j)$ where A is a scaling factor. This is the donor step in Fig. 5.4. The $\tilde{\mathbf{c}}_n$ member in the trial population \tilde{P} is found by replacing L consecutive values from the original $\tilde{\mathbf{c}}_n$ with values from \mathbf{v}_n . The length of L is given by a Poissonian distribution with mean $Cr = 0.97$ and a minimum length of one. The starting point of this replacement is random. This is the crossover step in Fig. 5.4. In our simulations, A is linearly decreased from $A = 0.4$ to $A = 0.1$ over the generations in order to promote early exploration and later exploitation.

5.3.2 Results

We first applied Global-Local on two control problems relating to manipulation of Bose-Einstein condensates trapped on atom chips. As in the previous chapter the dynamics of the condensate are well described by



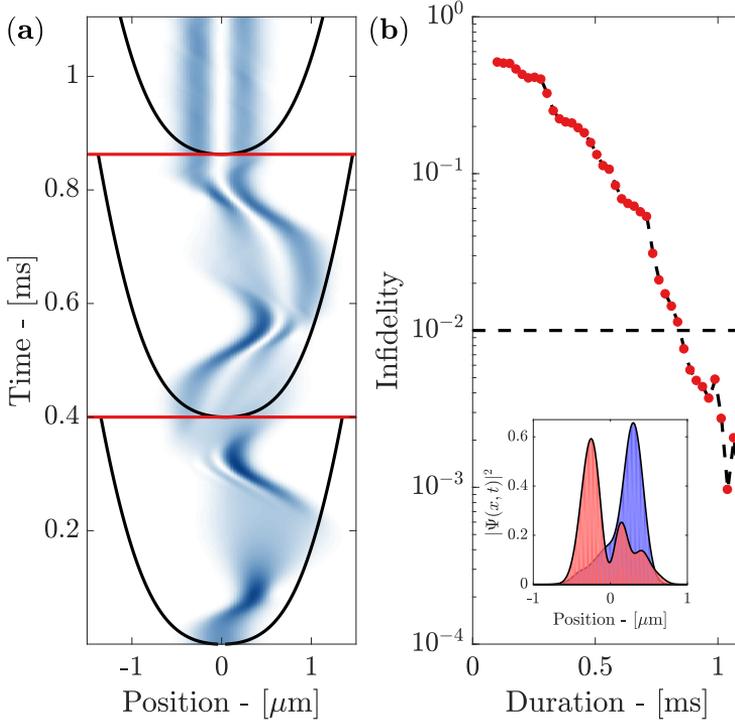


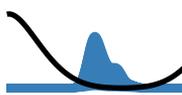
Figure 5.5: Results for condensate driving. **(a)** The density of the condensate ($|\psi(x, t)|^2$) when propagated along the best control at duration $T = 0.89\text{ms}$. Above the top line, the potential is held constant and the state is seen to be stationary. The associated potentials are drawn on top for $t = 0, t = 0.4T$ and $t = T$ where T is the process duration. **(b)** The best infidelity at different durations from the Global-Local algorithm shown using red dots. An insert shows the final density of the condensate at $T = 0.28\text{ms}$ (blue) and $T = 0.43\text{ms}$ (red) above and below the kink at $T = 0.30\text{ms}$.

an effective one-dimensional Gross-Pitaevskii equation,

$$i\frac{\partial\psi}{\partial t} = (\hat{H} + \beta|\psi|^2)\psi, \quad (5.42)$$

The first control problem is the *condensate driving* problem where a condensate must be excited from the ground state to the first excited state. The potential is given by

$$V(x, u(t)) = p_2(x - u(t))^2 + p_4(x - u(t))^4 + p_6(x - u(t))^6, \quad (5.43)$$



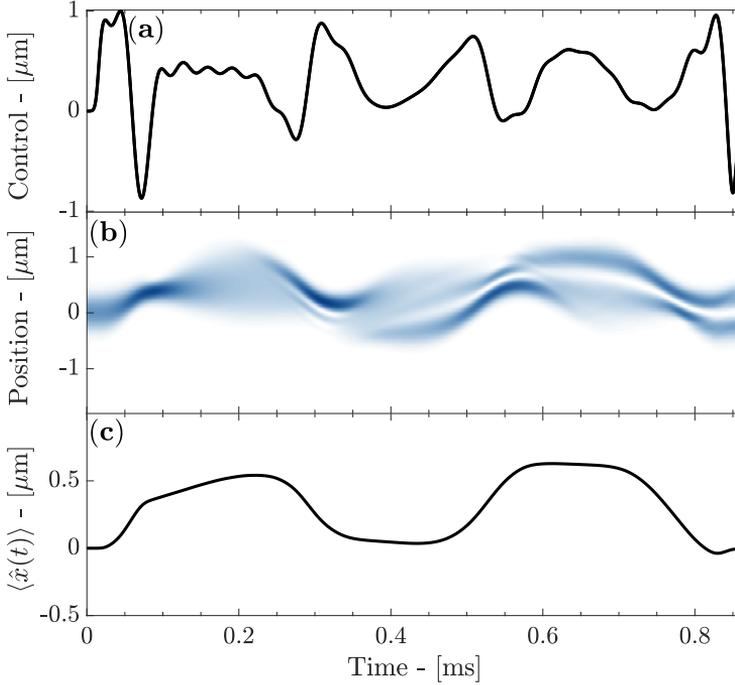


Figure 5.6: (a) The fastest optimal control for condensate driving found using Global-Local optimization. (b) The density for the condensate when propagated along the optimal control for comparison with the control. (c) The expectation value of the position as a function of time $\langle \hat{x}(t) \rangle$. This expectation value has a clear main oscillation component, which is close to the frequency difference between the single particle ground state and the first excited state.

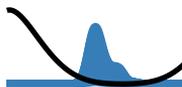
This problem was also studied in the last chapter and a detailed account is given in Sec. 4.1. The second problem is *condensate splitting* where a condensate is coherently split into two. The potential is created using the magnetically dressed states from Sec. 2.3. The potential is given by

$$V(\mathbf{r}) = m'_F g_F \sqrt{\Delta^2(\mathbf{r}) + \Omega^2(\mathbf{r})}, \quad (5.44)$$

with

$$\Delta(\mathbf{r}) = |\mathbf{B}_S(\mathbf{r})| - \frac{\omega}{|g_F \mu_B|}, \quad \Omega(\mathbf{r}) = \frac{B_{\text{RF}} B_z}{2|\mathbf{B}_S(\mathbf{r})|}, \quad (5.45)$$

where \mathbf{B}_S is the static Ioffe field with z -component B_z . The radio frequency field has magnitude B_{RF} and frequency $\omega = 625$ kHz. $g_F = 1/2$ is the Landé g -factor and μ_B is the Bohr magneton. In the experimental



realization. Here we discuss the specific experimental procedure outlined in Refs. [46, 74, 110] where $\beta = 2\pi\hbar \cdot 500\text{Hz} \mu\text{m}$. The magnitude of the radio frequency field can be adjusted as $B_{\text{RF}} = (0.5 + 0.3u(t))\text{G}$. For $u = 0$ the potential is a single well and it dynamically changes into a double well as u changes into $u = 1$, which is displayed in Fig. 2.4(b). The initial and target states are the ground states for $u = 0$ and $u = 1$ respectively.

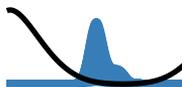
Condensate Driving

Initially I discuss the results from applying Global-Local on the *condensate driving* control problem. I perform the optimizations over a range of durations from $T = 0.1\text{ms}$ to $T = 1.09\text{ms}$ and the resulting best infidelity is shown in Fig. 5.5(b). A numerical upper bound on the quantum speed limit is $T_{\text{QSL}}^{\text{num}} = 0.86\text{ms}$, which is the shortest duration with $F \geq 0.99$. This result is faster than the 1.09ms reported in Ref. [47] where traditional multistarting and gradient-free optimization was used. It is also an improvement over Refs. [81, 110, 129] that used other gradient-based methods.

In Ref. [47] a double \sin^2 -behavior is found, which was interpreted to indicate that the solutions had mapped out the true quantum speed limit. Unexpectedly, our Global-Local does not follow a \sin^2 -behavior, indicating that a \sin^2 -behavior does not always imply that the numerical results have identified the quantum speed limit. Following the discussion at the end of Sec. 5.2.2, we attribute this to the fact that our optimal control pulse cannot be described with an effective two-level system. Note, that we cannot exclude the existence of even better solutions and therefore we cannot exclude the existence of a better $F(T)$ -curve with a \sin^2 -behavior.

The $F(T)$ -curve has a distinct kink around $T = 0.30\text{ms}$. For durations below $T = 0.30\text{ms}$ the final states are more reminiscent of a displaced Gaussian than a double peak structure. Above $T = 0.30\text{ms}$ the final state is more similar to a partial transfer into the first excited state as shown in the insert of Fig. 5.5(b). For short durations the displacement offers higher infidelities but it is limited at higher durations whereas the opposite is true for the partial transfer. We interpret the kink as the optimal limit between these two processes. A similar kink was observed in Ref. [47].

The optimal control $u(t)$ and the resulting $\langle \hat{x}(t) \rangle$ are shown in Fig. 5.6. The control is highly complex with many frequency components unlike $\langle \hat{x}(t) \rangle$ that clearly has a main dominant frequency component. A Fourier transformation of $\langle \hat{x}(t) \rangle$ shows that this oscillation's frequency is close to



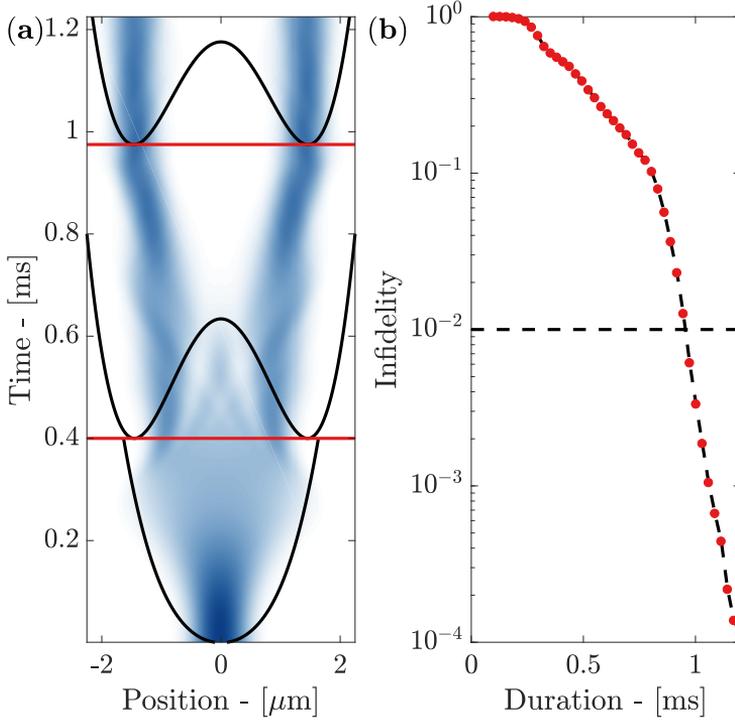
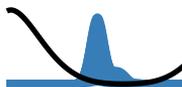


Figure 5.7: Results for condensate splitting. **(a)** The condensate density ($|\psi(x,t)|^2$) along the best control at a duration of $T = 0.97\text{ms}$. The potential is held constant above the top line, which shows that the state is in an eigenstate. The potential is shown for durations $t = 0, t = 0.4T$ and $t = T$. **(b)** Infidelity shown for different durations obtained using the Global-Local algorithm displayed with red dots.

the energy difference between the single particle ground state and the first excited state. Hence, the resulting control may partially be understood as resonant driving.

Condensate Splitting

In condensate splitting we applied Global-Local over a range of durations from $T = 0.1\text{ms}$ to $T = 1.2\text{ms}$ with the results being displayed in Fig. 5.7(b). The resulting numerical upper bound on the condensate splitting is $T_{\text{QSL}}^{\text{num}} = 0.97\text{ms}$. To our knowledge there is no $F(T)$ -curve in the literature for comparison but our results are faster than the 2.0ms reported in Ref. [46, 110]. Neither in the case of condensate splitting do we find a \sin^2 -behavior as seen in Fig. 5.7(b).



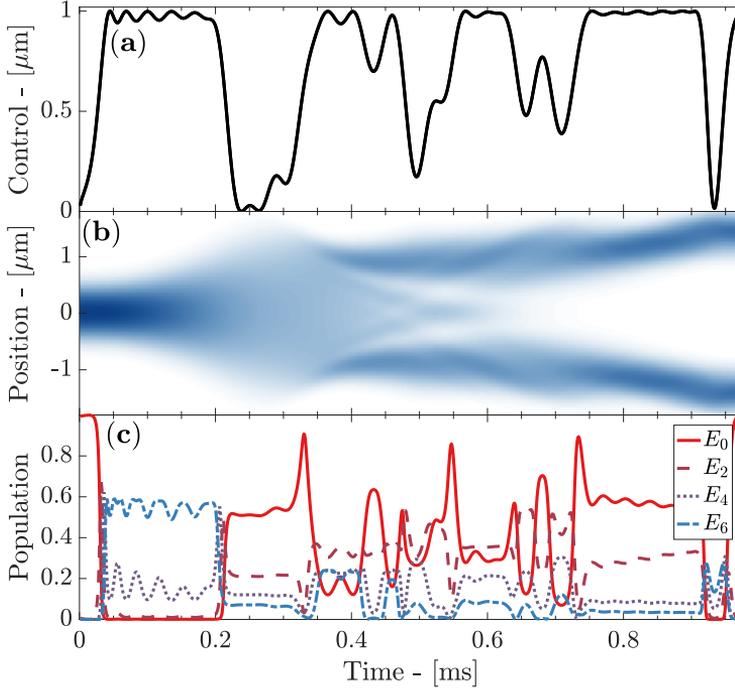
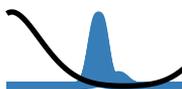


Figure 5.8: **(a)** The fastest optimal control for condensate splitting found using Global-Local optimization. **(b)** The condensate’s density when propagated along the control. **(c)** The population in the instantaneous energy eigenstates as a function of time. Only the even levels are populated due to the symmetry of the initial state and the potential.

The optimal control curve is displayed in Fig. 5.8(a). The control spends a considerable amount of time with values close to minimally and maximally allowed values $u(t) = 0$ and $u(t) = 1$. This is similar to bang-bang type solutions that within quantum optimal control are typically found using purely analytical methods [52, 160, 183].

In Sec. 5.1, we introduced the concept of the geometric quantum speed limit that states the instantaneous rate of change in the Wooters’ distance $\theta = 2 \arccos(\sqrt{F})$ is $d\theta/dt = 2\Delta E(t)$ [Eq. (5.15)] along some predefined state trajectory. We also discussed how one should expect a pattern of initial excitation and then refocusing into the ground state when moving quickly between two ground states. This pattern can be observed in Fig. 5.7(c) that plots the population in the instantaneous linear energy eigenstates. Note that only the even states are populated since the potential is invariant under parity $x \rightarrow -x$ for any values of $u(t)$.



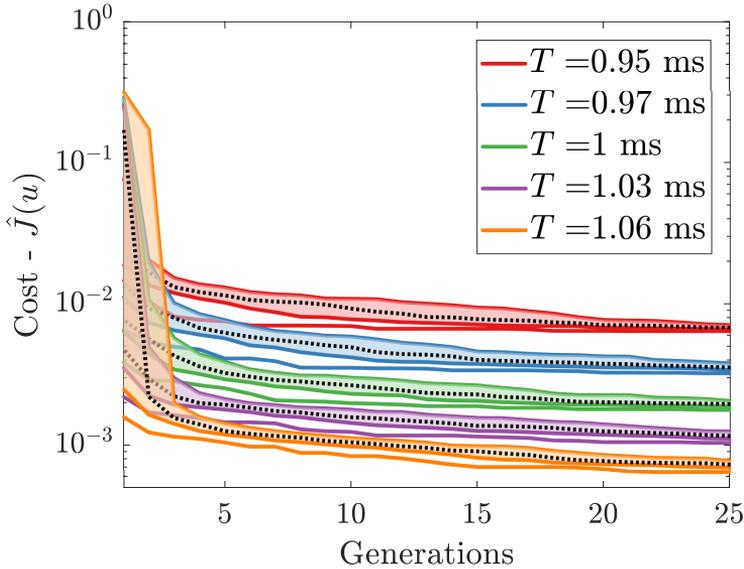


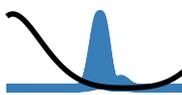
Figure 5.9: Learning in the Global-Local algorithm in condensate splitting at different durations (see legend). The cost is shown as a function of the generations for the population. The dotted line is the median infidelity in the population and the shaded area shows the 25%- and 75%-quartiles. The solid line shows the lowest cost in each generation.

The initial pulse with $u(t) \simeq 1$ excites the condensate by applying the maximally allowed splitting for the first $1/5$ of the control duration. The importance of control constraints on the quantum speed limit is especially clear here, since the excitation process could be completed faster if a larger double-well splitting was allowed. After an intermediate period there is a drop in the control just below $T = 1.0\text{ms}$ removing remaining excitations.

This optimal control is highly diabatic and clearly differs from adiabatically inspired solutions where the splitting is gradually turned on. The control found here is quite different from that in e.g. Ref [110] highlighting Global-Local’s ability to probe the control landscape far from the adiabatic regime.

5.3.3 Learning

Finally we discuss the learning in the Global-Local optimization. The distribution of cost values within each generation is shown as a function of generations for condensate splitting at a number of different durations



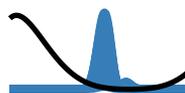
in Fig. 5.9. In each generation there is a total of $N = 100$ members. The figure shows that the Global-Local algorithm gradually decreases the median infidelity and thus learns a better solution strategy. The initial population is wide and it narrows as a function of generations. This shows that the optimization does an early exploration phase and subsequently spends successive generations on refining the current best members. However, this plot also suggest that the learning in algorithm should be considerably improved. Around the quantum speed limit, we observe a factor of two reduction in the cost between the final and initial distribution, indicating that most of results are due to the local optimization. With these early results, we also cannot rule out that with sufficient fine-tuning of the local optimization algorithms it might be possible to achieve similar results. The relative merit of improving local and global search methodologies represents an interesting avenue of future research.

5.4 Conclusion

In this section, I have introduced the time-optimal state-to-state problem. I have discussed the connection between the optimal time T_{\min} and quantum speed limits, which capture the physical principle that rapid motion in Hilbert space can only be achieved with a large energy E or energy variance ΔE with respect to the instantaneous Hamiltonian. In this sense, quantum speed limits may be understood as a refinement of Heisenberg's time-energy uncertainty relation.

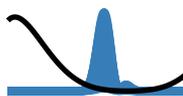
Later, I discussed how time-optimal control problems can be solved analytically using Pontryagin's Maximum Principle. I presented the analytic solutions to the Landau-Zener problem from Ref. [52, 160], which revealed more connections between the optimal time and the quantum speed limit.

However, in most cases an analytic approach is not feasible. When numerically optimizing for the quantum speed limit the quantum control problem may loose its benign properties and become complex as illustrated in the Optimization Landscape Hypothesis. In complex optimization problems it is important to balance exploration-exploitation trade-off and as we have seen in Chap. 4 it may also be beneficial to use a good problem parametrization. In this chapter, I proposed to use a combination of global and local optimization to give a better exploration-exploitation balance than the preferred multistarting approach. I applied this method to a set of control problems relating to manipulation of Bose-Einstein condensates, where we observed improvements to existing estimates of



the quantum speed limit.

The combined optimization scheme is directly applicable to other problems in quantum control where it might also give improvements in the quantum speed limit. However, the current rate of learning is limited and it should be improved. One possibility is to modify the Differential Evolution algorithm or using another explorative algorithm such as Covariance Matrix Adaptation Evolutionary Strategy [176]. Within Differential Evolution, exploration can be promoted by changing the scheme for the generation of the donor vector using a method like SaDE [175]. Finally, it would also be very interesting to do a more in depth comparison with multistarting in order to better quantify the advantages and disadvantages of Global-Local methods. At the time of writing, I am exploring these details and hope to present more results on this in the near future.



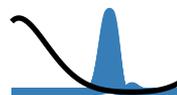
6 Optimal Control of Single Atoms

The introduction is based on Ref. [7]. Most of Sec. 6.1 is copied from my part A. Sec. 6.2 and the first part of Sec. 6.3 is copied from Ref. [3] with some modifications to the text in order to give a cohesive presentation. Secs. 6.3.1-6.4 are copied from Ref. [7]. The chapter's conclusion is partially based on the conclusions from Refs. [3, 7].

In modern scientific research, high-tech applications such as quantum computation [184] require exquisite levels of control while taking into account increasingly complex environmental interactions [185]. This necessitates continual development of optimization methodologies used in quantum control. Both the GROUP and Global-Local methods are steps in this direction. In the last chapter, I briefly introduced the concept of the optimization landscape as the graph for the cost-function [186], which constitutes a unifying mathematical framework for search problems from natural [124, 187–189] to social sciences [190, 191].

As discussed in the last chapter, search in a landscape can be approached with local or global optimization methods. Local solvers such as GRAPE, Krotov's method and GROUP are efficient and analogous to greedy hill climbers. In complex landscapes, however, they might get trapped locally and cannot reach the global optimum. The global methods attempt to escape these traps by taking larger stochastic steps typically at the cost of major increases in the runtime. Achieving the proper balance between these types of searches is the exploration-exploitation trade-off [170, 171].

Much effort in computer science is therefore focused on developing algorithms that exploit the topology of the landscape to adapt search strategies and make better-informed jumps [188, 192]. Machine learning algorithms have achieved success across numerous domains. However,



among researchers pursuing truly domain-general artificial intelligence, there is a growing call to rely on insights from human behavior and psychology [193, 194]. Thus, emphasis is currently shifting towards the development of human-machine hybrid intelligence [195, 196].

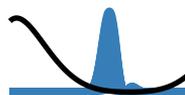
Recruiting potentially thousands of laymen in a science project is known as *citizen science* [197]. This emerging field provides a promising way to investigate and harness the unique problem-solving abilities humans possess [198]. In recent years the creativity and intuition of non-experts using gamified interfaces have enabled scientific contributions across different fields such as quantum physics [199], astrophysics [200] and computational biology [201–203]. Here, citizen scientists often seem to jump across very rugged landscapes and solve complex optimization problems efficiently using search methodologies that are difficult to quantify and encode in a computer algorithm.

In this chapter, we investigate if i) citizen scientists, provided with an appropriate interface, can help solve complex quantum control problems and ii) how they explore the optimization landscape. In order to present the optimization problem in an approachable manner by non-experts we turn the optimization problem into a game. In this chapter, we discuss our game Quantum Moves that gamifies quantum control problems relating to control of single atoms. At the time of writing more than 250.000 unique users have played our game.

Quantum Moves is part of the *scienceathome.org* project that attempts to bring citizen science to a new level with massive online citizen science games involving not only quantum physics research problems but also social sciences. This project has goals as broad as open characterization of user behavior, supplying simulation tools, and teaching.

This chapter is organized as follows: Sec. 6.1 introduces the physical motivation for the Quantum Moves game and in Sec. 6.2 a particular optimal control problem, or level, from Quantum Moves is analyzed in detail. In Sec. 6.3 I discuss the optimal control landscape in the Quantum Moves level using visualization techniques from machine learning. These insights and techniques are later extended to the experimental problem of Bose-Einstein condensate production in Sec. 6.4. The chapter ends with discussing a critique of our Quantum Moves analysis and a conclusion in Secs. 6.5-6.6.

The content of Sec. 6.2 was published in Ref. [3] with me as first author and I therefore contributed to the entire project. The content of Sec. 6.3 was published in Ref. [7] with me as co-author where I was a central contributor. Ref. [7] has a primary focus on citizen scientists optimizing trap geometries for Bose-Einstein condensation, and Sec. 6.3 focuses on my main contribution to Ref. [7] about optimal control landscapes in a



Quantum Moves problem.

6.1 Basic Architecture

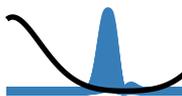
Quantum mechanics holds the promise of large advances in the field of computations using the quantum computer. In a quantum computer the fundamental bits consist of two-level systems or qubits, which can be in a superposition of $|0\rangle$ and $|1\rangle$ that unlike a classical bit must be either 0 or 1 [204]. This leads to enormous speedups for problems such as prime number factorization [22], searching in lists [205] and solving systems of linear equations [206]. In order to ensure scalable operations, the quantum operations must be executed almost perfectly with fidelities above a given threshold F_{th} necessary for quantum error correction. In general F_{th} will depend on the particular error source and correction scheme [185]. In this work we will apply $F_{\text{th}} = 0.999$. There exists several possible architectures for the realization of a quantum computer with promising candidates being superconducting systems [33], but here we discuss an alternative proposal based on neutral atoms from Ref. [207]. Ultracold atomic systems are a prime candidate for quantum simulation [26] meaning that investigating the limits of controllability in these systems is warranted.

Recall from the discussion in Sec. 2.1 that an atom in a detuned light field is affected by the conservative potential $V_{\text{dip}} \propto I(\mathbf{r})$ where $I(\mathbf{r})$ is the intensity of the light field. As explained in Sec. 2.2, two counterpropagating laser beams form a standing wave of light that creates a potential of the form $V(x) \propto \cos(kx)^2$ with a spacing of $\lambda/2$. In a deep lattice, a single atom will occupy each potential well within the lattice, which is the so-called Mott insulator state. Each of these atoms may be used as a qubit by addressing the internal hyperfine states of rubidium-87

$$|0\rangle = |F = 1, m_F = 1\rangle, \quad |1\rangle = |F = 2, m_F = 2\rangle. \quad (6.1)$$

Quantum computation requires both single and two qubit operations [206]. The single qubit operations can be performed using direct microwave transitions with a speed of $30 \mu\text{s}$ and infidelity $1.4 \cdot 10^{-3}$ [208]. As proposed in Ref. [77] the two qubit operations can be performed using controlled ultracold collisions between the atoms. This requires the ability to shuttle atoms from one site to another where the entangling collision may be performed. The atoms can be shuttled using optical tweezers, which consist of off-resonant Gaussian light beams. We focus on the dynamics along the x -axis with potential

$$V = V_0(t) \exp\left(-\frac{2(x - x_0(t))^2}{w_0^2}\right), \quad (6.2)$$



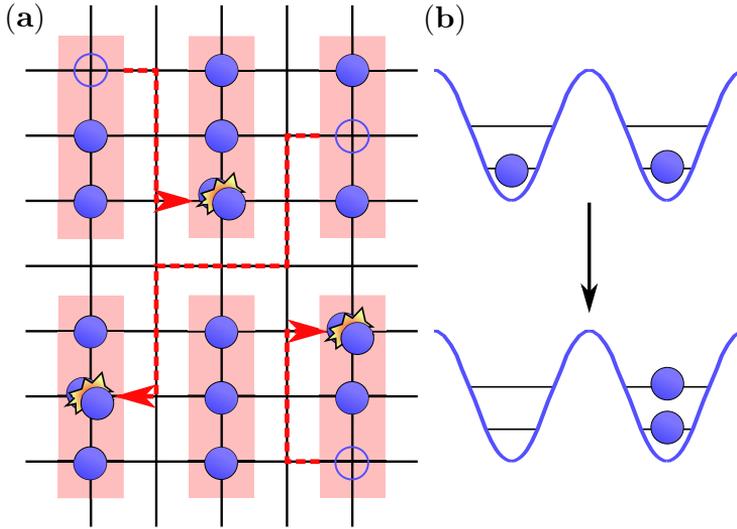


Figure 6.1: **(a)** The proposed lattice seen from above where every second column and one central row has been depopulated, which allows unhindered transport of the atoms between the different sites. The qubit gates are realized through collisional interactions. The blue lines show the trajectories of optical tweezers moving the atoms. **(b)** Top (bottom) is initial (final) state in the gate process.

where $V_0(t)$ is the amplitude of the beam and $x_0(t)$ is the position and w_0 is the beam waist.

The shuttling requires space between the atoms. This is created starting from a single plane Mott insulator by removing every second column and one row that allows for moving the atoms between different sites - see Fig. 6.1(a). As seen in Fig. 6.1(a) multiple qubit operations may be performed in parallel using several optical tweezers.

A collision based entangling gate can be executed on two atoms that are nearby. This is a $\sqrt{\text{SWAP}}$ gate performing the operation

$$|0\rangle|0\rangle \rightarrow e^{-i\pi/4}|0\rangle|0\rangle, \quad |0\rangle|1\rangle \rightarrow \frac{1}{\sqrt{2}}\left(|0\rangle|1\rangle - i|1\rangle|0\rangle\right), \quad (6.3a)$$

$$|1\rangle|1\rangle \rightarrow e^{-i\pi/4}|1\rangle|1\rangle, \quad |1\rangle|0\rangle \rightarrow \frac{1}{\sqrt{2}}\left(-i|0\rangle|1\rangle + |1\rangle|0\rangle\right). \quad (6.3b)$$

This operation is achieved by merging two atoms into the same well with one atom in the vibrational ground state and the other in the vibrational first excited state - see Fig. 6.1(b). When the two atoms are in the collective well, the so-called exchange interaction causes the population



of hyperfine states to evolve in time. If the atoms are allowed to interact for the appropriate amount of time a $\sqrt{\text{SWAP}}$ gate is realized.

Let us discuss this collision gate in more detail. The interaction is modeled by a contact interaction $V(\mathbf{r}_1 - \mathbf{r}_2) = g\delta^3(\mathbf{r}_1 - \mathbf{r}_2)$ as described in Sec. 2.4.2. In the merged well the computational basis is given by the states $|0_g, 0_e\rangle, |0_g, 1_e\rangle, |1_g, 0_e\rangle, |1_g, 1_e\rangle$ where the subscript refers to ground state and first excited states of the two atoms. These states are not the energy eigenstates, which are instead the symmetrized singlet and triplet states

$$|\Psi_S\rangle = \varphi_S(\mathbf{r}_1, \mathbf{r}_2)|S\rangle = \frac{1}{\sqrt{2}}\left(|0_g, 1_e\rangle - |1_g, 0_e\rangle\right), \quad (6.4a)$$

$$|\Psi_T^0\rangle = \varphi_T(\mathbf{r}_1, \mathbf{r}_2)|T^0\rangle = \frac{1}{\sqrt{2}}\left(|0_g, 1_e\rangle + |1_g, 0_e\rangle\right), \quad (6.4b)$$

$$|\Psi_T^+\rangle = \varphi_T(\mathbf{r}_1, \mathbf{r}_2)|T^+\rangle = |1_g, 1_e\rangle, \quad (6.4c)$$

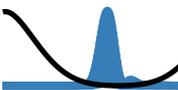
$$|\Psi_T^-\rangle = \varphi_T(\mathbf{r}_1, \mathbf{r}_2)|T^-\rangle = |0_g, 0_e\rangle, \quad (6.4d)$$

where $\varphi_T(\mathbf{r}_1, \mathbf{r}_2)$ and $\varphi_S(\mathbf{r}_1, \mathbf{r}_2)$ are the symmetric and antisymmetric combination of a single atom in the ground state and the first excited state. In the limit of vanishing interactions these states are $\varphi_{S/T}(\mathbf{r}_1, \mathbf{r}_2) \propto \varphi_e(\mathbf{r}_1)\varphi_g(\mathbf{r}_2) \pm \varphi_g(\mathbf{r}_1)\varphi_e(\mathbf{r}_2)$. The other states are the singlet and triplet states of the internal spin degree of freedom $|S\rangle = (|0\rangle|1\rangle - |1\rangle|0\rangle)/\sqrt{2}$, $|T^0\rangle = (|0\rangle|1_e\rangle + |1\rangle|0\rangle)/\sqrt{2}$, $|T^+\rangle = |1\rangle|1\rangle$, and $|T^-\rangle = |0\rangle|0\rangle$. Let $E_{S/T}$ be the eigenenergies of the wave functions $\varphi_{S/T}(\mathbf{r}_1, \mathbf{r}_2)$. These states are split by the positive difference $\Delta E = E_T - E_S$. The antisymmetric states are zero at $\mathbf{r}_1 = \mathbf{r}_2$ causing their energy to be unaffected by the interaction unlike the symmetric states whose energy are shifted by the point like interaction. Let us consider the initial state $|\psi(t=0)\rangle = |0_g, 1_e\rangle = (|\Psi_S\rangle + |\Psi_T^0\rangle)/\sqrt{2}$. Shifting this state by phase $\exp(-i(E_S + E_T)t/2)$ gives the time evolution

$$\begin{aligned} |\psi(t)\rangle &= \frac{1}{\sqrt{2}}\left(e^{-\frac{i\Delta Et}{2}}|\Psi_T^0\rangle + e^{i\frac{i\Delta Et}{2}}|\Psi_S\rangle\right) \\ &= \cos\left(\frac{\Delta Et}{2}\right)|0_g, 1_e\rangle - i\sin\left(\frac{\Delta Et}{2}\right)|1_g, 0_e\rangle. \end{aligned} \quad (6.5)$$

At time $t = T_{\sqrt{\text{SWAP}}} = \pi/(2\Delta E)$ this realizes one of the $\sqrt{\text{SWAP}}$ operations from Eq. (6.3a). Note, at $t = 2T_{\sqrt{\text{SWAP}}}$ the two atoms swap from ground state to excited state. The name $\sqrt{\text{SWAP}}$ comes from the fact that it is “halfway” to a full swap. With similar arguments one may be convinced that $T_{\sqrt{\text{SWAP}}}$ realizes a $\sqrt{\text{SWAP}}$ for all the states in Eqs. (6.3a)-(6.3b).

The architecture described here requires optimal control for two tasks.



Algorithm 5 Telescoping Optimization

- 1: Initialize a control $\mathbf{u}(t)$ at $T = T_1$.
 - 2: Select a $0 < \alpha < 1$.
 - 3: **while** $T > T_{\min}$ **do**
 - 4: Replace $\mathbf{u}(t)$ by the optimized control.
 - 5: Contract $\mathbf{u}(t)$ in time by $\mathbf{u}(t) \leftarrow \mathbf{u}(\alpha t)$.
 - 6: $T \leftarrow \alpha T$.
-

T.1 Efficient pick up and shuttling of the atoms.

T.2 Efficient merger of two wells such that the final state is a superposition of ground and excited state (Fig. 6.1(b)).

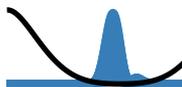
A key challenge is that optimal solutions with $F \geq 0.99$ is not sufficient. We seek to execute the operations as fast as possible due to unavoidable decoherence effects. Typically, it is desired to perform multiple gate operations, so the duration should be significantly shorter than the decoherence time. Therefore, we are really attempting to solve the time-optimal state-to-state control problem as introduced in Chap. 5. In this thesis we will only be concerned with the first of these two tasks. At the time of writing, I am involved in a collaboration with younger students on the second task where we have applied GROUP optimization and the early results are encouraging.

6.2 Bring Home Water

BringHomeWater is an optimal control problem addressing the pickup and shuttling task **T.1**. Here an initially empty controllable optical tweezer must be moved to collect an atom from another static tweezer. Hereafter the atom and tweezer must be brought back to a region around the starting point known as the target area. At the end of the protocol the atom must be in the ground state - see Fig. 6.2(a)-(c) for an graphical illustration of the problem. In this section we use units $\hbar = m = \lambda/2 = 1$ where time and energy are measured in units of $t_u = m\lambda^2/(4\hbar) \simeq 0.38\text{ms}$ and $E_u = \hbar^2 k^2/m\pi^2 = 2E_r/\pi^2$. In terms of these units the potential is

$$V(x, \mathbf{u}(t)) = -V_s \exp\left(-\frac{2(x-x_s)^2}{w_0^2}\right) - u_2(t) \exp\left(-\frac{2(x-u_1(t))^2}{w_0^2}\right), \quad (6.6)$$

where the first term is the static tweezer and the second term is the controllable tweezer. Note, that the control is two-dimensional in this problem.



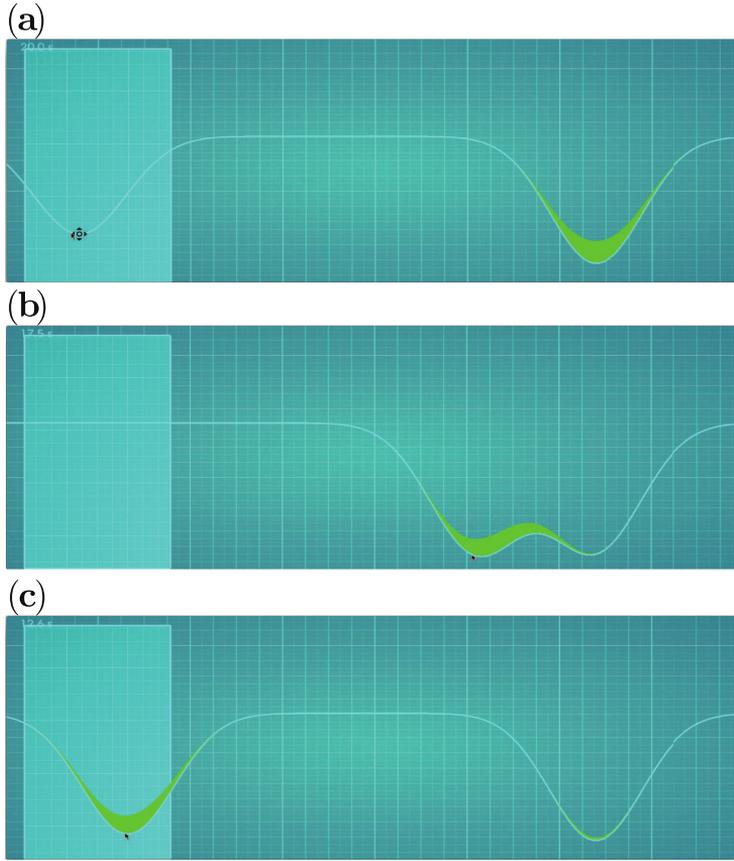
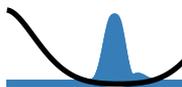


Figure 6.2: The BringHomeWater challenge as seen by the player. The atom is represented by the square of its wave function, $|\psi(x, t)|^2$, shown as a green liquid. The blue curve represents the potential felt by the atom. **(b)** The controllable tweezer is initially on the left and the atom is trapped in the right static potential. The player controls the optical tweezer by moving a computer cursor, picks up the atom **(b)**, and drags it back to the target area **(c)**, marked by a cyan rectangle, to collect points.

Initially we attempted to solve this problem using Krotov's method. We optimized 2,400 seeds at $T = 0.4$ of the form

$$u_n(k\Delta t) = w_n(k\Delta t) + \sum_{n=0}^{N-1} X[n] \sin\left(\frac{n\pi k}{N}\right), \quad (6.7)$$

where Δt is the time discretization, N is the number of time slices, and k is the index for the discretized control. The values of $X[n]$ were chosen stochastically and $\mathbf{w}(t)$ is a motion towards the atom and back into the

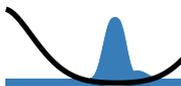


target area at constant speed and amplitude. These optimizations were performed at a duration of $T = 0.40$. In order to obtain an estimate of the quantum speed limit we performed a telescoping optimization, which iteratively optimizes and shortens a control in time - see Algorithm 5 for details. The telescoping optimization was performed for all the 2,400 seeds and repeated until a duration of $T_{\min} = 0.07$ in steps of Δt . The Krotov optimization ran for 800 iterations to ensure proper convergence. The fastest solutions located with this approach was at $T_{\text{QSL}}^{\text{num}} = 0.286$. The individual telescoping optimizations and best result are shown in Fig. 6.3.

Examination of our solutions lead us to suspect that better solutions existed. As an unconventional approach we created the computer game Quantum Moves, which enables citizen scientists to help solve quantum optimal control problems such as BringHomeWater. The player view of BringHomeWater is shown in Fig. 6.2(a)-(c) where the potential created by the tweezer is shown. The player controls the position and amplitude of the tweezer with the cursor. When the player moves the cursor the potential changes and the wave function is updated according to the time-dependent Schrödinger equation. When the player releases the cursor a score is calculated based on the final fidelity between the resulting wave function and the target state. We encourage the players to seek for the quantum speed limit by penalizing long durations in the game score system. The players were trained in a series of introductory levels before reaching BringHomeWater. In Fig. 6.3(a) we see that the players trace out a region very close to $T_{\text{QSL}}^{\text{num}}$, despite the fact that the computer optimizations used approximately 100,000 more function evaluations than the players. Some players do find faster solutions than the computer but with imperfect fidelity. This lead us to to introduce a hybrid computer-human optimization, where the players' solutions are used as the starting point (seed) for the Krotov optimization. We applied a telescoping optimization to about 70% of the player solutions with a duration shorter than $T = 0.40$, these results are shown in Fig. 6.3(b). The optimized player solutions provide a refined estimate of the quantum speed limit at $T_{\text{QSL}}^{\text{num}} = 0.20$. In Fig. 6.3(b) the optimized player solutions solutions are seen to cluster in bands.

To better understand how the optimized player solutions outperformed the pure computer optimization we investigated the clustering of the optimized player solutions using machine learning techniques. We introduce a distance measure between two solutions as

$$D_{jk} = \frac{1}{T} \int_0^T \langle f_{jk} | f_{jk} \rangle dt, \quad (6.8)$$



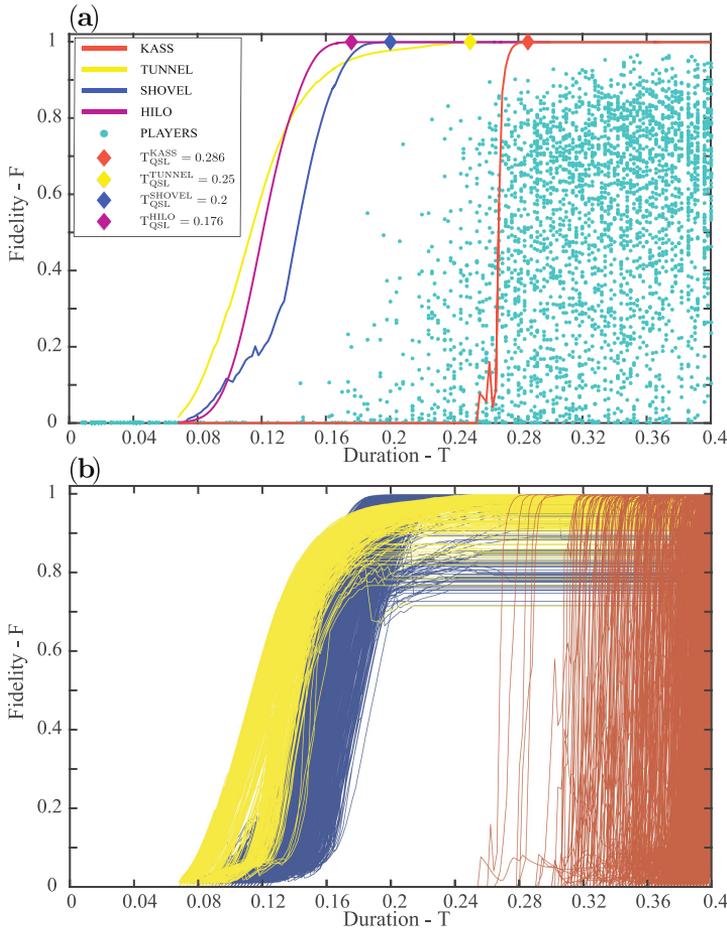
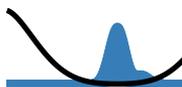


Figure 6.3: **(a)** A subset of the solutions found by the players as turquoise dots. The curves show the best solutions found by computer optimizations for each duration. The optimizations shown are the telescoped Krotov optimizations (red), shoveling (blue), tunneling (yellow), and HILO (purple). The diamonds mark the shortest duration with optimal fidelity ($F \geq 0.999$) for each optimization method. **(b)** shows the telescoped Krotov optimizations using seeds generated by players (yellow and blue) and computers (red). Player solutions divide into shoveling (blue) and tunneling (yellow) clans.

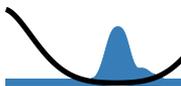
where $|f_{jk}\rangle = |\psi_j(x, t)\rangle - |\psi_k(x, t)\rangle$ is the unnormalized difference between two wave functions evaluated along the controls $\mathbf{u}_j(t)$ and $\mathbf{u}_k(t)$. In Fig. 6.4(a) the distance between all the player optimized solutions is shown for $T = 0.17$. The matrix with the pairwise distance between all the solutions is the distance map, which has some arbitrary ordering of



the solutions. The distance map in Fig. 6.4(a) has an ordering such that the $(n+1)$ 'th solution is the closest to the n 'th solution. Unexpectedly the solutions bunch into two clusters with low pairwise distance as Fig. 6.4(a) the solutions bunch into two clusters with low pairwise distance. We denote these clusters as “clans”. A clan must have 200 consecutive solutions with a distance smaller than 0.05. Fig. 6.4(b) shows the average control where the position is $u_1(t)$ and the amplitude is $u_2(t)$. Here it is seen that the solutions marked with yellow collect the atom by tunneling from the static to the controllable tweezer. Note, that around $t \simeq 0.06$ all the ~ 500 yellow solutions move to a particular tweezer position, which we interpret as the position maximizing the tunneling matrix element from tweezer to tweezer. Surprisingly, the amplitude of the tweezer is only weakly correlated with the final fidelity permitting the large fluctuations in amplitude shown in Fig. 6.4(b). This tunneling strategy is no longer successful for $T \leq 0.22$, since there it not enough time to complete the tunneling. In the clan marked with blue in Fig. 6.4 the tweezer is moved past the atom and back to the starting position. The overlapping tweezers form a steep potential gradient that quickly pushes the atom towards the target area. As shown in Fig. 6.3 the shoveling clan has a better quantum speed limit than the tunneling clan. Fig. 6.4(b) illustrates that both clans shake the tweezer when approaching the target area in order to remove unwanted excitations.

A priori it was not clear that two distinct physical solution strategies should exist for this problem. Furthermore, we were only able to identify the solutions by using the data from the players, who have no or little knowledge about quantum mechanics. This is remarkable since it shows that non-experts were able to provide seeds giving a clear physical understanding of the quantum dynamics. All the player solutions were also located in another cluster than the pure Krotov optimizations. This indicates that the human players navigate the optimization landscape in a different way than the numerical optimization.

Recall the discussion of the geometric quantum speed limit from Sec. 5.1 stating that the rate of change in the Wooters' distance is $d\theta/dt = 2\Delta E$ Eq. (5.15). We used this to argue that one should observe a pattern of initial excitation facilitating rapid motion in Hilbert space and then deexcitation as the target state is approached. In Fig. 6.4(c) the population in the different energy eigenstates is shown as a function of time for a solution in both clans. In Fig. 6.4(c) we observe this pattern as the optimal solutions do spread the wavefunction into many energy states in order to obtain high ΔE and thus fast transport. This also highlights that the solutions are strongly non-adiabatic. Fig. 6.4(c) shows that the shoveling solution has a higher ΔE than the tunneling solution,



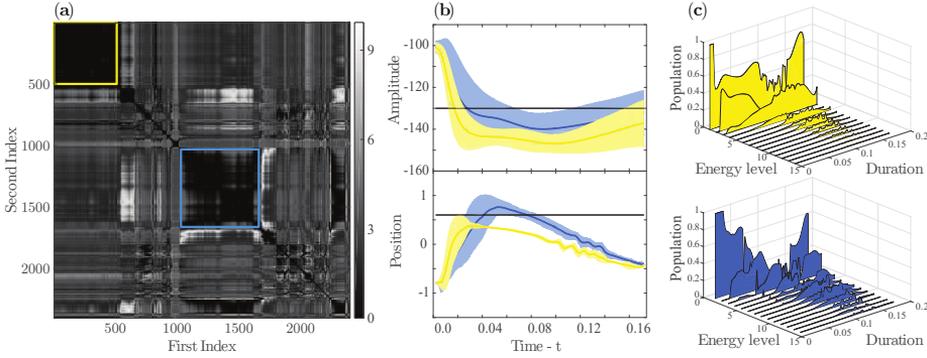
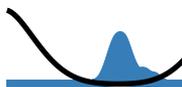


Figure 6.4: The shoveling (tunneling) clan is marked blue (yellow) throughout. (a) is a distance map, showing the distances between optimized player solutions as defined by Eq. (6.8). Boundaries of the clans are marked with colored squares. (b) shows the average control followed by the clans as thick lines, and a single standard deviations thereof as translucent areas. Controls are divided into the tweezer amplitude (top) and tweezer position (bottom). A horizontal line shows the position of the static tweezer. (c) displays the population in the different instantaneous energy eigenstates for the best solution from the tunneling (shoveling) clan at the top (bottom).

which explains why the shoveling clan has a lower quantum speed limit.

One long-term vision of our work is to circumvent the need for gamification by learning how players form their successful low-dimensional heuristic strategies, and to incorporate this into autonomous optimization algorithms. As a first step in this direction, we introduce here a Heuristically Initialized Local Optimizer (HILO) algorithm. This method parametrizes the player seeds in a low-dimensional subspace, whilst retaining the main features of good seeds. A local search algorithm can then move beyond this subspace to find optimal solutions. More specifically, inspired by the player solutions we constructed a three-dimensional parameterization consisting of moving the tweezer i) right, ii) slowly left, and iii) quickly left. Paths from this three-dimensional space for $T = 0.15$ were used as seeds for the Krotov's method and iteratively applied to shorter and longer durations with telescoping optimization (purple curve in Fig. 6.3(a)). Seeds were taken from the low-dimensional subspace using a simple direct search. HILO finds the lowest quantum speed limit at $T_{\text{QSL}}^{\text{num}} = 0.176$, outperforming even the best player optimized strategies.

In general, any optimization problem is simple if a proper seed



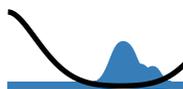
parametrization can be found. A core contribution of this work is that it hints at a generic, generalizable approach to finding such seed parametrizations by acquiring data and using clustering techniques from machine learning. One could argue that if a low dimensional successful seed parametrization exists, then perhaps one should have been able to guess it in advance. However, we believe that this is beside our main point, because it is much more scalable to investigate methodologies that allow automatic dimensionality reduction.

The work in this chapter predated our work on the GROUP method as presented in the previous chapters. However, the conclusion that search in a proper parametrization is important, was a direct inspiration for all the work presented in chapter 6.

6.3 Optimization Landscape Structure

As motivated in this chapter's introduction, the difficulty of a particular optimization problem can often be assessed with knowledge of the topology or ruggedness of the so-called optimization or fitness landscape. If many global optima are distributed across the landscape, local gradient-based search is often sufficient to finding the highest peak. On the other hand, if the landscape is very rugged and contains many local maxima, local methods will in general fail. In order to understand BringHomeWater within this terminology we seek to perform a visualization of the optimization landscape based on the player seeded solutions and the pure Krotov optimizations.

High dimensional datasets can be visualized using dimensionality reduction techniques that map the high dimensional data to two dimensions. Here we attempt to reduce the dimensions of the entire (reduced) cost-function $\hat{J}(u)$ Eq. (3.36) for illustrative purposes, which is a fundamentally different type of dimensionality reduction than when reparametrizing the problem or seed space as discussed in GROUP and HILO. In order to perform the visualizations, the algorithms cannot guarantee that all relative distances are preserved from the high-dimensional to the low-dimensional data. Therefore the visualizations do not reflect the actual structure in the high-dimensional data and one must be cautious about drawing conclusions directly from the visualizations. There are a number of algorithms for finding these visualizations such as Multidimensional Scaling [209] or Autoencoders [210]. These methods calculate the distances between points p_{ij} in a high dimensional dataset and then place points in lower dimensions with distance q_{ij} such that the q_{ij} are similar to the p_{ij} . A simple algorithm is Multidimensional Scaling that directly



minimizes the the error function

$$E = \sum_i \sum_{j \neq i} |q_{ij} - p_{ij}|. \quad (6.9)$$

This error function can be minimized using Nelder-Mead or gradient descent. The problem with this method is that it has difficulties in preserving local structure i.e. clusters in the datasets.

A recent and very popular method is t-Distributed Stochastic Neighbor Embedding (t-SNE) [211]. In this method the distance between points in the high dimensional dataset are rescaled according to a Gaussian probability distribution

$$\tilde{p}_{ij} = \frac{r_{ij} + r_{ji}}{2}, \quad r_{ij} = \frac{\exp(-p_{ij}/2\sigma_i^2)}{\sum_{k \neq i} \exp(-p_{ik}/2\sigma_i^2)}, \quad (6.10)$$

where the σ_i values are found using a binary search related to the number of neighbours [211]. This distribution favors close related points. In the low-dimensional space the distances are rescaled using a Student's t-distribution

$$\tilde{q}_{ij} = \frac{(1 + q_{ij})^{-1}}{\sum_{k \neq i} (1 + q_{ik})^{-1}}, \quad \tilde{q}_{ii} = 0. \quad (6.11)$$

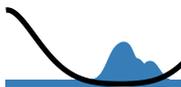
This distribution has a long tail increasing the distance between points far apart. This makes it easier to spot clusters in the low dimensional dataset. The goal in a dimensionality reduction is to make \tilde{p}_{ij} as similar to \tilde{q}_{ij} as possible. Due to the normalizations, the Eqs. (6.11)-(6.10) may be understood as probability distributions. A distance measure between probability distributions is the relative entropy or Kullback-Leibler divergence [212],

$$D_{\text{KL}}(P||Q) = \sum_i \sum_{j \neq i} p_{ij} \ln \left(\frac{\tilde{p}_{ij}}{\tilde{q}_{ij}} \right). \quad (6.12)$$

The idea in t-SNE is to minimize Eq. (6.12) using gradient descent. It is straightforward to compute the analytic gradient of Eq. (6.12).

Having introduced some methods for dimensionality reduction, let us return to problem of visualizing the player optimized solutions and pure Krotov optimizations constituting the high dimensional dataset. Both Multidimensional Scaling and t-SNE require a definition of distance in this high dimensional dataset

$$p_{ij} = \sum_i \int_0^T |u_i^j - u_i^l| dt, \quad (6.13)$$



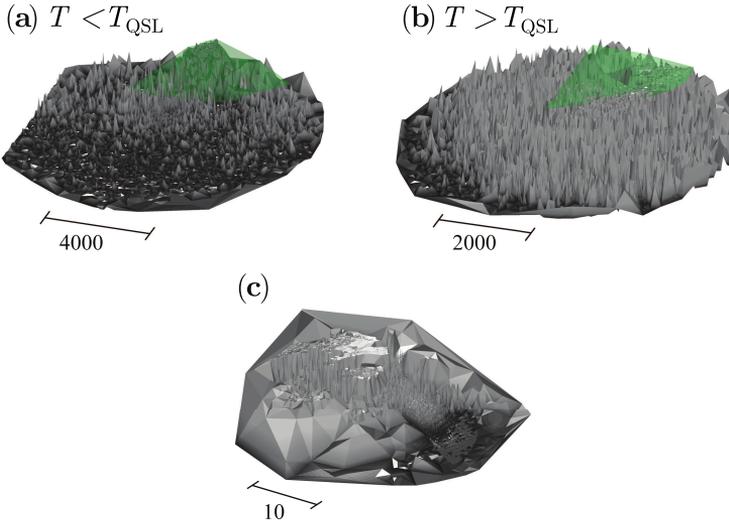
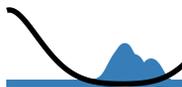


Figure 6.5: (a) and (b) show the two-dimensional rendering of the high-dimensional optimization landscape for process durations $T = 0.17$ and $T = 0.40$, respectively. Green areas mark the space probed by the optimized player solutions. (c) is the low-dimensional HILO landscape.

where $u_i^{j(l)}$ are the tweezer positions and amplitudes rescaled to intervals of unit length. The low dimensional distances q_{ij} are the Euclidean distances in the two-dimensional plane. We applied a slightly modified version of Multidimensional Scaling using Nelder-Mead to find the low dimensional landscapes, which are visualized in Figs. 6.5(a) and 6.5(b) for durations $T = 0.17$ and $T = 0.40$, above and below $T_{\text{QSL}}^{\text{num}}$. The multitude of spikes appears because gradient based optimization leads to nearly vertical lines in the landscape. This highlights the failure of local optimization algorithms to explore extended parts of the global landscape. For the longer duration (Fig. 6.5(b)), global maxima spread across the optimization landscape, explaining the success of Krotov-based methods. For the shorter duration (Fig. 6.5(a)) all high-fidelity solutions lie in the green player based region. This explains the failure of the Krotov-based methods and difficulty in locating the true T_{QSL} , since the global optima are no longer spread across the landscape. These figures directly corroborate the Optimization Landscape Hypothesis illustrated in Fig. 5.2 and highlight the dramatic change from benign to complex optimization landscape as the duration is decreased. It is worth stressing that the player optimized region is tiny in the high-dimensional landscape and therefore easily missed by different seeding strategies. We also applied this dimensionality reduction to the solutions from HILO shown in Fig. 6.5(c).



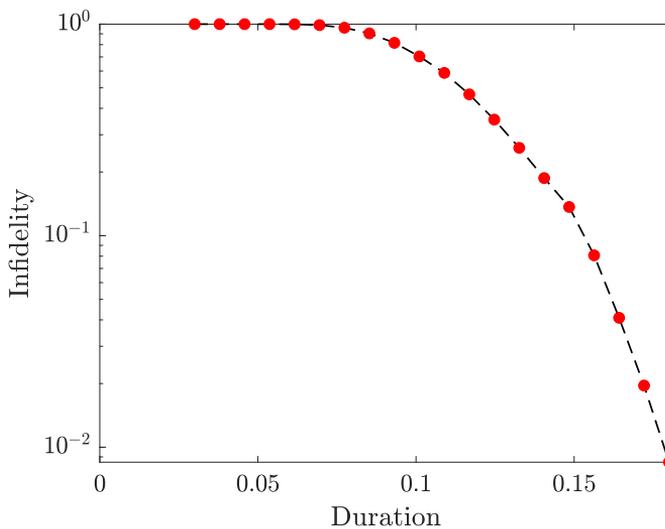


Figure 6.6: $F(T)$ -curve from applying Global-Local optimization to BringHomeWater. The global optimization was Differential Evolution as in Chap. 5 and the local optimization was an improved Krotov algorithm with variable step size and fast linesearch.

Initialized in a low-dimensional seed space, HILO efficiently explores a smaller but more optimal volume of the global optimization space. This landscape is much more benign than the two other in Fig. 6.5 explaining the success of this method.

The main conclusion, is that the much better solutions may be found if the local Krotov search is guided by some global component. In the language of Chap. 5 we stated this conclusion as: the exploration-exploitation trade-off could be improved by using a better exploration method than multistarting. In this sense, the player optimized solutions directly inspired the Global-Local optimizations presented in Chap. 5. In fact in unpublished results, we tried to apply the Global-Local algorithm to BringHomeWater. The results are shown in Fig. 6.6 and locates a quantum speed limit very close to HILO. In these optimizations we also used Differential Evolution and an improved version of Krotov algorithm with variable step size and fast linesearch. In general, the algorithm converged very fast on this problem so we decided to continue to the more complex Bose-Einstein problems in Chap. 5.



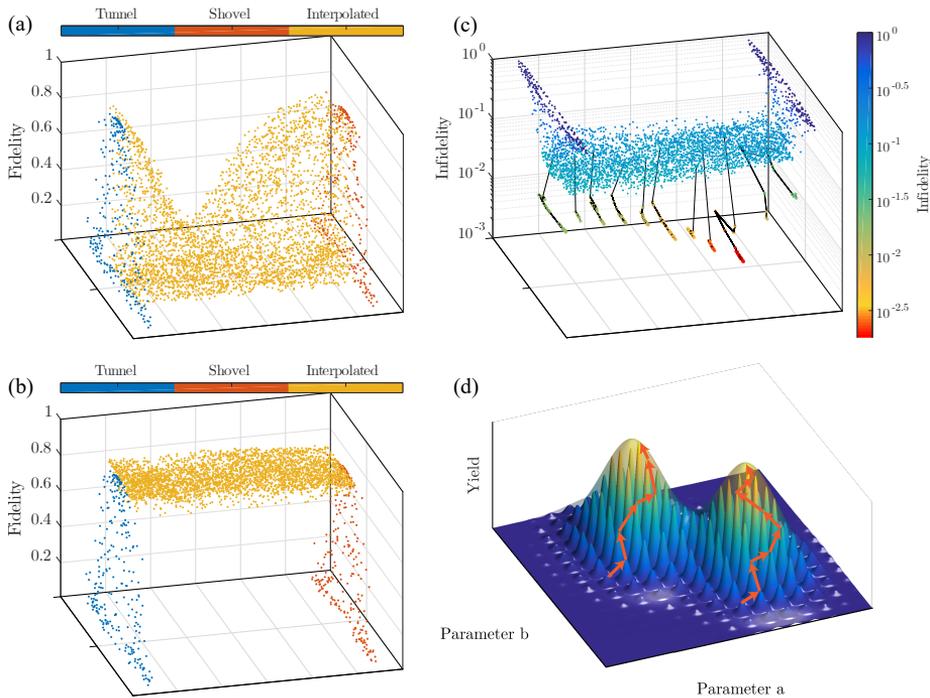
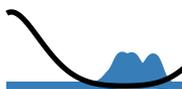


Figure 6.7: Visualization of a *bridge* between the two clans of solutions in the BringHomeWater challenge. The process time was set to $T = 0.19$, which is below the estimated quantum speed limit for both tunneling and shoveling clans ($T = 0.25$ and $T = 0.20$ respectively). The colors in (a) and (b) denote the type of each solution. Note that for this representation, the dimensionality has been reduced to two. Therefore, distances on the two horizontal axes should not be seen as absolute and axes labeling is omitted. (a) The result of using convex combinations of solutions and local perturbations of them to establish a connection between the two clans is shown. (b) Local Krotov optimization was applied to the yellow marked points in (a) and close-to-optimal solutions are attained. (c) Infidelity of resulting optimized solutions. Starting from the established bridge from (a), individual points were optimized using the combined Nelder-Mead and Krotov scheme (black lines). (d) Illustration of the concept of the *superlandscape*. Whereas the underlying optimization landscape consists of densely lying local optima, the superlandscape is defined as the smooth envelope function spanned on top of them. The two orange paths visualize the optimization with the combined Nelder-Mead and Krotov scheme by moving along the local optima of the underlying landscape towards an extremum in the superlandscape.



6.3.1 Uniqueness of Strategies

In the remainder of this section I investigate to what extent the tunneling and shoveling strategies are truly distinct. That is if these clans really represent physically distinct strategies in the sense that no mixed-strategy, high-yield solutions exist. Exhaustive search between the clans is not feasible due to the high dimension of the control problem. Instead, we investigated the topology of the landscape spanned by linear interpolation between the individual controls of representatives from the two clans. Given the interpolation parameter $\alpha \in [0, 1]$, the interpolated control is defined as

$$\mathbf{u}_{\text{int}}(\alpha) = \alpha \mathbf{u}_1 + (1 - \alpha) \mathbf{u}_2, \quad (6.14)$$

where \mathbf{u}_1 and \mathbf{u}_2 are the two controls used for the interpolation.

We perform a dimensionality reduction of interpolated solutions and local random perturbations to the two-dimensional plane using t-SNE. Here the high dimensional distances p_{ij} are given by Eq. (6.8) and the low dimensional distances q_{ij} are the Euclidean distances. The resulting landscape is shown in Fig. 6.7(a). The rapid decline in fidelity of the interpolated points and the multitude of points yielding zero fidelity, suggests that the clans can be seen as distinct regions of nearly optimal solutions. According to this interpretation, one would expect local optimization of these solutions to drag them towards either the shoveling or the tunneling solutions and thereby yield a region of attraction for each clan. Instead, local optimization using the Krotov algorithm results in the high fidelity *bridge* shown in Fig. 6.7(b). Using the distance metric Eq. (6.8), the displacement of each numerically optimized solution from the initial seed is relatively small. This implies that the optimization of the points in Fig. 6.7(a) leads to nearly vertical lines in the visualized landscape. This means that as long as the right region of the landscape is explored, a better solution lies very close to the non-perfect trial solution. Additionally, each initial seed converged to a different optimum, i.e., new distinct solutions have been found as illustrated by the yellow points of Fig. 6.7(b). Thus, the landscape is locally very rugged, but rich in optima. Each peak has a narrow width since any minor perturbation to a control at some instant will in general lead to a decreased fidelity. However, most minor decreases in the fidelity can be compensated by another carefully chosen perturbation at another time. This leads to many closely spaced locally optimal solutions.

Given the density of locally optimal points, we now define the *super-landscape* as the approximately smooth envelope function spanned by the optimal points. If \mathcal{O} is a local optimization algorithm then this landscape



is the composition $\tilde{J} = \hat{J}[\mathcal{O}[u(t)]]$. Fig. 6.7(d) illustrates a simple, generic superlandscape. The underlying landscape consists of a dense collection of individual peaks with smoothly varying heights.

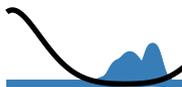
If the superlandscape can be evaluated with sufficient speed (using efficient local optimization), then we propose to perform completely deterministic global optimization using a local search algorithm. Here, we use Nelder-Mead search combined with Krotov optimization. The Global-Local algorithm from Sec. 5.3 may also be understood as optimizing \tilde{J} just with Differential Evolution instead of Nelder-Mead. Depending on the structure of \tilde{J} it may be an advantage to use Nelder-Mead or Differential Evolution. The Nelder-Mead search of \tilde{J} is illustrated in Fig. 6.7(d).

We start this optimization of \tilde{J} at a number of interpolated solutions along the identified bridge. This results in significantly improved solutions as shown in Fig. 6.7(c). These solutions are found at a duration of $T = 0.19$. The best optimized solutions from this combined search reached $F = 0.998$ in fidelity. When we inspect the actual solution, it is clearly seen to be a combination of the tunneling and shoveling strategies, since it places the transport tweezer on top of the atom rather than to the left or right of it.

The fact that a single one-dimensional line scan identifies a bridge illustrates that low dimensional search is sufficient once a good search heuristic is determined [213]. In the previous section, we constructed these search spaces explicitly in HILO using parametrizations that emerged from data analysis of the player solutions. The search along convex linear combinations of existing solutions [Eq. (6.14)] may provide a computationally inexpensive methodology to identify good solutions in a multi-dimensional landscape.

6.4 Experimental Landscape Structure

Inspired by the bridge between the solution strategies that emerged from the convex linear combinations, we sought to investigate if a similar structure in the optimization landscape could be observed elsewhere. In this section we search along the linear combinations in the experimental problem of Bose-Einstein condensate production [214]. In our experimental setup an increased condensate atom number, N_{BEC} , will provide significantly improved initial conditions for subsequent quantum simulation experiments using optical lattices [26]. Extensive optimization has been applied to the condensate creation problem over the past decade by employing global closed-loop optimization strategies using genetic algorithms [215–217]. However, little effort has been devoted to the char-

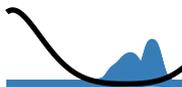


acterization of the underlying landscape topology and thereby assessing the fundamental difficulty level of the optimization problem. It has not been discussed if the landscape is more similar to single distinct peaks as in Fig. 6.8(a) or many closely connected solutions as in Fig. 6.8(b). Recent experiments [218] indicated a convex and thereby simple underlying landscape. However, this study did not explicitly optimize N_{BEC} and operated within a severely restricted subspace.

In our experiment [219], we capture Rb atoms in a trap made of two orthogonal, focused 1064nm laser beams and a superimposed quadrupolar magnetic field which creates a magnetic field gradient at the position of the atoms and thereby forms a magnetic trap. We evaporatively cool the atoms past the phase transition to a condensate by lowering the intensity of the laser beams as well as the magnetic field gradient. Then, the traps are turned off, and the atoms are imaged with resonant light. Image analysis yields the total and condensed atom numbers N_{tot} and N_{BEC} .

This setup allows for evaporative cooling into two widely used trap configurations. First, making use of only the laser beams, a purely optical trap can be created; this is commonly known as a *crossed dipole trap* (CDT) [69]. Second, a single laser beam can be combined with a weak magnetic gradient to form a *hybrid trap* [220]. In both cases, the traps are initially loaded from a pure tight magnetic trap. Conventionally, one of two geometrically different loading methods is pursued. The first method is loading into a large volume trap that is nearly spatially mode-matched with the initial magnetic trap [214]. The second method is loading into a small volume trap that only has a small spatial overlap with the initial trap. The latter leads to a “dimple” type loading in which a smaller but colder atom cloud is produced [221]. We can directly control the effective volume of the trap by translating the focus position of one of the dipole trap beams. This inspired us to identify four initial “conventional” trap configurations or strategies: i) a small volume “narrow” crossed dipole trap (NCDT), ii) a large volume “wide” counterpart (WCDT), iii) a dimple type hybrid trap (HT), and iv) a “wide” hybrid trap (WHT).

We first optimize the system by applying a simple standard experimental approach. Starting from the set of control variables associated with a known strategy, we iteratively perform 1D scans of single variables until a specified level of convergence is reached. The 1D scans yield four distinct strategies. We find that the third strategy HT has the best performance. This hints at the landscape topology sketched in Fig. 6.8(a). Further systematic studies would then proceed to scans of two or more parameters simultaneously. However, scanning many simultaneous parameters lead to prohibitively many combinations that cannot be explored in reasonable time. Instead, we search along the convex linear combinations of these



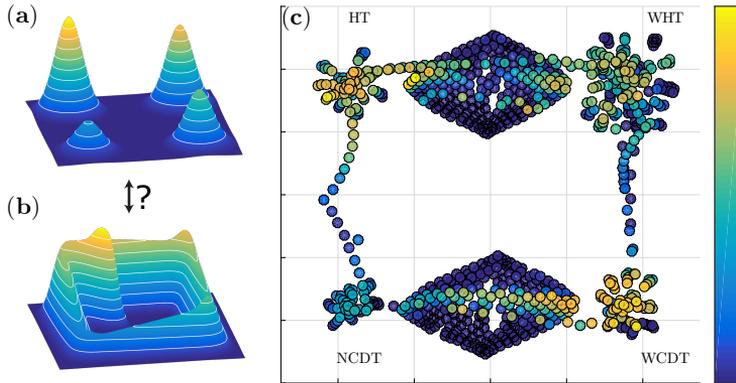
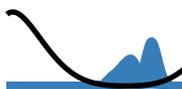


Figure 6.8: (a) Illustration of the apparent global landscape topology for condensate production after performing 1D parameter scans. It seems to contain distinct local optima. However, as (b) illustrates, connecting *bridges* were found both between some of the conventional strategies and to novel high-yield solutions in the high-dimensional search space. (c) 2D tSNE [211] representation of the landscape showing the variety of different trap configurations that are accessible in our experiment. The plot contains data of the four main configurations, which were scanned and optimized by 1D and 2D parameter scans as described in the text.

strategies as in Eq. (6.14) from the previous section.

Both the low-yield NCDT strategy and the WCDT are types of crossed dipole traps with different effective volumes. However, a simple linear interpolation of all the available parameters between the NCDT and the WCDT fails to locate a bridge. However, a bridge can be located by performing extended 2D-interpolations that also include effective trap volume as an independent parameter. Changing the trap volume also changes the trap depth, which must be compensated by a quadratic increase in the laser intensities. The emergence of a bridge by adding more parameters disproves the initial assumption that each strategy was distinct as illustrated in Fig. 6.8(a).

To visualize the actual landscape, we created a dimensionality-reduced illustration using t-SNE [211] from the parameter scans (Fig. 6.8(c)). The four initial strategies are represented by the four clusters in the corners. The data points forming the bridge between NCDT and WCDT lie in the diamond shape at the bottom. A few other 1D and 2D interpolations between other pairs of strategies are shown, but none form a bridge. In an attempt to locate a bridge between NCDT and HT, extended 3D scans are performed. These scans identify a novel optimum away from the four

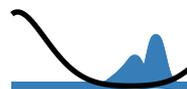


initially defined experience-based trap configurations. This demonstrates HT strategy, which was initial for a global optimum, is not even a local optimum when appropriate parameter sets are investigated. In conclusion the landscape is more similar to what is depicted in Fig. 6.8(b) than in Fig. 6.8(a). There are two differences, the first that the four conventional strategies are now connected with bridges and the second that there exists one other solution with higher-yield than the four initial strategies.

This work shows how dimensionality reduction can be used to gain a qualitative insight into the structure of the optimization landscape and how searching among the strategies may reveal novel solutions. The solutions found here are qualitatively different from well-known strategies for producing Bose-Einstein condensates.

We further investigated this condensate production problem using the popular closed-loop method Nelder-Mead with CRAB. In Ref. [7] we compared this numerical search algorithm with teams of citizen scientists that also attempted to solve the Bose-Einstein condensate production problem. This study was much more controlled than Quantum Moves, allowing us to draw conclusions about how the citizen scientists perform their search. In this thesis, I will only describe the main conclusion and for further details the reader is referred to Ref. [7].

The players conduct an adaptive search, i.e. if one had identified a good solution compared to the other solutions visible to the player, the player tended to make small adjustments in the next attempt. In contrast, if the solution found by a player was far behind the best solution, the player tended to engage in more substantial adjustments to their previous solution. Advancing previous studies [222, 223], we were also able to identify that players engaged in similar types of adaptive search, when engaged in social learning. The nature of adaptive search leads to a heterogeneous human search “algorithm” that combines local search with a global component. This search is prevented from stopping too early as the poorly performing individuals engage in a more explorative search unlike the top performing players, who are more exploitative. This type of search is very different from the local nature of Nelder-Mead with CRAB algorithm, which only makes small incremental changes in either positive or negative directions. This shows that citizen scientist find their own balance between exploration and exploitation when searching in complex landscapes. More studies on how people search holds the promise of encoding these principles in an algorithm that could be transferred to other domains.



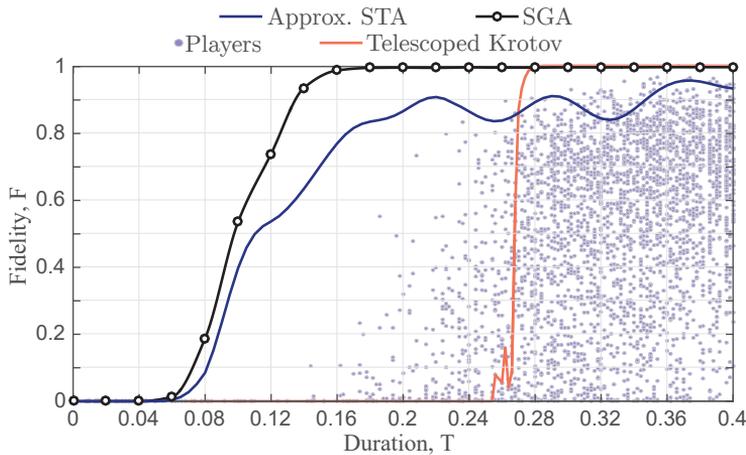
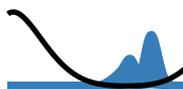


Figure 6.9: Fidelity for different protocol durations. The fidelity found using the approximation shortcuts-to-adiabaticity shown with blue. The results from the stochastic gradient descent is shown with black dots. This figure is copied from Ref. [224].

6.5 Discussion of D. Sels' Results

Before concluding this chapter, I find it necessary to include a comment on the very direct and confrontational paper by D. Sels in Ref. [224]. D. Sels raises two main points. The first point constituting most of Ref. [224] is a set of analytical solutions obtained by using a formalism known as shortcuts-to-adiabaticity. The second point is a set of solutions obtained using a numerical method. I address the viability of these two points separately.

The first point raised by D. Sels is a set of analytical solutions found using the shortcuts-to-adiabaticity formalism, which is also presented in Chap. 7. D. Sels states that the BringHomeWater control problem cannot be solved within this formalism since it is too complicated. However, D. Sels circumvents this problem by an extension to the shortcut formalism, which allows for producing approximate shortcuts from Ref. [225]. Note that this extension was published after our work leaving us unable to perform a similar analysis. The result of this analysis is shown with a blue curve in Fig. 6.9. D. Sels also includes an analysis of a simplified BringHomeWater problem with a single tweezer, which I do not include since it is another problem. Based on the blue curve D. Sels concludes that the approximate shortcuts “grossly outperforms any Quantum Moves player”. However, if the reader carefully inspects Fig. 6.9 this is wrong as there are clearly several player points that lie above the blue curve. More



Algorithm 6 Stochastic Gradient Ascent from Ref. [224]

-
- 1: Initialize a random starting protocol $\mathbf{u} = (x_N, x_{N-1}, \dots, x_2, x_1)$.
 - 2: Initialize $n = 0$ and $F^0 = 0$.
 - 3: **while** $F^n > F^{n-1}$ **do**
 - 4: Let \mathbf{I} be a random permutation of all $(1, 2, \dots, N - 1, N)$.
 - 5: **for** each W in \mathbf{I} **do**
 - 6: Compute $\langle \varphi_{W+1} | = \langle \varphi | \prod_{i=W+1}^N \hat{U}_{k_i}$.
 - 7: Compute $|\psi_{W+1}\rangle = \prod_{i=i}^{W-1} \hat{U}_{k_i} |\psi_0\rangle$
 - 8: Compute the fidelity for all $F_k = |\langle \varphi_{W+1} | \hat{U}_k | \psi_{W-1} \rangle|^2$.
 - 9: Select the optimal k_j .
 - 10: Replace $\mathbf{u}_W = x^{k_j}$.
 - 11: $n \leftarrow n + 1$.
 - 12: Compute $F^n = F(\mathbf{u})$.
-

importantly, the approximate shortcuts do not give a better estimate of the quantum speed limit than our original telescoped Krotov solutions despite the fact they outperform a slightly larger number of players. For these reasons the first point by D. Sels has no particular relevance. Most of Ref. [224] is purely devoted to the shortcut formalism. The second and actually somewhat interesting point is unfortunately hidden in Ref. [224].

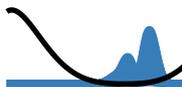
The second point is a comparison of a stochastic gradient ascent algorithm introduced by D. Sels and the Quantum Moves players. I start by giving a brief account of this algorithm. D. Sels starts by removing the tweezer amplitude degree of freedom setting $u_2(t)$ to the minimal value in Eq. (6.6). The only remaining degree of freedom is the tweezer position. This position is discretized in $M = 128$ different steps x^k . The author then computes the 128 time evolution operators associated to each possible value of the control

$$\hat{U}_k = \exp(-i\hat{H}_k\Delta t). \quad (6.15)$$

The total time evolution is the product

$$\hat{U}(T) = \prod_{i=1}^N \hat{U}_{k_i}. \quad (6.16)$$

The fidelity is computed by $F = |\langle \psi_t | \hat{U} | \psi_0 \rangle|^2$ where $|\psi_t\rangle$ and $|\psi_0\rangle$ are the target and initial states. The fidelity is then optimized using the stochastic ascent algorithm in Algorithm 6. Note that despite sharing the same name this is not the same as the stochastic ascent algorithm known from



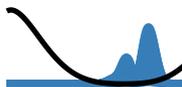
machine learning. In machine learning this method refers to training some statistical estimator using gradient descent over stochastically chosen subsets or batches of a full dataset [226]. D. Sels' method is quite efficient, since it computes all allowed tweezer positions without having to recompute the entire time evolution. This method also explores the control landscape efficiently since it probes all the allowed values of the control. In comparison Krotov's method only makes a small perturbation to each control value.

In Ref. [224] this method was applied at different durations and the result is shown in Fig. 6.9. This locates an estimated quantum speed limit around $T = 0.17$. If any reader only compares the black numerical results in Fig. 6.9 with the grey player points and red Krotov optimizations within the same figure then the reader would conclude as D. Sels that stochastic gradient ascent (not approximate shortcuts) "falsify the thesis of Sørensen and co-workers".

Unfortunately, this figure is very misleading. If the reader compares Fig. 6.9 with our original Fig. 6.3(a) then it is clear that D. Sels intentionally has removed all the player optimized solutions and only included the bare numerical optimizations and player solutions. In this way D. Sels only compares his best results to our worst results making his own results seem relevant. In fact, the figure and the entire paper completely leaves out any mentioning of player optimized solutions and of the fact that our HILO-method finds a quantum speed limit similar to the stochastic gradient ascent algorithm by D. Sels. D. Sels also redefines the problem such that the atom is initialized in the combined tweezer and thereby the entire pick up phase is avoided. Therefore it is not possible to perform an honest one-to-one comparison between his results and ours. In private communications, we have made D. Sels aware of the problems with his presentation. Unfortunately he chose to not update his presentation and instead submit it to Physical Review A. In this way D. Sels has been quite successful in promoting his work at the expense of ours as Ref. [224] was an editors' suggestion.

An honest way of updating the work by D. Sels would be: *If we had compared the player based optimizations to pure numerical optimization with stochastic gradient ascent then we would have observed no significant improvement in the estimate of the quantum speed limit.*

A main point in our work was stated at the end of Sec. 6.2. The point being that if one naively uses optimal control such as Krotov's method then it will often fail. It is typically necessary to use some kind of tweaks or heuristics such as the good seeding strategy from HILO, the convex linear combinations or the bridges in the condensate optimization, the parametrization in GROUP or the Global-Local optimization. In fact,

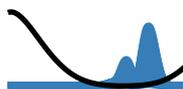


before mentioning stochastic gradient ascent, we have already seen two separate ways of solving BringHomeWater. The first way is Krotov with the HILO seeds and the second is the Global-Local optimization. D. Sels also tweaks his approach by i) redefining the problem, ii) removing the tweezer position, iii) discretizing the control in an appropriate number of steps, and iv) updating the fidelity in a clever way. This constitutes tweaking or heuristics, whether D. Sels agrees or not. Our work addresses what do to if one does not have access to such heuristics. In this case new heuristics may be found by analyzing the player solutions with machine learning.

6.6 Conclusion

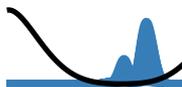
We have developed the first citizen science game in quantum mechanics. The game Quantum Moves gamifies quantum optimal control problems. We discussed the BringHomeWater control problem in detail where users with little or no training in quantum physics provided high quality seeds. The player solutions allowed us to identify the tunneling and shoveling strategies using clustering methods from machine learning. In subsequent work, we showed that these strategies are not distinct in the sense that they are connected by a high fidelity bridge. An analysis of the player solutions allowed us to find high quality seeds from a three dimensional space (HILO), which gave the best estimate of the quantum speed limit. The clustering analysis and dimensionality reduction techniques from the Quantum Moves project was later extended to our optimized Bose-Einstein condensate experiments. In this project we saw that the simple heuristic of searching among convex linear combinations could be transferred from BringHomeWater to an experimental problem. We also had citizen scientists optimize the production of Bose-Einstein condensates. The controlled nature of this study allowed us to gain insight into the collective adaptive search performed by the players.

In the beginning of this chapter, one goal of the Quantum Moves project was stated as being: To explore to what extend citizen science could be used to navigate complex optimization landscapes. The citizen scientists do search the optimization landscape in a fundamentally different way from the local optimization algorithms. There were indications in both Quantum Moves and the experimental condensate production that the players could perform jumps away from local traps. However, it is also clear that the players cannot perform the high level of fine-tuning needed to converge to high fidelity or N_{BEC} . This highlights that the next versions of Quantum Moves should directly include the optimization.



Another point is also the fact that Quantum Moves' dynamic interface, tracking the cursor as a function time, makes it very difficult to make iterative improvements and thereby achieve high fidelities. An interface focusing more on the control could improve the quality of the player data.

At the time of writing, we at the scienceathome.org project are planning a new version of Quantum Moves that will include many of these features. This will allow us to explore to what extent the player generated data is useful on a wider class of problems with access to different types of optimization.

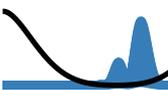


7 Combining Shortcuts-To-Adiabaticity and Control Theory

Sec. 7.3 and Sec. 7.4 are copied directly from Ref. [4]. There are a few fragments in the other sections within this chapter that are also copied from Ref. [4].

In the last chapters, we have seen how to solve the state-to-state problem using different optimization algorithms minimizing the infidelity. In this chapter, we take another approach and discuss an analytic method guaranteeing perfect transfers. The starting point in the optimization methods is the control that by the Schrödinger equation define a state trajectory those final state gives the fidelity. The formalism in this chapter inverts this process, it starts by choosing the desired state trajectory and then engineers a control that transport the system along this trajectory. This formalism is known as shortcuts-to-adiabaticity (STA) and it encompasses several closely related methods [227–231]. This chapter discuss the so-called transitionless driving [228] where the state trajectory is along the instantaneous eigenstates of some reference Hamiltonian \hat{H}_0 .

However, guaranteed perfect fidelity is indeed too good to be true and this method has a number of limitations. Firstly, it is non-trivial to apply shortcuts-to-adiabaticity for non-diagonalizable systems [232]. Secondly, even for diagonalizable systems the control following the desired state trajectory may not exist. Finally, the proposed control might also require too much energy or be unstable against perturbations in the control. In this chapter, we discuss a combination of optimal control and shortcuts-to-adiabaticity called Combining STA and Control Theory (COSTACT). The idea in this method is to optimize the state trajectory for



optimal energy consumption and robustness. This technique is applied to control population transfer in a three-level Λ -system.

This chapter is organized as follows, Sec. 7.1 introduces the shortcuts-to-adiabaticity formalism and discusses an example with the Landau-Zener system. Sec. 7.2 introduces the experimentally relevant Λ -system and the widespread Stimulated Rapid Adiabatic Passage methodology. It is discussed how shortcuts may be utilized to speed up this method. Sec. 7.3 introduces the COSTACT method and demonstrates how it can be applied in the Λ -system to find solutions almost twice as fast those reported in [233], while still satisfying the experimental requirements. The chapter ends with a conclusion and outlook on COSTACT.

The content of Sec. 7.3 was published in Ref. [4] with me as second author. Although the main work was carried out by the first author, I made significant contributions to all parts of the project as the effective day-to-day supervisor.

7.1 Shortcuts-To-Adiabaticity

The shortcuts-to-adiabaticity formalism is based on the adiabatic theorem, which states that a system always follow the instantaneous ground state of a sufficiently slowly varied Hamiltonian. Before discussing the shortcut formalism we present a brief discussion of the adiabatic theorem.

7.1.1 The Adiabatic Theorem

Consider a Hamiltonian with instantaneous eigenstates $\hat{H}(t)|\varphi_n(t)\rangle = E_n(t)|\varphi_n(t)\rangle$ and some fixed reference basis $|n\rangle$. We may then define an unitary operator \hat{U} that maps from the eigenbasis into the fixed diagonal basis

$$\hat{U} = \sum_n |n\rangle\langle\varphi_n(t)|. \quad (7.1)$$

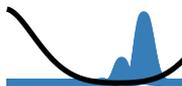
Let us consider the unitary transformed states $|\tilde{\psi}\rangle = \hat{U}|\psi\rangle$. Taking the time-derivative of the transformed states and using the Schrödinger equation shows they have associated Hamiltonian

$$\tilde{H} = \hat{U}\hat{H}\hat{U}^\dagger + i(\partial_t\hat{U})\hat{U}^\dagger \quad (7.2)$$

A straightforward calculation gives

$$\hat{U}\hat{H}\hat{U}^\dagger = \sum_n E_n|n\rangle\langle n| \quad (7.3)$$

$$i(\partial_t\hat{U})\hat{U}^\dagger = i\sum_{m,n}\langle\partial_t\varphi_n|\varphi_m\rangle|n\rangle\langle m|, \quad (7.4)$$



where the time-dependence has been suppressed for brevity. Combining these two results and differentiating the definition of an eigenstate $\hat{H}(t)|\varphi_n(t)\rangle = E_n(t)|\varphi_n(t)\rangle$ gives

$$\tilde{H} = \sum_n \left(E_n(t) - i\langle\varphi|\partial_t\varphi\rangle \right) |n\rangle\langle n| + i \sum_{\substack{n,m \\ n \neq m}} \frac{\langle\varphi_n|\partial_t\hat{H}|\varphi_m\rangle}{E_n - E_m} |n\rangle\langle m|, \quad (7.5)$$

where we have assumed a non-degenerate spectrum. The above expression is exact. In the *adiabatic approximation* the Hamiltonian is slowly changing compared with the instantaneous energy gap, i.e. $\langle\varphi_n|\partial_t\hat{H}|\varphi_m\rangle \ll E_n - E_m$ for all n and m . This approximation allows us to neglect the second term. If the state is initially in an eigenstate, it will remain in this state and only gain a phase

$$|\psi(t)\rangle = e^{i\xi_0(t)}|\varphi_n(t)\rangle, \quad (7.6)$$

where

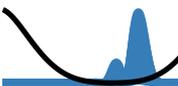
$$\xi_0(t) = - \int_0^t E_0(t)dt + i \int_0^t \langle\varphi_0(t)|\partial_t\varphi_0(t)\rangle dt. \quad (7.7)$$

The first term is known as the dynamical phase whereas the second term is the geometric phase or Berry's phase [161]. This result is also known as the adiabatic theorem. The principle in adiabatic control is to slowly change the system Hamiltonian and the state follows the instantaneous eigenstate. Adiabatic protocols also have the advantage that they are naturally robust against perturbations in the control as only the initial and final Hamiltonians matter. The main disadvantage is that they are naturally slow and therefore vulnerable to decoherence.

7.1.2 Transitionless Driving

Shortcuts-to-adiabaticity is a set of related analytic methods that speed up adiabatic protocols. In this thesis we will discuss the transitionless driving protocol [228].

The goal in transitionless driving is to design a Hamiltonian such that the dynamics follow the eigenstates of a reference Hamiltonian \hat{H}_0 i.e. the time-evolution follows Eq. (7.6) with eigenstates $\hat{H}_0(t)|\varphi_n(t)\rangle = E_n(t)|\varphi_n(t)\rangle$. Transitions would occur among the instantaneous eigenstates $|\varphi_n(t)\rangle$ if \hat{H}_0 is changed so fast the adiabatic condition is violated. All these transitions are due to the second term in Eq. (7.5) originating from Eq. (7.4). The idea in transitionless driving is to add a term to \hat{H}_0 known as the counter diabatic Hamiltonian \hat{H}_{CD} that exactly cancels



this term whereby all unwanted transitions are removed. The counter-diabatic Hamiltonian is defined by

$$\hat{U} \hat{H}_{\text{CD}} \hat{U}^\dagger = -i(\partial_t \hat{U}) \hat{U}^\dagger - i \sum_n \langle \varphi | \partial_t \varphi \rangle |n\rangle \langle n|. \quad (7.8)$$

Consider the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_{\text{CD}}$ and transform this with the unitary from Eq. (7.1). Using the results from Eqs (7.2)-(7.4) and the definition (7.8) the transformed Hamiltonian becomes

$$\tilde{H} = \sum_n \left(E_n(t) - i \langle \varphi | \partial_t \varphi \rangle \right) |n\rangle \langle n|. \quad (7.9)$$

The dynamics of this Hamiltonian will exactly follow Eq. (7.6) and there are no transitions among the instantaneous eigenstates. In this presentation the counter-diabatic Hamiltonian may be interpreted as the Hamiltonian that cancels the extra term that appears when shifting into the adiabatic frame. These calculations show that the second term in Eq. (7.8) gives rise to the geometric phase contribution. This term only affects to phase and not the populations in the adiabatic frame and for this reason it is often omitted for simplicity. In calculations one typically rearranges Eq. (7.8) into

$$\hat{H}_{\text{CD}} = i \sum_n |\partial_t \varphi\rangle \langle \varphi_n| - \langle \varphi_n | \partial_t \varphi \rangle | \varphi_n \rangle \langle \varphi_n|. \quad (7.10)$$

The sum $\hat{H} = \hat{H}_0 + \hat{H}_{\text{CD}}$ is known as the *physical Hamiltonian*. Although the dynamics of the physical Hamiltonian (\hat{H}) follows the adiabatic states of the reference Hamiltonian (\hat{H}_0), the dynamics is not adiabatic with respect to the physical Hamiltonian. A central challenge in the shortcuts-to-adiabaticity formalism is to what extent the physical Hamiltonian is actually realizable in concrete applications.

7.1.3 Shortcut-To-Adiabaticity in the Landau-Zener Model

As an initial example we will discuss how to implement a shortcut-to-Adiabaticity in the Landau-Zener model from Sec. 3.1.1,

$$\hat{H}_{\text{LZ}} = \begin{pmatrix} \Delta(t) & \Omega_R(t) \\ \Omega_R(t) & -\Delta(t) \end{pmatrix} = \Omega_R(t) \sigma_x + \Delta(t) \sigma_z. \quad (7.11)$$

Recall from the discussion in Sec. 3.1.1 that the eigenstates are,

$$|\lambda_+(t)\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} |1\rangle, \quad |\lambda_-(t)\rangle = \sin \frac{\theta}{2} |0\rangle - \cos \frac{\theta}{2} |1\rangle, \quad (7.12)$$



where the rotation angle is $\theta(t) = \arctan(\Omega_R(t)/\Delta(t))$ for $\Delta \geq 0$ and $\theta(t) = \pi + \arctan(\Omega_R(t)/\Delta(t))$ for $\Delta < 0$. In order to calculate the counter-diabatic Hamiltonian in Eq. (7.10) we need to know the derivative of the eigenstates,

$$|\dot{\lambda}_+\rangle = -\frac{\dot{\theta}}{2}|\lambda_-\rangle, \quad |\dot{\lambda}_-\rangle = \frac{\dot{\theta}}{2}|\lambda_+\rangle, \quad (7.13)$$

where the derivative of the angle is

$$\dot{\theta} = \frac{\dot{\Omega}_R\Delta - \dot{\Delta}\Omega_R}{\Delta^2 + \Omega_R^2}, \quad (7.14)$$

Given these time derivatives may now calculate \hat{H}_{CD} explicitly using Eq. (7.10),

$$\begin{aligned} \hat{H}_{\text{CD}} &= \frac{i\dot{\theta}}{2} \left(|\lambda_+\rangle\langle\lambda_-| + \langle\lambda_-|\lambda_+\rangle|\lambda_-\rangle\langle\lambda_-| - |\lambda_-\rangle\langle\lambda_+| - \langle\lambda_+|\lambda_-\rangle|\lambda_+\rangle\langle\lambda_+| \right) \\ &= \frac{i\dot{\theta}}{2} \left(|\lambda_+\rangle\langle\lambda_-| - |\lambda_-\rangle\langle\lambda_+| \right). \end{aligned} \quad (7.15)$$

With the definition of the eigenstates Eq. (7.12) it is straightforward to show that the matrix representation of the counter diabatic Hamiltonian is

$$\hat{H}_{\text{CD}} = \frac{1}{2} \begin{pmatrix} 0 & -i\dot{\theta} \\ i\dot{\theta} & 0 \end{pmatrix} = -\Omega_a\sigma_y, \quad (7.16)$$

where we in accordance with the literature have introduced the angle $\Omega_a = -\dot{\theta}/2$. In the original Hamiltonian Eq. (7.11) we had only assumed control over σ_x and σ_z and the counter diabatic Hamiltonian ostensibly requires control over the remaining σ_y Pauli matrix. This illustrates the recurring problem that there is no guarantee that the physical Hamiltonian $\hat{H}_0 + \hat{H}_{\text{CD}}$ is realizable. Fortunately, the counter diabatic term can be absorbed into the original Hamiltonian

$$\hat{H} = \hat{H}_{\text{LZ}} + \hat{H}_{\text{CD}} = \begin{pmatrix} \Delta & \sqrt{\Omega_R^2 + \Omega_a^2}e^{i\gamma} \\ \sqrt{\Omega_R^2 + \Omega_a^2}e^{-i\gamma} & -\Delta \end{pmatrix}, \quad (7.17)$$

where $\gamma(t) = \arctan(\Omega_a/\Omega_R)$. The Hamiltonian can be transformed by the unitary

$$\hat{U} = \exp \begin{pmatrix} -i\gamma(t)/2 & 0 \\ 0 & i\gamma(t)/2 \end{pmatrix}. \quad (7.18)$$

A brief calculation gives the result

$$\hat{H}_{\text{STA}} = \begin{pmatrix} \Delta_{\text{STA}} & \Omega_{\text{STA}} \\ \Omega_{\text{STA}} & -\Delta_{\text{STA}} \end{pmatrix}, \quad (7.19)$$



where $\Delta_{\text{STA}}(t) = \Delta(t) + \dot{\gamma}(t)/2$ and $\Omega_{\text{STA}}(t) = \sqrt{\Omega_R^2(t) + \Omega_a^2(t)}$. The populations in the different states are not affected by this transformation since \hat{U} is diagonal and unitary. Hence, solutions using this shortcut Hamiltonian are also solutions to the original problem. In this example we have shown that the instantaneous eigenstates of the reference Landau-Zener Hamiltonian defined by some $\Omega_R(t)$ and $\Delta(t)$ are precisely followed if we instead use the physical Hamiltonian in Eq. (7.19). In the next section, Eq. (7.19) is applied on an experimentally relevant problem.

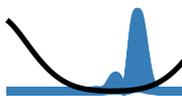
7.2 The Λ -system

In this section, we are going to consider a quantum mechanical system with three different energy levels $|1\rangle$, $|2\rangle$ and $|3\rangle$, where the goal is to transfer the initial state from $|1\rangle$ to $|3\rangle$. However, the second level $|2\rangle$ is subject to a fast decay. A direct coupling between level $|1\rangle$ and $|3\rangle$ is not feasible due to e.g. selection rules or experimental limitations. This type of system is shown in Fig. 7.1. The Λ -system has Hamiltonian,

$$\hat{H} = \frac{1}{2} \begin{pmatrix} 0 & \Omega_P & 0 \\ \Omega_P & 2\Delta & \Omega_S \\ 0 & \Omega_S & 0 \end{pmatrix}, \quad (7.20)$$

where Ω_P and Ω_S are two the Rabi frequencies that couple $|1\rangle$ to $|2\rangle$ and $|3\rangle$ to $|2\rangle$. Δ is the detuning of the Rabi frequencies. The initial state is $|1\rangle$ and the target state is $|3\rangle$. The detuning is kept fixed meaning the Rabi frequencies are the controls.

This system can be realized and controlled in a multitude of physical systems and for an overview I refer the reader to the seminal review [234]. Motivated by a later experimental application of shortcuts-to-adiabaticity, we discuss a realization of the Λ -system in the internal energy structure of cold ^{87}Rb atoms [233]. Here the two lower levels are the hyperfine state $|1\rangle = |F = 1, m_F = 0\rangle$ and $|3\rangle = |F = 2, m_F = 0\rangle$ within the $|5^2S_{1/2}\rangle$ manifold, while the excited state is $|2\rangle = |5^2P_{3/2}\rangle$. These different levels can be coupled with a laser. The laser from $|1\rangle$ to $|2\rangle$ is the pump laser and from $|3\rangle$ to $|1\rangle$ is the Stokes laser. This excited level can decay spontaneously and it has a lifetime of a few nanoseconds. This time scale must be compared with the time scale for the laser interactions, which is on the order of a few milliseconds. This implies that any population in the second level effectively vanishes immediately. Hence, the goal is to design a state transfer avoiding any population in the excited state. Within the dipole approximation and the rotating wave approximation one obtains the effective Hamiltonian in Eq. (7.20) [235].



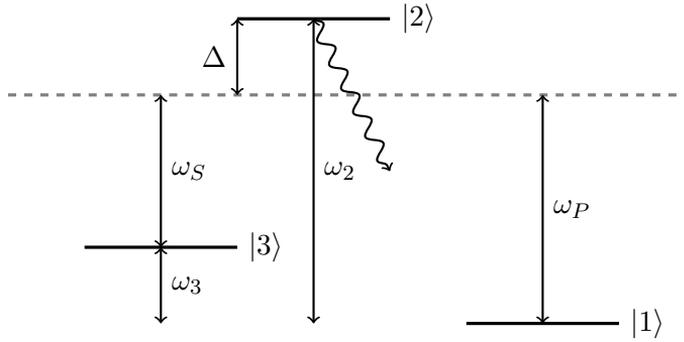


Figure 7.1: This figure shows the energy level diagram in the Λ -system. The pump laser couple $|1\rangle$ and $|2\rangle$ with angular frequency ω_P whereas the Stokes laser couple $|2\rangle$ and $|3\rangle$ with ω_S . Both of the two fields have the same detuning Δ from the second level. The level $|2\rangle$ is subject to fast decay. The state $|1\rangle$ has zero energy and the other states have energies corresponding to ω_3 and ω_2 . Note, there is a two photon resonance meaning $\omega_P = \omega_S + \omega_3$.

7.2.1 Control and Diagonalization of the Λ -System

The special structure of the Hamiltonian in Eq. (7.20) enables an adiabatic transfer from $|1\rangle$ to $|3\rangle$ without any population in the second state. The adiabatic nature of this transfer requires us to know the eigenstates of Eq. (7.20). The eigenstates can be found by a rotation into spherical coordinates,

$$|a^+\rangle = \sin \Theta \sin \Phi |1\rangle + \cos \Phi |2\rangle + \cos \Theta \sin \Phi |3\rangle \quad (7.21a)$$

$$|a^0\rangle = \cos \Theta |1\rangle - \sin \Theta |3\rangle \quad (7.21b)$$

$$|a^-\rangle = \sin \Theta \cos \Phi |1\rangle - \sin \Phi |2\rangle + \cos \Theta \cos \Phi |3\rangle \quad (7.21c)$$

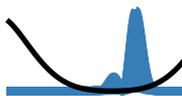
where the angles Θ and Φ are defined by

$$\tan \Theta = \frac{\Omega_P}{\Omega_S}, \quad \text{and} \quad \tan 2\Phi = \frac{\sqrt{\Omega_P^2 + \Omega_S^2}}{\Delta}, \quad (7.22)$$

The eigenstate $|a^0(t)\rangle$ can be written as

$$|a^0(t)\rangle = \frac{1}{\sqrt{\Omega_P^2(t) + \Omega_S^2(t)}} \left(\Omega_S(t) |1\rangle - \Omega_P(t) |3\rangle \right). \quad (7.23)$$

This eigenstate has eigenvalue zero and no component along $|2\rangle$. If the system is in this state during the entire state transfer then it never has a



component along the lossy $|2\rangle$ state. The initial state is $|1\rangle$ and the target state is $|3\rangle$. Note that if only the Stokes laser is turned on then $\Omega_P = 0$ and $\Omega_S \neq 0$ implying $|a^0\rangle = |1\rangle$. In the opposite case where only the pump laser is turned on then $\Omega_P \neq 0$ and $\Omega_S = 0$ implying $|a^0\rangle = -|3\rangle$. If we slowly turn on the pump laser, while slowly turning off the Stokes laser the system will remain in the $|a^0(t)\rangle$ state by the adiabatic theorem. If the Stokes laser is completely turned off while the pump is on then the eigenstate becomes $-|3\rangle$ and the state transfer has happened without populating the $|2\rangle$ state. This is known as Stimulated Raman Adiabatic Passage (STIRAP), which is any control function fulfilling the boundary conditions

$$\lim_{t \rightarrow -\infty} \arctan\left(\frac{\Omega_P(t)}{\Omega_S(t)}\right) = 0, \quad \lim_{t \rightarrow \infty} \arctan\left(\frac{\Omega_P(t)}{\Omega_S(t)}\right) = \frac{\pi}{2}, \quad (7.24)$$

while also satisfying the adiabatic condition during the transfer. This transfer is called the counter-intuitive pulse sequence since the pulses are turned on in the reverse order. This type of transfer has been demonstrated in many experimental applications with some examples being, control of molecular orientation [236], population transfers in rare earth crystals [237], control of ultracold atomic gasses [233] and manipulation of quantum gates [238]. A recent review of the many applications is given in Refs. [234, 239]. However, in a number of applications STIRAP is still fundamentally limited by the amount of time needed to complete the transfer. In the next section, we will discuss how to up speed up STIRAP using shortcuts-to-adiabaticity.

7.2.2 Shortcuts-To-Adibaticity in the Λ -System

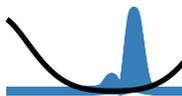
It is possible to reduce the Λ -system to an effective two-level system in the large single-photon detuning limit $\Delta \gg \Omega_S, \Omega_P$. It is possible to apply the shortcut formalism on the three-level Hamiltonian Eq. (7.20) [227], but we will later benchmark with the experiment described in Ref. [233] operating in the large single-photon detuning limit.

For a large detuning it is natural to assume that the population in $|2\rangle$ is constant. Adiabatically eliminating the second level (setting $\dot{c}_2 = 0$ where $|\psi\rangle = \sum_{i=1}^3 c_i |i\rangle$ and solving for c_2 in the Schrödinger equation) gives an effective two-level system [234],

$$\hat{H}_{\text{eff}} = -\frac{1}{2} \begin{pmatrix} \Delta_{\text{eff}} & \Omega_{\text{eff}} \\ \Omega_{\text{eff}} & -\Delta_{\text{eff}} \end{pmatrix} \quad (7.25)$$

where

$$\Delta_{\text{eff}} = \frac{1}{4\Delta} (\Omega_P^2 - \Omega_S^2), \quad \Omega_{\text{eff}} = \frac{\Omega_P \Omega_S}{4\Delta}. \quad (7.26)$$



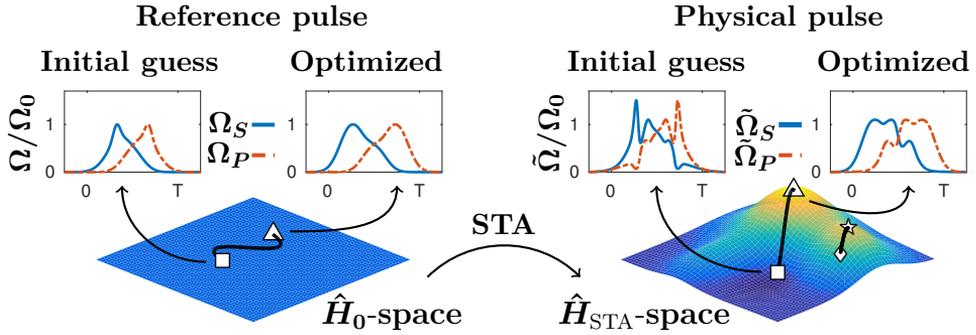
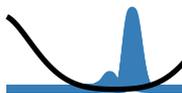


Figure 7.2: Overview of the optimization strategy. An initial guess for the reference Hamiltonian \hat{H}_0 , represented by the laser pulse amplitudes $\Omega_{S/P}$, is mapped by STA to the physical Hamiltonian, \hat{H}_{STA} , and corresponding physical pulses, Eqs. (7.27,7.28). The physical Hamiltonian is then given a score according to the cost functional, Eq. (7.32), designed to reflect the experimental constraints. We search through \hat{H}_{STA} -space to find reference pulses which correspond to a local optimum (white triangle) in \hat{H}_{STA} -space. The details of the optimization process is described in section 7.3.2. The pulses shown result from a single optimization at $T = 0.4$ ms. The white diamond and white star represent an optimization process, where a Gaussian shape was used as an initial guess for the reference pulses. The optimization reaches a local but not a global optimum illustrating the need to do global optimization by multistarting with a range of initial guesses.

The subsequent derivations are performed with this effective two-level system, but the physical system is the three level system and therefore all results are calculated within the full three level system. Note, that Eq. (7.25) is actually the Landau-Zener Hamiltonian from the previous example (7.11) implying that the associate shortcut is given by Eq. (7.19). We now look for the physical pulses, $\tilde{\Omega}_P(t)$ and $\tilde{\Omega}_S(t)$, applied to the original three-level system, that realize \hat{H}_{STA} . We thus solve for the values of $\tilde{\Omega}_P$ and $\tilde{\Omega}_S$ that yield Eq. (7.26) with Ω_{eff} , Δ_{eff} replaced by Ω_{STA} and Δ_{STA} . The result is

$$\tilde{\Omega}_P(t) = \sqrt{2\Delta \left(\sqrt{\Delta_{STA}^2 + \Omega_{STA}^2} + \Delta_{STA} \right)} \quad (7.27)$$

$$\tilde{\Omega}_S(t) = \sqrt{2\Delta \left(\sqrt{\Delta_{STA}^2 + \Omega_{STA}^2} - \Delta_{STA} \right)}. \quad (7.28)$$



For any choice of the reference pulses, $\Omega_P(t)$ and $\Omega_S(t)$, that fulfill the conditions, Eq. (7.24), we can calculate $\Omega_{\text{STA}}(t)$ and $\Delta_{\text{STA}}(t)$ and therefore the physical pulses $\tilde{\Omega}_P(t)$ and $\tilde{\Omega}_S(t)$. As long as the elimination of the excited state remains valid, subjecting the three-level atom to these physical pulses will yield the perfect transfer between the ground states in any finite time interval.

7.3 Combining Shortcuts and Control Theory

In Ref. [233] Eqs. (7.27) and (7.28) were used to find transfer protocols starting from a Gaussian reference pulses $\Omega_P(t)$ and $\Omega_S(t)$. However from the boundary condition (7.24) it is clear that there exist many different reference pulses. In this section we describe Combining STA and Control Theory (COSTACT) exploiting the freedom in choosing the reference pulse and thereby the state trajectory. In this work the control is reference pulses $\Omega_P(t)$ and $\Omega_S(t)$ that are optimized. The method is visualized in Fig. 7.2, where each point in the \hat{H}_0 -space represents a realization of $\hat{H}_0(t)$, mapped by STA to a time-dependent Hamiltonian $\hat{H}_{\text{STA}}(t)$ shown as a point in the \hat{H}_{STA} -space. The cost function is indicated as the vertical dimension in \hat{H}_{STA} -space. We search in \hat{H}_0 -space for the trajectory, leading to the physical control Hamiltonian in \hat{H}_{STA} -space that optimizes the cost. This type of approach has been applied in a two-level system, where the state trajectory was parameterized in order to find optimally robust solutions [240]. Previous work has also combined STA and optimal control theory for studying e.g. atom transport [42, 241, 242]. We shall primarily be concerned with the duration and the energy requirements associated with application of strong control fields [243, 244].

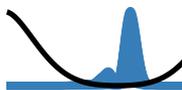
7.3.1 Cost functional

For any choice of the reference pulses that fulfill the boundary conditions, STA provides the corresponding physical pulses through Eq. (7.27) and (7.28). However, the physical pulses might violate constraints set by the experiment. The constraints considered here include the peak intensity of the lasers and robustness against a scaling of the control parameters and are based on the experiment reported in Ref. [233].

The peak intensity of the laser can be included in the cost functional as the dimensionless quantity

$$\tilde{\Omega}_{\text{peak}} = \max\{\tilde{\Omega}_{S/P}(t)\}/\Omega_0, \quad (7.29)$$

where we choose $\Omega_0 = 2\pi \cdot 5$ MHz to define the scale. Minimizing $\tilde{\Omega}_{\text{peak}}$ is equivalent to minimizing the peak intensity since $I_{S/P} \propto \tilde{\Omega}_{S/P}^2$.



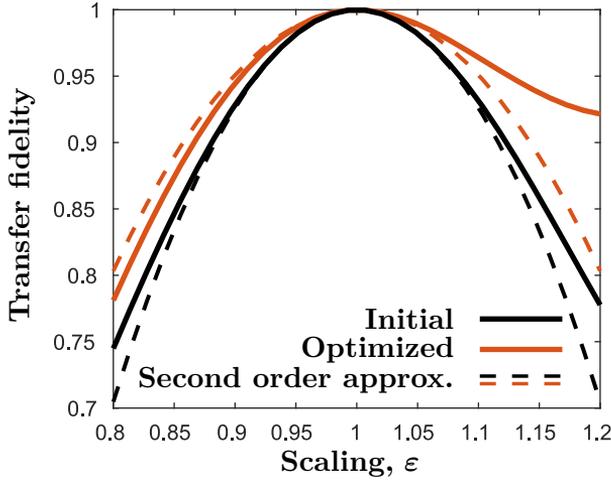


Figure 7.3: Transfer fidelity shown as a function of the scaling parameter ε on the field amplitudes $\tilde{\Omega}_{P/S} \rightarrow \varepsilon\tilde{\Omega}_{P/S}$. The numerical and second order approximation to the fidelity (see Eq. (7.30)) is plotted with solid and dashed lines. Optimization reduces the effect of scaling on the transfer fidelity, cf. the difference between the (lower) black and the (upper) red curves.

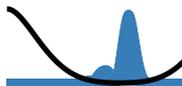
In experiments with many atoms, the spatial laser profile causes different atoms to experience different laser powers depending on their location. Effectively, this corresponds to a random scaling of the Rabi frequency, $\tilde{\Omega} \rightarrow \varepsilon\tilde{\Omega}$, where $\varepsilon \approx 1$. In the adiabatic limit, this scaling does not alter the state transfer. However, when we apply STA the values of the time dependent Rabi frequencies of the physical pulses are important and the scaling reduces the transfer fidelity, as shown in Fig. 7.3. We thus seek solutions which are robust against this perturbation. The sensitivity towards amplitude scaling can be quantified by perturbation theory [240]. To second order in $\varepsilon - 1$, the correction to the transfer fidelity is found to be

$$F \approx 1 - 4q(\varepsilon - 1)^2 \quad (7.30)$$

where we define the *sensitivity*,

$$q = \left| \int_0^T e^{i[\xi_0(t) - \xi_-(t)]} \langle a_-(t) | \hat{U}^\dagger(t) \hat{H}_{\text{STA}}(t) \hat{U}(t) | a_0(t) \rangle dt \right|^2, \quad (7.31)$$

where $\hat{U}(t)$ is given by Eq. (7.18), the states $|a_i(t)\rangle$ are the eigenstates of the Λ -system Eqs. (7.21a)-(7.21c) and ξ_i is the given by Eq. (7.7). By minimizing q we minimize the sensitivity for variations in intensity.



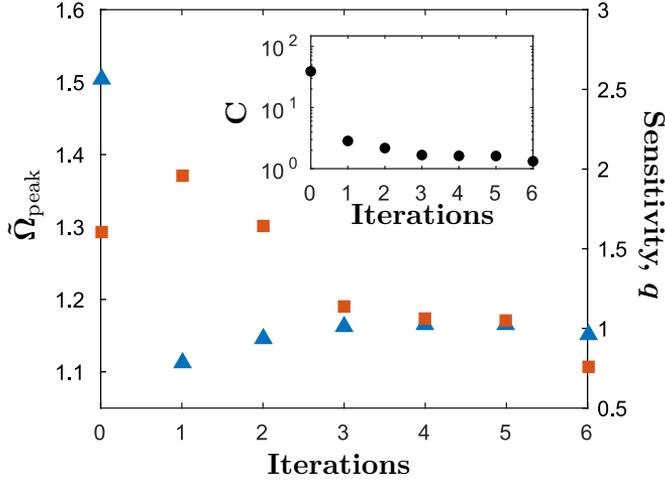


Figure 7.4: Optimization process for $T = 0.4$ ms. The sensitivity q is represented by the squares and the right axis while the triangles and the left axis represents the peak Rabi frequency $\tilde{\Omega}_{\text{peak}}$. The cost functional in Eq. (7.32) (insert) is monotonically decreasing with each iteration, while $\tilde{\Omega}_{\text{peak}}$ and q settle towards a balanced minimum.

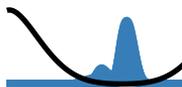
To penalize solutions with large peak Rabi frequency, $\tilde{\Omega}_{\text{peak}}$, and large values of the sensitivity, q , we introduce a cost functional. We define our goal based on the peak value and sensitivity found in Ref. [233]. Here, two Gaussians are used for the reference pulses at a process duration of $T = 0.4$ ms, and for such Gaussians we have $\tilde{\Omega}_{\text{peak}} = 1.14$ and $q = 1.59$. We seek to match these values at the lowest possible duration. We heuristically find that a cost functional defined as

$$C = \exp \left[10 \left(\tilde{\Omega}_{\text{peak}} - 1.14 \right) \right] + \exp [2 (q - 1.59)] \quad (7.32)$$

represents a balanced minimum of $\tilde{\Omega}_{\text{peak}}$ and q in accordance with our goal. For any initial reference pulses (control) Ω_P and Ω_S we can then readily find the corresponding STA pulses using Eqs. (7.27)-(7.28) and then calculate the cost using Eq. (7.32).

7.3.2 Parametrization of a family of reference pulses

As any two functions fulfilling the boundary conditions, Eq. (7.24), can be chosen for the reference pulses, it is difficult to search the entire \hat{H}_0 -space. In Ref. [233] the authors parameterize the reference pulses as partially overlapping Gaussians. We hence restrict the search to smooth and time



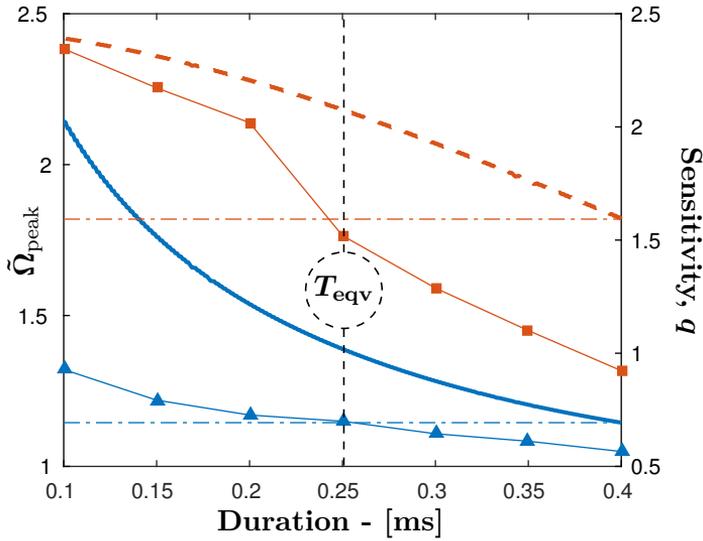


Figure 7.5: This figure presents the result of optimizing the STA pulse sequences according to the cost functional, Eq. (7.32). Red squares show the optimized sensitivity and blue triangles show the optimized peak Rabi frequency. The dashed-dotted lines show the target values of $\tilde{\Omega}_{\text{peak}} = 1.14$ and $q = 1.59$ obtained in Ref. [233] by using a single Gaussian for the reference pulses at $T = 0.4$ ms. We look for the fastest solutions with $\tilde{\Omega}_{\text{peak}}$ and q equal to (or below) these values and denote the corresponding duration as T_{eqv} . Such a solution is found at $T_{\text{eqv}} = 0.25$ ms by the optimization procedures described in Section 7.3.2. The thick blue line and red dashed line show for comparison the results from using a Gaussian reference pulse at different durations.

symmetric solutions $\Omega_P(t) = \Omega_S(T - t)$. We choose to define

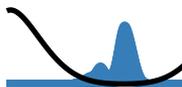
$$\Omega_P(t) = \Omega_0 f(t - T/2 - T/10), \quad (7.33)$$

$$\Omega_S(t) = \Omega_0 f(-t + T/2 - T/10). \quad (7.34)$$

The parametrization in Ref. [233] is extended by choosing the parametrization function, $f(t)$, as a sum of Gaussians,

$$f(t) = A \left(e^{-t^2/(T/6)^2} + \sum_{n=1}^N a_n e^{-(t-t_n^0)^2/(w_n)^2} \right), \quad (7.35)$$

where A is chosen such that $\max\{f(t)\} = 1$. Here, the amplitude, offsets and widths ($\{a_n, t_n^0, w_n\}$) are the control parameters. We found $N = 4$ to offer good solutions.



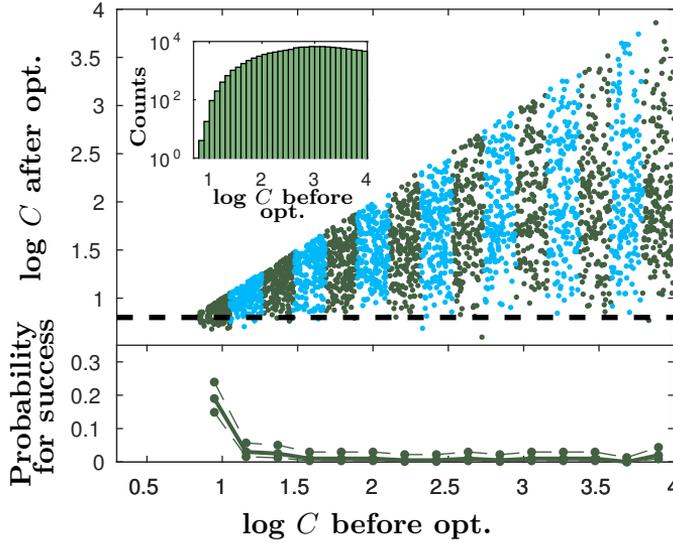
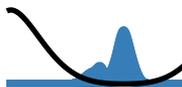


Figure 7.6: The value of the cost function before and after optimization of 3000 seed reference pulses of 0.25 ms duration. Blue and green colors indicate sorting of the data in 15 bins with 200 seeds in each bin. We define successful solutions as having a cost value below $C_{\text{success}} = 0.8$ (fat dashed line) after optimization. The probability for finding a successful solution is substantially higher for reference pulses with a low initial cost. The insert shows a histogram of cost values before optimization for 1 million seeds. The seeds are generated by randomizing the control parameters, $\{a_n, t_n^0, w_n\}$, in a suitable interval. The low-cost seeds are seen to be rare. However, the generation and evaluation of a new seed is computationally inexpensive compared to an optimization. This suggests a trade-off between the total number of generated reference pulses and the fraction that we choose to optimize.

At this point the problem is reduced to finding the set of control parameters, $\{a_n, t_n^0, w_n\}$, that minimizes the cost functional, C . This is done by locally optimizing several initial guesses, or seeds, for the parameters. The seeds are constructed by choosing random values in a suitable interval. An optimization routine is then employed to iteratively update the control parameters until C is locally minimized, as shown in Fig. 7.4. The optimal physical pulses are then constructed from the optimized control parameters. It is necessary to generate and optimize multiple seeds, since the optimization landscape often contains multiple local optima, as illustrated in Fig. 7.2.

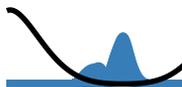


The CRAB, GRAPE and Krotov algorithms are widely used as local optimizers on multiple seeds in traditional quantum optimization approaches [55, 104, 115], and have also been used to solve control problems in three-level systems [245]. However, for these algorithms the cost functional involves the transfer fidelity in addition to the terms from the experimental constraints. This makes the evaluation of the cost functional computationally expensive, as it requires the Schrödinger Equation to be numerically solved. This is not required in our approach, as unit transfer fidelity is guaranteed by the STA formalism. Hence, the cost functional, C , can be quickly evaluated and optimized by a standard black-box optimization routine. In our applications, we used the FMINUNC-routine from MATLAB that performs a quasi-Newton optimization with gradients found using finite-differences.

7.3.3 Results

The result of the optimization is presented in Fig. 7.5. The triangles and squares mark $\tilde{\Omega}_{\text{peak}}$ and q respectively, obtained by minimizing the cost functional at the given process duration. The dashed-dotted lines mark the target values of $\tilde{\Omega}_{\text{peak}} = 1.14$ and $q = 1.59$, which are the values from Ref. [233] obtained using Gaussian reference pulses, $f(t) = \exp(-t^2/(T/6)^2)$, with $T = 0.4$ ms. We search for the lowest process duration where we can find equivalent values of $\tilde{\Omega}_{\text{peak}}$ and q . Both $\tilde{\Omega}_{\text{peak}}$ and q decrease as the process duration increases and the lowest duration where our target is met is found to be $T_{\text{eqv}} = 0.25$ ms. That is, we find solutions that are nearly twice as fast compared to Ref. [233] without compromising energy consumption or robustness. The thick and thick-dashed lines show $\tilde{\Omega}_{\text{peak}}$ and q obtained using Gaussian reference pulses for process durations lower than $T = 0.4$ ms. The optimized pulses perform significantly better at all durations compared to the Gaussian pulse.

In Fig. 7.6 we plot cost values before and after optimization for 3000 seeds. The figure shows that seeds with low cost yield good results more often than seeds with high cost. A seed is taken to be successful if the cost after optimization is below the criterion value, $\log C_{\text{success}} = 0.8$. The probability for finding a successful seed for each bin suggests that only seeds with low cost should be optimized. Such low-cost seeds are rare (see insert in Fig. 7.6). However, generating and evaluating a seed is computationally inexpensive, and to produce the data points in Fig. 7.5 we chose to generate 5 million seeds and optimized only the 1000 of them with lowest cost.



7.4 Conclusion

We have proposed an optimization strategy for solving state-to-state quantum control problems. Our strategy combines the shortcut-to-adiabaticity formalism and minimization of a cost functional incorporating resource requirements and a perturbative expression for the robustness. Unlike traditional quantum optimal control algorithms our cost functional does not include the transfer fidelity and is therefore computationally inexpensive to evaluate and optimize. We have demonstrated the capability of our strategy on a control problem in the three-level Λ -system. Here we find solutions that are almost twice as fast as the solution reported in Ref. [233], while still obeying experimental constraints. The calculation leading to the perturbative expression for the robustness, Eq. (7.30), can be carried out for any perturbation, and, in principle, to any order [240]. This makes our approach especially well-suited for finding solutions that are robust against perturbations. Traditional optimal control algorithms do not immediately offer these solutions, since unit fidelity is required to construct such perturbative expressions.

Our strategy can be applied and is efficient for any system where STA is applicable, i.e., in systems where the adiabatic eigenstates and the counter diabatic driving Hamiltonian can be easily evaluated and applied. A related method makes use of the so-called Lewis-Riesenfeld invariant to determine the time dependence of Hamiltonian parameters that lead with certainty to a desired final state [229]. In analogy with the mapping in our Fig. 7.2, we may use the parametrization of the invariant $I(t)$ as our starting point in an ' I -space' and as long as the associated $H(t)$ is easy to obtain, we may efficiently try several candidate $I(t)$ and optimize for robustness and energy costs. Examples of such systems include two- and three-level systems [42, 233, 246], atoms in harmonic traps [229, 241, 247], quantum many-body systems [248, 249] and quantum heat engines [250–252].

We believe that our approach is a valuable addition to the arsenal of quantum optimal control algorithms, and especially for control problems that require robust solutions.

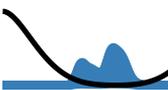


8 Measurement Based Control

The introduction in this chapter and Sec. 8.1 have several fragments copied from [1]. All the text in Sec. 8.2 and Sec. 8.3 is identical to Refs. [1] and [9] respectively. The chapter conclusion in Sec. 8.4 is a merger between the conclusions from Refs. [1] and [9].

In the previous chapter we have discussed how solve the state-to-state problem by engineering the Hamiltonian. Tailored manipulation of the system Hamiltonian is known as *unitary control* due to the unitarity of the dynamics. This type of control can produce any state when the system is controllable [32]. Controllability may be assessed using the Lie-algebra rank condition from Sec. 3.8 showing that state preparation is feasible in a bang-bang fashion by switching between at least two non-commuting Hamiltonians as depicted in Fig. 8.1(a). It is also possible to control a system by decomposing the unitary transformation into two-level unitaries [253] Fig. 8.1(b). Finally, it is possible to design a control pulse using one of the many quantum control algorithms discussed in the previous chapters.

However, in quantum mechanics a state can also be affected by the back-action induced by quantum measurements. This suggests a radically different control paradigm where the state is manipulated by the quantum back-action instead of the system Hamiltonian. Within this paradigm unitary control is no longer required giving rise to the name *control-free control* [254]. This type of control does not have any classical analogue. There has been a number of proposals for using control-free control to manipulate qubit systems [254–258]. Most of these papers prepare a quantum state by either measuring a sequence of varying operators known as Multiple Evenly Distributed Observables (MEDO) or alternate between two measurement operators in bang-bang manner known as



Mutually Unbiased Measurements (MUM). In many-body systems it has been proposed to induce spin correlations [259] by measuring a sequence of standing wave probes [260].

In Sec. 8.1 quantum measurements and control-free control is briefly reviewed. Sec. 8.2 introduces the method *Fixed Unitary Dynamics and Measurements* (FUMES) that combines measurement based control and unitary dynamics. FUMES shows improved performance when benchmarked against MEDO and MUM and it is utilized for preparation of Schrödinger cat states. Sec. 8.3 introduces a combination of FUMES and quantum Zeno dynamics that achieves an exponential speedup compared to regular FUMES. This section also discusses the preparation of a Mott state using FUMES and quantum Zeno dynamics.

The content of Sec. 8.2 was published in Ref. [1] that I co-authored. I had a secondary role in the project and contributed with developing the method and discussing the results. The content of Sec. 8.3 was published in Ref. [9] where I am first author. The paper is accepted for publication in Physical Review A. In this project, I made major contributions to all aspects.

8.1 Control-Free Control

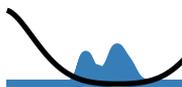
The goal in control-free control is to manipulate a quantum state using only measurements. Physical measurements are associated in quantum mechanics with special Hermitian operators \hat{Q} known as observables. An observable is always diagonalizable due to the spectral theorem

$$\hat{Q} = \sum_j q_j |\varphi_j\rangle\langle\varphi_j|, \quad (8.1)$$

where the q_j s are real numbers corresponding to different possible measurement outcomes. If we measure a q_i , then the state is projected into the eigenspace of \hat{Q} with eigenvalue q_i

$$|\psi\rangle \rightarrow \frac{\hat{P}_{q_i}|\psi\rangle}{\sqrt{\langle\psi|\hat{P}_{q_i}|\psi\rangle}}, \quad (8.2)$$

where \hat{P}_{q_i} is a projector onto the eigenspace with eigenvalue q_i . This change is discontinuous jump of the wave function known as the wave function collapse. This is consistent with the Born rule, which states that eigenvalue q_i is observed with probability $P(\hat{Q} = q_i) = |\langle\psi|\hat{P}_{q_i}|\psi\rangle|^2$. In the special case where an observable has a non-degenerate spectrum, i.e. $q_i \neq q_j$, then a measurement outcome of q_i corresponds to a single



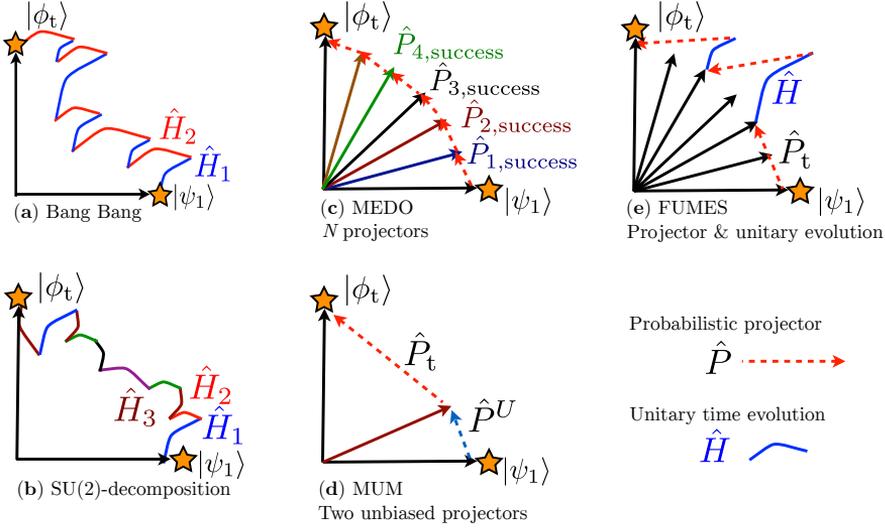


Figure 8.1: (color online) Paradigms for quantum state engineering. (a) Bang-Bang control via two non-commuting Hamiltonians. (b) SU(2)-decomposition of the unitary evolution into $d(d-1)/2$ two-level unitaries. (c) MEDO: multiple measurements successively bring the state closer to the target. (d) MUM: two mutually unbiased observables are measured alternately. (e) FUMES alternates the unitary evolution with measurements at optimized moments in time.

state projection $\hat{P} = |\varphi_i\rangle\langle\varphi_i|$. This is taught in introductory quantum mechanics courses [161].

The dynamics introduced by Eq. (8.2) is non-unitary unlike the time-evolution governed by the Schrödinger equation. The idea in control-free control is to steer the quantum state using repeated projective measurements. One approach uses multiple evenly distributed measurements (MEDO) [256, 258]. The idea is to gradually steer the system into the target state by projecting on a series of intermediate states $\{|\varphi_j\rangle\}_{j=1}^r$, which is illustrated in Fig. 8.1(c). The intermediate states are defined by

$$|\varphi_j\rangle = \cos\left(\theta - \frac{j\theta}{r}\right)|\psi_t\rangle + \sin\left(\theta - \frac{j\theta}{r}\right)|\psi_t^\perp\rangle, \quad (8.3)$$

where θ is the angle between the target state and initial state. $|\psi_t^\perp\rangle$ is a state orthogonal to both the initial state and target state. These projections may be realized with measurement operators of the form $\hat{Q}_j = q_j|\varphi_j\rangle\langle\varphi_j| + q_j^\perp|\varphi_j^\perp\rangle\langle\varphi_j^\perp|$, where $|\varphi_j^\perp\rangle$ is a state orthogonal to $|\varphi_j\rangle$ and we assume no degeneracy $q_j \neq q_j^\perp$. The transfer is successful if we at



each step project into $|\varphi_j\rangle$, which has the total probability

$$P_{\text{MEDO}} = \prod_{j=1}^r |\langle \varphi_j | \varphi_{j-1} \rangle|^2 = \prod_{j=1}^r \cos^2\left(\frac{\theta}{r}\right) = \cos^{2n}\left(\frac{\theta}{n}\right). \quad (8.4)$$

This probability always converges to unity as $n \rightarrow \infty$ independently of the size of the underlying Hilbert space. If the observables are not necessarily equally distributed, the success probability becomes $P = \prod_{j=1}^r \cos(v_j)^2$ with $\sum_{j=1}^r v_j = \theta$. This probability is maximized for $v_j = \theta/r$, which is the MEDO protocol [1]. It was shown in Ref. [256] for a two-level system that MEDO is the optimal measurement sequence for state preparation.

An alternative to MEDO is Mutually Unbiased Measurements (MUM) [261]. Here we only consider two observables with a non-degenerate spectrum $\hat{Q}_1 = \sum_{j=1}^d q_j |\psi_j\rangle\langle\psi_j|$, where $|\psi_t\rangle \in \{|\psi_j\rangle\}$ and d is the size of the Hilbert space. The other observable is $\hat{Q}_2 = \sum_{j=1}^d p_j |\varphi_j\rangle\langle\varphi_j|$. In MUM we alternate between measuring \hat{Q}_1 and \hat{Q}_2 until the target state is reached Fig. 8.1(d). As we now argue the optimal choice is that $\{|\psi_j\rangle\}$ and $\{|\varphi_j\rangle\}$ are mutually unbiased i.e. $|\langle\psi_j|\varphi_k\rangle|^2 = 1/d$ for all j and k . Consider starting from some initial state $|\psi_j\rangle \neq |\psi_t\rangle$. The probability that MUM converges after first measuring \hat{Q}_1 and then \hat{Q}_2 is

$$\sum_k |\langle\psi_j|\varphi_k\rangle|^2 |\langle\varphi_k|\psi_t\rangle|^2 = \mathbf{v}_j \cdot \mathbf{v}_t, \quad (8.5)$$

where $\mathbf{v}_j = (|\langle\psi_j|\varphi_1\rangle|^2, |\langle\psi_j|\varphi_2\rangle|^2, \dots, |\langle\psi_j|\varphi_d\rangle|^2)$. For additional measurements more similar terms appear. In order to maximize the transfer probability, the vector product should be as large as possible, which occurs if the two vectors are parallel. This leads to the conclusion that $|\langle\varphi_k|\psi_j\rangle|^2 = |\langle\varphi_k|\psi_l\rangle|^2 = 1/d$ for all k, j, l , since the entries in the vectors must be positive and sum to unity [261]. The probability of success after each measurement is $1/d$ and therefore the probability of success after n measurements is

$$p_{\text{MUM}}(r, d) = 1 - \left(1 - \frac{1}{d}\right)^r, \quad (8.6)$$

i.e. the probability of r consecutive failures $1 - p(r)$ decays exponentially on a scale defined by d . Unlike Eq. (8.4) this probability does depend on the size of the Hilbert space.

In both MUM and MEDO there is assumed no system dynamics between the measurements. In the next section we discuss an alternative that combines measurements and unitary dynamics.



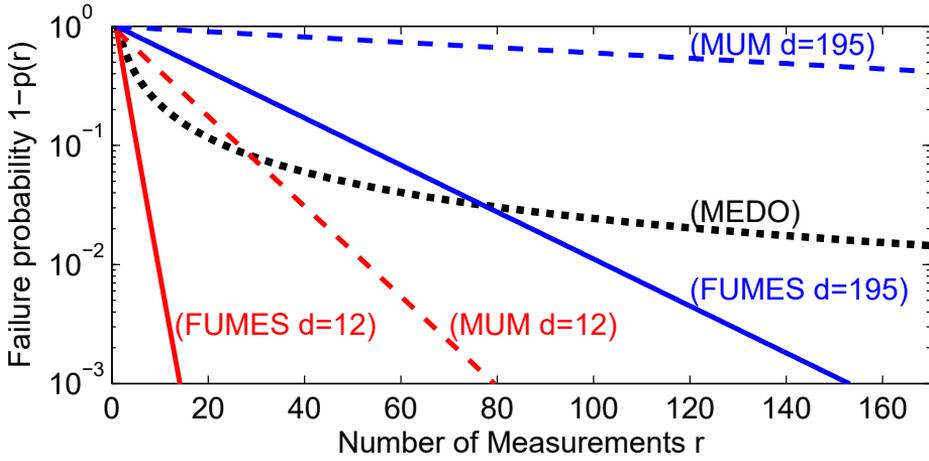


Figure 8.2: Probability for not reaching a randomly chosen target state after r measurements. MEDO (black dotted line) performs independently of dimensionality. Solid lines depict the average trajectory of 100,000 random Hamiltonians with FUMES for dimension $d = 12$ (red) and $d = 195$ (blue); dashed lines denote MUM for $d = 12$ (red) and $d = 195$ (blue).

8.2 Fixed Unitary Dynamics and Measurements

In this chapter, we discuss a hybrid of unitary and control-free control by investigating quantum state engineering via optimized *Fixed Unitary Evolution and Measurements* (FUMES), which imposes fairly weak requirements on the experimental infrastructure: The system Hamiltonian \hat{H} of finite dimension d is fixed and uncontrollable. The only other means to steer the system is provided by a unique fixed measurement operator. The measurement operator signals whether the target subspace is reached, or into which other subspace the state was projected. A crucial prerequisite for this protocol is the ability to precisely model the unitary many-body evolution and to precisely determine time that maximize the success probability of the subsequent measurement. Even provided with these minimal tools, desired states can be produced under mild assumptions on the measurement operator and the system Hamiltonian. FUMES is benchmarked with MEDO and MUM in this chapter using randomly chosen Hamiltonians.

Let the d eigenstates of the constant Hamiltonian \hat{H} be denoted by $|\psi_1\rangle, \dots, |\psi_d\rangle$. We assume access to some observable \hat{Q} that at least enables us to measure the projection on the target state $\hat{P}_t = |\psi_t\rangle\langle\psi_t|$.



The application of the projector at a random moment in time has a low success probability [262]. Therefore we numerically optimize the moment in time at which a measurement is performed, by maximizing the probability to populate a desired eigenstate of \hat{P}_t . A measurement that does not yield the desired outcome described by \hat{P}_t has *failed* and projects the state one of the M possible outcomes indicating failure $1 \leq M \leq d$. The procedure is repeated until success, i.e. for each failed measurement, a new optimal waiting time is chosen before the next measurement is applied.

Below we will also discuss the more general case of measuring some K -dimensional projector $\hat{P}_t = \sum_{k=1}^K |\varphi_k\rangle\langle\varphi_k|$, where $\hat{H}|\varphi_k\rangle = E_k|\varphi_k\rangle$. For $K = 1$, we retrieve the target state projector $\hat{P}_t = |\psi_t\rangle\langle\psi_t|$ and for $K = d - 1$, the aim is to prepare any state orthogonal to some $|\varphi^\perp\rangle$. For a *binary* measurement ($M = 1$), failure leads to state collapse onto an eigenstate of the $d - K$ -dimensional projector $\hat{Q} = \mathbb{1} - \hat{P}_t$, and only marginal information is gained. A fully *granular* measurement with $M = d - K$ reveals the precise state $|\eta_m\rangle$ that the system collapsed onto. In general, any target state can be reached by FUMES if and only if a failed measurement does not project the system into an eigenstate of the system Hamiltonian.

To prove the general applicability of FUMES independent of a specific physical systems, we benchmark it against MEDO and MUM for randomly chosen initial $|\psi_0\rangle$ and target states $|\psi_t\rangle$. We will discuss an application to a concrete physical system below; for the moment, we choose the Hamiltonian in an unbiased way by sampling from the Gaussian unitary ensemble, which ensures uniform distribution in the space of Hamiltonians with a particular dynamical time-scale [263]. These Hamiltonians have the form $H = A^\dagger + A$ where $A_{ij} = a + ib$ with a, b being drawn from a normal distribution. Random Hamiltonians can be used to model systems such as chaotic systems of single and many particles [264]. In our case, random Hamiltonians provide a system for which the measurement basis is completely unbiased with respect to the eigenstates of the Hamiltonian. A particular example for such a case is the measurement of a Fock-state in the superfluid regime of the Bose-Hubbard Model, as explained below. Our figure of merit is the probability to prepare the target state $|\psi_t\rangle$ after r measurements, $p(r)$. In order to rule out the pathological case in which the target state is very close to an eigenstate, we require $|\langle\psi_m|\psi_t\rangle|^2 > 0.1^d$ for all m , which amounts to neglecting 0.3% of all Hamiltonians for $d = 2$, and even fewer for $d \geq 3$.

The probability of r consecutive failures $1 - p(r)$, i.e. the probability of not yielding the desired target state after r steps, is shown for MEDO, MUM and FUMES in Fig. 8.2. The two latter clearly show exponential scaling,



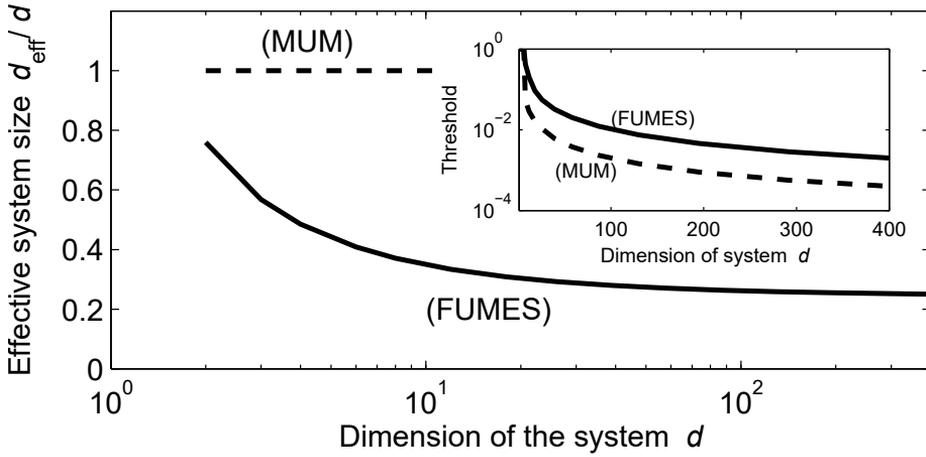


Figure 8.3: Average normalized effective system size Eq. (8.7) for MUM (dashed) and FUMES (solid) from 100,000 random Hamiltonians. (inset) Threshold failure probability below which MUM (dashed) and FUMES (solid) outperform MEDO.

and FUMES consistently outperforms MUM: To understand why, consider a random time-evolution after which the projector \hat{P}_t is applied [262]. Such a procedure yields, on average, a target-state occupation probability $1/d$, just as for MUM. For FUMES, however, each state $|\eta_j\rangle$ characterized by a failed measurement is accompanied by an optimal waiting time that maximizes the probability p_j to populate the target state, which on average is larger than the probability for MUM $p_j > 1/d$. $1 - p_{\text{FUMES}}(r)$ features an exponential decay since the target state can be populated after each measurement with finite probability. This observation allows us to define the *effective dimensionality* d_{eff} via

$$p_{\text{FUMES}}(r) = p_{\text{MUM}}(r, d_{\text{eff}}), \quad (8.7)$$

which quantifies the performance of FUMES with respect to MUM. For a system of dimension d , FUMES requires as many measurements as MUM does for the effective (and smaller) dimension d_{eff} . The effective dimension d_{eff}/d decreases with the system dimension d - see Fig. 8.3(b), because optimizing in a higher-dimensional space is more likely to yield a target-state occupation probability $p_j \gg 1/d$. In the unoptimized case with randomly chosen projection times, the effective dimension becomes $d_{\text{eff}}/d=1$, thus reducing the performance to that of MUM. In other words, even in the presence of substantial errors in the timing of the projection,



FUMES does not perform worse than MUM. This implies that FUMES could be applied to a physical system where the modeling does not need to be exact.

Thanks to the r projectors onto states of the form (8.3), the performance of MEDO is independent of the system dimension. As illustrated in Fig. 8.2, the exponential character of FUMES guarantees that it always outperforms MEDO if the target failure probability is below a certain threshold. For large system dimensions $d > 100$, the target failure probability needs to be below 10^{-2} before FUMES becomes favorable, for which, however, the actual feasibility of MEDO becomes improbable.

8.2.1 Fock-State Generation

We aim at engineering quantum states of multi-mode systems of ultracold bosons described by the Bose-Hubbard Hamiltonian, which was introduced in Sec. 2.5

$$\hat{H} = -J \sum_{j=1}^{L-1} (\hat{a}_j^\dagger \hat{a}_{j+1} + \hat{a}_{j+1}^\dagger \hat{a}_j) + \frac{U}{2} \sum_{j=1}^L \hat{n}_j (\hat{n}_j - 1), \quad (8.8)$$

where J is the inter-well tunnelling and U is the collisional interaction strength - see Fig. 8.4(a). The number of bosons N and lattice sites L determine the system dimension

$$d = \binom{N + L - 1}{N}. \quad (8.9)$$

Here we consider $N = L = 6$, which yields $d = 462$.

As an initial demonstration of control, we assume to be equipped with quantum non-demolition measurements of the local atom-numbers $\{\hat{n}_1, \hat{n}_2, \dots, \hat{n}_L\}$ [260, 265–268] - see Fig. 8.5(c). This leads to a collapse on a single Fock-state. For non-vanishing tunneling $J/U > 0$, Fock-states are not eigenstates of the Hamiltonian Eq. (8.8), and they can be target states for FUMES.

A method for experimentally realizing such single atom detections is by a combination of atomic fluorescence and Raman sideband cooling [269–274]. The concept of Raman sideband cooling is displayed in Fig. 8.4(b) for the $F = 1$ states used in experiments with ^{87}Rb . Besides the internal level structure the system, the vibrational states are modeled by a simple harmonic oscillator. An external magnetic field splits the energy levels of $m_F = -1, 0, 1$ such that the vibrational states with $m_F = 1$ becomes degenerate with $m_F = 1$ and $m_F = -1$. In ordinary Raman sideband cooling the goal is to cool the atoms into the vibrational ground state



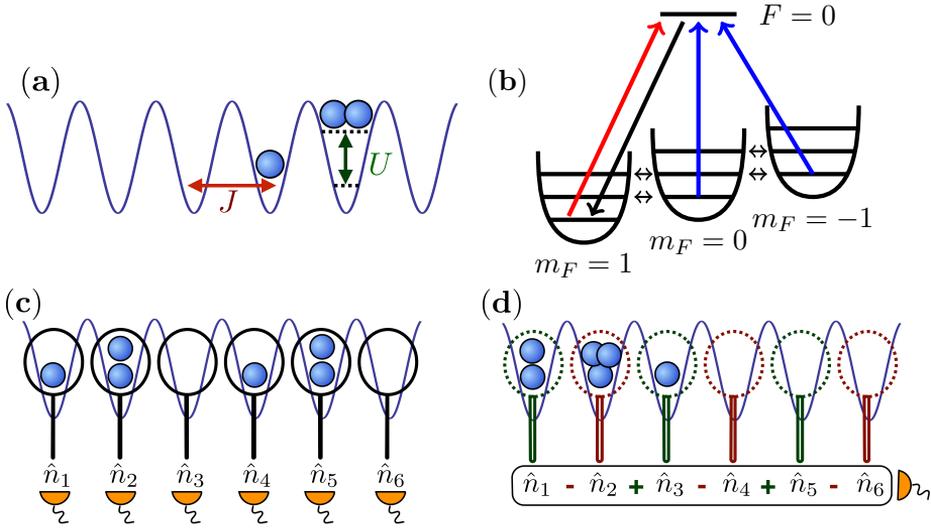


Figure 8.4: **(a)** An illustration of the Bose-Hubbard model with tunneling and on-site coefficients J and U . **(b)** Raman sideband cooling combined with fluorescent imaging. An external magnetic field splits the energy levels allowing for Raman transitions (vertical double arrows). Two pumping beams transition atoms into the bright $F = 0$ states that spontaneously decay back into the $F = 1$ states including the ground state (black arrows). A third beam (red arrow) brings the atoms back into fluorescence. **(c)** Measurement of a single Fock-state by simultaneous measurement of the atom number in each well. **(d)** Measurement of atomic number difference between the even and the odd sites.

for $m_F = 1$. The atom may leave the vibrationally excited states with $m_F = 1$ using Raman transitions (vertical double arrows). The states with $m_F = 0, -1$ are optically pumped (blue arrows) into the excited $|F = 0, m_F = 0\rangle$ state that may spontaneously decay (black arrow) back into the ground state of $m_F = 1$, which ordinarily is a dark state as no transitions away from this state are possible. In order to keep measuring a signal an additional imaging beam (red) brings the atom back to the bright state $|F = 0, m_F = 0\rangle$.

While a Hamiltonian taken from the Gaussian unitary ensemble is structureless, a natural hierarchy of Fock-states emerges for a system governed by the Bose-Hubbard Hamiltonian Eq. (8.8) via the distance



between two states $|\{n\}\rangle \equiv |n_1, \dots, n_L\rangle$ and $|\{m\}\rangle \equiv |m_1, \dots, m_L\rangle$,

$$D = \sum_{k=1}^L \left| \sum_{l=1}^k n_l - m_l \right|, \quad (8.10)$$

which counts the number of tunneling events required to obtain $|\{n\}\rangle$ starting from $|\{m\}\rangle$. Recall, that M denotes the possible number of failed outcomes of a measurement. The distance D motivates an intermediate level of granularity $M = D$ between binary ($M = 1$) and granular ($M = d - K$) measurements, which we refer to as *subspace* measurements.

In anticipation of the experimentally relevant infrastructure discussed below, we illustrate state engineering by FUMES for the target $|\psi_t\rangle = |0, 2, 0, 2, 0, 2\rangle$. The probability $1 - p(r)$ to remain unsuccessful after r measurements is shown in Fig. 8.5 for $J/U = 1.5$, for the three levels of granularity (binary: solid, subspace: dot-dashed, and granular: dashed). For all three granularities, numerical optimization yields an average time between measurements of $2/J$, which corresponds to 0.076 ms for ^{87}Rb in a conventional 512 nm optical lattice [66]. A finer granularity facilitates faster state-engineering, since it gives more detailed information about the current state of the system, which helps to choose the optimal time to apply \hat{P}_t . Cumulating a success rate of 99% with granular measurements thus takes approximately 30 ms.

To set the results into context, we also consider a random Hamiltonian with $d = 462$ and fully granular measurements (dotted). The structure of the Bose-Hubbard Hamiltonian implies that states with large distance D need many tunneling events to be connected, while random Hamiltonians typically couple every pair of states. Therefore, FUMES performs better for random Hamiltonians than for the Bose-Hubbard Hamiltonian.

We also implement bang-bang control [275], alternating the Hamiltonians $\hat{H}_1 = \hat{H}(J/U = 0)$ and $\hat{H}_2 = \hat{H}(J/U = 100)$ (squares). We count the number r of cycles $e^{-i\hat{H}_2 t_2} e^{-i\hat{H}_1 t_1}$ and interpret the fidelity of the target state preparation as the success probability $p(r)$. While the non-commutativity of the two Hamiltonians guarantees controllability, bang-bang control [275] requires more control cycles than measurement cycles needed for FUMES: A measurement in FUMES strongly perturbs the system, which facilitates the fast population of the target state. In practice, a bang-bang control sequence needs to be chosen in advance, i.e. before the actual start of the procedure, while FUMES allows just-in-time optimization: Every time a measurement fails to produce the target state, the timing for the next attempt is optimized.



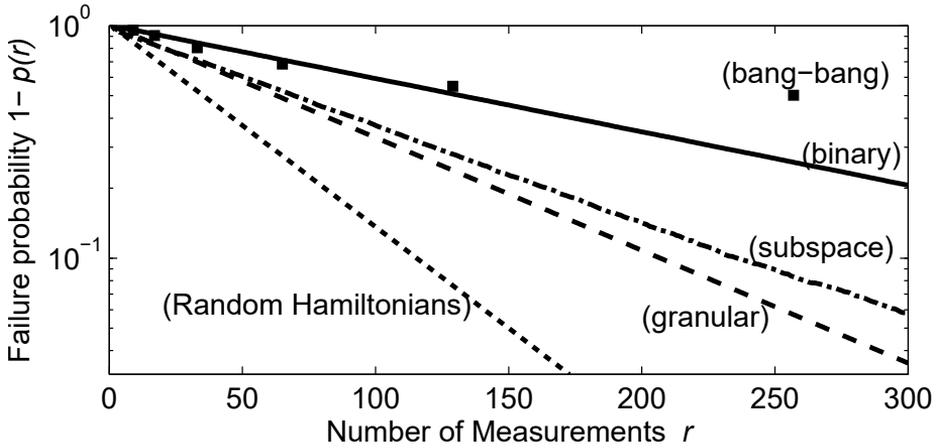


Figure 8.5: Probability of r consecutive failed measurements for the target state $|\psi_t\rangle = |0, 2, 0, 2, 0, 2\rangle$. The initial state is the ground state $|\psi_1\rangle$ of the Bose-Hubbard Hamiltonian Eq. (8.8) with $J/U = 1.5$. Three different granularities steer the state: Granular measurements (dashed) return a specific Fock-state. Subspace measurements (dot-dashed) yield the distance D [Eq. (8.10)] from the target state. Binary measurements (solid) only reveal “failure” and “success”. For comparison, we show the probability for r consecutive failures for a random Hamiltonian (dotted), and the fidelity of state-preparation via bang-bang control as a function of the number of control cycles r (squares).

8.2.2 Schrödinger Cat Generation

As a prominent application of FUMES, we demonstrate the near-deterministic generation of Schrödinger-cat states in the Bose-Hubbard model by the experimentally feasible measurements of the atomic number difference [Fig. 8.4(d)] between the even and odd sites [266, 267],

$$\hat{Z} = \frac{1}{2} \left| \sum_{j=1}^L (-1)^j \hat{n}_j \right|, \quad (8.11)$$

where \hat{n}_j counts the atoms in site j . For vanishing interaction $U \rightarrow 0$, macroscopic superpositions can be generated probabilistically [267], conditioned on a non-vanishing measurement result of \hat{Z} . However, this approach typically yields states of very low *macroscopicity* [276], as quantified by

$$N_{\text{eff}}(|\psi\rangle) = \frac{\max_{\mathbf{w}} (F(\mathbf{w}, |\psi\rangle))}{4L}, \quad (8.12)$$



Table 8.1: Macroscopicity N_{eff} of states sampled from the full Hilbert space and subspaces with specified Z for $N = d = 6$.

Space	Hilbert	$Z = 0$	$Z = 1$	$Z = 2$	$Z = 3$
$\langle N_{\text{eff}}(\psi) \rangle$	1.83(6)	1.6(1)	1.65(9)	2.64(3)	5.9(1)

where $F(\mathbf{w}, |\psi\rangle)$ is the quantum Fisher-information

$$F(\mathbf{w}, |\psi\rangle) = 4\left(\langle\psi|\hat{S}(\mathbf{w})^2|\psi\rangle - (\langle\psi|\hat{S}(\mathbf{w})|\psi\rangle)^2\right), \quad (8.13)$$

and $\hat{S}(\mathbf{w})$ is the measurement operator

$$\hat{S}(\mathbf{w}) = \sum_{j=1}^L w_j \hat{n}_j = \mathbf{w} \cdot \hat{\mathbf{n}}, \quad (8.14)$$

where we restrict \mathbf{w} to $\{\pm 1\}^L$, accounting the impossibility to directly measure superpositions of different particle numbers without auxiliary reservoirs [277]. The macroscopicity $N_{\text{eff}}(|\psi\rangle)$ is the minimal number of particles for which the validity of quantum mechanics is required for a description of the observed macroscopic fluctuations [276]. A Schrödinger-cat state fulfils $N_{\text{eff}} = N$, the ground-state of the Bose-Hubbard Hamiltonian for $J/U = 1.5$ and $d = N = 6$ gives $N_{\text{eff}} = 0.87 \approx N/7$, while single Fock-states do not carry any macroscopic entanglement and yield $N_{\text{eff}} = 0$.

The macroscopicity within subspaces of constant Z has a small spread and is given in Table. 8.1. Thus, finding $Z = N/2$ is necessary to achieve Schrödinger-cat-like macroscopicity characterised by $N_{\text{eff}} \approx N$. However, measuring $Z = N/2$ in the ground state of the Bose-Hubbard model is highly unlikely: For a superfluid state of 100 bosons in 100 sites, the probability is less than 10^{-30} , and even for $N = 6$ particles and $L = 6$ sites with $J/U = 1.5$, the probability remains less than 2%.

Using FUMES, we can exploit the tunnelling dynamics induced by the Hamiltonian to steer the initial state into the otherwise improbable subspace $Z = N/2$ [Fig. 8.6]. The target space $Z = N/2 = 3$ has dimension 56, which facilitates state engineering in comparison to Fock-state generation [Fig.8.5]. Larger values of J/U come with super-fluid-like eigenstates that mix more Fock-states than low values leading to Mott-insulator-like eigenstates. This makes larger values of J/U more favorable, consistent with Fig 8.6. Eight measurements then suffice to accumulate a success rate of more than 80% for $J/U = 1.5$, while the typical time



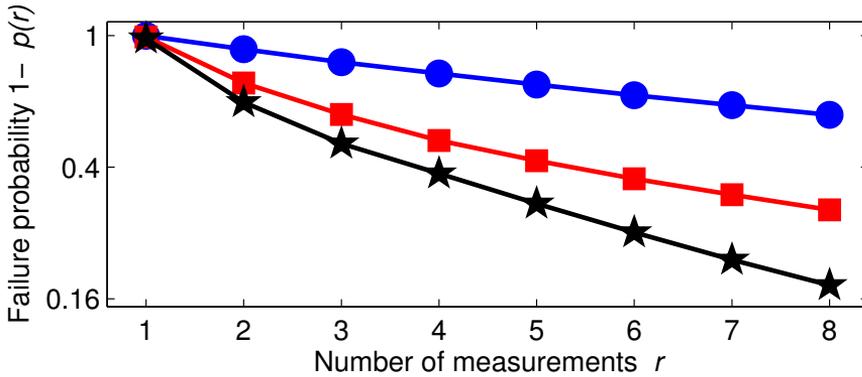


Figure 8.6: (color online) Probability not to yield a state in the Schrödinger cat target-subspace $Z = 3$ after r measurements with FUMES, for $J/U=0.25$ (blue circles), $J/U=0.50$ (red squares), $J/U=1.50$ (black stars).

between two measurements for $J/U = 1.5$ remains around $2/J$, as for the Fock-state target.

8.3 Measurement Based Control With Zeno dynamics

In the remainder of this chapter, we discuss an improvement to FUMES using the quantum Zeno effect describing that a quantum system subject to repeated observations may become frozen in time [278]. This effect is named after the ancient Greek philosopher Zeno due to his arrow paradox, which states that a flying arrow is at rest. Zeno's argument may succinctly be formulated as: At every moment in time the arrow is motionless at some position and the “sum” of such positions at rest is not motion [278]. This paradox is resolved with proper definitions of infinity and infinitesimals in calculus [278], but a similar effect can occur in quantum mechanics. Before discussing how to combine this with FUMES, let us have a brief discussion of the quantum Zeno effect.



Consider some initial state $|\psi_0\rangle$ being periodically projected by \hat{P}

$$\begin{aligned} |\psi(t)\rangle &= \hat{P}e^{-i\hat{H}t/N}\hat{P}e^{-i\hat{H}t/N}\hat{P}\dots\hat{P}e^{-i\hat{H}t/N}\hat{P}|\psi_0\rangle \\ &= \left[\hat{P} \left(\mathbb{1} - iH\frac{t}{N} + \mathcal{O}\left(\frac{1}{N^2}\right) \right) \hat{P} \right]^N |\psi_0\rangle \end{aligned} \quad (8.15)$$

$$\rightarrow \hat{P}e^{-i\hat{P}\hat{H}\hat{P}t}|\psi_0\rangle \quad \text{for } N \rightarrow \infty, \quad (8.16)$$

where it was utilized that a projector is idempotent $\hat{P}^2 = \hat{P}$. If \hat{P} is chosen such that the projected Zeno Hamiltonian $\hat{P}\hat{H}\hat{P}$ becomes diagonal then the state remains frozen in the initial state up to a phase factor [278]. As a simple example, consider a the Landau-Zener system with Hamiltonian

$$\hat{H} = \Omega_R\sigma_x + \Delta\sigma_z, \quad (8.17)$$

and an initial state in the $|0\rangle$ state. For repeated measurements of the projector $|0\rangle\langle 0|$, which is an eigenstate of σ_z , the time-evolved state becomes according to Eq. (8.16)

$$|\psi(t)\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \exp\left(\begin{matrix} -i\Delta t & 0 \\ 0 & 0 \end{matrix}\right) |0\rangle = e^{-i\Delta t}|0\rangle. \quad (8.18)$$

That is, the state is frozen in time up to a phase rotation that also disappears in the special case $\Delta = 0$.

In general, if we measure an observable with a degenerate spectrum we can project on a space spanned by multiple states. In this case, we can observe from Eq. (8.16) that the dynamics become constrained to the Zeno subspace $\hat{P}\mathcal{H}$ where \mathcal{H} is the full Hilbert space. The dynamics within this subspace is unitary and generated by

$$\hat{U}_Z(t) = \hat{P} \exp(-i\hat{P}\hat{H}\hat{P}t), \quad (8.19)$$

which is referred to as quantum Zeno dynamics [278]. It is a bit surprising that repeated non-unitary dynamics on the full Hilbert space induces unitary dynamics within a smaller subspace. The unitarity of Eq. (8.19) hinges on the $N \rightarrow \infty$ limit. If the measurements are performed with some finite rate then probability leaks out of the Zeno subspace with a rate proportional to $(t/N)^2$.

Let us return to the discussion of FUMES. In this method the unitary dynamics is given by a fixed static Hamiltonian while measurements, attempting to project the system into the target state, are performed at the times with highest success probability.

A projective measurement on a many-body system is typically realized by many individual (local) measurements. Even if the full projective



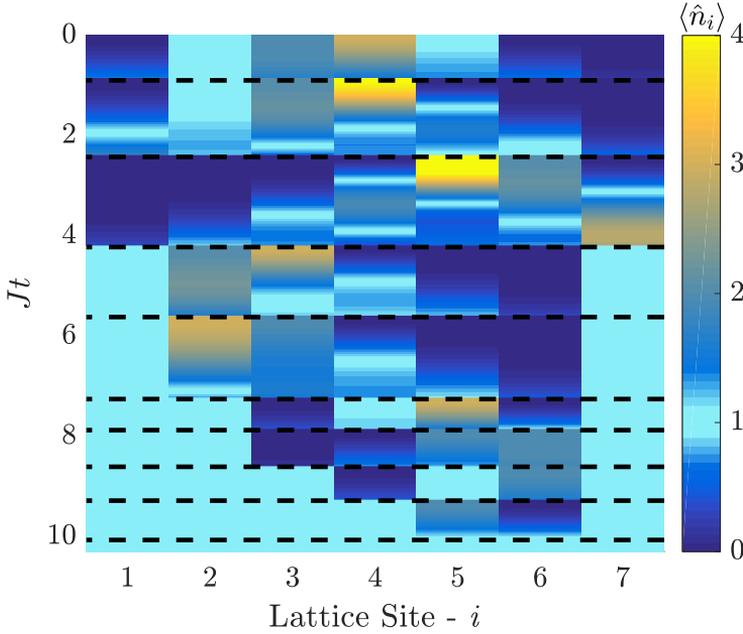


Figure 8.7: A single z-FUMES trajectory starting from the superfluid state and reaching the Mott state for a lattice with $N = 7$ particles and $L = 7$ sites. A measurement is performed at each of the dashed, horizontal lines. The colormap tracks the populations, $\langle \hat{n}_i \rangle$ on the individual sites.

measurement fails to produce the desired outcome, some of the individual measurements might still have succeeded. Despite exhibiting superior performance to MEDO and MUM, the FUMES strategy suffers from the drawback that it cannot maintain these partial successes. In this section, we propose to employ quantum Zeno dynamics to improve the FUMES strategy by freezing the state components prepared by each partial success [279]. This effectively confines the unitary dynamics to smaller Zeno subspaces similar to what was found in Refs. [280, 281]. We demonstrate that this gives an exponential speed-up relative to FUMES.

8.3.1 Mott State Preparation with Zeno Dynamics

In this section we will discuss a state transfer from a superfluid state into the Mott insulator state [119, 282, 283], which is the starting point for applications such as performing quantum logic gate operations [284–288], quantum simulations [289], and single atom transistors [290]. This transfer is difficult since the adiabatic time scales diverge close to the phase



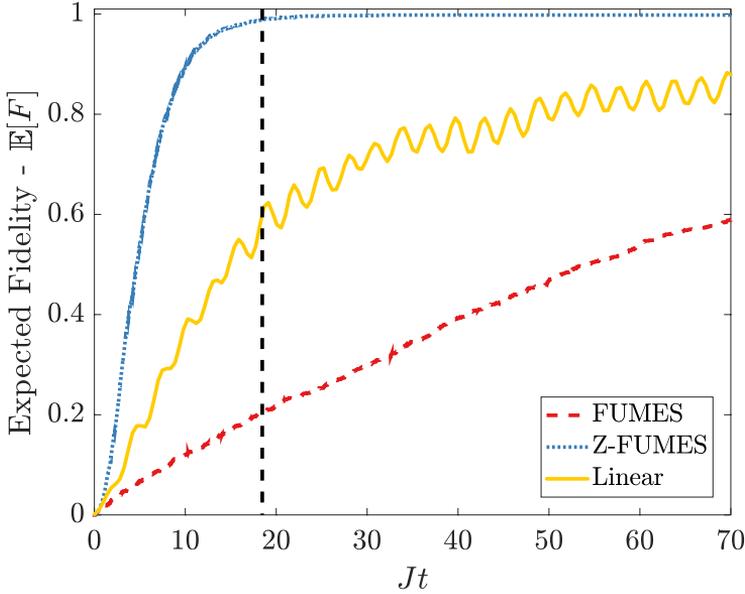


Figure 8.8: Fidelity for FUMES and z-FUMES averaged over 1000 simulated trajectories for $N = L = 7$. The solid, yellow curve shows the deterministic fidelity obtained using a linear ramp without measurements from $U/J = 0$ at Jt zero to $U/J = 30$ at time Jt . The rapid, coherent oscillations are due to the non-adiabatic excitation of the system. The vertical line indicates that z-FUMES reaches $F = 0.99$ at $Jt = 18.5$.

transition where the gap to the excited state closes in an infinite system [291]. There have been attempts to numerically optimize the transfer using optimal control theory and adiabatic ramp shapes [47, 119, 292], and the transition has also been studied and optimized experimentally [47, 282, 293]. Similar to the last section we assume access to quantum non-demolition measurements of the local atom-numbers $\{\hat{n}_1, \hat{n}_2, \dots, \hat{n}_L\}$ for both the projective measurement and the quantum Zeno dynamics - see Fig. 8.4(c).

FUMES is problematic in the sense that for a large lattice it becomes exponentially improbable to project directly into the Mott state. In order to remedy this effect, we propose z-FUMES, which is a combination of FUMES and quantum Zeno dynamics [279]. Quantum Zeno dynamics in optical lattices have been reported experimentally in Refs. [273, 274]. Although a projective measurement of all the sites may not have reached the Mott state some of the individual sites may still have the desired unit occupancy. In order to prevent these particles from tunneling away we



propose to trap them using quantum Zeno dynamics, i.e. by performing rapid repeated measurements of the on-site number operator [279]. *Zeno-locking* the number of particles on a site does not only ensure the correct occupancy, it also prevents particles from tunneling across that site. This implies that locked sites effectively decouple the lattice into smaller parts. However, the Mott state can only be reached if each of these sublattices contain the correct number of particles. Hence, we should only Zeno-lock a given site whenever it contains a single particle *and* the right and left sublattices have matching numbers of sites and particles.

We compare FUMES and z-FUMES for creating a Mott-state in this model for a system with size $N = L = 7$. Between the measurements the evolution of the state is governed by the Hamiltonian (8.8) with $U/J = 0$. The on-site density ($\langle \hat{n}_i \rangle$) during a z-FUMES trajectory is shown in Fig. 8.7. The state is initially in the superfluid state and discrete changes in the density are introduced by projective measurements of the on-site density at specific times marked by dashed lines. At the time $Jt = 4$, the two outer sites have been Zeno locked, creating a sublattice of length five. The edges of this sublattice are gradually Zeno locked in this trajectory, and after about $Jt = 10$, the system has converged to the Mott insulator state with $\langle \hat{n}_i \rangle = 1$ for $i = 1, 2, \dots, 7$. The gradual locking in the proper subspaces is the reason z-FUMES converges faster than FUMES.

In Fig. 8.8 the mean fidelity as a function of Jt is shown for both FUMES and z-FUMES. The curves are obtained by averaging the results of 1000 simulated trajectories. The figure shows that z-FUMES reaches an expected unit fidelity after about $Jt = 20$ while after $Jt = 70$ FUMES still only has a success rate of 60%. For comparison we also show the fidelity after a unitary linear ramp of the interatomic interaction strength from $U/J = 0$ to $U/J = 30$ during the same time-interval but in absence of any measurements. For each value of Jt , the ramp is thus performed with a different speed. The yellow line in Fig. 8.8 shows the final fidelity as a function of the total ramp time. For higher values of Jt the transfer is adiabatic and the fidelity will approach unity. The curve further exhibits characteristic rapid, coherent oscillations, which are due to the energy differences between populated eigenstates during the transfer. Although the fidelity from the linear ramp lies higher than FUMES on the curve, one should remember that the linear ramp never reaches a pure Mott state whereas FUMES leads to formation of the pure Mott state in 60% of the simulated runs. z-FUMES clearly performs better than both the linear ramp and FUMES.



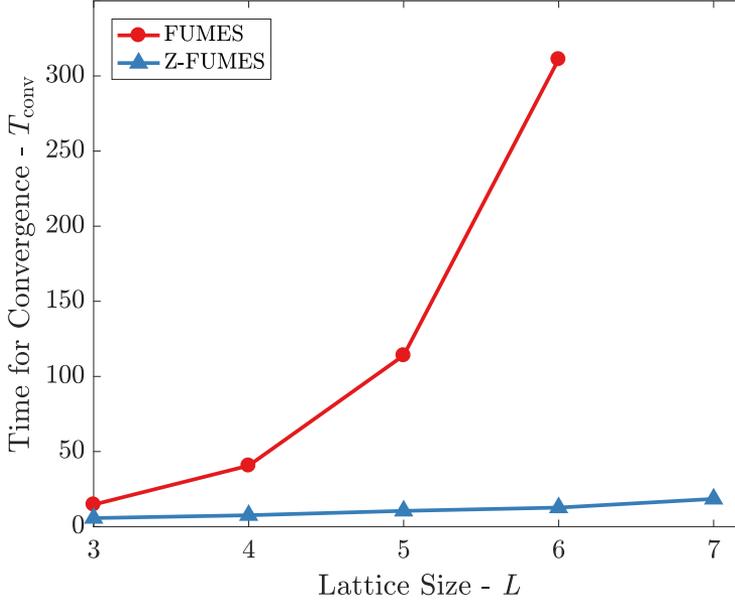


Figure 8.9: Expected Jt needed to reach a fidelity above 0.99 as a function of the lattice size for FUMES and Z-FUMES simulated under the Bose-Hubbard model (8.8). The results are averaged over 1000 trajectories.

8.3.2 Scaling with System Size

In this part, we discuss how FUMES and Z-FUMES scale with the lattice size in the Bose-Hubbard model (8.8). For this purpose, we define T_{conv} as the time where the mean fidelity reaches $\mathbb{E}[F] = 0.99$. The scaling of this quantity with the lattice size L is illustrated in Fig. 8.9 where T_{conv} is averaged over 1000 simulations for each value of L . FUMES scales poorly with the lattice size as it becomes exponentially improbable to project the system into the Mott state. The improved scaling in Z-FUMES is due to the fact that each time a site is locked, the lattice is divided into smaller sublattices each with a higher probability of measuring the desired outcome in subsequent measurements.

Due to the exponential growth of the Hilbert space, it is not possible to simulate Eq. (8.8) for large systems. However, it is possible to perform a toy model analysis by assuming a multinomial distribution. In section 2.5.1 it was discussed how for a superfluid state the number of particles on each site follow a Poissonian distribution with mean $n = N/L$ i.e. the



probability for observing n particles on site i is

$$P_i(n_i) = \frac{e^{-1}}{n_i!}, \quad (8.20)$$

where we assume unit occupancy $n = N/L = 1$. We assume that there is N_1, N_2, \dots, N_L mutually independent Poisson distributions on each site. The probability for the collective distribution for $\mathbf{n} = (n_1, n_2, \dots, n_L)$ is a product $P_{\mathbf{n}} = e^{-L} \prod_{i=1}^L 1/n_i!$. However, there is also the condition that the total number of particles is fixed at $N = L$. This can be included by calculating the conditional probability with the condition $\sum_{i=1}^L n_i = L$

$$\begin{aligned} P_{\mathbf{n}|\mathbf{n}}(N = L) &= \frac{P(\mathbf{n} \cap \sum_{i=1}^L n_i = N)}{P(N = L)} = \frac{e^{-L} \prod_{i=1}^L 1/n_i!}{e^{-L}/N!} \\ &= \frac{1}{n_1! n_2! \cdots n_L!} \frac{L!}{L^L}. \end{aligned} \quad (8.21)$$

We can recognize this result as the multinomial distribution. In each iteration of this toy model simulation of z-FUMES, the outcome of a measurement on each sublattice L_{sub} is sampled from Eq. (8.21) with $L = L_{\text{sub}}$. The first few iterations of one realization within this toy model simulation are shown in Fig. 8.10 where the bars represent the number of atoms on a given site and the white bars show the Zeno locked sites. Here, after just four measurements most of the sites have been Zeno locked.

We shall now discuss how to derive an upper estimate for the scaling of z-FUMES and FUMES with the number of sites using the uniform distribution. In FUMES the collective measurement either projects the state into the target state with $\mathbf{n} = (1, 1, \dots, 1)$ or not. The average number of measurements needed for convergence may then be estimated directly from Eq. (8.21) using Stirling's approximation $n! \simeq \sqrt{2\pi n}(n/e)^n$,

$$M^F = \frac{1}{P(1, 1, \dots, 1)} = \frac{L^L}{L!} \simeq \frac{e^L}{\sqrt{2\pi L}}. \quad (8.22)$$

Before discussing an estimate for the number of measurements needed in z-FUMES, we discuss the probability for fulfilling the condition to lock the the i 'th site P_i . This can be found directly from the simulations using Eq. (8.8). In Fig 8.11 P_i , are estimated from measurements in 600 trajectories with for $N = L = 7$. The numerical results are compared with the same probabilities found using the multinomial distribution in Eq. (8.21) and a uniform distribution where all Fock states are equally likely to be measured. The uniform distribution *underestimates* the probability for locking a specific site, which may be attributed to lockable



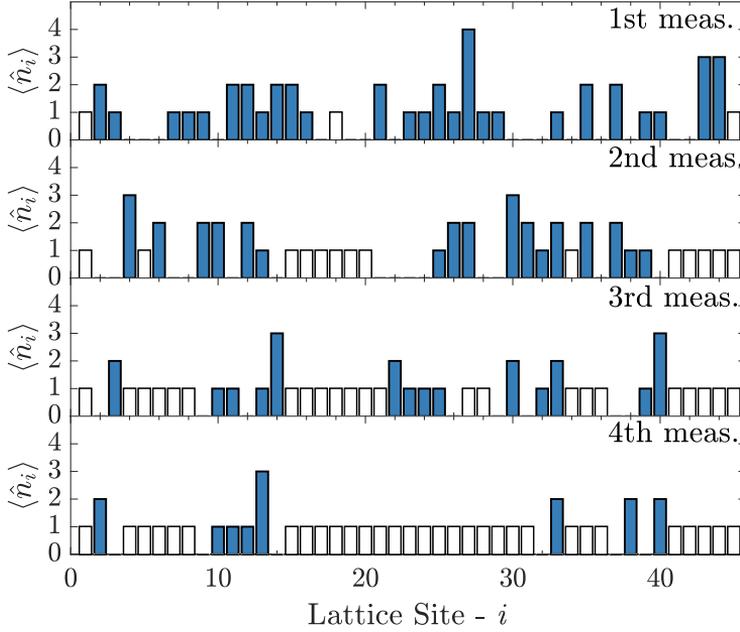


Figure 8.10: First four measurements of a toy model simulation of z-FUMES assuming a complete reshuffling between measurements. Zeno locked sites are shown with white. The distributions of particles, (n_1, n_2, \dots, n_L) after a measurement is sampled from Eq. (8.21) for each sublattice with $L = L_{\text{sub}}$ where L_{sub} is the sublattice length.

configurations with a low number of particles on each site being more probable as evident from Eq. 8.21.

Using the uniform distribution we find a conservative estimate for the scaling of z-FUMES with the number of sites. The probabilities in Fig. 8.11 for the uniform case may be found directly by counting the number of possible configurations allowing a site to be Zeno locked

$$P_i = \frac{C(i-1)C(L-i)}{C(L)}. \quad (8.23)$$

Here we assume unit filling $L = N$ and $C(N) = (2N)!/2N!^2$ is the number of ways to distribute N particles among N sites. In the special cases of $i = 1$ or $i = L$ Eq. (8.23) becomes $P_i = L/2(2L-1)$. The limit, $1 \ll i \ll L$ of Eq. (8.23) can be investigated using Stirling's approximation

$$P_i \simeq \frac{1}{8} \sqrt{\frac{L}{\pi(i-1)(L-i)}}. \quad (8.24)$$



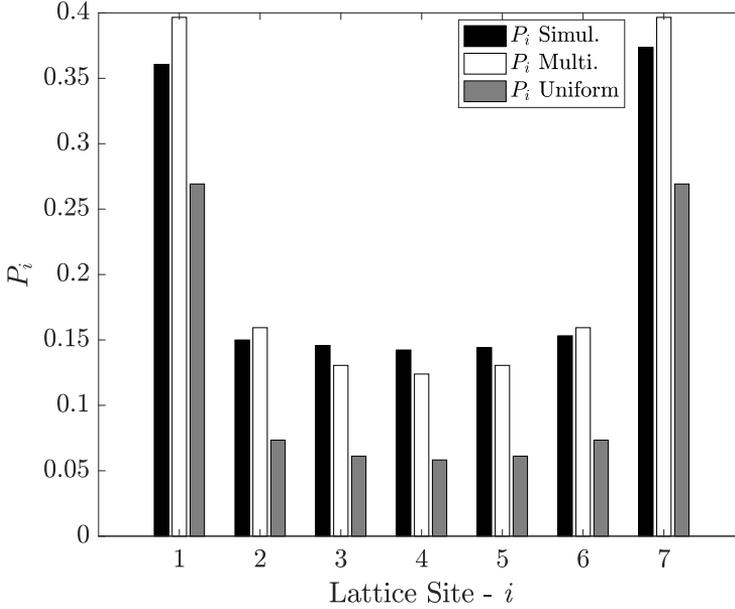


Figure 8.11: Probability distribution P_i for Zeno locking site i obtained from the multinomial distribution (white) Eq. 8.21 and uniform distribution (grey) found by summing over all configurations allowing Zeno locking. The black bars show estimated values of P_i by checking which sites could have been Zeno locked from measurements on 600 trajectories Eq. 8.8.

Seeking an upper estimate on the number of measurements needed for convergence, we compute the average probability for locking a site from Eq. (8.24)

$$P_{\text{avg}} \equiv \frac{1}{L} \sum_{i=1}^L P_i \simeq \frac{1}{2L-1} + \frac{1}{8L} \int_2^{L-1} \sqrt{\frac{L}{\pi(i-1)(L-i)}} di \quad (8.25)$$

where the first term is the edge contribution and the discrete sum is approximated by an integral. Calculating the integral and performing a large L expansion we obtain

$$P_{\text{avg}} = \frac{1}{8} \sqrt{\frac{\pi}{L}} + \frac{1}{2L} - \frac{1}{2\sqrt{\pi}L^{3/2}} + \dots \quad (8.26)$$

Assuming that the probability distribution for locking is binomial with success probability P_{avg} , the average number of sites locked in a chain of length K is $K P_{\text{avg}}$. An estimate on the number of measurements needed



for convergence is found by summing the average time for locking each site

$$M^Z \leq \sum_{K=1}^L \frac{8}{\sqrt{\pi K}} \simeq 16\sqrt{\frac{L}{\pi}}, \quad (8.27)$$

where the discrete sum is approximated by an integral and only the dominant first term from Eq. (8.26) is included. The non-exponential scaling clearly demonstrates the power of continually dividing the system into ever smaller decoupled sublattices. Eqs. (8.22) and (8.27) clearly show that z-FUMES achieves an exponential speed up compared to FUMES, resulting in the favorable scaling seen in Fig. 8.9. The trajectory in Fig. 8.7 simulated with Eq. (8.8) used nine measurements for convergence whereas Eq. (8.27) estimates $M^Z \leq 24$ showing that the bound is not tight.

8.3.3 Continuous Measurements

In the previous sections we have assumed that the measurements occur instantaneously. This is a valid assumption when the typical duration of a measurement is short compared to the dynamics of the measured system [294]. However, this is not generally true in practical applications with finite interaction strengths where the measurement record $I(t)$ is continuous [295]. As a step towards experimental realizability we investigate the performance of z-FUMES with weak continuous measurements.

Due to the backaction of the continuous measurements, the dynamics obeys a stochastic Schrödinger equation, which can be understood as the time evolution conditioned on the stream of measurement results in $I(t)$ [294, 296]. At each instant in time, $I(t)$ is dominated by stochastic noise, resulting in a state $|\psi(t)\rangle$, which evolves in a random manner. This time evolution exhibits quantum jumps if measurement outcomes occur at discrete point in times as in photo detection while for example homodyne detection leads to a diffusive trajectory [296]. Quantum jump trajectories have previously been studied in the context of the Bose-Hubbard model [267, 281, 297–300]. Here we consider diffusion-type measurements which imply a stochastic Schrödinger equation of the form

$$d|\psi(t)\rangle = dt \left[-i\hat{H}(t) + \sum_j -\frac{\gamma_j}{2}\hat{c}_j^\dagger\hat{c}_j + I_j(t)\hat{c}_j \right] |\psi(t)\rangle, \quad (8.28)$$

where the $\hat{c}_j = \hat{n}_j$ are the measurement operators and the γ_j corresponding measurement strengths, which determine the rate at which information is extracted [296, 301]. In the examples studied here, we assume uniform



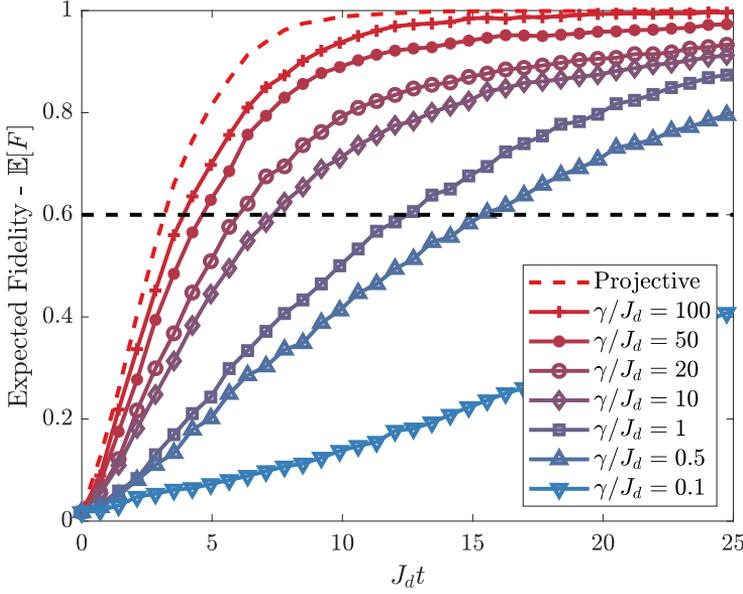


Figure 8.12: Expected fidelity in z-FUMES averaged over 1000 trajectories simulated using continuous homodyne measurements. The simulations are made for different measurement strengths γ and a lattice with $L = 5$ sites. The dashed curve is not simulated using continuous homodyne measurements but with unitary dynamics and projective measurements as in Fig. 8.8. The horizontal dashed lines indicates 60% expected fidelity.

strengths $\gamma_j \equiv \gamma$. A measurement becomes projective in the limit $\gamma \rightarrow \infty$ [294]. Note that Eq. (8.28) does not preserve the normalization, which is instead imposed explicitly in each time step. The measurement record $I_j(t)$ for the j 'th detector reflects the current state of the system,

$$I_j(t) = \gamma_j \langle \hat{c}_j^\dagger + \hat{c}_j \rangle(t) + \sqrt{\gamma_j} \xi_j(t), \quad (8.29)$$

where the $\xi_j(t) = dW_j(t)/dt$ are infinitesimal Wiener increments, representing white noise in the detection setup. Integration of the record allows one to determine the outcome of a measurement for large values of γ . In order to investigate the performance of FUMES and z-FUMES at low values of γ , we assume that it is possible to quench the lattice such $J(t) = J_d$ with $U(t) = 0$ in absence of measurements and $J(t) = 0$ with $U(t) = J_d$ during measurements where J_d is the characteristic energy scale. After a measurement some of the particles may be trapped by the Zeno effect using sufficiently strong measurements to suppress tunneling.



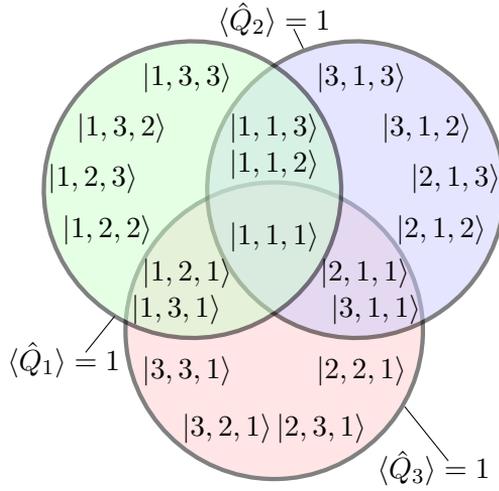


Figure 8.13: An illustration of how the q_j are distributed in a general context. The colored areas show the Zeno subspace corresponding to a value of $q_j^{(i)}$ for a particular \hat{Q}_i . Each Zeno-subspace contains a number of the basis states $|q_1, q_2, \dots, q_L\rangle$. A simultaneous measurement of all the Q_i yields a single state projection, which in this illustration is into the target state $|\psi_{\text{target}}\rangle = |1, 1, 1\rangle$. In this example we have $B = L = 3$.

In our simulations we used $\gamma_j/J_d = 1000$ for Zeno locking of the particle on site j .

In Fig. 8.12, z-FUMES is investigated for different values of the measurement strength in a system with $L = 5$ lattice sites. The figure also shows the expected fidelity for z-FUMES simulated using discrete projective measurements as in Fig. 8.8. As expected, the fidelity of the continuous measurement scheme converges towards that corresponding to projective measurements as the measurement strength becomes large. At the time $J_d t = 15$ more than 60% of the trajectories have converged with $\gamma/J_d = 0.5$, which means that z-FUMES is effective even with moderate values of the measurement strength.

8.3.4 General application of Z-FUMES

In this section we show that FUMES and z-FUMES can also be applied to prepare an arbitrary target state of a system, almost independently of its Hamiltonian evolution. In the last section we gave the example of preparing an individual Fock state by measuring the on-site populations.



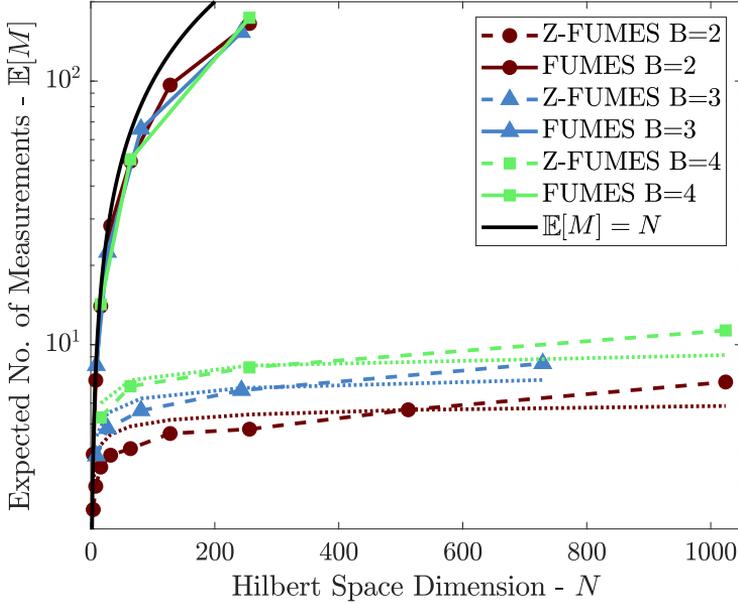


Figure 8.14: Expected number of measurements needed for convergence $\mathbb{E}[M]$ for FUMES and Z-FUMES for systems with different number of measurement operators L and outcomes B - see legend. The dashed line shows the estimate from Eq. (8.30). Each point is averaged over 1000 random Hamiltonians and initial states. The Hilbert space dimension is $N = B^L$ meaning that a larger N corresponds to more measurement operators.

If the measurements were instead performed in the Fourier basis (corresponding to momentum eigenstates) then many-body states with exotic phase correlations could be prepared [1, 259, 299]. Similar to Sec. 8.2, we demonstrate this general applicability in an unbiased way, we here assume random Hamiltonians drawn from the Gaussian Unitary Ensemble which ensures a uniform distribution in the space of Hamiltonians with a particular dynamical time-scale [263].

We assume the system can be manipulated through the backaction from measuring a set of commuting observables $\{\hat{Q}_1, \hat{Q}_2, \dots, \hat{Q}_L\}$, $[\hat{Q}_i, \hat{Q}_j] = 0$. The joint eigenstates of the measurement operators define a set of orthonormal states $|q_1, q_2, \dots, q_L\rangle$ and the target state $|\psi_{\text{target}}\rangle$ must be one of these eigenstates. A measurement of all \hat{Q}_i observables must uniquely determine a $|q_1, q_2, \dots, q_L\rangle$ state. Each measurement operator may be written as a linear combination $\hat{Q}_i = \sum_j q_j^{(i)} |\dots, q_j^{(i)}, \dots\rangle \langle \dots, q_j^{(i)}, \dots|$.



The eigenvalues $q_j^{(i)}$ should be constructed with a large degeneracy such that they define different subspaces - see Fig. 8.13. A measurement of a \hat{Q}_i is successful if it measures the same $q_j^{(i)}$ as for the target state. Zeno-locking this value confines the time evolution to a smaller subspace containing the target state. The subsequent time evolution within this subspace is given by Eq. (8.16) where \hat{P} is the projector on the locked Zeno-subspace [278]. A subspace should only be Zeno-locked if there is a sufficient coupling between the current state and the target state within that subspace, i.e. $|\langle \psi_{\text{target}} | \hat{U}_Z(\Delta t) | \psi(t) \rangle| > \varepsilon_o$ where ε_o is a predefined threshold.

These ideas constitute a direct generalization of the Bose-Hubbard control scheme discussed in the previous sections. There the orthonormal basis consists of the Fock-states and the individual on-site number operators can be written as linear combinations of projectors on these states. We have for instance, that the on-site density operator for the i 'th site is $\hat{n}_i = \sum_j n_j^{(i)} | \dots, n_j^{(i)}, \dots \rangle \langle \dots, n_j^{(i)}, \dots |$. The eigenspectrum of the number operator on a single site is clearly degenerate. Measuring the \hat{n}_i operators one-by-one gradually leads to a collapse onto a single Fock-state. The condition of only Zeno-locking in subspaces with sufficient coupling between the current state and the target state corresponds to exclusively locking sites with matching numbers of particles and sites on the left and right sublattices.

We have applied z-FUMES to perform general state transfers using random Hamiltonians. In order to show the average behavior, we have performed calculations for 1000 different random Hamiltonians and target states. We assume access to L different measurement operators with B different measurement outcomes. Fig. 8.14 compares the number of measurement needed for convergence $\mathbb{E}[M]$ for different number of measurement operators (L) and number of outcomes (B). As in the case of the Bose-Hubbard model, the z-FUMES curves converge much faster than FUMES, requiring about two orders of magnitude fewer measurements. All z-FUMES curves show a similar rate of convergence despite the system size differing by an order of magnitude.

The scaling with the system size may be understood by a simple model. Assuming that all measurement outcomes are equally likely, the probability of a measured \hat{Q}_i producing the desired outcome is B^{-1} . If the measurements are independent, the mean number of locked observables is K/B when measuring K observables. The number of measurements needed to converge may then be estimated by summing the average time



for locking each operator

$$M^Z = B \sum_{K=1}^L \frac{1}{K} \simeq B \ln L. \quad (8.30)$$

This logarithmic scaling is slower than the square root found for the Bose-Hubbard model in Eq. (8.27). In the Bose-Hubbard model there are fewer states in the Hilbert space, but only a few states are lockable due to the constraint that the left and right sublattice must have matching number of particles and sites. In combination this gives a lower probability for locking a site than in the unconstrained case. The value in Eq. (8.30) is compared with the simulations using random Hamiltonians in Fig. 8.14. Fig. 8.14 shows that it is favorable to have fewer possible measurement outcomes, which is also captured by Eq. (8.30). For larger systems, Eq. (8.30) seems to underestimate the number of measurements needed, which may be due to violation of the assumption that the observables are independent.

In FUMES the system either projects into the target state or not and assuming a uniform distribution the expected number of measurements is $M^F = B^L = N$. This linear scaling is plotted with a black line in Fig. 8.14. The FUMES curves follow a linear scaling depending only on N , but the slope is lower than unity. We attribute this to the fact that FUMES only performs the measurements at peaks in the fidelity, but also note that the scaling may change for larger Hilbert space dimensions.

8.4 Conclusion

The minimal requirements for unitary control are the availability of two non-commuting Hamiltonians \hat{H}_1 and \hat{H}_2 giving full controllability [32, 275]. We translated this setup into the realm of control-free engineering: A fixed Hamiltonian \hat{H} and a projector onto the desired target subspace \hat{P}_t permit FUMES to steer quantum states into a target subspace as long as the Hamiltonian dynamics gives an overlap with the target subspace after some finite time. In situations where the Hamiltonian is uncontrollable or no tailored measurements can be applied, FUMES emerges as a natural way to perform state engineering. Thanks to the optimization of the waiting time, it outperforms methods for which the success probability of the measurement scales as $1/d$ [261, 262]. Timing errors may affect the overall success probability of FUMES, since inaccuracies in timing will result in projections at non-optimal moments in time and, consequently, to a lower probability to reach the target. However, timing errors will



not affect the actual fidelity of the preparation, which is entirely defined by the fidelity of the projection operator.

In the second half of this chapter we discussed a combination of FUMES and quantum Zeno dynamics. Our protocol steers the evolution towards a target state by measuring a set of observables. Each observable is Zeno-locked when an appropriate outcome is obtained, which confines the time evolution to gradually shrinking Zeno-subspaces. Z-FUMES gives an exponential speed up compared to other measurement-based control protocols. We analyzed in detail the preparation of a Mott state using Z-FUMES. Here it is necessary to measure the density on each lattice site, which can be realized experimentally using strong fluorescence imaging combined with Raman side-band cooling as shown in Refs. [273, 274]. We demonstrate furthermore that Z-FUMES can be applied for preparing a Mott state in more realistic settings relying on weak continuous rather than projective quantum measurements. We also discussed how to implement Z-FUMES in a more general setting with random Hamiltonians.

Another possible extension, would be to combine FUMES and MEDO, i.e. to optimize, both, the time-evolution and the form of the next projector to be, under the given experimental limitations. Such hybrid protocol would likely saturate the general possibilities for control, exploiting all available resources.

One could also envision to combine the measurement based schemes from this section with the optimal control theory from the last sections whereby one more systematically could search for optimal projectors or projection times perhaps in a gradient-based manner.



9 Conclusion and Outlook

The aim of this work has been to gain insight into quantum optimal control of ultracold atoms especially for single atom transport, manipulation of Bose-Einstein condensates on atom chips and the Bose-Hubbard model. This topic is not only interesting theoretically, but also motivated by the numerous experimental applications in e.g. matter-wave interferometry [83] or quantum simulation [26].

In chapter 3, I give an overview of the main optimization algorithms used in the field. In particular, Gradient Optimization using Parameterization (GROUP) is introduced. It combines a chopped basis expansion with gradient based methods such as GRAPE. In chapter 4, the performance of this algorithm is investigated for a particular problem relating to optimal coherent excitation of a Bose-Einstein condensate. I shown that GROUP gives a faster rate of convergence than Nelder-Mead with CRAB and that GROUP is competitive with the traditional derivative based methods such as Krotov and GRAPE. On other quantum control problems it remains to be explored how GROUP compares to Krotov and GRAPE. Many of these methods are available in our open source C++ library QEngine [8]. We have spent considerable efforts on optimizing and polishing this code, so I am confident that future group members and hopefully other scientists will find it useful. The QEngine will also be the backbone of all quantum games at ScienceAtHome and it has already been included in the next version of Quantum Moves.

Most of these methods could most likely be considerably improved by directly computing the full Hessian and performing second order descent ensuring a quadratic convergence rate towards the minimum. This would most likely be very useful for control problems requiring ultra low infidelities where the last tweaking of the control is very important [125]. It is also possible to study the structure of the quantum control landscape using the Hessian's eigenvalue spectrum [302, 303].

Due to the greedy local nature of these methods they are inherently



prone to being trapped in local minima. Although the complexity of finding optimal controls is a subject of vivid debate, it is generally accepted that traps occur for control problems that are either constrained in time or allowed control amplitudes [124, 165, 166, 168]. This is formulated in terms of the Optimization Landscape Hypothesis in Chap. 5. This hypothesis states that the optimization landscape is benign or trap-free at arbitrarily long durations and it becomes increasingly complex as the duration is decreased. In Chap. 5, I discuss how searching in complex landscapes require the balancing of the exploration-exploitation trade-off. This trade-off motivated our proposal of adding a global search component to the optimization algorithms. This Global-Local algorithm allowed us to find improved solutions to a set of experimentally relevant problems relating to time-optimal control of Bose-Einstein condensates. I also discuss how a control problem with an effective two-level structure always fulfills the a universal \sin^2 -behavior below the quantum speed limit, which has been discovered numerically in a number of problems [140, 162]. I present how this may not be true for control problems involving multiple levels. Due to the numerical nature of this study, it does not definitively proof a non- \sin^2 -behavior. Such a proof could perhaps be found using the time-optimal form of Pontryagin's Maximum Principle for multilevel systems.

In chapter 6, I study optimal control of single atoms as part of an architecture for quantum computation in optical lattices. In this project we used the citizen science game Quantum Moves to recruit users that helped us find good seeds for the optimization. Clustering analysis of the optimized player solutions shows that they split into two different solution strategies - shoveling and tunneling. This shows that when provided with the correct interface, citizen scientists can help provide high quality seeds for the optimization and identify solution strategies or heuristics. The best results were found by numerically creating our own seeds from a low dimensional search space based on these solutions. This shows that naively applying multistarting on an optimization problem may fail and one should consider different heuristics for helping to solve the problem. Furthermore, citizen science games may be a method for finding such heuristics. The visualization and clustering techniques inspired our citizen science work on optimizing experimental Bose-Einstein production. In this work, we were also able to conduct more systematic studies on the players' search methodology, which shows that the players engage in an adaptive search that balances the exploration-exploitation trade-off.

This was my first work during my PhD studies and the conclusions from this work directly inspired my later work. The first conclusion that search in low dimensional spaces is important inspired GROUP. The second



conclusion that an global overview or heuristic is needed for initializing the local algorithms directly inspired the Global-Local algorithm in Chap. 5. Recently, clustering analysis of the optimal solutions have also been used to characterize the complexity of quantum control problems in the vicinity of the quantum speed limit [166].

In chapter 7, I combined shortcuts-to-adiabaticity and regular optimal control for finding shortcuts that maximize the robustness and energy consumption. This method has the advantage that there is no need to solve the Schrödinger equation whereby it becomes possible to research through hundreds of seeds in a few minutes. A natural extension of this work is to apply this method beyond the two level system. For instance, it is possible to design shortcuts for manipulating single atoms or Bose-Einstein condensates, which could complement the optimal control work in this thesis [241, 304, 305].

In chapter 8, I introduced FUMES that is a combination of unitary dynamics and measurements for controlling Bose-Hubbard type dynamics and random Hamiltonians. This method has a faster convergence rate than other measurement based control schemes when benchmarked using random Hamiltonians. This method is also for preparing Schrödinger cat states in a Bose-Hubbard model. We later extended this work using quantum Zeno dynamics that gave an exponential speed up compared to regular FUMES that is required for controlling quantum many-body systems. This speed up enables us to prepare Mott-insulator type states in a Bose-Hubbard model. In this work we did not utilize the mathematical framework of quantum optimal control and I believe significantly more advanced methods could be devised if this approach was taken.

9.1 Outlook

In this section, I describe two projects that are direct extensions of the work presented in this thesis. At the time of writing both these two projects are in preparation with some encouraging early results.

GROUP Optimization with Matrix Product States

The first project is to apply the GROUP optimization to a quantum many-body system. In Chap. 8, we simulated the many-body system using exact diagonalization. The Hilbert space of quantum many-body systems scales exponentially with the number of particles. Therefore simulating even small systems quickly becomes extremely demanding if not impossible with exact diagonalization. In one-dimensional lattice systems, the state can be decomposed in terms of so-called tensor-networks or matrix product



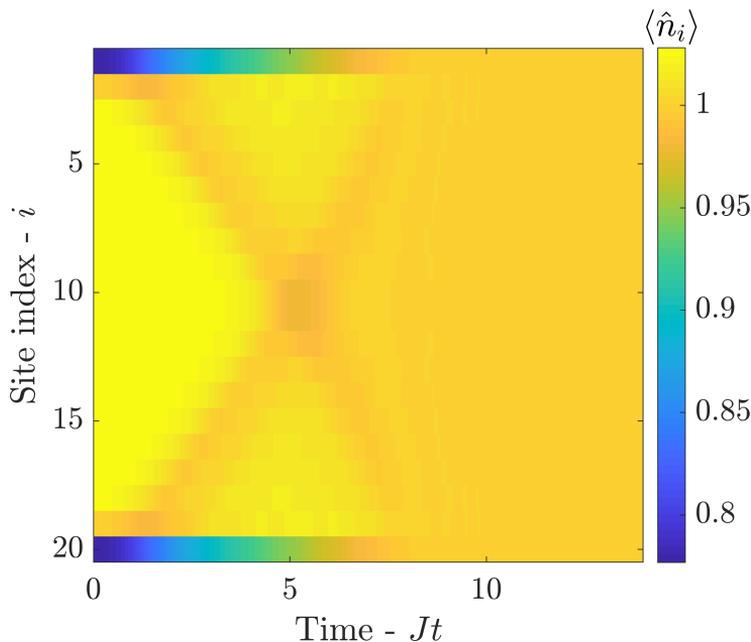


Figure 9.1: The density when propagated along the optimized control for a superfluid to Mott-insulator transfer in the Bose-Hubbard model. In the target Mott-insulator state there is unit occupancy on each site. This transfer has fidelity $F = 0.99$. The optimizations were performed using the full Hessian and gradients in GROUP combined with a matrix product states simulation.

states. These techniques are incredibly useful as they by construction put an upper bound on the entanglement, which allows most states in the Hilbert space to be neglected [306, 307]. This provides a low-entanglement effective theory that remains valid at long durations due to the finite spreading velocity of correlations [308, 309]. In this way matrix product states allow for simulation of systems that would be infeasible with exact diagonalization.

In this project we combine GROUP type quantum optimal control and matrix product states to optimize across the quantum phase transition from superfluid to Mott-insulator in the Bose-Hubbard model. These states are described in Sec. 2.5. This optimization has previously been performed using Nelder-Mead with CRAB where it was claimed that computation of the gradients with matrix product states was too resource demanding and thus impossible [119]. However, in this project we performed exactly this calculation. In order to find high quality gradients we



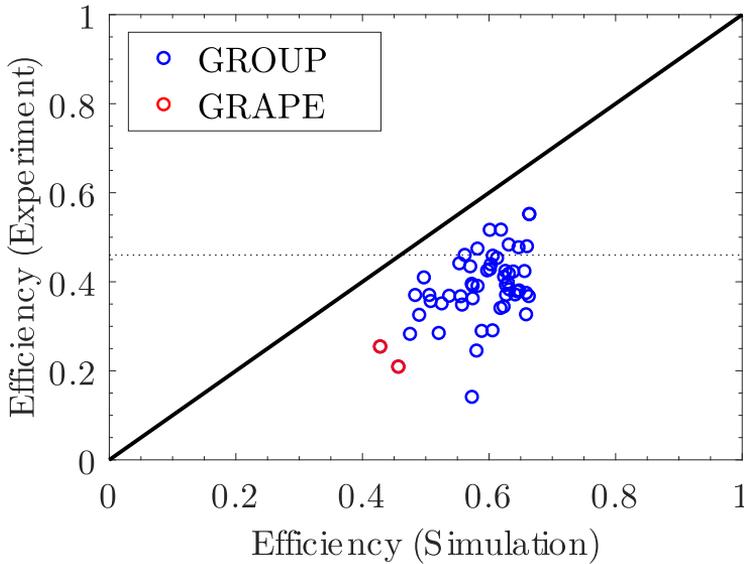


Figure 9.2: The ratio between the efficiency measured in experiments and calculated in theory. The blue dots are results from different GROUP optimizations. The red dots show the previously best results obtained using GRAPE by the bioNMR group. Most of the GROUP results have better efficiencies than those previously calculated.

need to include the temporal discretization, which gives slightly different expressions from those presented in Chap. 3. We are also able to find expressions for the second order gradients that allows for computation of the full Hessian. The structure of the Hessian expression even allows for efficient parallel computation. These analytic expressions allow us to perform efficient full Newton optimization using interior point methods that converge in a low number of iterations. A single optimized trajectory is shown in Fig. 9.1 with a superfluid state on the left and a Mott-insulator on the right. The figure shows a wave-front phenomenon as density is redistributed from the center to the edges. In this project we have already successfully integrated the matrix product states formalism and second order GROUP. We are in the process of fine-tuning the optimization for the best possible estimates of the quantum speed limit and characterizing the physics of optimally crossing a quantum phase transition. I plan to present a manuscript in the near future.



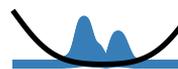
GROUP for Magnetic Nuclear Resonance Experiments

In this project, we optimize the quality of an experimental signal in a magnetic nuclear resonance experiment using GROUP. Magnetic nuclear resonance experiments can give many details at the molecular level. Most routine experiments are performed with spin-1/2 systems like ^1H or ^{13}C . However, most atoms have spins higher than spin-1/2, which may possess a quadrupolar moment due to their non-spherical symmetry. This complicates the interpretation of data from especially solid-state experiments. A possible solution is multiple quantum magic-angle spinning, which makes the data easier to interpret at the expense of lower efficiency. This efficiency can be improved using optimizing control pulses. In this work, we collaborate with the bioNMR group at Aarhus University who previously used GRAPE optimized pulses. Most of these pulses are great in the computer optimizations but perform very poorly in actual experiments. The bioNMR group attribute this to the experimental apparatus being unable to exactly follow controls with high-frequency components. In this project, we found new optimal controls with high efficiency and no high-frequency components using GROUP with a low-frequency chopped basis. Our initial results are shown in Fig. 9.2. These solutions already outperform previous experiments and we find a higher rate of transferability between theory and experiment. We still need to establish more conclusively that the experimental issue is the high frequency components and the project has been delayed due to various experimental difficulties. In this project, I also plan to present results in the near future.



Bibliography

- [1] M. K. Pedersen, J. J. W. Sørensen, M. C. Tichy, and J. F. Sherson, *Many-body state engineering using measurements and fixed unitary dynamics*, New Journal of Physics **16**, 113038 (2014).
- [2] J. J. Sørensen, M. K. Pedersen, and J. Sherson, *Spillere hjælper kvantecomputeren på vej*, Kvant **25**, 31 (2014).
- [3] J. J. W. Sørensen, M. K. Pedersen, M. Munch, P. Haikka, J. H. Jensen, T. Planke, M. G. Andreassen, M. Gajdacz, K. Mølmer, A. Lieberoth, *et al.*, *Exploring the quantum speed limit with computer games*, Nature **532**, 210 (2016).
- [4] H. L. Mortensen, J. J. W. Sørensen, K. Mølmer, and J. Sherson, *Fast state transfer in a Λ -system: a shortcut-to-adiabaticity approach to robust and resource optimized control*, New Journal of Physics **20**, 025009 (2018).
- [5] J. J. Sørensen, M. Aranburu, T. Heinzl, and J. Sherson, *Quantum optimal control in a chopped basis: Applications in control of Bose-Einstein condensates*, Physical Review A **98**, 022119 (2018).
- [6] J. J. Sørensen, M. Aranburu, T. Heinzl, and J. Sherson, *Approaching the Quantum Speed Limit with Global-Local Optimization*, arXiv preprint arXiv:1802.07521 (2018).
- [7] R. Heck, O. Vuculescu, J. J. Sørensen, J. Zoller, M. G. Andreassen, M. G. Bason, P. Ejlertsen, O. Eliasson, P. Haikka, J. S. Laustsen, *et al.*, *Remote optimization of an ultracold atoms experiment by experts and citizen scientists*, Proceedings of the National Academy of Sciences , 201716869 (2018).



- [8] J. J. Sørensen, J. Jensen, T. Heinzl, and J. Sherson, *QEngine: An open-source C++ Library for Quantum Optimal Control of Ultracold Atoms*, arXiv preprint arXiv:1807.11731 (2018).
- [9] J. J. Sørensen, M. Dalgaard, A. H. Kiilerich, K. Mølmer, and J. Sherson, *Quantum Control with Measurements and Quantum Zeno Dynamics*, arXiv preprint arXiv:1806.07793 (2018).
- [10] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, *Quantum cryptography*, Reviews of modern physics **74**, 145 (2002).
- [11] L.-M. Duan, M. Lukin, J. I. Cirac, and P. Zoller, *Long-distance quantum communication with atomic ensembles and linear optics*, Nature **414**, 413 (2001).
- [12] F. Caruso, S. F. Huelga, and M. B. Plenio, *Noise-enhanced classical and quantum capacities in communication networks*, Physical review letters **105**, 190501 (2010).
- [13] C. L. Degen, F. Reinhard, and P. Cappellaro, *Quantum sensing*, Reviews of modern physics **89**, 035002 (2017).
- [14] V. Giovannetti, S. Lloyd, and L. Maccone, *Advances in quantum metrology*, Nature photonics **5**, 222 (2011).
- [15] G. Kucsko, P. Maurer, N. Y. Yao, M. Kubo, H. Noh, P. Lo, H. Park, and M. D. Lukin, *Nanometre-scale thermometry in a living cell*, Nature **500**, 54 (2013).
- [16] V. Giovannetti, S. Lloyd, and L. Maccone, *Quantum-enhanced measurements: beating the standard quantum limit*, Science **306**, 1330 (2004).
- [17] D. P. DiVincenzo, *Quantum computation*, Science **270**, 255 (1995).
- [18] R. P. Feynman, *Simulating physics with computers*, International journal of theoretical physics **21**, 467 (1982).
- [19] S. Lloyd, *Universal quantum simulators*, Science , 1073 (1996).
- [20] D. Deutsch, *Quantum theory, the Church–Turing principle and the universal quantum computer*, Proc. R. Soc. Lond. A **400**, 97 (1985).
- [21] A. Steane, *Quantum computing*, Reports on Progress in Physics **61**, 117 (1998).



- [22] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM review **41**, 303 (1999).
- [23] C. H. Bennett and D. P. DiVincenzo, *Quantum information and computation*, Nature **404**, 247 (2000).
- [24] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, *et al.*, *Towards quantum chemistry on a quantum computer*, Nature chemistry **2**, 106 (2010).
- [25] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, *Elucidating reaction mechanisms on quantum computers*, Proceedings of the National Academy of Sciences , 201619152 (2017).
- [26] I. Bloch, J. Dalibard, and S. Nascimbene, *Quantum simulations with ultracold quantum gases*, Nature Physics **8**, 267 (2012).
- [27] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, *Quantum machine learning*, Nature **549**, 195 (2017).
- [28] P. Rebentrost, M. Mohseni, and S. Lloyd, *Quantum support vector machine for big data classification*, Physical review letters **113**, 130503 (2014).
- [29] V. Dunjko, J. M. Taylor, and H. J. Briegel, *Quantum-enhanced machine learning*, Physical review letters **117**, 130501 (2016).
- [30] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, *et al.*, *Training Schrödinger's cat: quantum optimal control*, The European Physical Journal D **69**, 279 (2015).
- [31] A. De Touzalin, C. Marcus, F. Heijman, I. Cirac, R. Murray, and T. Calarco, *Quantum manifesto. A new era of technology*, European Commission (2016).
- [32] D. d'Alessandro, *Introduction to quantum control and dynamics* (CRC press, 2007).
- [33] J. Clarke and F. K. Wilhelm, *Superconducting quantum bits*, Nature **453**, 1031 (2008).
- [34] F. Vatan and C. Williams, *Optimal quantum circuits for general two-qubit gates*, Physical Review A **69**, 032315 (2004).



- [35] S. Montangero, T. Calarco, and R. Fazio, *Robust optimal quantum gates for Josephson charge qubits*, Physical Review Letters **99**, 170501 (2007).
- [36] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. Heinzen, *Optimal frequency measurements with maximally correlated states*, Physical Review A **54**, R4649 (1996).
- [37] W. Dür, M. Skotiniotis, F. Froewis, and B. Kraus, *Improved quantum metrology using quantum error correction*, Physical Review Letters **112**, 080801 (2014).
- [38] J. Werschnik and E. Gross, *Quantum optimal control theory*, Journal of Physics B: Atomic, Molecular and Optical Physics **40**, R175 (2007).
- [39] A. Borzì, G. Ciaramella, and M. Sprengel, *Formulation and numerical solution of quantum control problems*, Vol. 16 (SIAM, 2017).
- [40] X. Wang, A. Bayat, S. G. Schirmer, and S. Bose, *Robust entanglement in antiferromagnetic Heisenberg chains by single-spin optimal control*, Physical Review A **81**, 032312 (2010).
- [41] M. H. Goerz, D. M. Reich, and C. P. Koch, *Optimal control theory for a unitary operation under dissipative evolution*, New Journal of Physics **16**, 055012 (2014).
- [42] A. Ruschhaupt, X. Chen, D. Alonso, and J. Muga, *Optimally robust shortcuts to population inversion in two-level quantum systems*, New Journal of Physics **14**, 093040 (2012).
- [43] H. Zhang and H. Rabitz, *Robust optimal control of quantum molecular systems in the presence of disturbances and uncertainties*, Physical Review A **49**, 2241 (1994).
- [44] T. Hornung, M. Motzkus, and R. de Vivie-Riedle, *Teaching optimal control theory to distill robust pulses even under experimental constraints*, Physical Review A **65**, 021403 (2002).
- [45] T. Nöbauer, A. Angerer, B. Bartels, M. Trupke, S. Rotter, J. Schmiedmayer, F. Mintert, and J. Majer, *Smooth optimal quantum control for robust solid-state spin magnetometry*, Physical Review Letters **115**, 190801 (2015).



- [46] U. Hohenester, P. K. Rekdal, A. Borzì, and J. Schmiedmayer, *Optimal quantum control of Bose-Einstein condensates in magnetic microtraps*, Physical Review A **75**, 023602 (2007).
- [47] S. van Frank, M. Bonneau, J. Schmiedmayer, S. Hild, C. Gross, M. Cheneau, I. Bloch, T. Pichler, A. Negretti, T. Calarco, *et al.*, *Optimal control of complex atomic quantum systems*, Scientific reports **6** (2016).
- [48] F. Dolde, V. Bergholm, Y. Wang, I. Jakobi, B. Naydenov, S. Pez-zagna, J. Meijer, F. Jelezko, P. Neumann, T. Schulte-Herbrüggen, *et al.*, *High-fidelity spin entanglement using optimal control*, Nature communications **5**, 3371 (2014).
- [49] D. Egger and F. K. Wilhelm, *Adaptive hybrid optimal quantum control for imprecisely characterized systems*, Physical Review Letters **112**, 240503 (2014).
- [50] U. Boscain and P. Mason, *Time minimal trajectories for a spin 1/2 particle in a magnetic field*, Journal of Mathematical Physics **47**, 062101 (2006).
- [51] B. Bonnard and M. Chyba, *Singular trajectories and their role in control theory*, Vol. 40 (Springer Science & Business Media, 2003).
- [52] G. C. Hegerfeldt, *Driving at the quantum speed limit: optimal control of a two-level system*, Physical Review Letters **111**, 260501 (2013).
- [53] D. Sugny, C. Kontz, and H.-R. Jauslin, *Time-optimal control of a two-level dissipative quantum system*, Physical Review A **76**, 023419 (2007).
- [54] A. Konnov and V. F. Krotov, *On global methods for the successive improvement of control processes*, Avtomatika i Telemekhanika **10**, 77 (1999).
- [55] D. M. Reich, M. Ndong, and C. P. Koch, *Monotonically convergent optimization in quantum control using Krotov's method*, The Journal of chemical physics **136**, 104103 (2012).
- [56] S. G. Schirmer and P. de Fouquieres, *Efficient algorithms for optimal control of quantum dynamics: the Krotov method unencumbered*, New Journal of Physics **13**, 073029 (2011).



- [57] S. Chu, *Nobel Lecture: The manipulation of neutral particles*, Reviews of Modern Physics **70**, 685 (1998).
- [58] W. D. Phillips, *Nobel Lecture: Laser cooling and trapping of neutral atoms*, Reviews of Modern Physics **70**, 721 (1998).
- [59] C. N. Cohen-Tannoudji, *Nobel Lecture: Manipulating atoms with photons*, Reviews of Modern Physics **70**, 707 (1998).
- [60] H. J. Metcalf and P. Van der Straten, *Laser cooling and trapping* (Springer Science & Business Media, 2012).
- [61] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. Van Druten, D. Durfee, D. Kurn, and W. Ketterle, *Bose-Einstein condensation in a gas of sodium atoms*, Physical Review Letters **75**, 3969 (1995).
- [62] C. C. Bradley, C. Sackett, J. Tollett, and R. G. Hulet, *Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions*, Physical review letters **75**, 1687 (1995).
- [63] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, *Observation of Bose-Einstein condensation in a dilute atomic vapor*, science **269**, 198 (1995).
- [64] B. DeMarco and D. S. Jin, *Onset of Fermi degeneracy in a trapped atomic gas*, Science **285**, 1703 (1999).
- [65] A. G. Truscott, K. E. Strecker, W. I. McAlexander, G. B. Partridge, and R. G. Hulet, *Observation of Fermi pressure in a gas of trapped atoms*, Science **291**, 2570 (2001).
- [66] I. Bloch, J. Dalibard, and W. Zwerger, *Many-body physics with ultracold gases*, Reviews of modern physics **80**, 885 (2008).
- [67] E. Raab, M. Prentiss, A. Cable, S. Chu, and D. E. Pritchard, *Trapping of neutral sodium atoms with radiation pressure*, Physical Review Letters **59**, 2631 (1987).
- [68] P. J. Ungar, D. S. Weiss, E. Riis, and S. Chu, *Optical molasses and multilevel atoms: theory*, JOSA B **6**, 2058 (1989).
- [69] R. Grimm, M. Weidemüller, and Y. B. Ovchinnikov, *Optical dipole traps for neutral atoms*, in *Advances in atomic, molecular, and optical physics*, Vol. 42 (Elsevier, 2000) pp. 95–170.
- [70] J. Miller, R. Cline, and D. Heinzen, *Photoassociation spectrum of ultracold Rb atoms*, Physical Review Letters **71**, 2204 (1993).



- [71] J. Miller, R. Cline, and D. Heinzen, *Far-off-resonance optical trapping of atoms*, Physical Review A **47**, R4567 (1993).
- [72] N. W. Ashcroft and N. D. Mermin, *Solid state physics (holt, rinehart and winston, new york, 1976)*, Google Scholar **403** (2005).
- [73] W. Kohn, *Analytic properties of Bloch waves and Wannier functions*, Physical Review **115**, 809 (1959).
- [74] I. Lesanovsky, T. Schumm, S. Hofferberth, L. M. Andersson, P. Krüger, and J. Schmiedmayer, *Adiabatic radio-frequency potentials for the coherent manipulation of matter waves*, Physical Review A **73**, 033619 (2006).
- [75] I. Lesanovsky, S. Hofferberth, J. Schmiedmayer, and P. Schmelcher, *Manipulation of ultracold atoms in dressed adiabatic radio-frequency potentials*, Physical Review A **74**, 033619 (2006).
- [76] D. E. Pritchard, *Cooling neutral atoms in a magnetic trap for precision spectroscopy*, Physical Review Letters **51**, 1336 (1983).
- [77] J. Weinstein and K. Libbrecht, *Microscopic magnetic traps for neutral atoms*, Physical Review A **52**, 4004 (1995).
- [78] J. Schmiedmayer, *A wire trap for neutral atoms*, Applied Physics B **60**, 169 (1995).
- [79] J. Reichel and J. H. Thywissen, *Using magnetic chip traps to study Tonks-Girardeau quantum gases*, in *Journal de Physique IV (Proceedings)*, Vol. 116 (EDP sciences, 2004) pp. 265–274.
- [80] B. M. Garraway and H. Perrin, *Recent developments in trapping and manipulation of atoms with adiabatic potentials*, Journal of Physics B: Atomic, Molecular and Optical Physics **49**, 172001 (2016).
- [81] R. Bücker, T. Berrada, S. Van Frank, J.-F. Schaff, T. Schumm, J. Schmiedmayer, G. Jäger, J. Grond, and U. Hohenester, *Vibrational state inversion of a Bose–Einstein condensate: optimal control and state tomography*, Journal of Physics B: Atomic, Molecular and Optical Physics **46**, 104012 (2013).
- [82] S. Hofferberth, I. Lesanovsky, B. Fischer, J. Verdu, and J. Schmiedmayer, *Radiofrequency-dressed-state potentials for neutral atoms*, Nature Physics **2**, 710 (2006).



- [83] T. Schumm, S. Hofferberth, L. M. Andersson, S. Wildermuth, S. Groth, I. Bar-Joseph, J. Schmiedmayer, and P. Krüger, *Matter-wave interferometry in a double well on an atom chip*, Nature physics **1**, 57 (2005).
- [84] C. J. Pethick and H. Smith, *Bose-Einstein condensation in dilute gases* (Cambridge university press, 2002).
- [85] A. L. Fetter and J. D. Walecka, *Quantum theory of many-particle systems* (Courier Corporation, 2012).
- [86] E. P. Gross, *Structure of a quantized vortex in boson systems*, II Nuovo Cimento (1955-1965) **20**, 454 (1961).
- [87] L. Pitaevskii, *Vortex lines in an imperfect Bose gas*, Sov. Phys. JETP **13**, 451 (1961).
- [88] M. P. Kennett, *Out-of-equilibrium dynamics of the bose-hubbard model*, ISRN Condensed Matter Physics **2013** (2013).
- [89] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller, *Cold bosonic atoms in optical lattices*, Physical Review Letters **81**, 3108 (1998).
- [90] R. J. Gordon and S. A. Rice, *Active control of the dynamics of atoms and molecules*, Annual review of physical chemistry **48**, 601 (1997).
- [91] C. Sayrin, I. Dotsenko, X. Zhou, B. Peaudecerf, T. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune, *et al.*, *Real-time quantum feedback prepares and stabilizes photon number states*, Nature **477**, 73 (2011).
- [92] R. Bücker, J. Grond, S. Manz, T. Berrada, T. Betz, C. Koller, U. Hohenester, T. Schumm, A. Perrin, and J. Schmiedmayer, *Twin-atom beams*, Nature Physics **7**, 608 (2011).
- [93] B. Bloom, T. Nicholson, J. Williams, S. Campbell, M. Bishof, X. Zhang, W. Zhang, S. Bromley, and J. Ye, *An optical lattice clock with accuracy and stability at the 10⁻¹⁸ level*, Nature **506**, 71 (2014).
- [94] B. Lücke, M. Scherer, J. Kruse, L. Pezzé, F. Deuretzbacher, P. Hyllus, J. Peise, W. Ertmer, J. Arlt, L. Santos, *et al.*, *Twin matter waves for interferometry beyond the classical limit*, Science **334**, 773 (2011).



- [95] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, *Atom-chip-based generation of entanglement for quantum metrology*, *Nature* **464**, 1170 (2012).
- [96] C. Gross, T. Zibold, E. Nicklas, J. Esteve, and M. K. Oberthaler, *Nonlinear atom interferometer surpasses classical precision limit*, *Nature* **464**, 1165 (2012).
- [97] M. Anderlini, P. J. Lee, B. L. Brown, J. Sebby-Strabley, W. D. Phillips, and J. Porto, *Controlled exchange interaction between pairs of neutral atoms in an optical lattice*, *Nature* **448**, 452 (2007).
- [98] D. Kielpinski, C. Monroe, and D. J. Wineland, *Architecture for a large-scale ion-trap quantum computer*, *Nature* **417**, 709 (2002).
- [99] J. P. Palao and R. Kosloff, *Quantum computing by an optimal control algorithm for unitary transformations*, *Physical Review Letters* **89**, 188301 (2002).
- [100] A. Kaiser and V. May, *Optimal control theory for a target state distributed in time: Optimizing the probe-pulse signal of a pump-probe-scheme*, *The Journal of chemical physics* **121**, 2528 (2004).
- [101] A. P. Peirce, M. A. Dahleh, and H. Rabitz, *Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications*, *Physical Review A* **37**, 4950 (1988).
- [102] M. H. Goerz, F. Motzoi, K. B. Whaley, and C. P. Koch, *Charting the circuit QED design landscape using optimal control theory*, *npj Quantum Information* **3**, 37 (2017).
- [103] C. P. Koch, J. P. Palao, R. Kosloff, and F. Masnou-Seeuws, *Stabilization of ultracold molecules using optimal control theory*, *Physical Review A* **70**, 013402 (2004).
- [104] T. Caneva, T. Calarco, and S. Montangero, *Chopped random-basis quantum optimization*, *Physical Review A* **84**, 022326 (2011).
- [105] D. Liberzon, *Calculus of variations and optimal control theory: a concise introduction* (Princeton University Press, 2011).
- [106] J.-F. Mennemann, D. Matthes, R.-M. Weishäupl, and T. Langen, *Optimal control of Bose–Einstein condensates in three dimensions*, *New Journal of Physics* **17**, 113027 (2015).



- [107] A. Borzi, *Quantum optimal control using the adjoint method*, Mathematics of Quantum Technologies **1**, 93 (2012).
- [108] G. Von Winckel and A. Borzi, *Computational techniques for a quantum control problem with H1-cost*, Inverse Problems **24**, 034007 (2008).
- [109] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, *Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms*, Journal of magnetic resonance **172**, 296 (2005).
- [110] G. Jäger, D. M. Reich, M. H. Goerz, C. P. Koch, and U. Hohenester, *Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Comparison of gradient-ascent-pulse-engineering and Krotov optimization schemes*, Physical Review A **90**, 033628 (2014).
- [111] J. Nocedal and S. J. Wright, *Numerical optimization 2nd* (2006).
- [112] J. R. Shewchuk *et al.*, *An introduction to the conjugate gradient method without the agonizing pain* (1994).
- [113] Y.-H. Dai and Y. Yuan, *A nonlinear conjugate gradient method with a strong global convergence property*, SIAM Journal on optimization **10**, 177 (1999).
- [114] W. W. Hager and H. Zhang, *A new conjugate gradient method with guaranteed descent and an efficient line search*, SIAM Journal on optimization **16**, 170 (2005).
- [115] S. E. Sklarz and D. J. Tannor, *Loading a Bose-Einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear Schrödinger equation*, Physical Review A **66**, 053619 (2002).
- [116] R. Eitan, M. Mundt, and D. J. Tannor, *Optimal control with accelerated convergence: Combining the Krotov and quasi-Newton methods*, Physical Review A **83**, 053426 (2011).
- [117] S. Lloyd and S. Montangero, *Information theoretical analysis of quantum optimal control*, Physical Review Letters **113**, 010502 (2014).
- [118] F. Motzoi, J. M. Gambetta, S. Merkel, and F. Wilhelm, *Optimal control methods for rapidly time-varying Hamiltonians*, Physical Review A **84**, 022307 (2011).



- [119] P. Doria, T. Calarco, and S. Montangero, *Optimal control technique for many-body quantum dynamics*, Physical Review Letters **106**, 190501 (2011).
- [120] I. Brouzos, A. I. Streltsov, A. Negretti, R. S. Said, T. Caneva, S. Montangero, and T. Calarco, *Quantum speed limit and optimal control of many-boson dynamics*, Physical Review A **92**, 062110 (2015).
- [121] R. E. Goetz, M. Merkel, A. Karamatskou, R. Santra, and C. P. Koch, *Maximizing hole coherence in ultrafast photoionization of argon with an optimization by sequential parametrization update*, Physical Review A **94**, 023420 (2016).
- [122] J. A. Nelder and R. Mead, *A simplex method for function minimization*, The computer journal **7**, 308 (1965).
- [123] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, *Convergence properties of the Nelder–Mead simplex method in low dimensions*, SIAM Journal on optimization **9**, 112 (1998).
- [124] H. A. Rabitz, M. M. Hsieh, and C. M. Rosenthal, *Quantum optimally controlled transition landscapes*, Science **303**, 1998 (2004).
- [125] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, *Tunable, Flexible, and Efficient Optimization of Control Pulses for Practical Qubits*, Physical Review Letters **120**, 150401 (2018).
- [126] D. Lucarelli, *Quantum optimal control via gradient ascent in function space and the time-bandwidth quantum speed limit*, Physical Review A **97**, 062346 (2018).
- [127] T. E. Skinner and N. I. Gershenzon, *Optimal control design of pulse shapes as analytic functions*, Journal of Magnetic Resonance **204**, 248 (2010).
- [128] F. Gerbier, *Quasi-1D Bose-Einstein condensates in the dimensional crossover regime*, EPL (Europhysics Letters) **66**, 771 (2004).
- [129] G. Jäger and U. Hohenester, *Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Consideration of filter effects*, Physical Review A **88**, 035601 (2013).
- [130] D. M. Reich, J. P. Palao, and C. P. Koch, *Optimal control under spectral constraints: enforcing multi-photon absorption pathways*, Journal of Modern Optics **61**, 822 (2014).



- [131] C. Gollub, M. Kowalewski, and R. de Vivie-Riedle, *Monotonic convergent optimal control theory with strict limitations on the spectrum of optimized laser fields*, Physical Review Letters **101**, 073002 (2008).
- [132] N. Rach, M. M. Müller, T. Calarco, and S. Montangero, *Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape*, Physical Review A **92**, 062343 (2015).
- [133] P. Rebentrost and F. K. Wilhelm, *Optimal control of a leaking qubit*, Physical Review B **79**, 060507 (2009).
- [134] F. C. Langbein, S. Schirmer, and E. Jonckheere, *Time optimal information transfer in spintronics networks*, in *Decision and Control (CDC), 2015 IEEE 54th Annual Conference on* (IEEE, 2015) pp. 6454–6459.
- [135] N. Khaneja, R. Brockett, and S. J. Glaser, *Time optimal control in spin systems*, Physical Review A **63**, 032308 (2001).
- [136] N. Khaneja, S. J. Glaser, and R. Brockett, *Sub-Riemannian geometry and time optimal control of three spin systems: quantum gates and coherence transfer*, Physical Review A **65**, 032301 (2002).
- [137] L. B. Levitin and T. Toffoli, *Fundamental limit on the rate of quantum dynamics: the unified bound is tight*, Physical Review Letters **103**, 160502 (2009).
- [138] S. Lloyd, *Ultimate physical limits to computation*, Nature **406**, 1047 (2000).
- [139] S. Deffner and E. Lutz, *Generalized Clausius inequality for nonequilibrium quantum processes*, Physical Review Letters **105**, 170402 (2010).
- [140] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro, *Optimal control at the quantum speed limit*, Physical Review Letters **103**, 240501 (2009).
- [141] B. Bonnard and D. Sugny, *Time-minimal control of dissipative two-level quantum systems: the integrable case*, SIAM Journal on Control and Optimization **48**, 1289 (2009).
- [142] U. Boscain, T. Chambrion, and G. Charlot, *Nonisotropic 3-level quantum systems: complete solutions for minimum time and minimum energy*, arXiv preprint quant-ph/0409022 (2004).



- [143] H. Yuan, C. P. Koch, P. Salamon, and D. J. Tannor, *Controllability on relaxation-free subspaces: On the relationship between adiabatic population transfer and optimal control*, Physical Review A **85**, 033417 (2012).
- [144] T. O. Reiss, N. Khaneja, and S. J. Glaser, *Time-optimal coherence-order-selective transfer of in-phase coherence in heteronuclear IS spin systems*, Journal of Magnetic Resonance **154**, 192 (2002).
- [145] K. Hoffmann, P. Salamon, Y. Rezek, and R. Kosloff, *Time-optimal controls for frictionless cooling in harmonic traps*, EPL (Europhysics Letters) **96**, 60015 (2011).
- [146] S. Deffner and S. Campbell, *Quantum speed limits: from Heisenberg's uncertainty principle to optimal quantum control*, Journal of Physics A: Mathematical and Theoretical **50**, 453001 (2017).
- [147] N. Margolus and L. B. Levitin, *The maximum speed of dynamical evolution*, Physica D: Nonlinear Phenomena **120**, 188 (1998).
- [148] L. Mandelstam and T. I., *The uncertainty relation between energy and time in non-relativistic quantum mechanics*, J. Phys. (USSR) **9**, 249 (1945).
- [149] G. N. Fleming, *A unitarity bound on the evolution of nonstationary states*, Il Nuovo Cimento A (1971-1996) **16**, 232 (1973).
- [150] K. Bhattacharyya, *Quantum decay and the Mandelstam-Tamm-energy inequality*, Journal of Physics A: Mathematical and General **16**, 2993 (1983).
- [151] J. Anandan and Y. Aharonov, *Geometry of quantum evolution*, Physical Review Letters **65**, 1697 (1990).
- [152] J. Uffink, *The rate of evolution of a quantum state*, American Journal of Physics **61**, 935 (1993).
- [153] M. G. Bason, M. Viteau, N. Malossi, P. Huillery, E. Arimondo, D. Ciampini, R. Fazio, V. Giovannetti, R. Mannella, and O. Morsch, *High-fidelity quantum driving*, Nature Physics **8**, 147 (2012).
- [154] A. Uhlmann, *An energy dispersion estimate*, Physics Letters A **161**, 329 (1992).
- [155] S. Deffner and E. Lutz, *Energy-time uncertainty relation for driven quantum systems*, Journal of Physics A: Mathematical and Theoretical **46**, 335302 (2013).



- [156] S. L. Braunstein and C. M. Caves, *Statistical distance and the geometry of quantum states*, Physical Review Letters **72**, 3439 (1994).
- [157] P. J. Jones and P. Kok, *Geometric derivation of the quantum speed limit*, Physical Review A **82**, 022107 (2010).
- [158] M. Gajdacz, K. K. Das, J. Arlt, J. F. Sherson, and T. Opatrny, *Time-limited optimal dynamics beyond the quantum speed limit*, Physical Review A **92**, 062106 (2015).
- [159] P. M. Poggi, F. C. Lombardo, and D. Wisniacki, *Quantum speed limit and optimal evolution time in a two-level system*, EPL (Europhysics Letters) **104**, 40005 (2013).
- [160] G. C. Hegerfeldt, *High-speed driving of a two-level system*, Physical Review A **90**, 032110 (2014).
- [161] J. J. Sakurai and E. D. Commins, *Modern quantum mechanics, revised edition* (1995).
- [162] T. Caneva, T. Calarco, R. Fazio, G. E. Santoro, and S. Montangero, *Speeding up critical system dynamics through optimized evolution*, Physical Review A **84**, 012312 (2011).
- [163] B. Russell, H. Rabitz, and R.-B. Wu, *Control landscapes are almost always trap free: a geometric assessment*, Journal of Physics A: Mathematical and Theoretical **50**, 205302 (2017).
- [164] D. V. Zhdanov and T. Seideman, *Role of control constraints in quantum optimal control*, Physical Review A **92**, 052109 (2015).
- [165] M. Bukov, A. G. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, *Reinforcement learning in different phases of quantum control*, Physical Review X **8**, 031086 (2018).
- [166] A. G. Day, M. Bukov, P. Weinberg, P. Mehta, and D. Sels, *The Glassy Phase of Optimal Quantum Control*, arXiv preprint arXiv:1803.10856 (2018).
- [167] M. Bukov, A. G. Day, P. Weinberg, A. Polkovnikov, P. Mehta, and D. Sels, *Broken symmetry in a two-qubit quantum control landscape*, Physical Review A **97**, 052114 (2018).
- [168] E. Zahedinejad, S. Schirmer, and B. C. Sanders, *Evolutionary algorithms for hard quantum control*, Physical Review A **90**, 032310 (2014).



- [169] F. Neri and C. Cotta, *Memetic algorithms and memetic computing optimization: A literature review*, Swarm and Evolutionary Computation **2**, 1 (2012).
- [170] G. Dueck, *New optimization heuristics*, Journal of Computational physics **104**, 86 (1993).
- [171] J. G. March, *Exploration and exploitation in organizational learning*, Organization science **2**, 71 (1991).
- [172] Z. Ugray, L. Lasdon, J. Plummer, F. Glover, J. Kelly, and R. Martí, *Scatter search and local NLP solvers: A multistart framework for global optimization*, INFORMS Journal on Computing **19**, 328 (2007).
- [173] J. Kennedy, *Particle swarm optimization*, in *Encyclopedia of machine learning* (Springer, 2011) pp. 760–766.
- [174] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, *et al.*, *Optimization by simulated annealing*, science **220**, 671 (1983).
- [175] S. Das and P. N. Suganthan, *Differential evolution: A survey of the state-of-the-art*, IEEE transactions on evolutionary computation **15**, 4 (2011).
- [176] N. Hansen, S. D. Müller, and P. Koumoutsakos, *Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES)*, Evolutionary computation **11**, 1 (2003).
- [177] E. Zahedinejad, J. Ghosh, and B. C. Sanders, *High-fidelity single-shot Toffoli gate via quantum control*, Physical review letters **114**, 200502 (2015).
- [178] R. Bartels, S. Backus, E. Zeek, L. Misoguti, G. Vdovin, I. Christov, M. Murnane, and H. Kapteyn, *Shaped-pulse optimization of coherent emission of high-harmonic soft X-rays*, Nature **406**, 164 (2000).
- [179] J. L. Herek, W. Wohlleben, R. J. Cogdell, D. Zeidler, and M. Motzkus, *Quantum control of energy flow in light harvesting*, Nature **417**, 533 (2002).
- [180] T. Weinacht, J. White, and P. Bucksbaum, *Toward strong field mode-selective chemistry*, The Journal of Physical Chemistry A **103**, 10166 (1999).



- [181] C. J. Bardeen, V. V. Yakovlev, K. R. Wilson, S. D. Carpenter, P. M. Weber, and W. S. Warren, *Feedback quantum control of molecular electronic population transfer*, *Chemical Physics Letters* **280**, 151 (1997).
- [182] R. J. Levis, G. M. Menkir, and H. Rabitz, *Selective bond dissociation and rearrangement with optimally tailored, strong-field laser pulses*, *Science* **292**, 709 (2001).
- [183] L. Viola and S. Lloyd, *Dynamical suppression of decoherence in two-state quantum systems*, *Physical Review A* **58**, 2733 (1998).
- [184] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, *Quantum computers*, *Nature* **464**, 45 (2010).
- [185] S. J. Devitt, W. J. Munro, and K. Nemoto, *Quantum error correction for beginners*, *Reports on Progress in Physics* **76**, 076001 (2013).
- [186] S. Wright, *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*, Vol. 1 (na, 1932).
- [187] S. Kryazhimskiy, G. Tkačik, and J. B. Plotkin, *The dynamics of adaptation on correlated fitness landscapes*, *Proceedings of the National Academy of Sciences* **106**, 18638 (2009).
- [188] K. M. Malan and A. P. Engelbrecht, *Fitness landscape analysis for metaheuristic performance prediction*, in *Recent advances in the theory and application of fitness landscapes* (Springer, 2014) pp. 103–132.
- [189] P. Palittapongarnpim, P. Wittek, E. Zahedinejad, S. Vedaie, and B. C. Sanders, *Learning in quantum control: High-dimensional global optimization for noisy quantum dynamics*, *Neurocomputing* **268**, 116 (2017).
- [190] D. A. Levinthal, *Adaptation on rugged landscapes*, *Management science* **43**, 934 (1997).
- [191] A. Acerbi, C. Tennie, and A. Mesoudi, *Social learning solves the problem of narrow-peaked search landscapes: experimental evidence in humans*, *Royal Society open science* **3** (2016).
- [192] O. Mersmann, B. Bischl, H. Trautmann, M. Preuss, C. Weihs, and G. Rudolph, *Exploratory landscape analysis*, in *Proceedings of the*



- 13th annual conference on Genetic and evolutionary computation* (ACM, 2011) pp. 829–836.
- [193] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, and S. J. Gershman, *Building machines that learn and think like people*, Behavioral and Brain Sciences **40** (2017).
- [194] G. Marcus, *Deep learning: A critical appraisal*, arXiv preprint arXiv:1801.00631 (2018).
- [195] E. Kamar, *Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence.*, in *IJCAI* (2016) pp. 4070–4073.
- [196] E. Baltz, E. Trask, M. Binderbauer, M. Dikovskiy, H. Gota, R. Mendoza, J. Platt, and P. Riley, *Achievement of sustained net plasma heating in a fusion experiment with the optometrist algorithm*, Scientific reports **7**, 6425 (2017).
- [197] J. Silvertown, *A new dawn for citizen science*, Trends in ecology & evolution **24**, 467 (2009).
- [198] R. Bonney, J. L. Shirk, T. B. Phillips, A. Wiggins, H. L. Ballard, A. J. Miller-Rushing, and J. K. Parrish, *Next steps for citizen science*, Science **343**, 1436 (2014).
- [199] The BIG Bell Test Collaboration, *Challenging Local Realism with Human Choices*, Nature **557**, 212 (2018).
- [200] C. Lintott, K. Schawinski, S. Bamford, A. Slosar, K. Land, D. Thomas, E. Edmondson, K. Masters, R. C. Nichol, M. J. Raddick, *et al.*, *Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies*, Monthly Notices of the Royal Astronomical Society **410**, 166 (2010).
- [201] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović, *et al.*, *Predicting protein structures with a multiplayer online game*, Nature **466**, 756 (2010).
- [202] J. Lee, W. Kladwang, M. Lee, D. Cantu, M. Azizyan, H. Kim, A. Limpaecher, S. Gaikwad, S. Yoon, A. Treuille, *et al.*, *RNA design rules from a massive open laboratory*, Proceedings of the National Academy of Sciences **111**, 2122 (2014).
- [203] J. S. Kim, M. J. Greene, A. Zlateski, K. Lee, M. Richardson, S. C. Turaga, M. Purcaro, M. Balkam, A. Robinson, B. F. Behabadi,



- et al.*, *Space-time wiring specificity supports direction selectivity in the retina*, *Nature* **509**, 331 (2014).
- [204] C. Monroe, *Quantum information processing with atoms and photons*, *Nature* **416**, 238 (2002).
- [205] L. K. Grover, *A fast quantum mechanical algorithm for database search*, in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (ACM, 1996) pp. 212–219.
- [206] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (2000).
- [207] C. Weitenberg, S. Kuhr, K. Mølmer, and J. F. Sherson, *Quantum computation architecture using optical tweezers*, *Physical Review A* **84**, 032322 (2011).
- [208] C. Weitenberg, M. Endres, J. F. Sherson, M. Cheneau, P. Schauß, T. Fukuhara, I. Bloch, and S. Kuhr, *Single-spin addressing in an atomic Mott insulator*, *Nature* **471**, 319 (2011).
- [209] J. W. Sammon, *A nonlinear mapping for data structure analysis*, *IEEE Transactions on computers* **100**, 401 (1969).
- [210] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, *Extracting and composing robust features with denoising autoencoders*, in *Proceedings of the 25th international conference on Machine learning* (ACM, 2008) pp. 1096–1103.
- [211] L. v. d. Maaten and G. Hinton, *Visualizing data using t-SNE*, *Journal of machine learning research* **9**, 2579 (2008).
- [212] S. Kullback and R. A. Leibler, *On information and sufficiency*, *The annals of mathematical statistics* **22**, 79 (1951).
- [213] A. Newell, H. A. Simon, *et al.*, *Human problem solving*, Vol. 104 (Prentice-Hall Englewood Cliffs, NJ, 1972).
- [214] *Making, Probing and Understanding Bose-Einstein Condensates*, in *Bose-Einstein Condensation in Atomic Gases: Varenna on Lake Como, Villa Monastero, 7-17 July 1998*, Proceedings of the International School of Physics "Enrico Fermi" No. 140 (IOS Press, Amsterdam ; Washington, DC, 1999) pp. 67–176.



- [215] W. Rohringer, D. Fischer, M. Trupke, J. Schmiedmayer, and T. Schumm, *Stochastic optimization of Bose-Einstein condensation using a genetic algorithm*, in *Stochastic Optimization-Seeing the Optimal for the Uncertain* (InTech, 2011).
- [216] I. Geisel, K. Cordes, J. Mahnke, S. Jöllenbeck, J. Ostermann, J. Arlt, W. Ertmer, and C. Klempt, *Evolutionary optimization of an experimental apparatus*, *Applied Physics Letters* **102**, 214105 (2013).
- [217] T. Lausch, M. Hohmann, F. Kindermann, D. Mayer, F. Schmidt, and A. Widera, *Optimizing quantum gas production by an evolutionary algorithm*, *Applied Physics B* **122**, 112 (2016).
- [218] P. B. Wigley, P. J. Everitt, A. van den Hengel, J. Bastian, M. A. Sooriyabandara, G. D. McDonald, K. S. Hardman, C. Quinlivan, P. Manju, C. C. Kuhn, *et al.*, *Fast machine-learning online optimization of ultra-cold-atom experiments*, *Scientific reports* **6**, 25890 (2016).
- [219] M. G. Bason, R. Heck, M. Napolitano, O. Eliasson, R. Müller, A. Thorsen, W.-Z. Zhang, J. J. Arlt, and J. F. Sherson, *Measurement-enhanced determination of BEC phase transitions*, *Journal of Physics B: Atomic, Molecular and Optical Physics* **51**, 175301 (2018).
- [220] Y.-J. Lin, A. R. Perry, R. L. Compton, I. B. Spielman, and J. V. Porto, *Rapid production of Rb-87 Bose-Einstein condensates in a combined magnetic and optical potential*, *Physical Review A* **79**, 063631 (2009).
- [221] P. Pinkse, A. Mosk, M. Weidemüller, M. Reynolds, T. Hijmans, and J. Walraven, *Adiabatically changing the phase-space density of a trapped Bose gas*, *Physical Review Letters* **78**, 990 (1997).
- [222] O. Vuculescu and C. Bergenholtz, *How to solve problems with crowds: A computer-based simulation model*, *Creativity and Innovation Management* **23**, 121 (2014).
- [223] S. Billinger, N. Stieglitz, and T. R. Schumacher, *Search on rugged landscapes: An experimental study*, *Organization Science* **25**, 93 (2013).
- [224] D. Sels, *Stochastic gradient ascent outperforms gamers in the Quantum Moves game*, *Physical Review A* **97**, 040302 (2018).



- [225] C. Jarzynski, S. Deffner, A. Patra, and Y. Subaşı, *Fast forward to the classical adiabatic invariant*, Physical Review E **95**, 032122 (2017).
- [226] M. Song, A. Montanari, and P. Nguyen, *A mean field view of the landscape of two-layers neural networks*, in *Proceedings of the National Academy of Sciences*, Vol. 115 (2018) pp. E7665–E7671.
- [227] X. Chen, I. Lizuain, A. Ruschhaupt, D. Guéry-Odelin, and J. Muga, *Shortcut to adiabatic passage in two- and three-level atoms*, Physical Review Letters **105**, 123003 (2010).
- [228] M. Berry, *Transitionless quantum driving*, Journal of Physics A: Mathematical and Theoretical **42**, 365303 (2009).
- [229] X. Chen, E. Torrontegui, and J. Muga, *Lewis-Riesenfeld invariants and transitionless quantum driving*, Physical Review A **83**, 062116 (2011).
- [230] E. Torrontegui, S. Ibáñez, S. Martínez-Garaot, M. Modugno, A. del Campo, D. Guéry-Odelin, A. Ruschhaupt, X. Chen, and J. G. Muga, *Shortcuts to adiabaticity*, in *Advances in atomic, molecular, and optical physics*, Vol. 62 (Elsevier, 2013) pp. 117–169.
- [231] L. Theis, F. Motzoi, S. Machnes, and F. Wilhelm, *Counteracting systems of diabaticities using DRAG controls: The status after 10 years (a)*, EPL (Europhysics Letters) **123**, 60001 (2018).
- [232] S. Deffner, C. Jarzynski, and A. del Campo, *Classical and quantum shortcuts to adiabaticity for scale-invariant driving*, Physical Review X **4**, 021013 (2014).
- [233] Y.-X. Du, Z.-T. Liang, Y.-C. Li, X.-X. Yue, Q.-X. Lv, W. Huang, X. Chen, H. Yan, and S.-L. Zhu, *Experimental realization of stimulated Raman shortcut-to-adiabatic passage with cold atoms*, Nature communications **7**, 12479 (2016).
- [234] N. V. Vitanov, A. A. Rangelov, B. W. Shore, and K. Bergmann, *Stimulated Raman adiabatic passage in physics, chemistry, and beyond*, Reviews of Modern Physics **89**, 015006 (2017).
- [235] K. Bergmann, H. Theuer, and B. Shore, *Coherent population transfer among quantum states of atoms and molecules*, Reviews of Modern Physics **70**, 1003 (1998).



- [236] R. Unanyan, L. Yatsenko, K. Bergmann, and B. Shore, *Laser-induced adiabatic atomic reorientation with control of diabatic losses*, Optics communications **139**, 48 (1997).
- [237] J. Klein, F. Beil, and T. Halfmann, *Robust population transfer by stimulated Raman adiabatic passage in a Pr $3+$: Y 2 SiO 5 crystal*, Physical review letters **99**, 113003 (2007).
- [238] N. Timoney, I. Baumgart, M. Johanning, A. Varón, M. Plenio, A. Retzker, and C. Wunderlich, *Quantum gates and memory using microwave-dressed states*, Nature **476**, 185 (2011).
- [239] K. Bergmann, N. V. Vitanov, and B. W. Shore, *Perspective: stimulated Raman adiabatic passage: the status after 25 years*, The Journal of chemical physics **142**, 170901 (2015).
- [240] M. Ndong, G. Djotyan, A. Ruschhaupt, and S. Guérin, *Robust coherent superposition of states by single-shot shaped pulse*, Journal of Physics B: Atomic, Molecular and Optical Physics **48**, 174007 (2015).
- [241] X. Chen, E. Torrontegui, D. Stefanatos, J.-S. Li, and J. Muga, *Optimal trajectories for efficient atomic transport without final excitation*, Physical Review A **84**, 043415 (2011).
- [242] Q. Zhang, X. Chen, and D. Guéry-Odelin, *Fast and optimal transport of atoms with nonharmonic traps*, Physical Review A **92**, 043410 (2015).
- [243] A. C. Santos and M. S. Sarandy, *Superadiabatic controlled evolutions and universal quantum computation*, Scientific reports **5**, 15775 (2015).
- [244] S. Campbell and S. Deffner, *Trade-off between speed and cost in shortcuts to adiabaticity*, Physical Review Letters **118**, 100601 (2017).
- [245] P. Kumar, S. A. Malinovskaya, and V. S. Malinovsky, *Optimal control of population and coherence in three-level Λ systems*, Journal of Physics B: Atomic, Molecular and Optical Physics **44**, 154010 (2011).
- [246] B. B. Zhou, A. Baksic, H. Ribeiro, C. G. Yale, F. J. Heremans, P. C. Jerger, A. Auer, G. Burkard, A. A. Clerk, and D. D. Awschalom, *Accelerated quantum control using superadiabatic dynamics in a solid-state lambda system*, Nature Physics **13**, 330 (2017).



- [247] D. Guéry-Odelin and J. Muga, *Transport in a harmonic trap: Shortcuts to adiabaticity and robust protocols*, Physical Review A **90**, 063425 (2014).
- [248] S. Campbell, G. De Chiara, M. Paternostro, G. M. Palma, and R. Fazio, *Shortcut to adiabaticity in the Lipkin-Meshkov-Glick model*, Physical review letters **114**, 177206 (2015).
- [249] H. Saberi, T. Opatrny, K. Mølmer, and A. del Campo, *Adiabatic tracking of quantum many-body dynamics*, Physical Review A **90**, 060301 (2014).
- [250] A. Del Campo, J. Goold, and M. Paternostro, *More bang for your buck: Super-adiabatic quantum engines*, Scientific reports **4**, 6208 (2014).
- [251] J. Deng, Q.-h. Wang, Z. Liu, P. Hänggi, and J. Gong, *Boosting work characteristics and overall heat-engine performance via shortcuts to adiabaticity: Quantum and classical systems*, Physical Review E **88**, 062122 (2013).
- [252] M. Beau, J. Jaramillo, and A. del Campo, *Scaling-up quantum heat engines efficiently via shortcuts to adiabaticity*, Entropy **18**, 168 (2016).
- [253] M. Reck, *M. Reck, A. Zeilinger, HJ Bernstein, and P. Bertani, Phys. Rev. Lett. 73, 58 (1994).*, Phys. Rev. Lett. **73**, 58 (1994).
- [254] S. Ashhab and F. Nori, *Control-free control: Manipulating a quantum system using only a limited set of measurements*, Physical Review A **82**, 062103 (2010).
- [255] K. Jacobs, *Feedback control using only quantum back-action*, New Journal of Physics **12**, 043005 (2010).
- [256] A. Pechen, N. Il'in, F. Shuang, and H. Rabitz, *Quantum control by von Neumann measurements*, Physical Review A **74**, 052102 (2006).
- [257] S. Tanaka and N. Yamamoto, *Robust adaptive measurement scheme for qubit-state preparation*, Physical Review A **86**, 062331 (2012).
- [258] L. Roa, M. L. de Guevara, A. Delgado, G. Olivares-Rentería, and A. Klimov, *Quantum evolution by discrete measurements*, in *Journal of Physics: Conference Series*, Vol. 84 (IOP Publishing, 2007) p. 012017.



- [259] P. Hauke, R. Sewell, M. W. Mitchell, and M. Lewenstein, *Quantum control of spin correlations in ultracold lattice gases*, Physical Review A **87**, 021601 (2013).
- [260] K. Eckert, O. Romero-Isart, M. Rodriguez, M. Lewenstein, E. S. Polzik, and A. Sanpera, *Quantum non-demolition detection of strongly correlated systems*, Nature Physics **4**, 50 (2008).
- [261] L. Roa, A. Delgado, M. L. de Guevara, and A. Klimov, *Measurement-driven quantum evolution*, Physical Review A **73**, 012322 (2006).
- [262] D. Burgarth, *Quantum state transfer with spin chains*, PhD thesis, University College London (2006).
- [263] M. R. Zirnbauer, *Riemannian symmetric superspaces and their origin in random-matrix theory*, Journal of Mathematical Physics **37**, 4986 (1996).
- [264] H.-J. Stöckmann, *Quantum chaos: an introduction*, (Cambridge University Press, Oxford, 2007) .
- [265] D. Hume, I. Stroescu, M. Joos, W. Muessel, H. Strobel, and M. Oberthaler, *Accurate atom counting in mesoscopic ensembles*, Physical Review Letters **111**, 253001 (2013).
- [266] I. B. Mekhov, C. Maschler, and H. Ritsch, *Cavity-Enhanced Light Scattering in Optical Lattices to Probe Atomic Quantum Statistics*, Phys. Rev. Lett. **98**, 100402 (2007).
- [267] I. B. Mekhov and H. Ritsch, *Quantum nondemolition measurements and state preparation in quantum gases by light detection*, Physical Review Letters **102**, 020403 (2009).
- [268] I. de Vega, J. I. Cirac, and D. Porras, *Detection of spin correlations in optical lattices by light scattering*, Physical Review A **77**, 051804 (2008).
- [269] J. F. Sherson, C. Weitenberg, M. Endres, M. Cheneau, I. Bloch, and S. Kuhr, *Single-atom-resolved fluorescence imaging of an atomic mott insulator*, Nature **467**, 68 (2010).
- [270] D. Greif, M. F. Parsons, A. Mazurenko, C. S. Chiu, S. Blatt, F. Huber, G. Ji, and M. Greiner, *Site-resolved imaging of a fermionic Mott insulator*, Science **351**, 953 (2016).



- [271] M. F. Parsons, F. Huber, A. Mazurenko, C. S. Chiu, W. Setiawan, K. Wooley-Brown, S. Blatt, and M. Greiner, *Site-resolved imaging of fermionic $Li\ 6$ in an optical lattice*, Physical Review Letters **114**, 213002 (2015).
- [272] A. Bergschneider, V. M. Klinkhamer, J. H. Becher, R. Klemt, G. Zürn, P. M. Preiss, and S. Jochim, *Spin-resolved single-atom imaging of $Li\ 6$ in free space*, Physical Review A **97**, 063613 (2018).
- [273] Y. Patil, S. Chakram, and M. Vengalattore, *Measurement-induced localization of an ultracold lattice gas*, Physical Review Letters **115**, 140402 (2015).
- [274] Y. Patil, S. Chakram, L. Aycok, and M. Vengalattore, *Nondestructive imaging of an ultracold lattice gas*, Physical Review A **90**, 033422 (2014).
- [275] L. Viola, *L. Viola and S. Lloyd, Phys. Rev. A 58, 2733 (1998).*, Phys. Rev. A **58**, 2733 (1998).
- [276] F. Fröwis and W. Dür, *Measures of macroscopicity for quantum spin systems*, New Journal of Physics **14**, 093039 (2012).
- [277] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, *Reference frames, superselection rules, and quantum information*, Reviews of Modern Physics **79**, 555 (2007).
- [278] P. Facchi and S. Pascazio, *Quantum Zeno dynamics: mathematical and physical aspects*, Journal of Physics A: Mathematical and Theoretical **41**, 493001 (2008).
- [279] B. Misra and E. G. Sudarshan, *The Zeno's paradox in quantum theory*, Journal of Mathematical Physics **18**, 756 (1977).
- [280] T. J. Elliott and V. Vedral, *Quantum quasi-zeno dynamics: Transitions mediated by frequent projective measurements near the zeno regime*, Physical Review A **94**, 012118 (2016).
- [281] W. Kozłowski, S. F. Caballero-Benitez, and I. B. Mekhov, *Non-hermitian dynamics in the quantum zeno limit*, Physical Review A **94**, 012123 (2016).
- [282] S. Braun, M. Friesdorf, S. S. Hodgman, M. Schreiber, J. P. Ronzheimer, A. Riera, M. del Rey, I. Bloch, J. Eisert, and U. Schneider, *Emergence of coherence and the dynamics of quantum phase*



- transitions*, Proceedings of the National Academy of Sciences **112**, 3641 (2015).
- [283] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, *Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms*, nature **415**, 39 (2002).
- [284] D. Jaksch, H.-J. Briegel, J. Cirac, C. Gardiner, and P. Zoller, *Entanglement of atoms via cold controlled collisions*, Physical Review Letters **82**, 1975 (1999).
- [285] O. Mandel, M. Greiner, A. Widera, T. Rom, T. W. Hänsch, and I. Bloch, *Controlled collisions for multi-particle entanglement of optically trapped atoms*, Nature **425**, 937 (2003).
- [286] G. K. Brennen, C. M. Caves, P. S. Jessen, and I. H. Deutsch, *Quantum logic gates in optical lattices*, Physical Review Letters **82**, 1060 (1999).
- [287] D. Jaksch, J. Cirac, P. Zoller, S. Rolston, R. Côté, and M. Lukin, *Fast quantum gates for neutral atoms*, Physical Review Letters **85**, 2208 (2000).
- [288] J. K. Pachos and P. L. Knight, *Quantum computation with a one-dimensional optical lattice*, Physical Review Letters **91**, 107902 (2003).
- [289] E. Jané, G. Vidal, W. Dür, P. Zoller, and J. I. Cirac, *Simulation of quantum dynamics with quantum optical systems*, Quantum Information & Computation **3**, 15 (2003).
- [290] A. Micheli, A. Daley, D. Jaksch, and P. Zoller, *Single atom transistor in a 1d optical lattice*, Physical Review Letters **93**, 140408 (2004).
- [291] F. M. Cucchietti, B. Damski, J. Dziarmaga, and W. H. Zurek, *Dynamics of the Bose-Hubbard model: Transition from a Mott insulator to a superfluid*, Physical Review A **75**, 023603 (2007).
- [292] J. Zakrzewski and D. Delande, *Breakdown of adiabaticity when loading ultracold atoms in optical lattices*, Physical Review A **80**, 013602 (2009).
- [293] S. Rosi, A. Bernard, N. Fabbri, L. Fallani, C. Fort, M. Inguscio, T. Calarco, and S. Montangero, *Fast closed-loop optimal control of*



- ultracold atoms in an optical lattice*, Physical Review A **88**, 021601 (2013).
- [294] K. Jacobs and D. A. Steck, *A straightforward introduction to continuous quantum measurement*, Contemporary Physics **47**, 279 (2006).
- [295] R. Yamamoto, J. Kobayashi, K. Kato, T. Kuno, Y. Sakura, and Y. Takahashi, *Site-resolved imaging of single atoms with a Faraday quantum gas microscope*, Physical Review A **96**, 033610 (2017).
- [296] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control* (Cambridge university press, 2009).
- [297] G. Mazzucchi, W. Kozłowski, S. F. Caballero-Benitez, T. J. Elliott, and I. B. Mekhov, *Quantum measurement-induced dynamics of many-body ultracold bosonic and fermionic systems in optical lattices*, Physical Review A **93**, 023632 (2016).
- [298] G. Mazzucchi, W. Kozłowski, S. F. Caballero-Benitez, and I. B. Mekhov, *Collective dynamics of multimode bosonic systems induced by weak quantum measurement*, New Journal of Physics **18**, 073017 (2016).
- [299] W. Kozłowski, S. F. Caballero-Benitez, and I. B. Mekhov, *Quantum state reduction by matter-phase-related measurements in optical lattices*, Scientific reports **7**, 42597 (2017).
- [300] I. B. Mekhov, C. Maschler, and H. Ritsch, *Probing quantum phases of ultracold atoms in optical lattices by transmission spectra in cavity quantum electrodynamics*, Nature Physics **3**, 319 (2007).
- [301] H. Carmichael, *An Open Systems Approach to Quantum Optics*, Lecture Notes in Physics Monographs (Springer, Berlin 1993), Vol. 18 .
- [302] Z. Shen, M. Hsieh, and H. Rabitz, *Quantum optimal control: Hessian analysis of the control landscape*, The Journal of chemical physics **124**, 204106 (2006).
- [303] J. Roslund, M. Roth, and H. Rabitz, *Laboratory observation of quantum control level sets*, Physical Review A **74**, 043414 (2006).
- [304] Q. Zhang, J. G. Muga, D. Guéry-Odelin, and X. Chen, *Optimal shortcuts for atomic transport in anharmonic traps*, Journal of Physics B: Atomic, Molecular and Optical Physics **49**, 125503 (2016).



- [305] J.-F. Schaff, X.-L. Song, P. Capuzzi, P. Vignolo, and G. Labeyrie, *Shortcut to adiabaticity for an interacting Bose-Einstein condensate*, EPL (Europhysics Letters) **93**, 23001 (2011).
- [306] U. Schollwöck, *The density-matrix renormalization group in the age of matrix product states*, Annals of Physics **326**, 96 (2011).
- [307] J. Eisert, M. Cramer, and M. B. Plenio, *Colloquium: Area laws for the entanglement entropy*, Reviews of Modern Physics **82**, 277 (2010).
- [308] S. Bravyi, M. Hastings, and F. Verstraete, *Lieb-Robinson bounds and the generation of correlations and topological quantum order*, Physical review letters **97**, 050401 (2006).
- [309] J. Eisert and T. J. Osborne, *General entanglement scaling laws from time evolution*, Physical review letters **97**, 150404 (2006).

