Optimal Quantum Engineering

Dissertation zur Erlangung des Doktorgrades Dr. rer. nat.
der Fakultät für Naturwissenschaften der Universität Ulm

Vorgelegt von:
Jonathan Emanuel Zoller
aus Burghausen

2018
Amtierender Dekan
Prof. Dr. Peter Dürre

Erstgutachter
Prof. Dr. Simone Montangero

Zweitgutachter
Prof. Dr. Matthias Freyberger

Betreuer
Prof. Dr. Simone Montangero

Tag der Promotion
24.7.2018
Abstract

Optimal control of complex quantum systems in theory and experiment is the focus of this thesis. The theoretical aspect of our work was concerned with the development of algorithms that are capable of precisely tailoring the shapes of control pulses to meet various control goals while respecting all physical constraints, both fundamental and problem specific. Having these effective tools at hand, we computed control inputs for a range of physical systems using simulation-based time propagation. We extended the dressed chopped random basis (dCRAB) optimization method, previously developed in our group, such that boundary conditions and other constraints that arise from experimental setups of quantum systems can be regarded. Therefore, we wrote a software suite that enables experimentalists to improve their control pulses by directly interfacing the lab with our control algorithm.
Contents

Abstract iii

1 Introduction 1

I QUANTUM OPTIMAL CONTROL THEORY 3

2 Theoretical background 5
  2.1 Introduction ........................................ 5
  2.2 Optimal control problem formulation and control objectives .... 6
  2.3 Quantum Optimal Control Algorithms ........................ 8
    2.3.1 A first outline .................................. 9
    2.3.2 GRAPE algorithm .................................. 12
    2.3.3 Krotov algorithm .................................. 14
    2.3.4 Chopped-random-basis algorithms ..................... 15
    2.3.5 Experimental closed-loop control ..................... 18
  2.4 Analytically solvable systems ............................. 21
    2.4.1 Harmonic Oscillator side shifting .................... 22
    2.4.2 Optimal two level quantum evolutions ................ 27
  2.5 Optimal Gaussian focusing ................................ 33
    2.5.1 Derivation of equation of motion .................... 33
    2.5.2 Discussion ........................................ 35
    2.5.3 Definition of the focusing control problem ............ 36
    2.5.4 Solution via optimal control theory .................. 37
  2.6 Appendix .............................................. 39
    2.6.1 A greedy one step algorithm for QOC .................. 39
    2.6.2 Fixed-point iteration method for QOC .................. 48

3 Fundamental Limits of Quantum Optimal Control 53
  3.1 Controllability ....................................... 53
  3.2 Quantum Speed Limit ................................... 56
3.3 Information Theoretical Limit ........................................ 59
3.4 Control Landscapes ..................................................... 62

II OPEN & CLOSED-LOOP CONTROL APPLICATIONS 65

4 Amplification of the Dynamical Casimir Effect 67
  4.1 Introduction to the dynamical Casimir effect ...................... 67
  4.2 Time-dependent Rabi Hamiltonian ................................... 68
  4.3 Amplification of the parametric dynamical Casimir effect via optimal control ......................................................... 71
    4.3.1 Abstract .......................................................... 71
    4.3.2 Introduction ..................................................... 71
    4.3.3 The model ....................................................... 73
    4.3.4 Results ........................................................... 74
    4.3.5 Conclusions ..................................................... 79
    4.3.6 Appendix A: Optimized Bang-bang strategy ..................... 81

5 Optimized BEC production 83
  5.1 Introduction into BEC condensation .................................. 83
  5.2 Remote optimization of an ultra-cold atoms experiment .......... 88
    5.2.1 Abstract ........................................................ 88
    5.2.2 Introduction ..................................................... 88
    5.2.3 Initial landscape investigations ................................ 90
    5.2.4 RedCRAB optimization ......................................... 94
    5.2.5 Citizen science optimization ................................... 97
    5.2.6 Supporting Information ......................................... 101

6 Optimal control of a spin qubit in NV color center in diamond 121
  6.1 Introduction into NV centers ......................................... 121
  6.2 Autonomous Calibration of Single Spin Qubit Operations .......... 124
    6.2.1 Abstract ........................................................ 124
    6.2.2 Introduction ..................................................... 124
    6.2.3 Results ........................................................... 127
    6.2.4 Discussion ....................................................... 128
    6.2.5 Methods .......................................................... 131
    6.2.6 Supplementary Material ......................................... 133
## 7 Central Moment Dynamics

7.1 Introduction into approximate quantum simulation methods . . . . . . 141
7.2 Quantized Hamilton Dynamics and Quantal Cumulant Dynamics . . . 142
7.3 Central Moment Dynamics . . . . . . . . . . . . . . . . . . . . . . . . 144
7.4 Proof of recursion formula . . . . . . . . . . . . . . . . . . . . . . . . 147
7.5 Central moment dynamics of continuous variables quantum systems . . 151
  7.5.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 151
  7.5.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 151
  7.5.3 CANDY method . . . . . . . . . . . . . . . . . . . . . . . . . . 154
  7.5.4 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 155
  7.5.5 Effective dimension . . . . . . . . . . . . . . . . . . . . . . . . 157
  7.5.6 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 158
  7.5.7 Supplementary Material . . . . . . . . . . . . . . . . . . . . . . 159
7.6 Appendix . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 170
  7.6.1 Appendix A: Alternative closure scheme . . . . . . . . . . . . . 170
  7.6.2 Appendix B: Energy fluctuation in terms of moments . . . . . . 173

## 8 Robust optimal control of single qubit operations

8.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 175
8.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 175
8.3 Problem definition . . . . . . . . . . . . . . . . . . . . . . . . . . . . 176
8.4 Controllability analysis . . . . . . . . . . . . . . . . . . . . . . . . . . 177
8.5 Optimal control strategy . . . . . . . . . . . . . . . . . . . . . . . . . 179
8.6 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 182
  8.6.1 Optimization of eight normal distributed detunings . . . . . . . 182
  8.6.2 Robust control of equidistant detunings of various ranges . . . . 184
8.7 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 186
8.8 Appendix . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 187
  8.8.1 Derivation of qubit dynamics in angular variables . . . . . . . . 187
  8.8.2 Optimized pulse shapes . . . . . . . . . . . . . . . . . . . . . . . 189

## 9 Conclusions

## A Quantum Optimal Control Software Projects

A.1 Quantum Toolbox in Python . . . . . . . . . . . . . . . . . . . . . . . . 193
A.2 RedCRAB . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 194

List of Figures
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>203</td>
</tr>
<tr>
<td>Bibliography</td>
<td>205</td>
</tr>
<tr>
<td>List of publications</td>
<td>243</td>
</tr>
</tbody>
</table>
1 Introduction

Quantum physics showed a stunning development over the course of the 20th century. When Max Plank discovered the law behind black body radiation in 1900, he realized that energy exchange must be quantized [1]. Building on this finding, Albert Einstein described the photoelectric effect in 1905 and showed that the relation between incoming electromagnetic radiation and the energy of emitted electrons scales according to the fundamental Planck constant $\hbar$. These developments mark the point in history at which humanity started to perceive the universe as essentially being quantized [2]. In the upcoming years, the theory of wave mechanics was developed [3–9]. Two of the main concepts thereof, superposition and entanglement, form the basic ingredients for the theory of and applications in the field of quantum information processing, which has prospered remarkably over the past three decades. While superposition describes the phenomenon of a quantum mechanical system to be in two (or more) states at the same time, entanglement is a fundamental manifestation of non-classical correlations. Therefore, we can store and process exponentially many configurations (and hence information) combining only a linear amount of physical resources (such as qubits). Going one step further, these phenomena can be exploited using algorithms for quantum computers – such as the Grover search algorithm [10] and Shor’s algorithm for integer factorization [11] – by making use of their inherent parallelism. Quantum parallelism stems from the ability to perform an operation (such as a function evaluation) on the exponentially many states all at once. However, the physical realization of quantum devices, i.e. applications that exploit quantum effects, is extremely challenging because the essential building blocks, such as realizations of stable two-level systems (qubits), are extremely sensitive to a variety of sources of noise. This fact prevented the scientific community from making rapid progress towards widely used real world applications. However, in recent years this began to change: with the advent of a series of developments in manufacturing, calibration and control of platforms such as neutral atoms in optical lattices [12], trapped ions [13] or superconducting devices [14], more and more qubits could be combined in a quantum circuit [15].

Common to all of these novel technologies such as quantum sensing, simulation, communication and computing [16–20], but also across all physical platforms, is the requirement
to engineer these systems. Precisely tailored laser pulses, radio-, or microwave radiation are inevitably needed for any technological application. That is the realm of quantum control: this technology enables high precision steering of dynamical processes at the atomic or molecular scale [21]. The necessary background and its applications – in theory and experiment – are presented here.

This thesis is a summary of the work over the past four years carried out at Ulm University and has been partially published in peer reviewed journals. In the first two chapters after the introduction an overview of quantum optimal control theory is provided. Special emphasis is put on the algorithmic part as well as opportunities and pitfalls with regard to experimental control. Chapter 4 gives an insight into how to enhance photon generation from the vacuum as described by the dynamical Casimir effect. These systems are nowadays experimentally accessible and our computed control pulses might serve as an adequate starting point as we also affirmed our solution to be robust against disturbances. In the following two chapters we directly applied our control algorithm on two different experiments. Optimizing the evaporative cooling sequence in a BEC setup is reported in Chapter 5 and the fidelity of single qubit operations of NV centers in diamond has been optimized in the work described in Chapter 6. A method targeting and addressing some relevant degrees of freedom of continuous variables quantum systems via central moments of the wave function is introduced in Chapter 7. In addition, the relation with optimal control is discussed. Finally, we developed a hybrid optimal control method to generate robust pulses with respect to detuning fluctuations of qubits which is reported in the next Chapter. Chapter 9 concludes the work and an overview of our remote experimental control software suite is given in the appendix.
Part I

QUANTUM OPTIMAL CONTROL THEORY
2 Theoretical background

Quantum optimal control theory, which attracts lots of attention these days, forms the basis of this thesis. We present how it enables a multitude of physical processes, which would, in some cases, otherwise not be possible.

2.1 Introduction

The origins of optimal control theory (OCT) dates back much before the human perception of quantum phenomena. Johann Bernoulli’s brachistochrone curve problem, introduced in 1696, can be considered as one of the earliest examples of the field [22]. In the following two centuries to come, calculus of variations has been introduced as an elegant way to tackle dynamical optimal control problems. Famous mathematicians were involved in the development of the theory, most notably the Bernoulli brothers, Euler, Lagrange, Legendre, Jacobi, Weierstrass, Hilbert, and Carathéodory [23; 24]. Going much further ahead in time, an important milestone is marked by the formulation of the Maximum Principle by L. S. Pontryagin [25]. It provides necessary conditions for (constrained) optimal control and has proven its usefulness over all the decades until today [26].

Quantum optimal control theory (QOCT) has been introduced in the late 1980s [27–30] when control strategies for the design of, in most cases, electromagnetic field profiles were developed. In the following years, the field started to mature and a series of sophisticated techniques were presented. One important example is the Krotov method, still being actively explored, which has been introduced by Tannor and coworkers to the field [31; 32]. Its foundations are based on Krotov’s initial work. A widely adopted variant is the method introduced by Zhu and Rabitz [33] in the late 1990s. At this stage, quantum optimal control theory was mainly concentrating on optimizing state transfers and extremizing the expectation value of an operator. With the advent of major discoveries in the field of Quantum Information Processing (QIP), such as Shore’s and Grover’s algorithm [10; 11; 34], around and shortly after the turn of the millennium,
QOCT received more and more interest. On the one hand, challenges such as highly precise quantum gate synthesis, and, on the other hand, ground-breaking progress on the realization and control of experimental apparatus across various platforms [35–39], opened up new horizons and frontiers. Nowadays, the theoretical description as well as open-loop algorithms are believed to be well developed, while the major obstacles that prevent us from developing further quantum technologies lie in our limited experimental capabilities in scaling up system sizes while maintaining reasonable error thresholds. Due to the fact that we lack of favorably scaling numerical methods to simulate arbitrary quantum systems on a classical computer that would allow us to study, predict and also generate an open-loop solution, closed-loop control is getting into the focus. More and more interest was attracted in the recent years also owing to quite some break-throughs that have been reported on the experimental realization of controllable large-scale quantum technologies platforms such as superconducting qubits [40; 41] and Rydberg atoms [42].

In this chapter, we give an introduction to common problem formulations in quantum optimal control in Sec. 2.2 and present some successful algorithms to solve them in Sec. 2.3. Special emphasis is also placed on experimental control. In addition, some rare analytically solvable systems are discussed in Sec. 2.4. In Sec. 2.5 the optimal focusing of Gaussian wavepackages is demonstrated and some preliminary results on novel control strategies are presented in the appendix of this chapter (Sec. 2.6).

2.2 Optimal control problem formulation and control objectives

Optimal control problems are formulated in terms of the extremization (maximization or minimization) of some functional $J$ that maps the resulting state trajectory, denoted by the time-dependent density operator $\rho(t)$, (considering either all instances in time or only a finite set) of a system driven by (a set of) control pulses $u(t)$ to the $\mathbb{R}^1$. In most of the cases, the initial state is given and $J$ is an (explicit) function of the final state (=evolved state after some transfer time $T$), only. Furthermore, in almost all the control literature the control linear approximation is applied: $H(u(t)) = H_d + u(t)H_c$. Common alternative notations for drift and control Hamiltonian are $H_d = H_0$ and $H_c = H_1$, respectively. In the case of a pure state transformation, the control problem is to find a control pulse $u(t)$ that drives the system from an initial state $|\psi_0\rangle$ to a goal
2.2 Optimal control problem formulation and control objectives

State $|\phi\rangle$ in time $T$. Hence the problem reads

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H(u(t)) |\psi(t)\rangle$$

(2.1)

$$|\psi(t = 0)\rangle = |\psi_0\rangle$$

(2.2)

$$|\psi(t = T)\rangle = |\phi\rangle$$

(2.3)

A great magnitude of different control goals $J$ are doable and have been studied. In the following paragraphs the most common ones are presented. All pure state notations are straight-forwardly extendable towards mixed states, too. For example, if the density matrix describing a mixed state is denoted by $\rho = \sum \omega_i |\psi_i\rangle \langle \psi_i|$ where the $\omega_i$ give the fraction of the ensemble to be in $|\psi_i\rangle$, the expectation value of an observable $O$ reads

$$\langle O \rangle = \sum \omega_i \langle \psi_i | O | \psi_i \rangle = \text{tr}(\rho O).$$

State to state transfer One of the most widely applied control problems is the state to state transformation from an initial state $|\psi_0\rangle$ to some goal state $|\phi\rangle$. Often, the total transfer time $T$ is given. The control problem might be formulated as a maximization of the functional $J$ being the fidelity

$$\max_{u(\cdot)} J = |\langle \psi(T) | \phi \rangle|^2.$$  

(2.4)

In rare cases the square is dropped from the fidelity overlap and sometimes $J$ is formulated as $\text{Re}\{\langle \psi(T) | \phi \rangle\}$ which is also equivalent to $1 - \frac{1}{2} ||\psi(T) - \phi||_2^2$. The latter term significantly changes the control objectives as the global phase is fixed, too. Besides different algorithmic convergence properties, the controllability criterion (see Sec. 3.1) changes, too. Often additional constraints on the pulse or accessible states are posed such as maximum pulse amplitude, maximum bandwidth of the pulse or highest populated energy.

Observable optimization The optimization of the expectation value of some observable $O$ can be formulated as an extension of the state transfer from Eq. (2.4). It might either be evaluated at the final time $T$ (Eq. (2.5)) or averaged over the whole time interval $[0, T]$ (Eq. (2.6)):

$$J = \langle \psi(T) | O | \psi(T) \rangle$$

(2.5)

$$J = \frac{1}{T} \int_0^T \langle \psi(t) | O | \psi(t) \rangle \, dt$$

(2.6)
**Unitary gate optimization**  For quantum gates, even more different options for the yield functional $J$ have been studied so far [43–45]. Two commonly used ones are

$$J = \frac{1}{N} \text{Re}\{\text{Tr}(V^\dagger U(T))\} \quad (2.7)$$

$$J = \frac{1}{N^2} \left| \sum_{i=1}^{N} \langle \xi_i | V^\dagger | \psi_i(T) \rangle \right|^2, \quad (2.8)$$

where $V$ is the target unitary, $U(T)$ the total time $T$ evolved unitary and the integer $N$ is the size of the basis. $U(T)$ also connects initial $\{|\xi_i\rangle\}$ and final states $\{|\psi_i(T)\rangle\}$ via $|\psi_i(T)\rangle = U(T) |\xi_i\rangle$. The unitary $U(t)$ obeys the Schrödinger Equation (2.1) $\forall t$, just as the states, but with an initial value of $U(t = 0) = 1_{[N \times N]}$.

**Entanglement maximization**  Quantum entanglement describes the phenomenon of states that are not separable such that the state of one subsystem (e.g. a single atom or photon) cannot be independently described of the state of another subsystem. To measure the degree of entanglement multiple notations exist [15; 46; 47]. In general a measure shall be invariant with respect to local transformations and is often defined to be zero for separable states. Here, we introduce the von Neumann entropy which is a common measure [48]. Entropy quantifies the amount of uncertainty in the state and is in general attributed to a probability distribution. In classical information theory, the Shannon entropy $\mathbb{H}$ associated to a probability distribution $\{p_i\}$ is defined as

$$\mathbb{H}(p_1, ..., p_n) = -\sum_i p_i \log p_i. \quad (2.9)$$

Borrowing this idea and applying it to a mixed state $\rho$ with fraction $\omega_i$ in the pure state $|\psi_i\rangle$ defining an ensemble, we get for the von Neumann entropy:

$$S(\rho) = -\sum_i \omega_i \log \omega_i = -\text{tr}(\rho \log(\rho)). \quad (2.10)$$

For the control objective, we can now define $J = S(\rho(T))$.

### 2.3 Quantum Optimal Control Algorithms

Calculating optimal control pulses $u^*(t)$ that meet the previously defined control goals (Eq. (2.4) – (2.10)) while respecting all constraints require in most cases iterative numerical algorithms. The most successful ones are described in this section.
2.3 Quantum Optimal Control Algorithms

2.3.1 A first outline

The field of QOCT might be divided into open- and closed-loop control. While the former is attributed to the simulation-based design of control pulses of dynamical quantum systems that try to model real systems as close as possible, the latter’s objective is the optimization of a dynamical process directly in the lab and is sometimes denoted as experimental control. In the literature various terms for the same discrimination were introduced. Some of the most commonly used synonyms for open-loop (closed-loop) control are \textit{ex situ (in situ)}, offline (online) optimization, simulation based (feedback) control [49–52]. More details about closed-loop control can be found in 2.3.5. In order to visualize these concepts and their interdependences a brief sketch is shown in Fig. 2.1.

\textbf{Figure 2.1:} Open-loop versus closed-loop control. Open-loop control optimizes the simulation of a physical system represented by the Hamiltonian $H$, Lindbladian $\mathcal{L}$ and possible noise modeling $\mathcal{E}$. The outcome from open-loop optimal control might serve as an initial guess for subsequent closed-loop optimization (dashed line pointing eastbound).

With the advent of new developments in QIP, advanced methods have been developed for both open- and closed-loop optimal control. We stress that some methods that have been originally designed to tackle open-loop control problems are applicable for closed-loop control, too. In turn, algorithms for closed-loop control are always applicable for open-loop considerations. We will elaborate more on this point in the next section.
Some of the main QOCT methods might be categorized as sketched in the following chart which includes a (non exhaustive) list of examples. While there has been some attempt to compare different optimal control algorithms in terms of their performance, computational cost and applicability \([43; 44; 53–55]\), according to the \textit{no free lunch theorem} \([56]\) an ultimate answer to the question “which algorithm is best” cannot be given in general. Typically, any comparison can only make statements that are valid for certain classes of problems. Anyway, it is commonly believed that we will eventually have several tools for several problems just as there is a separate hardware tool for every screw, nut and nail.
2.3 Quantum Optimal Control Algorithms

<table>
<thead>
<tr>
<th>Direct methods</th>
<th>Indirect methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformation of the optimal control problem to a standard optimization problem of constrained nonlinear multivariable functions</td>
<td></td>
</tr>
<tr>
<td>Indirect methods</td>
<td></td>
</tr>
<tr>
<td>Solving the optimality conditions as an outcome from calculus of variations</td>
<td></td>
</tr>
<tr>
<td>• GRAPE (Sec. 2.3.2)</td>
<td></td>
</tr>
<tr>
<td>• Krotov (Sec. 2.3.3)</td>
<td></td>
</tr>
<tr>
<td>• Other gradient based approaches [59]</td>
<td></td>
</tr>
<tr>
<td>• Shooting methods [60; 61]</td>
<td></td>
</tr>
<tr>
<td>• ...</td>
<td></td>
</tr>
</tbody>
</table>

Function expansion

Expansion of the control functions in a suitable basis and subsequent optimization of coefficients

- Fully discretized [57]
- Partially discretized [58]
- Direct multiple shooting (Ch. 8)
- ... |

Miscellaneous

- Lyapunov control [67–69]
- Geometric control theory [70–73]
- Pontryagin maximum principle [26; 74; 75] and (Sec. 2.4.2)
- QOC via Quantum Brachistochrone [76]
- Invariant based control [77]
- QOC with Floquet theory [78]
- QOC as Newton-Raphson root finding [79]
- Genetic Algorithm (Sec. 2.3.5)
- Stochastic processes and statistical gradients (Sec. 2.3.5)
- Many more, e.g. [51; 80–86]

Open-Loop only

Closed-Loop capability

We stress that this categorization is neither definite nor complete. Direct methods, despite being relatively rarely used in quantum optimal control, offer the great advantage that all the knowledge from centuries of research in optimization and its powerful algorithms can be used because of the reduction to a static (ordinary) optimization problem. These algorithms might be in the class of Newton methods, interior point methods, trust region methods, methods exploiting the Karush–Kuhn–Tucker conditions, simplex method (not to be confused with the simplex from the Nelder-Mead algorithm), greedy algorithm, sequential quadratic programming, differential evolution, genetic algorithms, or simulated annealing just to name some of the most popular ones. However,
costate variables (denoted by $|\lambda(t)\rangle$ in this thesis), containing physical insights on an optimized process, are not being employed and hence cannot be used for interpretation and further developments. In contrast, **indirect methods** do have these capabilities and therefore are still being used most frequently in the QOCT community. As already pointed out in Sec. 2.1, the Krotov method has been one of the first algorithms being used in the field. This algorithm, together with the so called **GRadient Ascent Pulse Engineering** (GRAPE) algorithm will be briefly introduced in the next sections. As opposed to the indirect methods that require analytical gradients to be computed beforehand, methods in the class of **function expansion** are based on a, at a first glance, completely distinctive design principle: control pulses are expanded in an often physically motivated basis that ensures experimental feasibility. As a consequence, the only thing that remains is the minimization of the coefficients of the basis functions which can be done with any standard method. Although this methodology appears to be completely different from the indirect methods, one could argue that any numerical method of the latter class will have to discretize the control pulses and hence introduce some sort of basis. In that regard, these two classes are much more similar than widely perceived. A bit offside the main playground, a vast variety of other successful ansatzes have been developed which we summarize under the term **miscellaneous** (misc). With the rising interest in experimental control, especially in the BEC community, quite some research has been devoted to genetic algorithms in the closed-loop setting. We will elaborate a bit more on this in Section 2.3.5. Nevertheless, some of the other approaches in misc, such as Lyapunov control and the Pontryagin Maximum Principle (PMP), are still widely appreciated and subject to active research. To the misc class also belongs an approach for a greedy algorithm whose rough idea and preliminary results are sketched in the appendix (Sec 2.6.1).

### 2.3.2 GRAPE algorithm

The central idea of the **GRadient Ascent Pulse Engineering** (GRAPE) algorithm [87] is very straight forward: The yield functional $J(u)$ is Taylor expanded to first order in $u$

$$J(u + \Delta u) \approx J(u) + \Delta u \frac{dJ}{du},$$

and the update (improvement direction) can be derived to be

$$\Delta u = u^{(i+1)}(t) - u^{(i)}(t) = -\frac{2\alpha}{\hbar} \text{Im} \{ \langle \lambda^{(i)}(t) | \frac{\partial H}{\partial u} | \psi^{(i)}(t) \rangle \}.$$ 

(2.12)
Here, $i$ is the iteration number and the parameter $\alpha > 0$ is the search length. The update $\Delta u$ is then being applied on the previous pulse for all times at once which makes it a concurrent algorithm. The state trajectory $|\psi^{(i)}(t)\rangle$ is obtained from a regular (forward) propagation of the system with the previous pulse $u^{(i)}$, while the $|\lambda^{(i)}(t)\rangle$ can be obtained when considering further optimality conditions that follow from calculus of variations of the extended (via Lagrange multipliers that can be identified as costates $|\lambda(t)\rangle$) yield functional $\tilde{J} = J + 2\text{Re}\{\int_0^T \langle \lambda(t) | \left[ \frac{i}{\hbar} H(t) + \frac{\partial}{\partial t} \right] |\psi(t)\rangle \, dt\}$. In order to obtain $0 = \delta \tilde{J}$, we get the following equations (in addition to $\Delta u = 0$ from Eq. (2.12)) in the case of state fidelity maximization:

\begin{align*}
    i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle &= H(t) |\psi(t)\rangle \\
    i\hbar \frac{\partial}{\partial t} |\lambda(t)\rangle &= H(t) |\lambda(t)\rangle \\
    |\lambda(T)\rangle &= \langle \phi |\psi(T)\rangle |\phi\rangle,
\end{align*}

where $|\phi\rangle$ is the goal state, and $|\langle \phi |\psi(T)\rangle|^2$ shall be maximized. Hence $|\lambda(T)\rangle$ can be obtained from Eq. (2.15) and the whole costate trajectory $|\lambda^{(i)}(t)\rangle$ by propagating backwards in time using Eq. (2.14).

In order to reach higher precisions and faster convergence rates, second order information can also be taken into account. Expanding the functional $J(u)$ from Eq. (2.11) up to second order gives

\begin{equation}
    J(u + \Delta u) \approx J(u) + \Delta u \frac{dJ}{du} + \frac{1}{2} \Delta u^T h(u) \Delta u,
\end{equation}

where $h(u)$ is the Hessian containing second derivative information. However, for most practical scenarios the (exact) computation of $h(u)$ becomes prohibitively computationally complex. As a good compromise between accuracy and speed, methods in the class of quasi-Newton methods might be employed. The idea is to efficiently compute an approximate Hessian based on the obtained gradient information $\Delta u^{(i)}$ in previous iterations of the algorithm. One of the most popular methods is based on the Broyden, Fletcher, Goldfarb, Shanno (BFGS) formula [88].

As an intermediate method between the pure GRAPE and quasi-Newton methods, the so called conjugate gradients can be exploited [89], where the gradient information from the previous step, together with the current gradient, is taken into account.
The two main branches of Krotov algorithms that have been developed, and also formed the de facto numerical tools in QOCT in the first years \cite{90,91}, are the one from Tannor \textit{et al.} \cite{32} and the one from Zhu and Rabitz \cite{33}. The main design principle of the former is to sweep forward in time, updating controls as well as the state trajectory \( |\psi(t)\rangle \) simultaneously which makes it a sequential algorithm (for a good visualization see Fig. 1 in \cite{43}). Hence, the central equation for updating the controls reads:

\[
u^{(i+1)}(t) = u^{(i)}(t) + \alpha S(t) \text{Re}\{\langle \lambda^{(i)}(t) | \frac{\partial H}{\partial u} | \psi^{(i+1)}(t) \rangle\}, \tag{2.17}
\]

where \( i \) gives the iteration number. The states \( |\lambda^{(i)}(t)\rangle \) and \( |\psi^{(i)}(t)\rangle \) are solutions to the Schrödinger Equation. The parameter \( \alpha > 0 \) is a search length parameter which greatly influences convergence speed and \( S(t) \) is a shape function \( 0 \leq S(t) \leq 1 \) which allows smooth turning on and off at the boundaries. In addition, the formulation of Zhu and Rabitz applies this kind of sweeping procedure for the backward propagation on the adjoint equation, too. Furthermore, we can recognize the structure of the expression \( \Delta u \) from the GRAPE method.

The beauty of the Krotov method is that it has the so called monotonic convergence property, which gives a purely positive increase in the yield functional at every iteration, if a regularization term \( R \) is added to the objective \( \bar{J} = J + R \). In the first years the following regularization was commonly applied \cite{53}

\[
R = -\frac{\beta}{2} \int_{0}^{T} u^2(t) dt, \tag{2.18}
\]

where \( \beta > 0 \) is a constant parameter that balances the need for small control amplitudes and high fidelity values. The problem with this regularization from Eq. (2.18) is that stationary points from \( \bar{J} \) (where the algorithm hopefully converges to) do not coincide with the ones from \( J \). This can be seen from

\[
\frac{\delta \bar{J}}{\delta u(t)} = \frac{\delta J}{\delta u(t)} - \beta u(t) = 0 \Rightarrow u(t) = \frac{1}{\beta} \frac{\delta J}{\delta u(t)}. \tag{2.19}
\]

On the other hand, a necessary condition for \( J \) to have an extremum is \( \frac{\delta J}{\delta u(t)} = 0 \). Hence the only shared critical point is the function \( u(t) = 0 \forall t \), which is often not optimal.
To overcome this issue, an alternative regularization has been introduced

\[ R = -\frac{\beta}{2} \int_0^T (u^{(i)}(t) - u^{(i-1)}(t))^2 dt, \] (2.21)

based on the difference in pulse amplitudes between the previous and current iteration. A proof for monotonic convergence can be found in the appendix of [91]. Despite this 'convergence' proof, (true) convergence cannot be guaranteed. In fact, for many practical scenarios it turns out that the improvement rates \( \Delta \bar{J} = \bar{J}^{(i)} - \bar{J}^{(i-1)} \) drop exponentially fast in \( i \) and we are left over with a suboptimal solution.

Closing this subsection, we like to add some remarks: The search length parameter \( \alpha \) might also be chosen automatically to speed-up convergence [44]. The Krotov algorithm might as well be extended to second or higher orders [92]. Further constraints on states [93] and pulses [94] can be respected via suitable extensions.

### 2.3.4 Chopped-random-basis algorithms

In contrast to the Krotov and GRAPE algorithms, the Chopped RAndon Basis (CRAB) algorithm [95; 96] can natively solve open- and closed-loop control problems, does not require any analytical computation beforehand nor any backward propagation and has a great ease of use. Because of these favorable properties we have chosen it to be at the core of our RedCRAB program suite (see Sec. A.2). The main idea is to expand the control field in a suitable truncated function basis \( \{f_i(t)\} \)

\[ u(t) = \sum_{i=1}^{N_f} c_i f_i(t), \] (2.22)

where the real coefficients \( c_i \) are subject to optimization via any method of choice. A common choice is the Nelder-Mead (NM) algorithm [97], however the type of optimization method is explicitly not limited to direct searches via NM. One peculiarity of CRAB is that the basis functions \( f_i \) are usually picked from an orthonormal basis and then being randomized. As a consequence, the effective search volume in terms of the function space is enlarged with respect to a regular basis. A mathematical reasoning for the effectiveness of the function expansion from Eq. (2.22) is presented in Sec. 3.3, which is based on the idea that the information content of typical quantum optimal control problems is very much limited. Further evidence for this argument is provided from the outcome of the analysis presented in Chapter 7.

Initial guesses \( u_0(t) \) on the pulse can be considered by extending \( u(t) \) from Eq. (2.22)
to the total pulse $u_t(t)$

$$u_t(t) = \zeta u_0(t) + (1 + \zeta_x(u_0(t) - 1))u(t),$$

(2.23)

where the two binaries $\zeta_+, \zeta_x \in \{0, 1\}$ control which type of update based on the initial guess should be employed. Most common are the options $(\zeta_+, \zeta_x) = (1, 0)$ and $(\zeta_+, \zeta_x) = (1, 1)$. For the option $(\zeta_+, \zeta_x) = (0, 1)$ the algorithm is not able to shift or remove any of the roots that $u_0(t)$ might exhibit.

The choice of the function basis is completely arbitrary but is often physically motivated. For instance, if the optimized pulses must meet bandwidth constraints as being posed by an experimental apparatus, the Fourier basis with respective limited highest frequency proportional to $N_f$ can be selected [52; 66]. Other bases could be Hermite [79] or Chebyshev polynomials, Gaussian bell curves [62; 98], bang-bang switches [99] or simply ramps or low order polynomials [100; 101].

Limits on the control pulse amplitudes, such as $b_l \leq u(t) \leq b_u \forall t$, are frequently occurring and can be respected using different approaches. One option is to simply chop any excursion beyond $b_l$ or $b_u$ to the respective limit which has the disadvantage of introducing high frequencies in the pulse due to the emerging edges. Therefore usually a better option is to take some smooth function $\varphi(u)$ such as $\tanh(u)$ that maps from $\mathbb{R} \rightarrow [-1, 1]$ and compute the limited pulse $\bar{u}(t)$ based on the unconstrained pulse $u(t)$:

$$\bar{u}(t) = \left(\frac{b_u + b_l}{2}\right) + \left(\frac{b_u - b_l}{2}\right)\varphi(u(t))$$

(2.24)

Finally, when using the Nelder-Mead algorithm to optimize the coefficients $c_i$, constrained pulses can be generated by attributing a large penalty cost (and also skipping its figure of merit evaluation) for a respective pulse if the amplitude limits are violated. That way the algorithm will reject the respective pulse as an optimal candidate and hence deliver pulses within the boundaries, only.

Many other restrictions on pulse shapes and even on accessible states can be factored in either via penalty terms or pulse wise by employing suitable transferfunction $\tilde{\varphi}(u)$ that ensure the desired properties. As a final remark we like to point out that the gradient methods are not completely flexible in terms of usable figure of merit functionals $J$ as there are formulations where analytical gradients cannot be computed even theoretically. The algorithms of the class function expansion do, when direct search is applied, not suffer from this restriction as only figure of merit evaluations are required.

One potential drawback of the (non dressed) CRAB algorithm is its vulnerability with respect to so called false traps [102]. These are suboptimal points where the algorithm
might get stuck at and are characterized by
\[
\frac{dJ}{dc_i} = 0 \forall i \land \exists du : \frac{dJ}{du} \neq 0.
\]
(2.25)

As a cure, the optimization can be halted at that point while the pulse information is kept and only resumed after employing a different set of basis functions \{f_i(t)\} subject to search. This method has been introduced in [103] under the term dressed CRAB (dCRAB) optimization. The search per set of basis functions is referred to as a single super-iteration. Hence, the expansion scheme reads
\[
u(t) = \sum_{j=1}^{N_{SI}} \sum_{i=1}^{N_f} c^*_{i,j} f_{i,j}(t),
\]
(2.26)

where the index \(i\) runs over the basis functions and the index \(j\) runs over the super-iterations (total number \(N_{SI}\)). The asterisk at the \(c^*\) indicates that they have been optimized.

In summary, the algorithm works with two nested layers where the lower layer optimizes in a fixed functional subspace and the upper layer is responsible for periodical restarts of the lower layers, where each new restart comes with a new (and hence different) set of basis functions, while the previous optimization efforts will always be kept. The effect of the basis change at the succeeding super-iteration is visualized in Fig. 2.2, where the unlocking of a false trap is exemplified. Consequently, dCRAB becomes basically equivalent to the gradient methods in terms of its capability to follow the instantaneous gradient of the control landscape [103] while having a number of advantages as mentioned at the beginning of the section. Although seemingly straight-forward to be implemented, some interesting alternative variants [63; 104] emerged that did not strictly follow the recipe from the literature. Moreover, recently there have been proposals in merging GRAPE with the CRAB ansatz [63; 105] and an approach where also a function expansion is employed and the time evolution propagator together with its gradient is utilized to obtain an improvement direction in the coefficients [62].

Mentioning only three of the latest (d)CRAB open-loop applications (for closed-loop applications see Sec. 2.3.5), the algorithm has been used to amplify the dynamical Casimir effect (see Chapter 4 and [106]), prepare different many-body states of Rydberg atoms [107] and greatly influence the relaxation speed of a qubit [108].
2.3.5 Experimental closed-loop control

Here we give an overview over some of the most important concepts, problems and methods in the field of optimal control of quantum experiments. Some software suites capable of solving these problems are described in appendix A.

Closed-loop quantum optimal control is quite distinct in comparison to open-loop control: for open-loop optimization, the outcome of the propagation of the system for a specific pulse is usually stable and readily reproducible, if not, the expected outcome can usually be determined very accurately and fast by simulating noise. The reason for this is, on the one hand, that using analytical or numerical calculations, in general, less sources of errors are taken into account. On the other hand, if averaging over several realizations is needed, nowadays parallel clusters allow to do this efficiently. Furthermore, errors, disturbances, and additional time consumption due to initial state preparation and final state read out and subsequent evaluation (often by the use of fits and regressions) are not an issue with simulations.

Apart from those rather practical points, there are also some fundamental reasons that prohibit methods and insights from open-loop control to be straight-forwardly extended towards closed-loop control: in many experimental setups such as, but not exclusively,
manipulation of objects on a microscopic scale, we do, if existent at all, only have very rough models of the relevant degrees of freedom and their interaction (e.g. the system Hamiltonian) at hand. In quantum control, especially the unavoidably present system-environment interaction is never perfectly known. Even though we have seen quite some rise in interest in the study and control of open quantum systems, we have a long way to go for a comprehensive picture. Especially the control of non-Markovian evolution is still largely uncharted territory with few studies carried out so far [109–112]. Furthermore, the laws of quantum mechanics do not allow to measure the system’s state at every instance in time without disturbing it: The collapse of the wave function on measurement does not allow unaffected further propagation without reinitialization. In a sense, this makes the study and control of quantum systems cumbersome. Note that the measurement difficulty does not only occur within quantum control problems: Batch reactors in chemical engineering may not allow for pressure or temperature measurements during a reaction process, large scale biological systems such as the number of species in a habitat or the number of bacteria in a solution might be practically not determinable or variables of the human body during a surgical operation might be inaccessible without causing risks to lives. Even though we could have access to the experimental state trajectory, quite some open-loop optimal control methods would still be not applicable as they require a propagation backwards in time, which we believe is not possible at all in the real world.

For some classes of systems, however, some successful applications of previously open-loop optimized pulses have been reported [66; 85; 113; 114]. However, there is only a very limited number of literature about subsequent open- and closed-loop optimization. In [65] they introduce a combination of open-loop GRAPE optimization and subsequent closed-loop ’Nelder-Mead control’. Turning towards tailoring pulses via closed-loop optimization, some natural difficulties arise: On one hand, there are in principle infinite degrees of freedom to be adjusted to achieve a certain control task. On the other hand, there is often only a very limited number of pulse evaluations experimentally achievable. However, the principally contributing number of relevant degrees of freedom is rather small for most realistic scenarios [115; 116]. Hence, the control goal within finite precision can be met as there are errors unavoidably occurring at any experimental stage. Such remaining degrees of freedom can be addressed through a suitable parametrization of the pulse (see also Chapter 7). In fact, taking this route, quite some remarkable results have been demonstrated: in [117] the authors employ an evolutionary algorithm to optimize different stages of a cold atom experiment. Here, the parametrization are constant.
functions for the magnetic fields and the detuning of the cooling laser frequency. Going one step further, in [118] the authors combine an evolutionary algorithm with a differential evolution step to enhance performance of an ultra-cold quantum gas experiment. The control functions are laser detunings, laser intensities and coil currents which are parametrized using (linear) ramps. Another BEC experiment, which employs machine learning in contrast to evolutionary algorithms, has been reported recently [49]. In detail, they model the experimental parameters as a stochastic process and try to learn from previous evaluations. By doing so, the authors claim to need ten times fewer experimental evaluations compared to a Nelder-Mead search.

Other types of parameterizations have been routinely applied, ranging from bang-bang method in suppressing decoherence [99], polynomial function parametrization for evaporative cooling [119], to parametrization with linear ramps. Once the reduction of parameters has been found, an optimization routine is selected. Local search algorithms, such as steepest descent and Newton method, or methods belonging to a class of exhaustive algorithms are usually not preferable. They tend to get trapped in parametrization induced sub-optimal local optima. To overcome this problem, combinations of global and local searches are available. One may use a simulated annealing method combined with the metropolis algorithm [120], or a local search with Levy flights [121]. Especially if the underlying super-landscape emerged from the chosen parametrization has a smooth and trap-free shape, a genetic programming method can be successful [122]. It combines previously relative fit candidates with random influences to adapt the evolution.

Within closed-loop control, a distinction is made between ‘measurement-based feedback control’ where the control is determined based on classical information and ‘coherent feedback’ where all the information along the loop stays quantum; however various naming conventions are being used in literature and there is still a lot of ambiguity. One of the first studies in this direction is reported in [123], where they theoretically shown that coherent feedback might perform tasks such as entanglement transfer that are impossible using classical feedback. In [50] an experimental realization of that type is shown. Using a combination of pi and pi/2 gates the authors do quantum feedback control to prolong coherence time as it is needed for quantum computing and sensing. Although, as we have seen in the previous paragraphs, there are some challenges with closed-loop control, it has been shown that the theoretical control landscape analysis and its encouraging conclusions (see also Sec. 3.4) seem to hold at least for some problem classes: in [124] the researchers predict that unconstrained searches over quantum control landscapes are in principle devoid of traps. This leads to a so called ‘inherent monotonicity’ of searches towards a global optimum. This conclusion was achieved from
2.4 Analytically solvable systems

A perturbative analysis of sub-optimal local minima yielding trapping heights that are below the respective signal-to-noise ratio. The same authors also introduced in [125] a 'statistical gradient' method, whose basis are randomly sampled points (e.g. from a Gaussian distribution) in the landscape to get an estimate on the real underlying gradient. Furthermore, they derive by means of a theoretical argument the convergence rate and experimentally validate its superiority over some 'standard' genetic algorithm. Landscape analysis based experimental control has also been demonstrated in [126], where the control landscape of excitations in potassium atoms is shown.

Among all these options on closed-loop experimental control, we identified the dCRAB algorithm as one of the most suitable platforms for general purpose quantum control. It is insensitive to local traps, at least for unconstrained problems, and needs only a limited numbers of search parameters per pulse [95; 96; 103]. Based on the conclusions from Rabitz’s work as presented previously, an effective search in many scenarios towards a global optimum should be possible. Moreover, experimental constraints can be respected easily and dCRAB can even be applied when there are no theoretical models available. In fact, a series of successful experimental applications with (d)CRAB as the core optimal control algorithm have been reported: the superfluid to Mott insulator transition was optimized [127] as well as the state preparation for a BEC [128] and it has been used to precisely control spins in NV beyond the rotating wave approximation [113]. Furthermore, a BEC trapped on an atom chip was used to create arbitrary superposition states [114]. The two quantum processes – BEC state manipulation and crossing quantum phase transition in a cold atoms in optical lattices experiment – were optimized [129] and single qubit gate synthesis for the NV has been reported in [52]. Recently, dCRAB has been applied to enhance the BEC creation processes [130] and to optimize atomic beam splitter sequences [131].

2.4 Analytically solvable systems

Apart from adiabatic quantum evolutions [132], for general problems, analytically solvable quantum control problems are extremely rare. On top of the two presented cases of the harmonic oscillator and single qubit operations, there are only very few problems reported where analytical control solutions can be derived. For example, the derivation of the minimum time of evolution for operations on short spin chains [133] has been shown and for the Landau–Zener system globally optimal controlled dynamics is possible [134].
2.4.1 Optimal shifting of a harmonic trapping potential

Problem definition  The control problem is about to move a particle with mass $m$ in a harmonic potential (which could possibly be generated by an optical tweezer)\[ V(x,u(t)) = \frac{1}{2}m\omega^2(x - u(t))^2 \] with frequency $\omega$ from an initial state $|\psi_0\rangle$ to a goal state $|\phi\rangle$ in transfer time $T$. Both, initial and goal state, are the respective groundstates of $V(x,u = x_0)$ and $V(x,u = x_g)$. The control $u(t)$ determines the minimum of the potential. In the limit of large transfer times $T$ (compared to the energy scale of the system $\frac{1}{2}\hbar\omega$), according to the adiabatic theorem [132] a valid strategy is to slowly move the potential from $u_0$ to $u_g$. As a consequence, the system will stay in its instantaneous eigenstate which is the groundstate in this scenario. Due to practical limitations, which are discussed in more detail in Sec. 3.2, we usually want to achieve fast transport times, orders of magnitudes faster than what would be required for adiabaticity. Setting the parameters to $m = 1$, $\omega = \sqrt{2}$, $u_0 = 0$, $u_g = 8$ and $T = 10$, we can formulate initial and goal state in position representation the following way:

\begin{align}
\psi_0(x) &= \left(\frac{\sqrt{2}}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\sqrt{2}}{2}x^2} \\
\phi(x) &= \left(\frac{\sqrt{2}}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\sqrt{2}}{2}(x-8)^2}.
\end{align}

These wave functions are depicted along with the respective potentials in Fig. 2.3. Hence the Hamiltonian reads

\[ H = \frac{p^2}{2m} + (x - u(t))^2 \]

\[ = \frac{p^2}{2m} + x^2 - 2xu(t) + u^2(t). \]  

The $u^2(t)$ term simply lifts the potential, only affecting an unphysical global phase. Because we accept any state $\psi(x,T) = e^{i\phi(x)} \forall \phi$ as a valid goal state, we can omit the last term in Eq. (2.30). Together with a simple variable transformation $\tilde{u}(t) = -2u(t)$, we arrive at

\[ H = \frac{p^2}{2m} + x^2 + x\tilde{u}(t) =: \frac{p^2}{2m} + \tilde{V}(x,t). \] 

Correspondence to classical mechanics  First, we will show that the expectation value of the position and momentum operators follow the classical trajectories for harmonic potentials. Then we argue why these two variables are sufficient for describing the controlled transport process and hence we can use the classical formalism to find a
Figure 2.3: Schematic drawing of probability amplitudes for initial wave function $|\psi_0(x)|^2 = |\Psi_1(x)|^2$ (red), some intermediate wave function $|\psi_t(x)|^2 = |\Psi_2(x)|^2$ (green) and goal wave function $|\phi(x)|^2 = |\Psi_3(x)|^2$ (blue) along with respective harmonic potentials (light blue).

According to the formulation for the time derivative of the expectation value of an operator\(^1\), we have

$$\frac{d}{dt} \langle O \rangle = i \frac{\hbar}{\hbar} \langle [H,O] \rangle + \langle \frac{\partial O}{\partial t} \rangle. \tag{2.32}$$

(Further details can be found in Chapter 7). Setting $\hbar = 1$, which is also the standard convention throughout this thesis, we apply this equation for $O = p$ and $O = x$. Starting with the former, we get

$$\frac{d}{dt} \langle p \rangle = i \langle [H,p] \rangle + 0 = i \langle \frac{p^2}{2m} + \bar{V}(x,t),p \rangle = i \langle [\bar{V}(x,t),p] \rangle. \tag{2.33}$$

---

\(^1\) which was in this more general form actually derived by Werner Heisenberg and not by Paul Ehrenfest.
Evaluating the commutator \([\tilde{V}(x,t),p] \) in position representation results in

\[
[\tilde{V}(x,t),i\nabla] = i(\nabla\tilde{V}).
\] (2.34)

Hence the time derivative of the expected momentum operator in position representation reads

\[
\frac{d}{dt} \langle p \rangle = -\langle \nabla\tilde{V}(x,t) \rangle .
\] (2.35)

Similarly, we get for the position operator

\[
\frac{d}{dt} \langle x \rangle = \frac{i}{2m} \langle \left[ H, x \right] \rangle + \frac{1}{m} \langle \left[ p, x \right] \rangle = \frac{1}{m} \langle p \rangle ,
\] (2.36)

where we used the commutation relation \([p,x] = -i\).

Looking at Eq. (2.35) and (2.36) it might appear that the expectation values obey classical equations of motion a la Newton. However, this is not the case in general as

\[
-\langle \nabla\tilde{V}(x) \rangle \neq -\nabla\tilde{V}(\langle x \rangle).
\]

Introducing the force \(F = -\nabla\tilde{V}\), we can Taylor expand the right hand side of Eq. (2.35) yielding

\[
\langle F(x) \rangle = \left( F(\langle x \rangle) + F'(\langle x \rangle)(x - \langle x \rangle) + \frac{1}{2} \cdot F''(\langle x \rangle)(x - \langle x \rangle)^2 + \mathcal{O}(\Delta x^3) \right)
\]

\[
= F(\langle x \rangle) + F'(\langle x \rangle)(x - \langle x \rangle) + \frac{1}{2} \cdot F''(\langle x \rangle)(x - \langle x \rangle)^2 + \mathcal{O}(\Delta x^3)
\]

\[
= F(\langle x \rangle) + \frac{1}{2} \cdot F''(\langle x \rangle) \Delta x^2 + \mathcal{O}(\Delta x^3).
\] (2.37)

This means that for harmonic potentials only the first term is non-zero. As a consequence, the expected position and expected momentum do exactly follow the classical trajectories.

It can be shown [135] that the probability distribution \(|\psi(x,t)|^2\) is a time-independent quantity for the kind of problem we are concerned with here. What remains is to discuss a potential difference in (relative) phase that would distinguish a quantum state transfer from the corresponding classical one. However, as we are aiming at goal state with zero expectation value of the velocity \(\langle p \rangle = 0\), the phase of the wave function at every position has to be the same (a shared global phase other than zero is
possible though). As a consequence, we can map the quantum optimal control problem previously introduced onto the classical control problem of transferring an object with mass $m$ from

$$x(t = 0) = x_0 = 0,$$
$$p(t = 0) = p_0 = 0$$

(2.38)
(2.39)

to

$$x(t = T) = x_g = 8,$$
$$p(t = T) = p_g = 0,$$

(2.40)
(2.41)

in a transfer time $T$, while obeying the dynamical constraints

$$\dot{x} = \frac{p}{m},$$
$$\dot{p} = -m\omega^2 x - \tilde{u}(t).$$

(2.42)
(2.43)

**Solution using Green’s function** The inhomogeneous system of first order linear differential equations (Eq. (2.42), (2.43)) has a general solution

$$x(t) = x_0 \cos(\omega t) + \frac{p_0}{m\omega} \sin(\omega t) - \int_0^t G(t - t')\tilde{u}(t')dt'.$$

(2.44)

The tricky part is usually to find a suitable Green’s function $G(\tau)$ for solving inhomogeneous differential equations. Once this is done, the inhomogeneous part can be taken into account with the convolution scheme as shown in Eq. (2.44). For this case the Green’s function $G(\tau) = c\frac{\sin(\omega\tau)}{\omega}$, where $c$ is a constant, is feasible.

Since the dimensionality of the system is two, a two parameter ansatz for the control function is reasonable because it allows to address the whole dynamics. We choose a straightforward ansatz; a linear ramp

$$\tilde{u}(t) = \beta - \gamma t$$

(2.45)

with free parameters $\beta$ and $\gamma$ (although (almost) any ansatz which allows for at least one root should work). Taking into account the boundary conditions at $t = 0$ (Eq. (2.38), (2.39)) and after some calculations, we arrive at the following condition at
Plugging in the boundary conditions Eq. (2.40), (2.41) at $t = T$, we can solve the system (Eq. (2.46) and (2.47)) for $\beta$ and $\gamma$. Since the variables only occur linearly, we can find the general formula

$$\beta = -\frac{T \omega p_g \csc^2(\omega T) + p_g \csc(\omega T) + \omega x_g \cot(\omega T) \csc(\omega T)}{2 \left( -1 - \cot^2(\omega T) - \csc^2(\omega T) + T \sqrt{2} \csc(\omega T) + 2 \cot(\omega T) \csc(\omega T) \right)}$$

$$\gamma = -\frac{\omega p_g - \omega p_g \cos(\omega T) - 2 x_g \sin(\omega T)}{2 \left( 1 + \sin^2(\omega T) - T \omega \sin(\omega T) + \cos^2(\omega T) - 2 \cos(\omega T) \right)},$$

where $\csc(\xi) = \frac{1}{\sin(\xi)}$ is the cosecant and $\cot(\xi) = \frac{1}{\tan(\xi)} = \frac{\cos(\xi)}{\sin(\xi)}$ is the cotangent.

Plugging in the parameters $m = 1, \omega = \sqrt{2}, T = 10$, we get for the optimal control function

$$\tilde{u}^*(t) = \beta^* - \gamma^* t = 1.325 - 1.865 t.$$  (2.50)

The optimized state trajectories $x^*(t), p^*(t)$ together with the optimal $\tilde{u}^*(t)$ and the resulting calculated net force on the system $F = -m \omega^2 x^*(t) - \tilde{u}^*(t)$ are plotted in Fig. 2.4.

This problem about the optimal allocation of a harmonic oscillator trap is in many regards educational: Not only allows it for full analytical solveability (either the way presented here or also purely quantum mechanically), it also shows that the effective dimension of this problem is very low and it exhibits the feature of a multitude of (equally) optimal solutions. In fact, the classical states $x(T)$ and $p(T)$ are invariant with respect to pulse changes according to

$$\tilde{u}^{**}(t) = \tilde{u}^*(t) + u^\perp.$$  (2.51)

Any pulse $u^\perp$, which is perpendicular to $G(T - t')$ can be added to the previous solution and the resulting pulse $\tilde{u}^{**}(t)$ will still be optimal. (Quantum mechanical propagation might exhibit a different global phase which, as we argued before, is irrelevant for us.) The existence of a multitude of optimal function classes is a typical feature of quantum optimal control problems.
2.4 Analytically solvable systems

Figure 2.4: Results for optimal transport problem of a single particle in a harmonic well.

2.4.2 Optimal two level quantum evolutions via Pontryagins Maximum Principle

The implications of this considerations are far reaching because single qubit operations are ubiquitous in many quantum technological setups and can explain some limiting cases of the problems discussed in Chapter 6 and Chapter 8. A multitude of publications [70; 136; 137] has been devoted on the optimal manipulation of two level systems. Some of them additionally employ time-optimal control [74; 76; 138–140] while others discuss the context of open quantum systems [141–143]. In the following we provide a brief introduction into the theory and the central concept, the Pontryagin’s maximum principle. The reader is referred to [70] for a more comprehensive treatment while the following description closely follows the manuscript from [74].

Pontryagin Maximum Principle  Before starting with the specific application of the qubit rotations, we briefly introduce the working principle of the Pontryagin maximum principle (PMP) using a general notation. The system dynamics shall be
given by

\[ \dot{x} = f(t,x(t),u(t)), \] (2.52)

where \( x(t) \) represents the (possibly high dimensional) state vector and \( u(t) \) the (also possibly high dimensional) control vector. The cost functional subject to minimization reads

\[ J(u,T) = g(x_u(T),T) + \int_0^T f^0(t,x_u(t),u(t))dt, \] (2.53)

where \( x_u(T) \) is the state at final time \( T \) as a result from propagation with controls \( u \). The end cost term \( g(x_u(T),T) \) is the so-called Mayer term and usually designed in a way to yield the distance from \( x_u(T) \) to a goal state \( x_g(T) \) in some metric. The second term in the cost functional, also referenced as Lagrange term, adds up running costs \( f^0(t,x_u(t),u(t)) \) over the whole time interval of the process. In total, the representation of Eq. (2.53) is referenced as Bolza form. The central object of the PMP formalism is the so-called (pseudo) control Hamiltonian \( \mathcal{H} \) denoted by

\[ \mathcal{H}(t,x,p,p_0,u) = p \cdot f + p_0 f^0, \] (2.54)

where the first term is the scalar product between some dynamical lagrangian multipliers \( p \) which form the so-called co-state or adjoint state, and the right hand side of Eq. (2.52). The second term is a multiplication of some lagrangian multiplier \( p_0 \in \mathbb{R}^1 \) and the integrand \( f^0 \) from Eq. (2.53). It can be shown that for time-invariant control problems, that is problems where neither \( f \) nor \( f^0 \) explicitly depend on time, \( \mathcal{H} \) has to be continuous along an optimal trajectory. The constant \( p_0 \) is taken to be negative and should be set such that \( p_0 \) and \( p \) never vanish simultaneously. Finally, the control strategy follows from the theorem that a necessary condition for an optimal solution \( (x^*,p^*,u^*) \) is that they must fulfill

\[ \mathcal{H}(t,x^*,p^*,p_0^*,u^*) = \max_{u \in \Omega} \mathcal{H}(t,x^*,p^*,p_0^*,u) \] (2.55)

for all times \( t \in [0,T] \) and \( u \) is in the feasible control domain \( \Omega \). The catch of the PMP is that, in general, the optimal trajectories \( x^*(t), p^*(t) \) must be known a priori which prevents PMP from being universally applied. However, for some limited problem classes, a dependence for optimal pulses as a function of state and co-state variables can be derived which leads to a solveable system of equation yielding optimal trajectories.
Problem definition and choice of coordinates  Introducing the underlying system in the context of nuclear magnetic resonance (NMR), we consider a spin $\frac{1}{2}$ particle interacting with a constant magnetic field $B_0$ along the z direction. Additionally, an external transverse magnetic field $B_1(t')$ is applied which is the control input. The goal is to create arbitrary state transfers as well as gate synthesis in minimal time. We focus on the case where $|B_0|$ is much bigger than $|B_1|$ and the control field oscillates at a frequency close to the Larmor frequency of the spin [144]. Hence the rotating-wave approximation can be applied and the Hamiltonian can be written as

$$H = \frac{1}{2} (\omega_x(t')\sigma_x + \omega_y(t')\sigma_y + \omega_0\sigma_z),$$

where $\omega_0$ is the detuning, considered to be constant, which corresponds to the frequency difference between $B_1$ and the Larmor frequency. Furthermore, we assume limited control input via a constraint $\omega_x^2 + \omega_y^2 \leq \omega_b^2$ (disc constraint). Next, we apply a canonical transformation by normalizing the prefactors in front of the three parts in the Hamiltonian $v_i = \omega_i/\omega_b$ and $\Delta = \omega/\omega_b$ and inverse transformation on the time $t = \omega_b t'$. Hence the Hamiltonian becomes

$$H = v_x(t)\sigma_x + v_y(t)\sigma_y + \Delta\sigma_z.$$  \hspace{1cm} (2.57)

Switching to polar coordinates, the control functions are rewritten as

$$v_x(t) = v(t) \cos(\mu(t))$$

$$v_y(t) = v(t) \sin(\mu(t)),$$  \hspace{1cm} (2.58, 2.59)

with amplitude $v(t)$ and phase $\mu(t)$. The dynamics is dictated by the Schrödinger operator equation

$$i\partial_t U(t) = H(t)U(t),$$

where $U(t)$ is a unitary matrix that connects the initial state with some final state via $\psi(T) = U(T)\psi_0$ or might directly represent a gate. A possible parametrization of $U(t)$ can be done in terms of co-called Hopf variables $\{\theta_1, \theta_2, \theta_3\}$. Notice that due to the unitary condition and non respecting of a global phase, the originally eight real
parameters for the $2 \times 2$ matrix can be reduced three, and $U$ can be written as

$$U(\theta_1, \theta_2, \theta_3) = \begin{pmatrix} \cos(\theta_1)e^{i\theta_2} & \sin(\theta_1)e^{i\theta_3} \\ -\sin(\theta_1)e^{-i\theta_3} & \cos(\theta_1)e^{-i\theta_2} \end{pmatrix}. \quad (2.61)$$

From Eq. (2.60) the dynamics is then recasted into

$$\begin{pmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ -\tan(\theta_1)u_2 - \Delta \\ \cot(\theta_1)u_2 - \Delta \end{pmatrix}, \quad (2.62)$$

with the respective control variables

$$u_1 = -v \sin(\mu + \theta_2 + \theta_3) \quad (2.63)$$
$$u_2 = -v \cos(\mu + \theta_2 + \theta_3), \quad (2.64)$$

which are still fulfilling the normalized disc constraints. The Hopf variables translate into the more frequently used Euler variables by

$$\psi = \theta_2 + \theta_3 \quad (2.65)$$
$$\theta = 2\theta_1 \quad (2.66)$$
$$\phi = \theta_2 - \theta_3. \quad (2.67)$$

We can recall the notation of the unitary matrices by $U(\psi, \theta, \phi) = e^{i/2\psi \sigma_z} e^{i/2\theta \sigma_y} e^{i/2\phi \sigma_z}$.

**Application of PMP on optimal qubit operations** Our goal is to achieve a certain transformation in variables in minimal time. Hence, the cost defining parameters in Eq. (2.53) should be set as $g = 0$ and $f_0 = 1$, reducing the total cost functional to $J(u,T) = T$. Plugging in the system variables, the PMP pseudo control Hamiltonian takes the form

$$\mathcal{H}(t, \bar{p}, p^0, u) = u_1p_1 + u_2(-p_2 \tan(\theta_1) + p_3 \cot(\theta_1)) - (p_2 + p_3)\Delta + p^0. \quad (2.68)$$
Respecting the amplitude limits \( u_1^2 + u_2^2 \leq 1 \), the optimal form (such that \( H \) would be maximal) of the controls can be expressed in terms of the other variables

\[
\begin{align*}
u_1 &= \frac{p_1}{N}, \\
u_2 &= -\frac{p_2 \tan \theta_1 + p_3 \cot \theta_1}{N},
\end{align*}
\]  

where

\[
N = \sqrt{p_1^2 + (-p_2 \tan \theta_1 + p_3 \cot \theta_1)^2}.
\]

We can already conclude from the fact that \( u_1, u_2 \) enter linearly in Eq. (2.68) that optimal controls will exploit full amplitude limits at all times. Furthermore, we can set \( p^0 = -1 \) and we have that the states and co-states fulfill the canonical equations

\[
\begin{align*}
\dot{x}^* &= \frac{\partial H}{\partial p}, \\
\dot{p}^* &= -\frac{\partial H}{\partial x}.
\end{align*}
\]

This, together with the normalized optimal controls, leads us to

\[
\begin{pmatrix}
\dot{\theta}_1 \\
\dot{\theta}_2 \\
\dot{\theta}_3
\end{pmatrix} =
\begin{pmatrix}
p_1 \\
p_2 \tan^2 \theta_1 - \Delta \\
-p_2 - \Delta
\end{pmatrix},
\]

\[
\begin{pmatrix}
\dot{p}_1 \\
\dot{p}_2 \\
\dot{p}_3
\end{pmatrix} =
\begin{pmatrix}
-p_2^2 \tan \theta_1 \sec^2 \theta_1 \\
0 \\
0
\end{pmatrix},
\]

where the co-state variables have been normalized to \( p_1' = p_1/N \) and the prime has been dropped for ease of reading. We can directly derive some constants of motion from Eq. (2.75):

\[
\begin{align*}
p_2(t) &= p_2, \\
p_3(t) &= 0.
\end{align*}
\]
The relation \( p_3(t) = 0 \) can directly be deduced from the expression for \( u_2 \) from Eq. (2.70). Since \( \theta_1(0) = 0 \), the term \( p_3 \cot \theta_1 \) would be infinite in \( t = 0 \) if \( p_3(0) \neq 0 \), which can be excluded since the controls are strictly amplitude limited. Here, we are focusing on the special case of zero detuning \( (\Delta = 0) \) in what follows. However this assumption is not necessary as the problem can also be solved in general. Due to the fact that optimal controls, as we have shown previously, lie on the boundary of the control domain, the amplitude \( v \) of the optimal control from Eq. (2.63) and (2.64) is already determined. The angular parameters can be summarized and translated into Euler variables according to

\[
\beta(t) = \mu(t) + \theta_2(t) + \theta_3 = \mu(t) + \psi(t) \quad \text{(see also Eq. (2.65))}.
\]

Using Eq. (2.65) – (2.67) as well as Eq. (2.76) and (2.77), we can rewrite Eq. (2.74) as

\[
\begin{pmatrix}
\dot{\psi} \\
\dot{\theta} \\
\dot{\phi}
\end{pmatrix} =
\begin{pmatrix}
p_2(\tan^2 \theta_1 - 1) \\
2p_1 \\
p_2 \sec^2(\theta_1)
\end{pmatrix}.
\tag{2.78}
\]

Moreover, using Eq. (2.63) and (2.64) as well as (2.69) and (2.70) together with the dynamics from Eq. (2.78), we get the following relations between the Euler variables \( \psi(t), \phi(t) \) and the essential control variable \( \mu(t) \):

\[
\dot{\mu} = 2p_2 \quad \tag{2.79}
\]

\[
\dot{\mu} + \dot{\psi} + \dot{\phi} = 0. \quad \tag{2.80}
\]

We already know that \( p_2 \) is constant from Eq. (2.76) and hence \( \mu \) has to be a linear function in time:

\[
\mu(t) = \mu(0) + 2p_2t. \quad \tag{2.81}
\]

It is possible to show that \( \beta(0) = -\frac{\pi}{2} \) and from this together with the fact that \( \phi(0) = -\psi(0) \) (since \( \theta_2(0) = 0 \)), we directly compute \( \mu(0) = \phi(0) - \frac{\pi}{2} \) by using the definition of \( \beta(t) \). Finally, we can deduce explicit expressions for the time evolution of the Euler variables:

\[
\theta(t) = \cos^{-1}(1 + \sin^2(\bar{\theta})(\cos(\eta(t)) - 1)), \quad \tag{2.82}
\]

\[
\phi(t) = \phi(0) + \sin(p_2)\frac{\pi}{2} + \tan^{-1}\left(\frac{\sin \eta(t)}{\cos \bar{\theta}(\cos(\eta(t)) - 1)}\right), \quad \tag{2.83}
\]

\[
\psi(t) = -2\phi(0) + \phi - 2p_2t, \quad \tag{2.84}
\]
where $\eta(t) = \frac{2t}{\sin \bar{\theta}}$, and $\bar{\theta} = \arctan(1/p_2)$.

### 2.5 Optimal Gaussian focusing

In this section, we show how to reduce the problem of the time evolution of a Gaussian wave function in a harmonic trap with controllable frequency to a one dimensional differential equation [145]. We then use this equation to optimize the problem of focusing a Gaussian wave function after some free evolution time after the trapping potential is released. This scenario has applications in implanting ions in solids. One example has been demonstrated for focusing down an ion beam consisting of single $^{40}\text{Ca}^+$ ions [146]. Furthermore, nitrogen might be implanted into diamonds to create a single defect center [147].

#### 2.5.1 An equation of motion for the width of a Gaussian wave function

We make an ansatz for the wave function (in arbitrary dimension) such that its probability amplitude remains of Gaussian shape and we will see that it fulfills the time dependent Schrödinger Equation. The ansatz reads:

$$
\psi(x,t) = \frac{1}{\sqrt{\prod_b b_k}} \psi_0(x/b_j, t) e^{\frac{im}{2\hbar} \sum_j x_j^2 \dot{b}_j} e^{-i\beta(t)}
$$

Here, $b_j(t)$ is a the time dependent spatial scaling

$$\bar{x}_j = \frac{x_j}{b_j},$$

and the index $j$ represents the dimension. Now we plug this ansatz in the (time dependent) Schrödinger Equation, evaluating it by parts. We begin with the time derivative

$$
i\hbar \frac{\partial \psi}{\partial t} = i\hbar \left( -\sum_j \frac{\partial \psi_0(x)}{\partial x_j} \frac{\ddot{b}_j}{b_j} x_j + \frac{\partial \psi_0(x,\tau)}{\partial \tau} - \sum_j \frac{\dot{b}_j}{2b_j} \psi_0(x) \right)$$

$$-i\beta \psi_0(x) + \frac{im \psi_0(x)}{2\hbar} \sum_j \dddot{x}_j \dot{b}_j b_j - \psi_0(x) \frac{i n}{2\hbar} \sum_j \dddot{x}_j^3 \dot{b}_j^2 \right).$$

$$+ \frac{1}{\sqrt{\prod_b b_k}} e^{\frac{im}{2\hbar} \sum_j x_j^2 \dot{b}_j} e^{-i\beta(t)},$$
which becomes

\[
\begin{align*}
\psi(\bar{x},\tau) = &\left(i\hbar \frac{\partial \psi_0(\bar{x})}{\partial \tau} - i\hbar \sum_j \frac{\dot{b}_j}{2\bar{b}_j} \psi_0(\bar{x}) - i\hbar \sum_j \frac{\partial \psi_0(\bar{x})}{\partial \bar{x}_j} \bar{x}_j \frac{\dot{b}_j}{b_j}\right) \\
+ &\hbar \dot{\beta} \psi_0(\bar{x}) - \frac{m\psi_0(\bar{x})}{2} \sum_j \bar{x}_j^2 \dot{b}_j b_j + \psi_0(\bar{x}) \frac{m}{2} \sum_j \bar{x}_j^2 \dot{b}_j^2 \right) \cdot \frac{1}{\sqrt{\prod_k b_k}} e^{\frac{i\imath}{\hbar} \sum_j \bar{x}_j^2 \dot{b}_j} e^{-i\beta(t)}.
\end{align*}
\]

The kinetic part reads

\[
\begin{align*}
-\frac{\hbar^2}{2m} \Delta \psi = &\left(\frac{-\hbar^2}{2m} \sum_j \frac{\partial^2 \psi_0(\bar{x})}{\partial \bar{x}_j^2} \frac{1}{\bar{b}_j^2} b_j^2 - i\hbar \sum_j \frac{\partial \psi_0(\bar{x})}{\partial \bar{x}_j} \bar{x}_j \dot{b}_j b_j - i\hbar \sum_j \frac{1}{2} \frac{b_j}{b_j} \psi_0(\bar{x}) + \sum_j \frac{m}{2} \bar{x}_j^2 \dot{b}_j \psi_0(\bar{x}) \right) \cdot \frac{1}{\sqrt{\prod_k b_k}} e^{\frac{i\imath}{\hbar} \sum_j \bar{x}_j^2 \dot{b}_j} e^{-i\beta(t)}.
\end{align*}
\]

All marked terms vanish: I cancels with I' and analogously for II and III. Hence the full Schrödinger equation in the ansatz is

\[
\begin{align*}
0 = &\ h \frac{\partial \psi}{\partial t} + \left(\frac{\hbar^2}{2m} \Delta - V(x)\right) \\
= &\ h \frac{\partial \psi}{\partial t} + \left(\frac{\hbar^2}{2m} \Delta - \frac{1}{2} m \sum_j \omega_j(t) \bar{x}_j^2\right) \\
= &\ \left(i\hbar \frac{\partial \psi_0(\bar{x},\tau)}{\partial \tau} - \frac{m\psi_0(\bar{x})}{2} \sum_j \bar{x}_j^2 \dot{b}_j b_j + \hbar \dot{\beta} \psi_0(\bar{x})\right) \\
+ &\ \frac{\hbar^2}{2m} \sum_j \frac{\partial^2 \psi_0(\bar{x})}{\partial \bar{x}_j^2} \frac{1}{\bar{b}_j^2} b_j^2 - \frac{1}{2} m \sum_j \omega_j^2(t) \bar{x}_j^2 \dot{b}_j^2 \psi_0(\bar{x}) \right) \cdot \frac{1}{\sqrt{\prod_k b_k}} e^{\frac{i\imath}{\hbar} \sum_j \bar{x}_j^2 \dot{b}_j} e^{-i\beta(t)}.
\end{align*}
\]
Since the exponential functions never yield zero we can consider the part in the brackets only

\[ 0 = i\hbar \frac{\partial \psi_0(\bar{x}, \tau)}{\partial \tau} - \frac{m \psi_0(\bar{x})}{2} \sum_j \bar{x}_j^2 \dot{b}_j b_j + \hbar \dot{\beta} \psi_0(\bar{x}) \quad (2.93) \]

\[ + \frac{\hbar^2}{2m} \sum_j \frac{\partial^2 \psi_0(\bar{x})}{\partial \bar{x}_j^2} \frac{1}{b_j^2} - \frac{1}{2} m \sum_j \omega_j^2(t) \bar{x}_j^2 b_j^2 \psi_0(\bar{x}). \quad (2.94) \]

The first term is zero due to the introduced scaling variables \( b \). Next we combine the second and fifth term

\[ 0 = -\frac{m}{2} \psi_0 \sum_j \left( \bar{x}_j^2 \left( \dot{\bar{b}}_j b_j + \omega_j^2(t) \dot{b}_j^2 \right) \right) + \hbar \dot{\beta} \psi_0(\bar{x}) + \frac{\hbar^2}{2m} \sum_j \frac{\partial^2 \psi_0(\bar{x})}{\partial \bar{x}_j^2} \frac{1}{b_j^2} \quad (2.95) \]

and plug in the static Schrödinger Equation \(-\frac{\hbar^2}{2m} \sum_j \frac{\partial^2 \psi_0}{\partial \bar{x}_j^2} = (E - \frac{m}{2} \sum_j \omega_j^2(0) \bar{x}_j^2) \psi_0\) into the last term

\[ 0 = -\frac{m}{2} \psi_0 \sum_j \left( \bar{x}_j^2 \left( \dot{\bar{b}}_j b_j + \omega_j^2(t) \dot{b}_j^2 \right) \right) + \hbar \dot{\beta} \psi_0(\bar{x}) + \sum_j \left( -E + \frac{m}{2} \omega_j^2(0) \bar{x}_j^2 \right) \frac{1}{b_j^2} \psi_0. \quad (2.96) \]

Some minor regrouping yields

\[ 0 = -\frac{m}{2} \psi_0 \sum_j \left( \bar{x}_j^2 \left( \dot{\bar{b}}_j b_j + \omega_j^2(t) \dot{b}_j^2 - \frac{\omega_j^2(0)}{b_j^2} \right) \right) + \left( \hbar \dot{\beta} - E \sum_j \frac{1}{b_j^2} \right) \psi_0. \quad (2.97) \]

In order to fulfill this equation, we choose \( \beta(t) \) such that the second summand is zero for all times. Thus we are left over with a differential equation for standard deviation of our wave function

\[ \Rightarrow \dot{\bar{b}}_j = \frac{\omega_j^2(0)}{b_j^3} - \omega_j^2(t) b_j. \quad (2.98) \]

### 2.5.2 Discussion of obtained equation

In one dimensional space (\( D = 1 \)), the differential equation for the standard deviation of the wave function reads

\[ \Rightarrow \dot{b} = \frac{\omega^2(0)}{b^3} - \omega^2(t) b. \quad (2.99) \]
When releasing the trap ($\omega^2(t) = 0$), together with the initial conditions $b(0) = 1$ and $\dot{b}(0) = 0$, a solution is given by

$$b(t) = \sqrt{1 + \omega^2(0)t^2}.$$  

(2.100)

Hence for long times, $b$ scales approximately linear with $t$.

**2.5.3 Definition of the focusing control problem**

The problem, to be discussed in the following, is about to minimize the standard deviation ($\sigma = b$) at some time $t_2$, which is after the wave function has propagated in free space ($V=0$) for $t_1 \leq t \leq t_2$. Between time zero and $t_1$, the trap frequency can be dynamically controlled. The outline of the problem is visualized in Fig. 2.5. Hence, we can formulate the problem as

$$\min J = b(t_2)$$

s.t. 

$$\ddot{b} = \frac{\omega^2(0)}{b^3} - u^2(t)b$$

(2.102)

$$b(0) = 1, \quad \dot{b}(0) = 0,$$

(2.103)

where $u(t) := \omega(t)$. 

![Controlled Sequence Free Evolution](image) 

**Figure 2.5:** Control input can only be applied for $0 \leq t \leq t_1$. Figure of merit: $\sigma(t_2) = b(t_2)$
2.5.4 Solution via optimal control theory

First, we are going to transform the second order differential equation to a system of first order, introducing an auxiliary variable $c(t)$

\[
\dot{b} = c \tag{2.104}
\]
\[
\dot{c} = \frac{\omega^2(0)}{b^3} - u^2(t)\theta' b. \tag{2.105}
\]

Here, $\theta'(t - t_1) := 1 - \theta(t - t_1)$, where $\theta$ is the Heaviside step function. This means that we can only apply control in the first sequence. Second, we introduce the control Hamilton function $\mathcal{H}$ as defined in standard literature [148; 149] and as it is also being used in the quantum context [74]:

\[
\mathcal{H} = \lambda_1 c + \lambda_2 \left( \frac{\omega^2(0)}{b^3} - u^2(t)\theta' b \right). \tag{2.106}
\]

Third, for the costate/ adjoint system dynamics, we get

\[
\dot{\lambda}_1 = -\frac{\partial \mathcal{H}}{\partial b} = 3\lambda_2 \frac{\omega^2(0)}{b^4} + \lambda_2 u^2 \theta' \tag{2.107}
\]
\[
\dot{\lambda}_2 = -\frac{\partial \mathcal{H}}{\partial c} = -\lambda_1. \tag{2.108}
\]

Since the states are not fixed at final time, the opposite is true for the costates at final time $t_2$

\[
\lambda_1(t_2) = \frac{\partial J(t_2)}{\partial b(t_2)} = 1 \tag{2.109}
\]
\[
\lambda_2(t_2) = \frac{\partial J(t_2)}{\partial c(t_2)} = 0. \tag{2.110}
\]

Finally, the optimality condition reads

\[
0 = \frac{\partial \mathcal{H}}{\partial u} = -\lambda_2 2u\theta' b. \tag{2.111}
\]

We are using the Gradient method to minimize the cost functional $J$ while respecting all constraints. Resulting trajectories are shown for different time fractions $\tau = t_1/(t_1 + t_2)$ in Fig. 2.6 (a)-(c). Each of the figures shows resulting standard deviations and control amplitudes reflecting the trap frequencies. One can clearly see, the more relative control time is given, the better the results get. For smaller $\tau$, the standard deviation shows a minimum in the free evolution section, while for larger $\tau$, the standard deviation is
steadily decreasing in $t_1 \leq t \leq t_2$.

Finally, we scanned time ratios $\tau \in [0.1, 0.9]$, to see the qualitative behavior of $J^*(\tau)$. ($J^*$ is a respective optimal cost value as an outcome of the optimal control algorithm.) In Fig. 2.6 (d) we can see the almost linear behavior between $\tau$ and $J^* = \sigma^*(t_2)$. However for larger $\tau$, it seems to flatten out.

**Figure 2.6:** Results for optimal Gaussian focusing.
2.6 Appendix

2.6.1 A greedy one step approach towards efficient QOC

Here we introduce a novel one step approach to optimized state and gate fidelities.

2.6.1.1 Introduction

Quantum optimal control of systems where the control amplitudes are constrained and 
the total transfer time is rather short (typically orders of magnitudes shorter than 
adiabatic time scales) lose their favorable property of the absence of local minima in the 
control landscape [84]. Hence, the result of any iterative optimal control method is highly 
dependent on the initial guess pulse, and might exhibit great difficulties in obtaining 
reasonable low errors in the cost functional [150]. As a consequence, avoiding tedious 
explorations of rugged and spiky landscapes as being done by iterative algorithms, a 
one step method - automatically taking into account the control amplitude constraints - 
yielding high levels of precision is highly demanded.

2.6.1.2 Idea and derivation of extremum conditions for controls

Idea In the context of fidelity maximization, that is the optimization of the quantum 
overlap between a time evolved state of the system and a goal state, we are usually 
considering complete controllable systems, only. Speaking in a more pictorial way, 
this means any direction of the unitary time propagator can be generated and hence 
accessed. Now the idea is quite straightforward: among all those directions, the 
direction that instantaneously maximizes the fidelity overlap while still respecting the 
control amplitude constraints should be chosen. We will refer to this as Condition 1. 
Going one step further, we can even select the instantaneously controls such that not 
only the fidelity of the system change is maximized, but also the (temporal) change 
rate is maximized (Condition 2).

Overview of derived greedy optimal control conditions Before actually going 
for the calculus, we give an overview in the following table, displaying what figure 
of merits and how many controllable parts in the Hamiltonian and hence number of 
control functions m are being considered.
### Extremum control conditions

**Condition 1 for state overlap maximization for one control (m=1)** Suppose the control problem reads:

\[
\begin{align*}
\text{max } J &= |\langle \phi | \psi(T) \rangle|^2 \\
&\text{s.t.} \\
\partial_t |\psi(t)\rangle &= -\frac{i}{\hbar} H_t |\psi(t)\rangle \\
H_t &= H_d + u(t) H_c \\
|\psi(t=0)\rangle &= |\xi\rangle \\
b_l \leq u(t) \leq b_u \forall t
\end{align*}
\]  

Here, \( H_t \) stands for total or full Hamiltonian, \( H_d \) for drift Hamiltonian i.e. non controllable parts for the Hamiltonian and \( H_c \) for control Hamiltonian. Instead of iteratively seeking for a control field \( u(t) \) that maximizes the objective, we are employing a field at every instance that maximizes

\[
\begin{align*}
\max_{u(\cdot)} |\langle \phi | \dot{\psi}(t) \rangle|^2
\end{align*}
\]  

until the precision hits a predefined threshold. Plugging in Schrödinger Equation yields:

\[
\begin{align*}
\max_{u(\cdot)} \left| -\frac{i}{\hbar} \langle \phi | H_t | \psi(t) \rangle \right|^2 \\
\max_{u(\cdot)} \left| -\frac{i}{\hbar} \langle \phi | H_d + u(\cdot) H_c | \psi(t) \rangle \right|^2
\end{align*}
\]
2.6 Appendix

After some algebra, we arrive at:

\[
\frac{1}{\hbar^2} \max_{u(\cdot)} \left\{ |\langle \phi | H_d | \psi(t) \rangle|^2 + 2u(\cdot) \Re \{\langle \phi | H_c | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle\} + u(\cdot)^2 |\langle \phi | H_c | \psi(t) \rangle|^2 \right\}.
\]

(2.121)

This extremum condition will be referenced as ‘Condition 1’ in the following. (Note that for each of the columns of the above Table there is a ‘Condition 1’).

**Condition 2 for state overlap maximization for one control (m=1)** Going one step further, we start off from Eq. (2.118) and require the time derivative of this expression to be maximized. That way, the rate of change towards the goal state shall be maximized.

\[
\max_{u(\cdot)} \frac{d}{dt} |\langle \phi | \dot{\psi}(t) \rangle|^2
\]

(2.122)

After some calculations, we arrive at:

\[
\frac{2}{\hbar^3} \max_{u(\cdot)} \left\{ - \Im \{\langle \psi(t) | H_d H_d | \phi \rangle \langle \phi | H_d | \psi(t) \rangle\} + u \Im \{\langle \psi(t) | H_c H_d | \phi \rangle \langle \phi | H_d | \psi(t) \rangle\} + u \Im \{\langle \psi(t) | H_d H_c | \phi \rangle \langle \phi | H_c | \psi(t) \rangle\} + u^2 \Im \{\langle \psi(t) | H_c H_c | \phi \rangle \langle \phi | H_c | \psi(t) \rangle\} + u^2 \Im \{\langle \psi(t) | H_d H_c | \phi \rangle \langle \phi | H_c | \psi(t) \rangle\} + u^2 \Im \{\langle \psi(t) | H_c H_d | \phi \rangle \langle \phi | H_d | \psi(t) \rangle\} + u^3 \Im \{\langle \psi(t) | H_c H_c | \phi \rangle \langle \phi | H_c | \psi(t) \rangle\} \right\}
\]

(2.123)

This condition shall be referenced as ‘Condition 2’.

**Condition 1 for gate synthesis for one control (m=1)** Going from state control to gate control, we do not only want to transfer a single state, but a whole basis which is described by a unitary matrix. We denote the desired transformation by \(V\). Therefore, we have to extend the problem description from Eqs. (2.112) - (2.117)
towards $N$ states (labeled by $n \in 1, 2, \ldots, N$):

$$\max J = \frac{1}{N^2} \left| \sum_{n=1}^{N} \langle \xi_n | V^\dagger | \psi_n(T) \rangle \right|^2$$

(2.124)

s.t. (2.125)

$$\partial_t |\psi_n(t)\rangle = -\frac{i}{\hbar} H_t |\psi_n(t)\rangle \quad \forall n$$

(2.126)

$$H_t = H_d + u(t) H_c$$

(2.127)

$$|\psi_n(t = 0)\rangle = |\xi_n\rangle \quad \forall n$$

(2.128)

$$b_l \leq u(t) \leq b_u \quad \forall t$$

(2.129)

Consequently, we are seeking to maximize:

$$\max J = \frac{1}{N^2} \left| \sum_{n=1}^{N} \langle \xi_n | V^\dagger | \dot{\psi}_n(t) \rangle \right|^2$$

(2.130)

The extremum condition for single states from Eq. (2.121) extends here (after some calculus) towards:

$$\frac{1}{N^2 \hbar^2 \max \{ }$$

$$\sum_{n=1}^{N} \langle \xi_n | V^\dagger H_d | \psi_n(t) \rangle \langle \psi_n(t) | H_d V | \xi_n \rangle$$

$$+ 2u \sum_{n=1}^{N} Re\{ \langle \xi_n | V^\dagger H_c | \psi_n(t) \rangle \langle \psi_n(t) | H_d V | \xi_n \rangle \}$$

$$+ u^2 \sum_{n=1}^{N} \langle \xi_n | V^\dagger H_c | \psi_n(t) \rangle \langle \psi_n(t) | H_c V | \xi_n \rangle$$

$$\sum_{n > m} 2Re\{ \langle \xi_n | V^\dagger H_d | \psi_n(t) \rangle \langle \psi_m(t) | H_d V | \xi_m \rangle \} +$$

$$u \sum_{n > m} 2Re\{ \langle \xi_n | V^\dagger H_c | \psi_n(t) \rangle \langle \psi_m(t) | H_d V | \xi_m \rangle \} +$$

$$u \sum_{n > m} 2Re\{ \langle \psi_m(t) | H_c V | \xi_m \rangle \langle \xi_n | V^\dagger H_d | \psi_n(t) \rangle \} +$$

$$u^2 \sum_{n > m} 2Re\{ \langle \xi_n | V^\dagger H_c | \psi_n(t) \rangle \langle \psi_m(t) | H_c V | \xi_m \rangle \}$$

(2.131)
Condition 2 for gate synthesis for one control \((m=1)\)  Condition 2 goes quite analogously: Eq. (2.122) from the state optimization now reads:

\[
\max J = \frac{1}{N^2} \left| \frac{d}{dt} \sum_{n=1}^{N} \langle \xi_n | V^\dagger | \dot{\psi}_n(t) \rangle \right|^2.
\]

(2.132)
After some pages of paper and pencil, we arrive at:

\[
-\frac{2}{N^2 \hbar^3} \max_{u(\cdot)} \{ \\
\sum_{n=1}^{N} \text{Im}\{\langle \psi_n(t)|H_dH_dV|\xi_n\rangle \langle \xi_n|V^\dagger H_d|\psi_n(t)\rangle\} \\
+ u \sum_{n=1}^{N} \text{Im}\{\langle \psi_n(t)|H_cH_dV|\xi_n\rangle \langle \xi_n|V^\dagger H_d|\psi_n(t)\rangle\} \\
+ u \sum_{n=1}^{N} \text{Im}\{\langle \psi_n(t)|H_dH_dV|\xi_n\rangle \langle \xi_n|V^\dagger H_c|\psi_n(t)\rangle\} \\
+ u^2 \sum_{n=1}^{N} \text{Im}\{\langle \psi_n(t)|H_cH_cV|\xi_n\rangle \langle \xi_n|V^\dagger H_d|\psi_n(t)\rangle\} \\
+ u^2 \sum_{n=1}^{N} \text{Im}\{\langle \psi_n(t)|H_dH_cV|\xi_n\rangle \langle \xi_n|V^\dagger H_c|\psi_n(t)\rangle\} \\
+ u^3 \sum_{n=1}^{N} \text{Im}\{\langle \psi_n(t)|H_cH_cV|\xi_n\rangle \langle \xi_n|V^\dagger H_c|\psi_n(t)\rangle\} \\
+ \sum_{n>m} \{-\text{Im}\{\langle \xi_n|V^\dagger H_dH_dV|\psi_n(t)\rangle \langle \psi_m(t)|H_dV|\xi_m\rangle\} \\
+ \text{Im}\{\langle \xi_n|V^\dagger H_dV|\psi_n(t)\rangle \langle \psi_m(t)|H_dH_dV|\xi_m\rangle\} \\
- u \text{Im}\{\langle \xi_n|V^\dagger H_cV|\psi_n(t)\rangle \langle \psi_m(t)|H_dV|\xi_m\rangle\} \\
+ u \text{Im}\{\langle \xi_n|V^\dagger H_dV|\psi_n(t)\rangle \langle \psi_m(t)|H_cV|\xi_m\rangle\} \\
- u \text{Im}\{\langle \xi_n|V^\dagger H_cV|\psi_n(t)\rangle \langle \psi_m(t)|H_dV|\xi_m\rangle\} \\
+ u \text{Im}\{\langle \xi_n|V^\dagger H_dV|\psi_n(t)\rangle \langle \psi_m(t)|H_cV|\xi_m\rangle\} \\
- u^2 \text{Im}\{\langle \xi_n|V^\dagger H_cV|\psi_n(t)\rangle \langle \psi_m(t)|H_dV|\xi_m\rangle\} \\
+ u^2 \text{Im}\{\langle \xi_n|V^\dagger H_cV|\psi_n(t)\rangle \langle \psi_m(t)|H_cV|\xi_m\rangle\} \\
- u \text{Im}\{\langle \xi_n|V^\dagger H_dV|\psi_n(t)\rangle \langle \psi_m(t)|H_cV|\xi_m\rangle\} \\
+ u \text{Im}\{\langle \xi_n|V^\dagger H_cV|\psi_n(t)\rangle \langle \psi_m(t)|H_dV|\xi_m\rangle\} \\
- u^2 \text{Im}\{\langle \xi_n|V^\dagger H_cV|\psi_n(t)\rangle \langle \psi_m(t)|H_cV|\xi_m\rangle\} \} \} \tag{2.133}
\]
Condition 1 for state overlap maximization for two control Hamiltonians ($m=2$) Extending the problem for just one Hamiltonian from (2.115), we have an extended total Hamiltonian $H_t$ of the following form:

$$H_t = H_d + u_1(t)H_1 + u_2(t)H_2.$$  \hspace{1cm} (2.134)

Here, two control functions $u_1$ and $u_2$ need to be determined. The rest, particularly the control objective, remains the same. Of course, we get a slightly different extremum condition as a new Condition 1:

$$\frac{1}{\hbar^2} \max_{u_{(\cdot)}} \left\{ \langle \phi | H_d | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \\
+ 2u_1 \mathrm{Re}\{ \langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \} \\
+ u_1^2 \langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_1 | \phi \rangle \\
+ 2u_2 \mathrm{Re}\{ \langle \phi | H_2 | \psi(t) \rangle \langle \psi(t) | H_1 | \phi \rangle \} \\
+ u_2^2 \langle \phi | H_2 | \psi(t) \rangle \langle \psi(t) | H_2 | \phi \rangle \\
+ 2u_1 u_2 \mathrm{Re}\{ \langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_2 | \phi \rangle \} \right\}$$  \hspace{1cm} (2.135)
Condition 2 for state overlap maximization for two control Hamiltonians (m=2) Condition 2 goes analogously, too. It reads:

\[
\frac{1}{\hbar^2} \max_{u_1, u_2} \left\{ -2 \text{Im}\{\langle \phi | H_d | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \} - 2u_1 \text{Im}\{\langle \phi | H_d | \psi(t) \rangle \langle \psi(t) | H_1 H_d | \phi \rangle \} - 2u_2 \text{Im}\{\langle \phi | H_d | \psi(t) \rangle \langle \psi(t) | H_2 H_d | \phi \rangle \} + 2u_1 \text{Im}\{\langle \phi | H_1 H_d | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \} + 2u_2 \text{Im}\{\langle \phi | H_1 H_d | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \} - 2u_1 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_d H_1 | \phi \rangle \} - 2u_2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_1 H_1 | \phi \rangle \} - 2u_1^2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_2 H_1 | \phi \rangle \} + 2u_2 \text{Im}\{\langle \phi | H_2 H_d | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \} + 2u_1 u_2 \text{Im}\{\langle \phi | H_2 H_1 | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \} + 2u_2^2 \text{Im}\{\langle \phi | H_2 H_2 | \psi(t) \rangle \langle \psi(t) | H_d | \phi \rangle \} - 2u_1 \text{Im}\{\langle \phi | H_2 | \psi(t) \rangle \langle \psi(t) | H_d H_1 | \phi \rangle \} - 2u_2 \text{Im}\{\langle \phi | H_2 | \psi(t) \rangle \langle \psi(t) | H_d H_1 | \phi \rangle \} - 2u_1^2 \text{Im}\{\langle \phi | H_2 | \psi(t) \rangle \langle \psi(t) | H_2 H_1 | \phi \rangle \} - 2u_2^2 \text{Im}\{\langle \phi | H_2 | \psi(t) \rangle \langle \psi(t) | H_2 H_2 | \phi \rangle \} + 2u_1 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_1 H_1 | \phi \rangle \} + 2u_2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_1 H_1 | \phi \rangle \} + 2u_1^2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_2 H_1 | \phi \rangle \} + 2u_2^2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_2 H_2 | \phi \rangle \} - 2u_1 u_2 \text{Im}\{\langle \phi | H_1 H_d | \psi(t) \rangle \langle \psi(t) | H_2 | \phi \rangle \} + 2u_1^2 u_2 \text{Im}\{\langle \phi | H_1 H_1 | \psi(t) \rangle \langle \psi(t) | H_2 | \phi \rangle \} - 2u_1 u_2^2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_1 H_2 | \phi \rangle \} + 2u_1 u_2^2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_2 H_2 | \phi \rangle \} - 2u_1 u_2^2 \text{Im}\{\langle \phi | H_1 | \psi(t) \rangle \langle \psi(t) | H_2 H_2 | \phi \rangle \} \right\}
\]

(2.136)
In the results section, we will consider a problem where we use exactly this control strategy.

### 2.6.1.3 Implementation

This method is implemented the following way: Propagating the system forward from $t \rightarrow t + \Delta t$ is done by employing the total Hamiltonian $H(t) = H_d + u_1(t)H_1 + u_2(t)H_2$, where the controls $u_1, u_2$ are achieved via a constraint (of type box constraint) maximization on these two parameters. Luckily, the extremum control condition functions are just a third order polynomial in $u_1, u_2$. Hence the gradient (as required for extremization) is of parabolic type only, which is easy to solve. Furthermore, note that the maximum does not have to be on the edges of the feasible region but might also be on the inside. This procedure is repeated until $t = 0, \Delta t, 2\Delta t, ..., T$ reaches the total time $T$. That way, we obtain the ‘greedy controls’ within just one single forward propagation (or sweep) of the system.

### 2.6.1.4 Results

To demonstrate the functionality of the derived strategy, we show an one step optimal control application for the following scenario:

- State transfer of a $n = 2$ qubit system.
- Goal: Maximize fidelity $|\langle \phi | \psi(t^\ast) \rangle|^2$ for a time $t^\ast \in [0,T]$, $T = 5$ (which is not to be specified in advance).
- System: Ising chain with drift Hamiltonian $H_d = -\frac{J}{2} \sum_{i=1}^{n-1} \sigma^z_i \sigma^z_{i+1} - h \sum_{i=1}^{n} \sigma^x_i$. We set $J = h = 1$.
- The control Hamiltonians $H_1$ and $H_2$ consist of simultaneous x and y rotations on all spins: $H_1 = \sum_{i=1}^{n} \sigma^x_i$, $H_2 = \sum_{i=1}^{n} \sigma^y_i$.
- The initial state is set to $|\xi\rangle = (0,0,1,0)$ and the goal state reads $|\phi\rangle = 0.5 \cdot (1, -1, 1, -1)$ in a computational basis $|\uparrow\rangle = (1,0)^T$, $|\downarrow\rangle = (0,1)^T$.
- Constraints on the control pulse amplitudes $u_1, u_2$ apply: $|u_i(t)| \leq 3 \ \forall t$, $i \in \{1,2\}$

As we show in Figure 2.7, a reasonable good fidelity of 0.9989 is achieved after $t = 3.98$. Note that the fidelity cannot always be increased over time but still reaches a good level after a certain time.
2.6.2 Fixed-point iteration as an optimal control algorithm

Here an idea for a fixed-point iteration method for optimal control of quantum systems in presented.

2.6.2.1 Idea

The core of the idea is pretty straight forward. Using the outcome from calculus of variations (see Eq. (2.12) – (2.15)) we can obtain an optimality condition where the control pulse $u(t)$ explicitly appears. Based on this condition we use a fixed-point iteration to converge from an arbitrary initial guess pulse $u_0(t)$ to an optimal one $u^*(t)$. Suppose we have a Hamiltonian of the (standard) form

$$H = H_d + u(t)H_c,$$  \hspace{1cm} (2.137)

an initial guess $u_0(t)$ and we want to maximize the overlap with some goal state $|\phi\rangle$ (gate synthesis is of course doable, too). Further suppose we have an optimality condition

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fidelity_vs_time.png}
\caption{Fidelity versus time for state transfer and $m=2$. The maximum achieved fidelity is 0.9989 at $t=3.98$.}
\end{figure}
2.6 Appendix

where the control pulse appears explicitly, that is a function $\kappa$ with the property

$$u^*(t) = \kappa(|\psi(t)\rangle, |\lambda(t)\rangle). \tag{2.138}$$

The idea is to use this relation (Eq. (2.138)) also for suboptimal pulses as a fixed-point equation by recognizing that $|\psi(t)\rangle$ and $|\lambda(t)\rangle$ can be considered as functions of $u(t)$. Hence the fixed-point equation simply reads

$$u(t) = \kappa(u(t)). \tag{2.139}$$

In detail, the time dependent trajectory $\kappa(t)$ is obtained for a certain $u(t)$ by propagating the states $|\psi(t)\rangle$ forward in time, computation of $|\lambda(T)\rangle$ via Eq. (2.15) from the standard GRAPE method (=outcome from calculus of variations) and subsequent propagation of $|\lambda(t)\rangle$ backwards in time via Eq. (2.14). What remains is to explain how to obtain $\kappa(|\psi(t)\rangle, |\lambda(t)\rangle)$ which we are doing in the next subsection.

2.6.2.2 Fixed-point Equation from the Optimality Condition

We take the opportunity and derive the subsequent equations based on the so-called control Hamilton function [74]. For readers that prefer to stay with the notation from the GRAPE method can of course do so without any restrictions.

The control Hamilton function for the previously introduced scenario reads

$$H = \frac{i}{\hbar} \langle \lambda(t) | \hat{H} | \psi(t) \rangle + \frac{-i}{\hbar} \langle \psi(t) | \hat{H} | \lambda(t) \rangle \tag{2.140}$$

$$= \frac{-2}{\hbar} \text{Im} \left\{ \langle \lambda(t) | \hat{H} | \psi(t) \rangle \right\}. \tag{2.141}$$

Variation with respect to the control $u$

$$\mathcal{H}_u = \frac{\partial \mathcal{H}}{\partial u} = \frac{-2}{\hbar} \text{Im} \left\{ \langle \lambda(t) | \frac{\partial \hat{H}}{\partial u} | \psi(t) \rangle \right\}. \tag{2.142}$$

Assuming a Hamiltonian of the form $\hat{H} = \hat{H}_d + u(t) \hat{H}_c$ leads to

$$\mathcal{H}_u = \frac{-2}{\hbar} \text{Im} \left\{ \langle \lambda(t) | \hat{H}_c | \psi(t) \rangle \right\}. \tag{2.143}$$

The optimality condition is obtained from $0 = \mathcal{H}_u(= \Delta u)$. Hence we can recognize perfectly what we already know from the GRAPE method where the same condition has been derived in Eq. (2.12).

Up to this point we can only work with the usual GRAPE scheme of trying to minimize
$H_u = \Delta u \ \forall t$ via slightly updating $u$ according to $\Delta u$. In order to obtain the aforementioned fixed-point equation we have to have an optimality condition (as the one from Eq. (2.143)) that does explicitly contain the control function $u(t)$ (which we then can solve for). Here comes the trick: By applying further time derivatives $\frac{d^n}{dt^n}$ to (both sides of) $0 = H_u$ we can hope that for some $n \geq 1$, $u(t)$ will appear and we can proceed.

**First time derivative (n=1)**

\[
0 = \frac{dH_u}{dt} = \frac{-2}{\hbar} \text{Im} \left\{ \langle \lambda(t)| \hat{H}_c |\psi(t)\rangle \right\} + \frac{-2}{\hbar} \text{Im} \left\{ \langle \lambda(t)| \hat{H}_c |\dot{\psi}(t)\rangle \right\} \\
= \frac{-2}{\hbar} \text{Im} \left\{ \frac{i}{\hbar} \langle \lambda(t)| \left( \hat{H}_d + u(t) \hat{H}_c \right) |\hat{H}_c |\psi(t)\rangle \right\} + \\
\frac{-2}{\hbar} \text{Im} \left\{ \frac{-i}{\hbar} \langle \lambda(t)| \hat{H}_c \left( \hat{H}_d + u(t) \hat{H}_c \right) |\psi(t)\rangle \right\} \\
= \frac{-2}{\hbar^2} \text{Re} \left\{ \langle \lambda(t)| \hat{H}_d \hat{H}_c |\psi(t)\rangle \right\} + \\
\frac{-2}{\hbar^2} \text{Re} \left\{ \langle \lambda(t)| \hat{H}_c \hat{H}_c |\psi(t)\rangle \right\} u(t) + \\
\frac{2}{\hbar^2} \text{Re} \left\{ \langle \lambda(t)| \hat{H}_c |\hat{H}_d |\psi(t)\rangle \right\} + \\
\frac{2}{\hbar^2} \text{Re} \left\{ \langle \lambda(t)| \hat{H}_c |\hat{H}_c |\psi(t)\rangle \right\} u(t) \\
= \frac{2}{\hbar^2} \text{Re} \left\{ \langle \lambda(t)| \left[ \hat{H}_c, \hat{H}_d \right] |\psi(t)\rangle \right\} \\
\tag{2.147}
\]

As it turns out, after the first time derivation ($n = 1$) we still do not have an explicit equation for $u$. Therefore we add another time derivative.
Second time derivative (n=2)

\[
0 = \frac{d^2 \mathcal{H}_u}{dt^2} = \frac{2}{\hbar^2} \text{Re} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right] | \psi(t) \rangle \right\} + \frac{2}{\hbar^2} \text{Re} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right] | \dot{\psi}(t) \rangle \right\}
\]

\[
= \frac{2}{\hbar^2} \text{Re} \left\{ \frac{i}{\hbar} \langle \lambda(t) | \left( \hat{H}_d + u(t) \hat{H}_c \right) \left[ \hat{H}_c, \hat{H}_d \right] | \psi(t) \rangle \right\} + \frac{2}{\hbar^2} \text{Re} \left\{ \frac{-i}{\hbar} \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right] \left( \hat{H}_d + u(t) \hat{H}_c \right) | \psi(t) \rangle \right\}
\]

(2.148)

\[
= -\frac{2}{\hbar^3} \text{Im} \left\{ \langle \lambda(t) | \hat{H}_d \left[ \hat{H}_c, \hat{H}_d \right] | \psi(t) \rangle \right\} + \frac{2}{\hbar^3} \text{Im} \left\{ \langle \lambda(t) | \hat{H}_c \left[ \hat{H}_c, \hat{H}_d \right] | \psi(t) \rangle u(t) \right\} + \frac{2}{\hbar^3} \text{Im} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right] \hat{H}_c | \psi(t) \rangle \right\} u(t)
\]

(2.149)

\[
= \frac{2}{\hbar^3} \text{Im} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right] , \hat{H}_d | \psi(t) \rangle \right\} + \frac{2}{\hbar^3} \text{Im} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right] , \hat{H}_c | \psi(t) \rangle \right\} u(t)
\]

(2.150)

Shortcutting notation by introducing another commutator results in:

\[
0 = \frac{d^2 \mathcal{H}_u}{dt^2} = \frac{2}{\hbar^3} \text{Im} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right], \hat{H}_d | \psi(t) \rangle \right\} + \frac{2}{\hbar^3} \text{Im} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right], \hat{H}_c | \psi(t) \rangle \right\} u(t)
\]

(2.151)

After the second time derivation we have an equation explicitly solvable for \( u(t) \) at hand. Indeed it is provable that for linear system the so-called order of singularity \( \text{here} = n \) has to be an even number [151]. The last equation has to be fulfilled for an optimal pulse just as the optimality condition for GRAPE will have to be met. After all we can rearrange the equation and finally obtain:

\[
u(t) = \frac{-\text{Im} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right], \hat{H}_d | \psi(t) \rangle \right\}}{\text{Im} \left\{ \langle \lambda(t) | \left[ \hat{H}_c, \hat{H}_d \right], \hat{H}_c | \psi(t) \rangle \right\}}
\]

(2.152)

We identify it as the \( \kappa(|\psi(t)\rangle, |\lambda(t)\rangle) \) we are looking for.

### 2.6.2.3 Discussion

At a first glance this method might look very similar to GRAPE but it has in fact a very much different update scheme. As a result, we expect significantly larger changes in the pulse from one iteration to another. Due to the fact that Newton method can be regarded as a fixed-point iteration, this approach compares much more to an excellent but less widely known work by Pierre de Fouquieres [79]. There, a Newton-Raphson
root finding approach has been selected to optimize coefficients of a function expansion (just as for CRAB (in the non-dressed variant)). As a result, the author could show that the method does greatly outperform Krotov and BFGS GRAPE as it shows double exponential convergence. We hope that similar convergence rates can be achieved with our approach while having a larger radius of convergence. One distinction will anyway be, that the presented approach does not require any function expansion in a non-dressed (fixed) basis which can induce false traps as discussed in Eq. (2.25).
3 Fundamental Limits of Quantum Optimal Control

In this chapter we provide an introduction into four of the most important concepts that pose some sort of boundary conditions on QOCT: quantum speed limit, which determines some minimal transfer time for a given process, an information theoretical limit, which states that the effective dimension of an optimal control problem is very much limited under mild reasonable assumptions. Moreover, controllability gives a mathematical framework to verify if all possible control goals can be achieved in the absence of resource limitations and finally the notation and use of control landscapes which can be understood also intuitively.

3.1 Controllability

Hereafter, only a primer on controllability in the quantum context is provided. For a more comprehensive treatment, the reader is referred to [152–155] and references therein. The central ideas of the elementary concepts from linear algebra and differential geometry presented here follow the description from [152].

We start off with a definition: a Lie algebra \( \mathcal{L} \) over a field \( \mathcal{F} \) is a vector space over \( \mathcal{F} \) with a operation \( \mathcal{L} \times \mathcal{L} \rightarrow \mathcal{L} \). This operation associates an ordered pair of elements \( \{x,y\} \) in \( \mathcal{L} \) an element \( z = [x,y] \). The expression in the square brackets is referenced as Lie bracket or commutator and must satisfy the following axioms:

1. Bilinearity
   \[
   [x + y, z] = [x, z] + [y, z], \quad [x, y + z] = [x, y] + [x, z],
   \]
   \[
   [\alpha x, y] = \alpha [x, y], \quad \forall \alpha \in \mathcal{F}.
   \]

2. \( [x,x] = 0, \quad \forall x \in \mathcal{L} \).
3. Jacobi identity \[ [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \] (3.4)

In the cases of interest for quantum physics, the field \( \mathcal{F} \) is a set of complex numbers \( \mathbb{C} \) and the condition (3.3) can be replaced by the skew-symmetry condition which reads \([x, y] = -[y, x]\). The structure of a Lie algebra is unambiguously determined by the commutation relations on a basis since these determine the value of the commutator of any pair of elements in the Lie algebra.

Given a set of vectors of a Lie algebra \( \mathcal{L} \), \( \{x_1, \ldots, x_n\} \), the set of all the (possibly repeated) commutators of \( \{x_1, \ldots, x_n\} \) spans a subalgebra of \( \mathcal{L} \) which is called the Lie algebra generated by \( \{x_1, \ldots, x_n\} \).

Before introducing the second important element borrowed from algebra, the Lie group, we briefly review the properties of groups. A group is a set of elements \( G \) together with an operation \( \ast \) that combines any two elements \( x \) and \( y \) by forming another element denoted by \( x \ast y \) or simply \( xy \) which satisfy the following four group axioms:

- **Closure** For all \( \{x, y\} \) in \( G \), the result of the operation \( xy \) is also contained in \( G \).
- **Associativity** For all \( \{x, y, z\} \) in \( G \), the following holds \((xy)z = x(yz)\).
- **Identity element** There exists an element \( e \) in \( G \) such that \( ex = xe = x \ \forall x \in G \).
- **Inverse element** An inverse \( x^{-1} \) exists for each element \( x \) in \( G \) such that \( x^{-1}x = xx^{-1} = e \).

If additionally all the elements of the group commute, i.e. \( xy = yx \), with one another, the group is said to be abelian. Subgroup of groups might be defined where the previous axioms hold for the same operation \( \ast \) also with a reduced number of elements.

A **Lie group** is a group which is also an analytic differentiable manifold and such that the group operations \( \{x, y\} \rightarrow xy \) and \( x \rightarrow x^{-1} \) are analytic.

We obtain the Lie algebra of the group by considering group elements which can be reached from the identity 1 by an infinitesimally tangent vector (for a mathematical derivation see [152] Chapter 3.1.2.3). Based on these elements, we can construct operators referenced as generators which gives us a unitary representation of the group. A representation of a group is a homomorphism between the group and a group of linear operators which act on a vector space. In our context, elements of that vector
space are the states (wave functions) of a quantum system\textsuperscript{1}. In detail, we obtain all the group elements which can be generated by differentiating curves at the identity at time \( t = 0 \). Various methods for obtaining the generators of a Lie group have been discussed also in the literature, see e.g. [153].

Suppose we have generators \( \{x_1,\ldots,x_n\} \) of a Lie group where the group is characterized by \( n \) parameters \( \alpha_i \). In case of hermitian generators, a unitary representation for an arbitrary group element \( U_a \) can be written as

\[
U_a = e^{-i\sum_{j=1}^{n} \alpha_j x_j}.
\]

It can be shown, that the \( x_j \) make up a Lie algebra [155]. Consequently they satisfy the algebraic equations

\[
[x_a,x_b] = i \sum_{j=1}^{n} c_{a,b}^j x_j,
\]

where the square brackets are the commutators and the \( c_{a,b}^j \) some constants. We can conclude that our Lie algebra has the property that commutators of any elements are just a linear combination between (all) the elements of the Lie algebra. A valid representation of a Lie algebra is therefore a set of matrices obeying the commutation relations.

The \textbf{unitary group} in \( n \) dimensions \( U(n) \) is the group of \( n \times n \) matrices \( U_a \) satisfying

\[
U_a^\dagger = U_a^{-1}.
\]

An arbitrary complex valued matrix with dimension \( n \) is given by \( 2n^2 \) real numbers. For unitary matrices this reduces to \( n^2 \) degrees of freedom and hence \( U(n) \) can be denoted by \( n^2 \) (real) parameters. In the simplest case, the group \( U(1) \) is characterized by just a phase factor \( e^{i\phi} \). An important case makes up the \textit{special unitary groups} \( SU(n) \) consisting of matrices with unit determinant. As a consequence there are \( n^2 - 1 \) free parameters. A series of further subgroups and their relations has been discussed in [156]. As an example we show how to construct the 2-dimensional representation of

\begin{footnote}
In many cases a finite dimensional \textit{matrix representation} of a group is used which is a homomorphism between the group and a group of matrices. Matrix representation is often simplified to the notation “representation”.
\end{footnote}
SU(2). The generators can be chosen to be the Pauli matrices

$$x_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.8)$$

They are satisfying the commutation relations

$$[x_a, x_b] = 2i\varepsilon_{ab} x_j. \quad (3.9)$$

Now we are turning towards how to determine whether or not a system is (completely) controllable. According to [157], the following applies: Definition: A quantum system $H = H_0 + \sum_{i=1}^{M} u_i(t) H_i$ is completely controllable if every unitary operator $U$ is accessible from the identity operator $1$ via a path $\gamma(t) = U(t,t_0)$, where $i\partial_t U(t,t_0) = HU(t,t_0)$. Using a theorem from [158; 159], a necessary and sufficient condition for complete controllability of a quantum system $H$ (as introduced previously) is that the Lie algebra $L_0$ has dimension $N^2$. $L_0$ is generated from the skew-Hermitian operators $\{iH_0, iH_1, \ldots, iH_M\}$.

Closing this section, we like to add some remarks. Although a system is controllable, finding a (set of) pulses that do a requested unitary transformation can become arbitrary hard. On the contrary, it might still be possible to identify a specific transformation after a given time although a system is not controllable. Hence the controllability criterion can only state that any transformation is possible in theory. As it turns out, controllability is given for many practical scenarios. The problem of determining the controllability can also be mapped onto the study of the connectivity of a graph, where the nodes are eigenstates of the system in some basis and the edges are the transition elements [160; 161].

### 3.2 Quantum Speed Limit

Optimized controlled processes are key to basically all cutting-edge quantum technologies as motivated in the introduction of this thesis (Chapter 1). One valid strategy to achieve this goal is to leave enough transition time and only vary dynamical parameters of the Hamiltonian slowly such that the instantaneous eigenstates are kept well under control. This approach is known as the adiabatic theorem [132; 162]. Unfortunately the strategy is often too slow for realistic scenarios to be feasible and initial and goal states might not
necessarily be eigenstates of the system. What remains to effectively fight decoherence is to drive the quantum system to its target as fast as possible. Time-optimal control theory provides recipes to achieve quantum operations with high accuracy and speed. Similarly, the transmission rate is a fundamental characteristic to assess the efficiency and the feasibility of real-world application of quantum cryptography [163–165]. Hence, the question about the maximal speed of a quantum evolution naturally arises. When doing experiments or computer-based simulations for every non-trivial quantum control problem, a lower bound on the transfer time occurs. This bound is not attributed to numerical or experimental subtleties nor our lack in finding a better solution. In fact it exists a fundamental bound referenced as quantum speed limit (QSL), or energy-time uncertainty relation, describing the fundamental maximum rate for quantum time evolutions [166–169]. The minimal time needed for a system to evolute from an initial state to an orthogonal one is denoted as \( \tau_{\text{QSL}} \). The first relation yields

\[
\tau \geq \tau_{\text{QSL}} = \frac{\pi \hbar}{2 \Delta E},
\]

where \( \Delta E \) is the energy variance \( \sqrt{\langle \psi_0 | (H - E)^2 | \psi_0 \rangle} \) and \( E = \langle \psi_0 | H | \psi_0 \rangle \). This inequality is known as the Mandelstam-Tamm bound [170]. Alternatively, there is the Margolus-Levitin bound [171] given by

\[
\tau \geq \tau_{\text{QSL}} = \frac{\pi \hbar}{2E},
\]

where \( E \) is the energy as previously defined. In general, the two speed limits are independent and the minimal evolution time therefore is given by

\[
\tau \geq \tau_{\text{QSL}} = \max \left\{ \frac{\pi \hbar}{2 \Delta E}, \frac{\pi \hbar}{2E} \right\} = \frac{\pi \hbar}{E + \Delta E - |E - \Delta E|}.
\]

A different approach has been shown in [76], where the problem of time optimal quantum control is casted into minimizing the quantum action \( S \) of a system and thus the problem can be interpreted as the quantum analogue of the classical brachistochrone. Exploiting the variational principle, a set of coupled non-linear equations is derived whose solution gives the required time-optimal Hamiltonian while satisfying constraints on the available control resources. The same authors also extended their scheme towards time-optimal target unitary operations [172].

Based on these results a generalization towards time-dependent system has been demonstrated e.g. in [173], where QSL are compared to numerical results from optimal control for a Landau-Zener transition and information transfer along a spin chain.
The idea is straight-forward: the essential quantities $\Delta E$ and $E$ from before are being replaced by the mean energy spread $\Delta \mathcal{E}$ and mean energy $\mathcal{E}$ respectively. Hence the QSL from Eq. (3.12) extends towards

$$\tau \geq \tau_{\text{QSL}} = \max \left\{ \frac{\pi \hbar}{2 \Delta \mathcal{E}}, \frac{\pi \hbar}{2 \mathcal{E}} \right\},$$

(3.13)

where

$$\Delta \mathcal{E} = \frac{1}{T} \int_0^T \sqrt{\langle \psi(t) | (H(t) - E(t))^2 | \psi(t) \rangle} dt$$

(3.14)

$$E(t) = \langle \psi(t) | H(t) | \psi(t) \rangle$$

(3.15)

$$\mathcal{E} = \frac{1}{T} \int_0^T E(t) dt.$$  

(3.16)

The relation with the time-independent situation becomes more apparent when considering $N$ equidistant time intervals of length $\Delta t = T/N$ where the Hamiltonian is approximately constant and the average of these becomes the respective integral in the limit of $N \to \infty$.

We like to stress the fact that methods from QOCT can be used to both identify the quantum speed limit and determine controls that drive the system at the quantum speed limit [173]. Nonetheless there is clearly lots of room for further developments as the tightness (gap between actually found quantum evolution in minimal time versus theoretical prediction) for various quantum systems could not yet be reached or additional arguments be given on how more evolved bounds look like [174–176]. One route is to look for approximate analytical solutions which has been demonstrated e.g. in [177]. The study of QSL is still an active field of research with lots of latest developments [139; 178–183].

Closing this section, we show for the analytically solvable system from Sec. 2.4.2 how for a paradigmatic case the QSL can be determined. We are searching for the minimal time $T^*$ to reach the unitary $U^* = e^{i/2a\sigma_z}e^{i/2b\sigma_y}e^{i/2a\sigma_z}$ with $(a,b) = (0, \frac{k\pi}{4})$,
where \( k \in \{1,2,3,4\} \). Following Eq. (2.82), we get in the situation \( p_2 = 0 \)

\[
\frac{k\pi}{4} = \theta(T^*) = \cos^{-1}\left( 1 + \cos\left( \frac{2T^*}{1} \right) - 1 \right) \tag{3.17}
\]

\[
\cos\left( \frac{k\pi}{4} \right) = \cos(2T^*) \tag{3.18}
\]

\[
T^* = \frac{k\pi}{8}. \tag{3.19}
\]

The respective trajectories are illustrated in Fig. 3.1.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{trajectory.png}
\caption{Projected trajectories for the optimal synthesis of \( U^* = e^{i/2a_s} e^{i/2b_s} e^{i/2a_s} \) with \((a,b) = (0, \frac{k\pi}{4})\), where \( k \in \{1,2,3,4\} \). The dots indicate the position of the target states. Taken from [74].}
\end{figure}

### 3.3 Information Theoretical Limit

QOCT can be viewed in terms of information flows: the control pulse encodes information which tells the system, starting off from an initial state, where to go within the pulse duration. Closely following the presentation in the papers [115; 116], we elaborate on how much information (i.e. number of bits) has to be encoded in the pulse to achieve a state transformation up to an error \( \varepsilon \) (e.g. the infidelity). The result is that the number of independent degrees of freedom in the pulse scales only polynomially with the dimension of some set of reachable states from an initial state. In order to get to this theorem, a number of assumptions are made that are justified from the fact that there are practical limitations unavoidably present in any actual physical realization such as limited control amplitude and transfer time.

We consider a quantum optimal control problem where the state \( \rho(t) \) of the system
follows some dynamical equation
\[ \dot{\rho} = \mathcal{L}(\rho, u(t)), \]  
(3.20)

where \( \rho \) is the density matrix describing a quantum system defined on a Hilbert space \( \mathcal{H} \), and \( \mathcal{L} \) is the Liouvillian operator with the unitary part generated by the Hamiltonian
\[ H = H_d + u(t)H_c, \]  
(3.21)

where \( H_d \) and \( H_c \) are the drift and control Hamiltonian, respectively. The goal is to find some control pulse \( u(t) \) that drives the system from an initial state \( \rho_0 \) to some goal state \( \bar{\rho} \).

At first we define the information content (number of bits \( b_u \)) carried by the control pulse \( u(t) \) as the product between the classical channel capacity \( C \) and the pulse duration \( T \). Following Hartley’s law, we get
\[ b_u = T \Delta \Omega \kappa_s, \]  
(3.22)

where \( \Delta \Omega \) is the bandwidth and \( \kappa_s = \log(1 + \Delta u/\delta u) \) is the bit depth of \( u(t) \), and \( \Delta u = u_{\text{max}} - u_{\text{min}} \) and \( \delta u \) are the maximal and minimal allowed variation of the field [184]. Given a uniform sampling rate of the signal \( \delta u \), then \( T \Delta \Omega = T/\delta t = n_s \), where \( n_s \) is the number of sampling points. Any numerical optimization algorithm will have the \( n_s \) variables as a basis and work on these until the control goal is fulfilled (see also discussion on function expansion methods in Sec. 2.3.1).

The dimension \( D \) of a quantum OC problem might be defined as follows: Given a dynamical law of the form of Eq. (3.20), an initial state \( \rho_0 \) and a goal state \( \bar{\rho} \), the dimension of the quantum OC problem is defined by the minimal number of independent degrees of freedom \( D \) in the OC field necessary to achieve the transformation up to an error \( \varepsilon \). Note that \( D \) might be the minimal number of sampling points \( n_s \) (i.e. \( D = T \Delta \Omega \)), referring to bang-bang switching points, or of frequencies present in the control field, or more general the dimension of the subspace of functions the control field has nonzero projection on.

Before turning towards considerations regarding the effective Hilbert space, we introduce some well motivated physical constraints: the control duration is finite, that is \( t \in [0, T] \). The Hamiltonians are bounded \( ||H_d|| = ||H_c|| = 1 \) in some operator norm. Also the control field is amplitude limited \( u(t) \in [u_{\text{min}}, u_{\text{max}}] \forall t \).

These physical constraints naturally introduce a set of time-polynomial reachable states \( \mathcal{W}^+ \) that can be reached with finite energy with precision \( \varepsilon \) in polynomial time as a
function of the Hilbert space dimension $N$. The dimensionality of these is defined as $\dim(W^+) = D_{W^+}(N)$. This class of states encompasses the practically relevant states that can be reached (exclusively) in the lab. The finite precision can be motivated by the fact that our capability in resolving a certain state will always be finite.

After having set the stage, we formulate the central theorem that allows us to draw conclusions about the effective OC complexity: the dimension $D$ of a quantum OC problem in $W^+$ up to precision $\varepsilon$ is a polynomial function of the dimension of the manifold of the time-polynomial reachable states $D_{W^+}$. In the following, the proof is briefly sketched. The principle idea is that the dimension of the problem is lower bounded by $D_{W^+}$ and upper bounded by a polynomial function of $D_{W^+}$.

Lower bound: The amount of information required to specify a state within a ball of radius $\varepsilon$ in the set of time-polynomial reachable states $W^+$ is given by $b_S = \log(\varepsilon^{-D_{W^+}})$. The information content of the OC field has to carry at least the same information content, i.e. $b_u \geq b_S$. Solving for $\varepsilon$ we obtain

$$\varepsilon \geq 2^{-\frac{T\Delta\Omega_{\kappa_s}}{D_{W^+}}}.$$ (3.23)

Setting a maximal precision (e.g. machine precision) expressed in bits $\kappa_\varepsilon = -\log_2(\varepsilon)$ results in $T\Delta\Omega_{\kappa_\varepsilon}/D_{W^+} = \kappa_\varepsilon$, and imposing $\kappa_\varepsilon = \kappa_s$, we finally arrive at

$$D \geq D_{W^+}.\quad (3.24)$$

Upper bound: Here we consider a path of finite length $L$ that connects the initial and goal state, which are both in $W^+$, in polynomial time. The maximum of information $b_S^+$ needed to fully describe the path is the number of balls $n_\varepsilon$ that make up the path times the information to describe each $\varepsilon$-ball (see lower bound). The former is given by

$$n_\varepsilon = L/\varepsilon \leq TV_{\max}/\varepsilon = \text{Poly}(D_{W^+})v_{\max}/\varepsilon,$$ (3.25)

where $v_{\max}$ is the maximal allowed velocity along the path which gives rise to bounded energy. Putting things together we obtain

$$b_S^+ = \frac{\text{Poly}(D_{W^+})v_{\max}}{\varepsilon} \log(\varepsilon^{-D_{W^+}}).$$ (3.26)
The information content describing the whole path through $\mathcal{W}^+$ hence marks an upper bound for $b_u$ and we finally get

$$\frac{\text{Poly}'(D_{W^+})v_{\text{max}}}{\varepsilon} \geq D,$$

(3.27)

where Poly’ is the polynomial function Poly with all powers increased by one.

In summary, we have shown that the effective dimension of an OC problem, at least for integrable systems, is polynomially bounded in terms of the resources (energy, time) given. The methodology used to achieve this result is based on dividing an effective set of states into $\varepsilon$-balls and applying intuitive geometric arguments to obtain the information content needed to select the target ball among this partition.

### 3.4 Control Landscapes

Quantum control landscapes are made up by the objective $J$ as a function of the control variables. The topology of the emerging control landscapes has important implications as it determines whether or not a local search algorithm converges to a global optimum with probability one. The key issue in answering that question is to resolve the existence of traps in the landscape.

Following the notation of standard literature [185; 186], control landscapes are denoted for state transfer problems as

$$J(u) = F(|\psi(T)\rangle).$$

(3.28)

(Approximate) gradient based searches look for critical points in the landscape, that is, points that fulfill

$$\delta J = \langle \nabla F(\psi(T))|\delta \psi(T)\rangle = 0 \forall \delta u.$$  

(3.29)

Here, the chain rule has been applied on Eq. (3.28) when calculating the variation of $J$. As a result, two constituents emerge: the fidelity gradient with respect to final state variation ($\nabla F/\delta \psi(T)$) and the final state variation as a function of the control variation $\delta \psi(T)/\delta u$. As a necessary condition for a global maximum, minimum or saddle point, the first gradient $\nabla F(\psi(T))$ has to vanish. However, for state overlaps as a figure of merit ($F \in [0,1]$), saddle points cannot occur because the second order variation is nonzero. Hence the condition corresponds to either the global maximum
or minimum. Unfortunately, the second part in Eq. (3.29) can also cause a vanishing variation at suboptimal points. If there does not exist a variation $\delta u$ that can generate arbitrary $|\psi(T)\rangle$ (which means, speaking pictorial, every possible escape direction can be explored) this point is being classified as a singular critical point.

There has been quite some debate (see e.g. [187–189]) whether or not there exist singular points that are (real) traps (i.e. (local) gradient algorithms are not able to escape). Numerical studies suggest that for controllable systems and in the absence of any constraints on the pulses (such as amplitude constraints), for non constant pulses, the landscape is devoid of traps if enough process time is given [186; 190–193]. Therefore, it is nowadays generally believed (although no rigorous proof has been conducted for general systems) that, given the aforementioned conditions are met, landscapes are trap free. As a remark to this discussion, we stress that for any realistic system (e.g. in the lab) at least the assumption of unlimited control amplitude availability can never be fulfilled. Moreover, as for most scenarios very short transfer times have to be realized to limit decoherence effects, the ‘sufficient enough time’ condition cannot always be ensured, too. For readers new to the field we show a visualization from one of the pioneering works [194] of a control landscape as a function of two control variables in Fig. 3.2. Here, all extrema correspond to the global maximum although exhibiting different shapes.

![Figure 3.2](image-url)

**Figure 3.2:** Transition probability for a state transfer as a function of two control variables. The transition probability does depend on many other control variables, too. Taken from [194].

Independently of the pure theoretical considerations, it is of high relevance to study the trapping behavior of real experimental landscapes. With the advent of more refined experiments that allow for a growing number of system sizes (e.g. number of individually addressable constituents such as atoms or spins) to be probed, we have already reached
a point were pure open-loop simulation on a classical computer is unfeasible for some scenarios. Thus, it is highly desirable to see whether possible traps emerging from closed-loop optimal control do exist or could be escaped from. This also forms the main driving force behind the publication presented in Chapter 5.
Part II

APPLICATIONS OF OPEN- AND CLOSED-LOOP OPTIMAL CONTROL
4 Amplification of the Dynamical Casimir Effect

4.1 Introduction to the dynamical Casimir effect

The phenomenon of an attractive force between two metal plates facing each other and which is rapidly falling off as their distance increases was named after the Dutch physicist Hendrik Casimir who was the first to theoretically predict the effect in 1948 [195]. In a purely classical description the force would not occur as no electromagnetic field exists between the plates when no voltage is applied. However, considering quantum field theory, there is still some zero-point energy due to the presence of the resonator. This energy affects then virtual photon that eventually make up a net force. Because of that, the sign and magnitude of the force is strongly dependent on the geometry of the boundaries.

As opposed to this often referenced effect as static Casimir effect, the term dynamical Casimir effect (DCE) has been introduced by Yablonovitch [196] and Schwinger [197] and is used today for the effect of the generation of real photons from the vacuum as a consequence of sudden (nonadiabatic) changes of the geometry of the boundaries or material properties. A broad categorization of the DCE into two general classes might be done according to [198]: The mirror induced DCE appears due to movement of the confining mirrors or changes of their material properties. It is noteworthy that for experimental detectability, mirror velocities close to the speed of light are needed, which, of course, poses difficulties. In this categorization, all other effects of photon creation that do not involve any manipulation of boundary conditions imposed by mirrors are summarized under the term parametric DCE. Inside this class, the first observation of the Dynamical Casimir effect has been reported in 2011 by researches at the Chalmers University of Technology [199]: a modified SQUID was used to change the effective length of the resonator in time by modulation of the inductance at high frequencies such that real photons could be created. Our publication reprinted in Sec. 4.3 is also based on the parametric DCE and any further mention of DCE shall, without specification,
4.2 Time-dependent Rabi Hamiltonian

A very straightforward and intuitive light-matter interaction model takes into account a single atomic transition which interacts with exactly one mode of the field. In the following, based on these two ingredients, we show how to derive the Rabi model and further simplified Jaynes-Cummings model. Owing to the simplicity and the existence of an analytical solution, especially the latter has been studied extensively and serves as a basis for a magnitude of studies [200–202]. The Hamiltonian for the more general Rabi model consists of three parts: A two-level atom (spin-1/2 particle) $H_{\text{atom}}$, a quantized cavity field $H_{\text{field}}$ and their interaction $H_I$. The former two are summarized as $H_0 = H_{\text{atom}} + H_{\text{field}}$ which form the rotating frame when going into the interaction picture. The essential building blocks are shown in Fig. 4.1. As our ultimate goal is to dynamically steer the system, we control the qubit transition frequency $\omega_q(t)$ and therefore obtain the time-dependent Hamiltonian

$$H_{\text{atom}}(t) = -\frac{1}{2} \omega_q(t) \sigma_z. \quad (4.1)$$

Here, $\sigma_i (i = x, y, z)$ are the Pauli matrices written in the $\{|g\rangle, |e\rangle\}$ basis. The frequency $\omega_q(t)$ might be further split into $\omega_q(t) = \omega_{q,0} + \dot{\Phi}(t)$, where $\dot{\Phi}(t)$ is a real valued (single)
control function. The field is modeled as a harmonic oscillator with frequency $\omega$

$$H_{\text{field}} = \omega(n + \frac{1}{2}),$$ \hspace{1cm} (4.2)

where $n = a^\dagger a$ is the photon number operator. The standard commutation relations, which will become important later on, hold:

$$[n,a] = -a$$ \hspace{1cm} (4.3)

$$[n,a^\dagger] = a^\dagger$$ \hspace{1cm} (4.4)

Moreover, when considering bosons, we have:

$$[a,a^\dagger] = 1$$ \hspace{1cm} (4.5)

$$[a,a] = [a^\dagger,a^\dagger] = 0$$ \hspace{1cm} (4.6)

Finally, the interaction is modeled as a dipole interaction $H_I = S \cdot E$, where the atom is coupled through its polarization $S \propto \sigma_x = \sigma_+ + \sigma_-$ to the electric field operator $E \propto a^\dagger + a$. The raising and lowering operators for the qubit are $\sigma_\pm = \frac{1}{2}(\sigma_x \mp i\sigma_y)$. Thus, we get

$$H_I = \lambda(\sigma_+ + \sigma_-)(a^\dagger + a) = \lambda(a^\dagger\sigma_+ + a\sigma_+ + a^\dagger\sigma_- + a\sigma_-).$$ \hspace{1cm} (4.7)

The terms in the interaction Hamiltonian describe the following physical process:
Simultaneous excitation (de-excitation) of the atom and the field is given by $a^\dagger\sigma_+ + a\sigma_-$. Excitation of the atom via absorption of a photon is given by $a\sigma_+$, while the opposite process is described by $a^\dagger\sigma_-$. The (real) coupling strength $\lambda$ shall be in the ultrastrong coupling regime $\lambda \lesssim \omega$ [203–205]. In the Heisenberg picture, the ladder operators evolve as:

$$a^\dagger(t) = a^\dagger e^{i\omega t}$$ \hspace{1cm} (4.8)

$$a(t) = ae^{-i\omega t}$$ \hspace{1cm} (4.9)

$$\sigma_\pm(t) = \sigma_\pm \exp(\pm i \int_0^t \omega_q(\tau)d\tau).$$ \hspace{1cm} (4.10)
Consequently, in the interaction picture (denoted by the tilde) we get for the relevant interaction part

\[ \tilde{H}_I = \lambda (a^{\dagger} \sigma_+ \exp(i(\omega_{q,0} + \omega)t + \Phi(t)) + \]
\[ a \sigma_- \exp(-i(\omega_{q,0} + \omega)t + \Phi(t))) + \]
\[ a \sigma_+ \exp(i((\omega_{q,0} - \omega)t + \Phi(t))) + \]
\[ a^{\dagger} \sigma_- \exp(-i((\omega_{q,0} - \omega)t + \Phi(t))) , \]

where \( \Phi(t) = \int_0^t \dot{\Phi}(\tau)d\tau \) is the accumulated phase. This model reduces to the standard Rabi model, when setting \( \Phi(t) = 0 \) and to the Jaynes-Cummings model if furthermore the counter-rotating terms \((a^{\dagger} \sigma_+, a \sigma_-)\) are dropped (Rotating Wave Approximation).

Since the goal is to amplify the dynamical Casimir effect, we are seeking to design a control function \( \Phi(t) \) to generate as many photons \( \langle n \rangle(T) \) from the vacuum as possible at a given time \( T \). As a side remark we like to point out, that this system might equally well be described in terms of conjugate variables, as it has been shown in [206]. Therefore, the moment framework, as introduced in Chapter 7, can by applied to simulate this system.
4.3 Amplification of the parametric dynamical Casimir effect via optimal control


4.3.1 Abstract

We introduce different strategies to enhance photon generation in a cavity within the Rabi model in the ultrastrong coupling regime. We show that a bang-bang strategy allows to enhance the effect of up to one order of magnitude with respect to simply driving the system in resonance for a fixed time. Moreover, up to about another order of magnitude can be gained exploiting quantum optimal control strategies. Finally, we show that such optimized protocols are robust with respect to systematic errors and noise, paving the way to future experimental implementations of such strategies.

4.3.2 Introduction

High-speed manipulation of quantum systems is key in quantum information processing: Quantum gates should operate on a time scale much smaller than the decoherence time, to allow efficient error correction and fault-tolerant architectures [20; 21; 173]. Similarly, the transmission rate is a fundamental characteristic to assess the efficiency and the feasibility of real-world application of quantum cryptography [163–165]. Circuit quantum electrodynamics [207; 208] might play a prominent role to speed up quantum protocols, since it allows to address the ultrastrong coupling regime [203; 209–212] of light-matter interaction, where the coupling strength $\lambda$ becomes comparable to, or even exceeds the resonator frequency $\omega$. This regime has also interesting properties on its own, such as the emergence of a strongly correlated light-matter ground state [211; 212].

A related problem is the detection of the dynamical Casimir effect (DCE), namely the generation of photons from the vacuum due to time-dependent boundary conditions or, more generally, as a consequence of the nonadiabatic change of some parameters of a system [198; 213; 214]. The case, addressed in this paper, where the photons are created by parametric amplification of vacuum fluctuations without moving boundaries is referred to as parametric DCE [198]. Since a rapid variation of the matter-field
coupling is needed to implement ultrafast quantum gates, the DCE appears as a fundamental limit to the implementation of high-speed quantum gates [215] and more generally to the development of ultrafast quantum technologies. Pioneering experimental demonstrations of the DCE have been reported in superconducting circuit quantum electrodynamics [199; 216]. However, it is of great interest to have the ability to either enhance or counteract [215] this effect: On the one hand, it improves our capability of investigating fundamental effects in nature, and, on the other hand, it enables us to push the limits of quantum information processing.

Here, we present three different strategies to amplify and thus improve the visibility of the DCE in a parametrically driven system via precisely tailored timing of the matter-field interaction. In detail, we consider on-off resonance sweeping of a parametrically driven qubit coupled to a single mode of the electromagnetic field. We model the qubit-field interaction by the Rabi Hamiltonian [217], with a time-dependent modulation of the qubit frequency. We use and compare two strategies to optimize the visibility of the DCE: a multi step heuristic method employing bang-bang switches of the qubit frequency from the off-resonance regime to the resonance regime and back out of resonance, and optimal control theory which has been proven to be able to successfully control circuit quantum electrodynamics processes [21; 185; 218–223]. In particular, we employ the dressed chopped random basis algorithm (dCRAB) which has been already applied successfully to various theoretical and experimental atomic and condensed matter control problems to meet various control goals, including state-transfer, gate synthesis, observable control, and fast quantum phase transition crossing [113; 127–129; 178; 224; 225]. For the problem studied here, the control function is the time-dependent modulation of the qubit frequency and the figure of merit is the expectation value of the number of photons that are generated in the cavity by parametric amplification of the DCE: This amounts to finding an optimal setup for the detection of the DCE. The results obtained from the dCRAB algorithm are compared to the bang-bang strategy and it is shown that dCRAB could identify pulses yielding up to about one order of magnitude more photons than the bang-bang strategy and up to two orders of magnitude more photons than an unoptimized protocol, which is a single sweep of the qubit to resonance.

The manuscript is organized as follows: First, we introduce the dynamical system model in Sec. 4.3.3. In Sec. 4.3.4 we describe the three control strategies employed to enhance the DCE, discuss the results and give an outlook towards their experimental implementation also investigating their robustness against some possible experimental imperfections. We conclude our work in Sec. 4.3.5.
4.3 Amplification of the parametric dynamical Casimir effect via optimal control

4.3.3 The model

Hereafter, we describe the interaction between a single qubit and a single mode of the quantized field by means of the Rabi Hamiltonian [217], with a time-dependent modulation:

\[
H(t) = H_0(t) + H_I,
\]

\[
H_0(t) = -\frac{1}{2} \left[ \omega_{q0} + \dot{\Phi}(t) \right] \sigma_z + \omega \left( a^\dagger a + \frac{1}{2} \right),
\]

\[
H_I = \lambda \sigma_+ (a^\dagger + a) + \lambda^* \sigma_- (a^\dagger + a),
\]

where the reduced Planck’s constant is set to \( \hbar = 1 \), \( \omega_{q0} \) being the reference frequency for the qubit, \( \sigma_i \) (\( i = x,y,z \)) are the Pauli matrices, written in the \( \{ |g\rangle, |e\rangle \} \) basis; \( \sigma_\pm = \frac{1}{2} (\sigma_x \pm i \sigma_y) \) are the raising and lowering operators for the qubit (so that \( \sigma_+ |e\rangle = |e\rangle, \sigma_- |g\rangle = 0 \), \( \sigma_- |e\rangle = 0, \sigma_+ |g\rangle = |g\rangle \)). The operators \( a^\dagger \) and \( a \) for the field create and annihilate a photon:

\[
a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad a |n\rangle = \sqrt{n} |n-1\rangle,
\]

\( |n\rangle \) being the Fock state with \( n \) photons. For the sake of simplicity, we consider a real coupling strength, \( \lambda \in \mathbb{R} \). The real function \( \dot{\Phi}(t) \) is the control field which allows to manipulate the system: the qubit frequency is modulated via \( \omega_q(t) = \omega_{q0} + \dot{\Phi}(t) \). Notice that we used the notation with the first derivative because, as it will become apparent in the next paragraph, the relevant quantity is the accumulated phase, i.e. \( \Phi(t) \).

The Rotating Wave Approximation (RWA) (valid for \( \lambda \to 0 \)) is obtained neglecting the term \( \sigma_+ a^\dagger \), which simultaneously excites the qubit and creates a photon, and \( \sigma_- a \), which de-excites the qubit and annihilates a photon. In this limit, the Hamiltonian (4.15) reduces to the Jaynes-Cummings Hamiltonian [217] with a time-dependent modulation. In the RWA the swapping time needed to transfer an excitation from the qubit to the field or vice versa (\( |e\rangle|0\rangle \leftrightarrow |g\rangle|1\rangle \)) is \( \tau_s = \pi/2\lambda \), and no DCE is possible since the total number of excitations in the system is conserved.

In the interaction picture, we first consider the unitary operator

\[
U(t) = \mathcal{T} \exp \left[ -i \int_0^t d\tau H_0(\tau) \right]
= \mathcal{T} \exp \left\{ i \left[ \omega_{q0} t + \Phi(t) \right] \sigma_z - i\omega t \left( a^\dagger a + \frac{1}{2} \right) \right\},
\]

where \( \mathcal{T} \) is the time-ordering operator and \( \Phi(t) = \int_0^t d\tau \dot{\Phi}(\tau) \) the accumulated phase.
The time-dependent Hamiltonian in the interaction picture then reads

\[
\hat{H}_I(t) = U^\dagger(t) H_I U(t)
\]

\[
= \lambda (a\sigma_- \exp\{-i[(2\omega - \Delta_0) t + \Phi(t)]\}
\]

\[
+ a\sigma_+ \exp\{i[-\Delta_0 t + \Phi(t)]\}
\]

\[
+ a^\dagger \sigma_- \exp\{-i[-\Delta_0 t + \Phi(t)]\}
\]

\[
+ a^\dagger \sigma_+ \exp\{i[(2\omega - \Delta_0) t + \Phi(t)]\}
\],

where we have defined the reference detuning \(\Delta_0 = \omega - \omega q_0\). The standard Rabi model is recovered for the time-independent Hamiltonian, \(\Phi(t) = 0\), and the Jaynes-Cummings model if we further neglect the counter-rotating terms at frequency \(2\omega\). From now on we will omit tildes and always refer to the interaction picture.

If we expand, in the interaction picture, the qubit-field state at time \(t\) as

\[
|\Psi(t)\rangle = \sum_{l,g,e} \sum_{n=0}^{\infty} C_{l,n}(t) |l,n\rangle,
\]

we obtain the equations that govern the evolution of the coefficients \(C_{l,n}(t)\):

\[
\begin{cases}
  i \dot{C}_{g,n}(t) &= \Omega_n e^{-i[-\Delta_0 t + \Phi(t)]} C_{e,n-1}(t) \\
                      &= +\Omega_{n+1} e^{-i[(2\omega - \Delta_0) t + \Phi(t)]} C_{e,n+1}(t), \\
  i \dot{C}_{e,n}(t) &= \Omega_{n+1} e^{i[-\Delta_0 t + \Phi(t)]} C_{g,n+1}(t) \\
                     &= +\Omega_n e^{i[(2\omega - \Delta_0) t + \Phi(t)]} C_{g,n-1}(t).
\end{cases}
\]

Here \(\Omega_n = \lambda \sqrt{n}\) are the Rabi frequencies, with \(n = 0,1,2,...\) (the terms \(C_{l,m}\) and \(\dot{C}_{l,m}\) must be set to zero when \(m < 0\)). In numerical simulations we will set the reference detuning \(\Delta_0 = 0\) to ensure that the off-resonance condition \(|\omega - \omega q| \ll \lambda\) holds in the ultrastrong coupling regime [203].

### 4.3.4 Results

We consider the time evolution of the qubit-oscillator system for an overall time interval \(T\). An initial and a final time interval \(\tau\), corresponding to off-resonance evolutions with detuning \(|\omega - \omega q(t)| \gg \lambda\), are excluded from manipulation to determine the figure of merit. Consequently, on-off resonance sweeps (governed by a time-dependent pulse
4.3 Amplification of the parametric dynamical Casimir effect via optimal control

\[ \omega_q(t) \] are possible during the intermediate time span \( \tau_p = T - 2\tau \). Initially both the field and the qubit are prepared in their ground state \( |\Psi(t = 0)\rangle = |g, 0\rangle \), so that within the RWA there is no generation of photons at all times \( \langle n \rangle = 0 \forall t \). On the other hand, when the terms beyond the RWA are taken into account, only a very weak photon generation is possible in the off-resonance regime, while a significant photon emergence is possible if the resonance condition is approached nonadiabatically. To quantify the strength of this manifestation of the DCE effect, we consider the figure of merit

\[ f = \bar{n}_f - \bar{n}_i, \tag{4.19} \]

where \( \bar{n}_{f/i} = \int \langle n(t) \rangle dt \) is the time-average of the mean photon number \( \langle n \rangle \) over the initial (i) and final (f) off-resonance time intervals of duration \( \tau \) each. Hereafter, we set the off-resonance condition to \( \omega_q = 4\omega \); the exact numerical value has, however, a very minor influence only.

4.3.4.1 On-off protocol

We begin by studying the efficiency of a straightforward protocol, namely a sudden switch to resonance \( \omega_q = \omega \) at time \( \tau \), followed by another instantaneous quench to the off-resonance condition at time \( T - \tau \). That is, the system is kept on resonance for a total time \( \tau_p \).

Typical examples of time evolutions of the instantaneous average number of generated photons \( \langle n \rangle \) are shown in Fig. 4.2 (left), in the ultra-strong coupling regime, for \( \lambda = 0.80 \) (solid line) and \( \lambda = 0.83 \) (dashed line): it is clearly visible that the number of photons remains quite small in the initial off-resonance regime \( t < \tau \), while \( \langle n \rangle \) grows rapidly after switching to the resonance condition (at time \( t = \tau = 4\tau_s \)). It can be clearly seen that at resonance \( \langle n \rangle \) does not grow indefinitely, but oscillates due to coherent generation (dynamical Casimir effect) and destruction of photons (anti-dynamical Casimir effect [226–228]). Finally, after the switch to the off-resonance regime (at time \( t = T - \tau = 16\tau_s \)), the average photon number keeps on oscillating around its value at \( t = T - \tau \) with smaller amplitude oscillations compared to the ones on resonance. In Fig. 4.2 (right) the results of this protocol are summarized, reporting the value of the figure of merit \( f \) given by Eq. (4.19) versus \( \lambda \). Note that this protocol reveals a strong sensitivity with respect to variations of the system parameters, especially in the regime \( \lambda \gtrsim 0.6 \) where a non negligible number of photons in the cavity are generated. Indeed, a slight change of the coupling strength \( \lambda \) induces a strong variation of the mean number of photons: For instance, in the cases reported in Fig. 4.2 (left) we obtain \( f \approx 0.37 \) for \( \lambda = 0.80 \) whereas \( f \approx 0.016 \) for \( \lambda = 0.83 \). Such a strong sensitivity suggests
that accurate control of the system parameters is needed for reliable implementation of quantum protocols in the ultrastrong coupling regime.

**Figure 4.2:** Results for the on-off protocol. Left: Dynamical evolution of the mean number of photons $\langle n \rangle$ at interaction strength $\lambda = 0.80$ (solid green line) and $\lambda = 0.83$ (dashed green line). Right: Figure of merit $f$ as a function of the coupling strength $\lambda$. For both figures, the initial and final off-resonance periods, as indicated by the vertical black lines, are set to $\tau = 4\tau_s$.

### 4.3.4.2 Optimized strategies

In order to amplify and improve the visibility of the DCE, more sophisticated pulses with multiple on-off resonance sweeps are required. Hereafter we will implement and compare two methods, a bang-bang strategy [70] and optimal control using dCRAB [95; 96; 103].

For the bang-bang strategy, the qubit is placed in-out resonance with instantaneous sweeps. The main idea of this strategy is pretty straightforward: the duration of the on-resonance intervals ($\omega_q = \omega$) is determined by the time needed to reach the first maximum in the number of photons, while the detuned intervals ($\omega_q = 4\omega$) last for a fixed time $\tau_O \ll \tau_p$ since gain in $f$ is mainly seen on resonance. This strategy already provides an improvement with respect to the on-off protocol (data not shown), however we are going one step further and employ a more sophisticated strategy, which makes use of multiple iterations steps, to obtain best possible on-resonance time interval lengths (for details see Appendix A).

The second approach to target the objective $f$ is by means of the dCRAB optimal control technique [103]. The two main ingredients of dCRAB are an expansion of the control functions over a randomized truncated basis, and iterative re-initialization of local searches allowing the algorithm to escape from false traps [95; 96; 103]: The basis functions, which define the subspace subject to search, are updated at each
re-initialization step and the new emerging search directions are explored by means of gradient-free minimization algorithms.

Figure 4.3: Expected photon number $\langle n \rangle$ versus time for bang-bang strategy (left, blue solid lines) and dCRAB (right, red solid lines) for coupling strengths $\lambda = 0.17$ (a,c) and $\lambda = 0.83$ (b,d). On the secondary y-axis, the control pulses (dashed lines, colors respectively) are displayed. Note that there is an initial and final time interval equally out of resonance for both strategies, used to compute the figure of merit $f$ (see Eq. (4.19)).

Two representative results for coupling strengths $\lambda = 0.17$ ($\lambda = 0.83$) are presented in Fig. 4.3a,c (Fig. 4.3b,d) comparing the dCRAB solution to the outcome of the bang-bang strategy: the expectation values of the photon number are reported along with the generating control pulses $\omega_q(t)/\omega$. From these exemplary cases, the working principle of the bang-bang strategy is clearly revealed: Once the photon number peaks to a new maximum, the qubit frequency is driven out of resonance to let the system settle for some pre-defined time until it is switched back on resonance (see, e.g., Fig. 4.3b). On the contrary, for the dCRAB solution the interplay between $\langle n \rangle$ and the control input is no longer apparent. However, it can be seen that, after some time, the number of photons generated via the dCRAB pulse always exceeds the bang-bang photons number (this is true for all cases tested). We like to stress the fact that the bang-bang strategy
is a heuristic while for the dCRAB algorithm optimality can be reached under some assumptions [103; 115; 160]. Notice that the non-negativity condition $(\omega_q(t) \geq 0)$ for the pulses is imposed at all times, preventing a swap in the computational basis of the qubit.

The results for different coupling strengths $\lambda$ are summarized in Fig. 4.4, where the protocol’s performances are clearly visible. Overall, in the ultrastrong coupling regime the improvement from the dCRAB optimizations over the bang-bang strategy is of up to one order of magnitude; compared to the unoptimized single on-off resonance protocol, the optimal solutions yield up to two orders of magnitudes more photons. We also tested the regime with coupling $\lambda \ll 1$: For $\lambda = 0.03$, dCRAB leads to a figure of merit $f \approx 12$, whereas the other two strategies perform very poorly ($f$ well below 0.2). Quite interestingly, the dCRAB strategy shows that amplification of the DCE with a significant photon production is also possible for $\lambda \ll 1$, at the price of a longer duration of the protocol, $T = 20\tau_\lambda \propto 1/\lambda$.

for the same model, assuming some specific resonance frequency. Although that analysis focuses only in the weak coupling regime,

![Figure 4.4:](image)

**Figure 4.4:** Created photon number $f$ versus coupling strength $\lambda$ after application of three different control strategies: dCRAB optimization (red circles), bang-bang strategy (blue squares) and unoptimized on/off resonance sweeps (green diamonds). Note that the dCRAB optimized solutions improve significantly compared to the other two strategies and hence pave the way towards optimal amplification of the DCE.

Finally, we point out that the optimal pulses identified here are expected to work equally well in a (reasonable) noisy environment, due to their intrinsic robustness
4.3 Amplification of the parametric dynamical Casimir effect via optimal control

against small variations, as it has been already theoretically and experimentally showed in many different scenarios [52; 80; 82; 115; 218; 229]. Moreover, if closed loop optimal control is employed, the optimization incorporates unknown and unpredictable drifts into the pulse design as well as makes the pulses robust against statistical disturbances (noise on the pulses and the figure of merit) [49; 52; 82; 122; 127; 230]. Indeed, we could confirm the robustness of the optimal strategies by numerical simulations of the system evolution steered by the optimized dCRAB pulse (from Fig. 3d) and additionally affected by either systematic or statistical errors. In the former case, we assume the presence of a systematic error in the coupling strength compared to one specific coupling strength that was used in the optimization. For the analysis shown in Fig. 4.5 (left), we choose $\lambda = 0.83$ as the reference strength for the optimization and show the outcome $f$ for different values of $\lambda \in (0.80, 0.86)$. We can conclude that for the tested range, the figure of merit $f$ still remains reasonably good (more than $2/3$ of the optimized $f$), and anyway much bigger than what could be obtained with the other strategies. Moreover, notice that the region of this particular coupling strength is highly sensitive in terms of photon generation as we could see for the on-off protocol in Fig. 4.2: For instance, going from $\lambda = 0.80$ to $\lambda = 0.83$, means a drop in $f$ by a factor of more than 20 for the unoptimized protocol.

Finally, we analyze the scenario where the optimal pulse $\omega_q(t)$ is affected by random noise $\xi(t)$ uniformly distributed in the interval $[-\frac{\delta\omega}{\omega}, \frac{\delta\omega}{\omega}]$. In Fig. 4.5 (right), we can see that the figure of merit $f$ (averaged over 100 noise realizations) is very stable in the range $\delta\omega/\omega = [0, 0.4]$ and was fitted to scale as $f(\delta\omega) = (9.62\pm 0.00494) - (0.174\pm 0.0633) \cdot \delta\omega^2$. The correlation time of the noise was set to be $\tau_c = 0.03 T$, however we checked over about three orders of magnitude that it does have an almost negligible effect only (data not shown).

4.3.5 Conclusions

We have applied different optimization strategies to a driven Rabi model Hamiltonian in the ultrastrong coupling regime. In particular, we have optimized the number of photons generated by the dynamical Casimir effect, in order to enhance its visibility in view of possible future experimental verifications. The huge amplification of the DCE given by the dCRAB optimization is quantified by an enhancement of the generated photons of up to about an order of magnitude with respect to a reference bang-bang strategy and of up to about two orders of magnitude with respect to the unoptimized on-off strategy.

We stress that our results are fully consistent with the general framework of the DCE
Figure 4.5: Figure of merit for the dCRAB pulse computed for the unperturbed $\lambda = 0.83$ case, when the system is exposed either to a systematic (left) or statistical (right) errors. In the former case, the actual value of $\lambda$ set for the propagation ranges in $(0.80, 0.86)$. In the latter, the qubit frequency $\omega_q(t)$ is affected by random noise uniformly distributed in the interval $[-\frac{\delta \omega_q}{\omega}, \frac{\delta \omega_q}{\omega}]$, with correlation time $\tau_c \approx 0.03 \, T$. That is, $\xi(t)$ is reset every time interval $\tau_c$. Error bars denote statistical errors for 100 realizations of noise per point, while the red straight line shows the outcome from a quadratic fit.

As a consequence of non-adiabatic modulation of the system [198; 214]. Indeed, if the system is not driven, like in the on-off protocol between the times of switching on and off resonance, then the number of photons is bounded and evolves quasi-periodically (as shown in Fig. 1 top of our paper). On the other hand, when the system is appropriately (and non-adiabatically) driven, like in the bang-bang protocol with optimized times for switch-off resonance, then the number of photons grows with time. Much better results can then be obtained if the non-adiabatic control is optimized numerically, as in dCRAB strategy.

The optimization performed in this paper is a valid proof of concept. Moreover, the parameters used in our simulations are within reach for present technology, where coupling strengths even exceeding the resonator frequency have been recently reported [211; 212]. Additionally, in circuit quantum electrodynamics experiments the populations of Fock and coherent states were measured by means of the Fourier transform of the time-dependent polarization signal of a probe qubit interacting with the field [231]; extensions of this method [232] would allow not only the detection of the number of photons but also the reconstruction of the exotic field states generated by the DCE [233]. More realistic models would need a detailed treatment of decoherence sources appropriate for a given implementation [234], and of experimental imperfections, e.g. in driving the
matter-field coupling or measuring the final state of the system. Hence, given that our results have been obtained for an idealized scenario, they are an estimate of the upper bound of the strength of the DCE in the driven Rabi model. However, as we have shown that optimal solutions are robust with respect to noise sources and systematic errors, we are optimistic about the possible experimental verification of the presented results in the future.

Finally, the same approach can be used to prepare and investigate different states and phenomena in the ultrastrong coupling regime, as for example, targeting a given Fock state of the field or generating squeezed field states.

### 4.3.6 Appendix A: Optimized Bang-bang strategy

![Figure 4.6](image)

**Figure 4.6:** Working principle of the bang-bang strategy shown for an initial pulse (a) and the first iteration of the algorithm (b), for $\lambda = 0.5$ and $\tau_O/\tau_s = 1.5$. During the time on resonance of the pulse (dashed line in a)) $\langle n \rangle(t)$ oscillates around an average value (solid line in a)). Upon switching back out of resonance at the maximum of the first oscillation (the corresponding pulse is the dashed line in b)), the values of $\langle n \rangle$ remains almost frozen around this maximum (solid line in b)). Reverting back to the resonance regime leads to the oscillatory behavior of $\langle n \rangle$. The initial fast increase after each switch on resonance results in the overall population increase of the cavity.

We build optimized pulses employing bang-bang switches from an off resonance regime, where $\omega_q = 4\omega$, to the resonance regime $\omega_q = \omega$ and back out of resonance after some fixed time $\tau_O$. The pulses have a total duration of $T = 20 \cdot \tau_s$, however we reserve an initial and final $\tau$ where the pulse is kept constant to compute $\bar{n}_{f/i}$ to determine the figure of merit $f$ in Eq. (5).

To optimize the pulses the system is firstly evolved according to an initial single bang-bang pulse $\omega_q^{(0)}_{r,ref}$ (dashed line in Fig. 4.6 a)), which is iteratively improved according to the following procedure:

1) Evaluate the time evolution for the pulse $\omega_q^{(i)}_{r,ref}$ and locate the first local maximum of
photon number $\langle n_M(t_M^{(i)}) \rangle$ after the last switch to resonance in the current pulse.

2) Generate a new pulse $\omega_{q,ref}^{(i+1)}$ based on $\omega_{q,ref}^{(i)}$ by switching out of resonance at $t_M^{(i)}$ (point in time of first occurring maximum in considered time interval). Remain out of resonance for some initially specified time $\tau_O$. Finally, at $t_M^{(i)} + \tau_O$, the pulse reverts back to resonance for the remaining $t \in (t_M^{(i)} + \tau_O, T - \tau)$.

3) At $t = T - \tau$ switch back off-resonance to allow for the computation of the figure of merit $f$.

4) Iterate steps 1) - 3) and stop when $t_M^{(i)} > T - \tau$.

In Fig. 4.6 we show a typical result of the first iteration of this strategy: It can be clearly seen that after going on resonance the photon number rapidly increase and stabilities to an higher value. The introduction of an non negligible off-resonance time period $\tau_O$ leads to generally better results by letting the system equilibrate to the new conditions.
5 Optimized BEC production

5.1 Introduction into BEC condensation

In 1924/25 S. Bose and A. Einstein predicted that a homogeneous ideal gas of particles of integer spin (known as bosons today) would condensate at very low temperatures close to absolute zero [235]. It took 70 years until the phenomenon could be experimentally observed. The first Bose-Einstein condensates (BEC) were produced by E. Cornell and C. Wieman at JILA (Boulder, Colorado) in June 1995 [236] and shortly after by W. Ketterle, K. Davis and M. Mewes at MIT (Cambridge, Massachusetts) [237]. In 2011 Cornell, Wiemann and Ketterle received the Nobel Prize in Physics for this achievement. These developments massively spurred research on the field of ultracold quantum gases with nowadays more than 200 groups involved worldwide [238].

BECs are a peculiar state of matter where a macroscopic portion of particles occupy the same quantum mechanical state and hence they might be described by a single collective wave function with a common quantum phase. The spatial extension usually ranges on the order of 1-100 $\mu$m which allows for relatively easy observation.

Not only detection but also manipulation of the atoms and their confinement is very flexible as an outcome of a great degree of sophistication of experimental tools. As a result tailored Hamiltonians can be realized which makes the system a viable quantum simulator. For example, the interaction between atoms can be controlled over a large order of magnitude via Feshbach resonances [239] allowing the study of superconductivity, superfluidity and ultracold molecules. Ultracold gases have also been studied by loading atoms in an optical lattice formed by counter-propagating laser beams [240]. Changing the laser amplitudes and phases, the ratio between on-site interaction and hopping probability can be influenced in a way to probe the phase transition from a superfluid phase to a Mott insulating phase [129]. This process might be described by the widely studied Bose-Hubbard model [241].

The creation of a Bose-Einstein condensate with neutral atoms requires a delicate multi-step process. A gas of neutral atoms with temperature $T$ and mass $m$ has a corresponding de Broglie wavelength of $\lambda = (\frac{2\pi\hbar^2}{mk_BT})^{\frac{1}{2}}$, where $k_B$ is the Boltzmann
constant. The density of atoms $n$ scales with respect to the average atom distance $d$ according to $n \propto d^{-\frac{1}{3}}$. For non-interacting particles, the BEC phase transition occurs at a critical temperature $T_c$ characterized by $n \lambda^3 \approx 2.612$. The usual starting point of an ultracold quantum gas experiment is a thermally activated source of neutral atoms. To reach the transition to a BEC, the temperature needs to be lowered by nine orders of magnitude and also the atomic density has to be raised by up to six orders of magnitude. An overview of a possible phase space trajectory towards a BEC is shown in Fig. 5.1. In order to achieve this extreme cooling process experimentally, a series of subsequent cooling methods are applied. The most important steps are summarized in the following: first, laser cooling via the so called Doppler cooling (Nobel Prize 1997) is applied. The principle idea is that laser light which is detuned with respect to an atomic transition frequency is shined in through pairwise counter propagating beams. That way, due to the Doppler shift, atoms of a targeted velocity can be addressed. These atoms get decelerated with respect to the direction of the incoming laser light on absorption, but show an omni-directional re-emission distribution which causes a net zero effect because momentum transfers are averaged out over time. As a consequence, average atom velocities decrease and hence the temperature of the gas drops. After that, conservative atom trapping and evaporative cooling are applied. The latter works via removal of the hottest atoms in the cloud via subsequent lowering of the depth $U$ of the trap. The velocity of an ensemble of atoms is distributed and can be modeled for a thermalized cloud well above the BEC transition according to the Maxwell-Boltzmann distribution. When $U$ got lowered a bit, the hottest atoms can escape which means that the high energy tail of the thermal distribution gets chopped off. After that, the thermal equilibrium is re-attained by elastic collisions of the remaining (cooler) atoms.

**Figure 5.1:** Schematic representation of the various steps needed to yield Bose-Einstein condensation in the experiment. Taken from [242].
and the trap can be lowered again, repeating the procedure until eventually a BEC emerges [243].

In the work presented in the next sections, we design optimized laser and magnetic field ramps for the evaporative cooling process to enhance the remaining atom number in the emerging BEC. A schematic drawing of the involved components of the crossed dipole trap (CDT) is depicted in Fig. 5.2. The setup is very flexible in terms of realizable trap geometries: Through the possibility to translate one of the laser beams and additionally to employ a magnetic quadrupole field, narrow and wide CDT can be studied as well as a hybrid trap which is a CDT where simultaneously a magnetic field can be switched on.

The detection of the atoms at the end of the whole process is done with time-of-flight absorption imaging. As a result, an image of the atomic density distribution can be extracted using computer controlled CCD cameras. A two-dimensional fit thereof gives an estimate of the involved number of atoms which is our figure of merit.

Prior to a more evolved and automated control of the pulse shapes, our collaborating
group in Aarhus\textsuperscript{1} has made a considerable effort in localizing local maxima of the control landscape, where an eight dimensional pulse parametrization has been employed. Based on this work, we used – for the first time – our RedCRAB control suite to iteratively improve the pulse shapes. As depicted in Fig. 5.3, the control loop worked as follows: Pulses were sent through an online interface to the BEC laboratory. There, the dipole trap is loaded using profiles based on the pulses, a picture is taken and the respective estimate on the atom number forming a BEC is fed back via the web connection. Our server would then process the gained information and calculate an updated set of pulses to be evaluated in the lab. Average cycle times were ranging between 35 and 45 seconds, allowing for roughly a thousand pulse evaluations (= loops) within twelve hours.

A photography of the BEC setup (taken on the 13th of September 2016) is shown in

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image.png}
\caption{Working principle of optimal control loop. Our RedCRAB server is interfaced via the web with Aarhus university BEC lab.}
\end{figure}

Fig. 5.4. The following principal components are labeled from 1–8 and represent the following parts of the experimental setup:

1. Rubidium source.

2. Transport stage. Atoms are loaded into a magnetic trap produced by attached coils (3). Coils and trapped atoms are mechanically moved to cube chamber (5).

3. Coils for trapping and transporting the atoms.

\textsuperscript{1} Led by J. Sherson. Involved collaborators can be found in list of authors of publication reprinted in Sec. 5.2
4. Transport rail (gray-silver item).

5. Final evaporation chamber where the BEC will emerge. After the transport the atoms are transferred in a first step to the magnetic quadrupole field of the fixed coils visible.

6. Optics for vertical imaging for detecting the atoms.

7. Optics for the transverse dipole trap beam for trapping the atoms.

8. Optics for the longitudinal dipole trap beam for trapping the atoms (major part outside the image).

Figure 5.4: Setup of BEC lab in Aarhus. The Rb atoms are transported from the left to the right and processed at various stages, see text.

As already mentioned, the lab control was limited to the laser powers and magnetic field currents. For future applications, it is interesting to see how results are affected when additional steps of the cooling process are taken into account. This can either be done sequentially (optimizing one step after another) or even simultaneously where all control variables are varied.
5.2 Remote optimization of an ultra-cold atoms experiment by experts and citizen scientists


Note, that there is a significantly more comprehensive version (v1) of the manuscript on arXiv [244].

5.2.1 Abstract

We introduce a novel remote interface to control and optimize the experimental production of Bose-Einstein condensates (BECs) and find improved solutions using two distinct implementations. First, a team of theoreticians employed a Remote version of their dCRAB optimization algorithm (RedCRAB), and second a gamified interface allowed 600 citizen scientists from around the world to participate in real-time optimization. Quantitative studies of player search behavior demonstrated that they collectively engage in a combination of local and global searches. This form of multi-agent adaptive search prevents premature convergence by the explorative behavior of low-performing players while high-performing players locally refine their solutions. In addition, many successful citizen science games have relied on a problem representation that directly engaged the visual or experiential intuition of the players. Here we demonstrate that citizen scientists can also be successful in an entirely abstract problem visualization. This is encouraging because a much wider range of challenges could potentially be opened to gamification in the future.

5.2.2 Introduction

In modern scientific research, high-tech applications such as quantum computation [20] require exquisite levels of control while taking into account increasingly complex environmental interactions [245]. This necessitates continual development of optimization methodologies. The fitness landscape [246] spanned by the possible controls and their corresponding solution quality forms a unifying mathematical framework for search
problems in both natural [150; 247–251] and social sciences [252; 253]. Generally, search in the landscape can be approached with local or global optimization methods. Local solvers are efficient and analogous to greedy hill climbers. In non-convex 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. That, however, typically increases the runtime dramatically compared to the local solvers. Achieving the proper balance between local and global methods is often referred to as the exploration/exploitation trade-off in both machine learning (ML) [254] and social sciences [255].

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 [249; 256]. ML 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 [257; 258]. Thus, emphasis is currently shifting towards the development of human-machine hybrid intelligence [259; 260].

At the same time quantum technology is starting to step out of university labs into the corporate world. For the realization of real-world applications, not only must hardware be improved but also proper interfaces and software need to be developed. Examples of such interfaces are the IBM Quantum Experience [261] and Quantum in the Cloud [262], both of which give access to their quantum computing facilities and have ushered in an era in which theoreticians can experimentally test and develop their error correction models and new algorithms directly [263]. The optimal development of such interfaces, allowing the smooth transformation of human intuition or experience-based insights (heuristics) into algorithmic advances, necessitates understanding and explicit formulation of the search strategies introduced by the human expert.

The emerging field of citizen science provides a promising way to investigate and harness the unique problem-solving abilities humans possess [264]. In recent years the creativity and intuition of non-experts using gamified interfaces have enabled scientific contributions across different fields such as quantum physics [150; 265], astrophysics [266] and computational biology [267–269]. Here, citizen scientists often seem to jump across very rugged landscapes and solve non-convex optimization problems efficiently using search methodologies that are difficult to quantify and encode in a computer algorithm.

The central purpose of this paper is to combine remote experimentation and citizen science with the aim of studying quantitatively how humans search while navigating the complex control landscape of Bose-Einstein condensate (BEC) production (Figure 5.5a). Before presenting the results of this Alice Challenge [270] we characterize
the corresponding landscape using a non-trivial heuristic that we derive from novel analysis of our previous citizen science work [150]. In addition, a team of experts using a state-of-the-art optimization algorithm explores the landscape of BEC production. This allows for a comparison of algorithmic and human search strategies.

5.2.3 Initial landscape investigations

On [www.scienceathome.org](http://www.scienceathome.org), our online citizen science platform, more than 250,000 people have so far contributed to the search for novel solutions to fast, 1D single-atom transport in the computer game Quantum Moves [150]. The solutions are investigated in terms of the fitness landscape: a spatial representation of a function $J(\vec{u}) \in \mathbb{R}$ describing the quality of a solution given by the control variables $\vec{u} \in S$, where $S$ is the set of possible solutions. Clustering analyses of player solutions were found to bunch into distinct groups with clear underlying physical interpretations. We refer to such a group of related solutions as a strategy.

A complete mapping of the fitness landscape topology would involve extensive numerical work such as the reconstruction of the full Hessian [271] or, as an approximation, random sampling of solutions around these strategies combined with methods like principal component analysis. Due to the high dimensionality of the problem, these approaches are difficult to scale. Instead, we introduce here the heuristics to explore the space between the identified strategies along the high-dimensional vector connecting the two as the strategy connecting heuristic (SCH). Perhaps surprisingly, we identify a narrow path of monotonically increasing high-fidelity solutions between the two strategies, which we denote a bridge. This demonstrates that for this problem a continuum of mixed-strategy solutions with no clear physical interpretation can be mapped out if all hundreds of control variables are changed synchronously in the appropriate way.

Moreover, due to its monotonicity, a correctly set up greedy algorithm would be able to realize (at least partially) global exploration of the landscape along such a bridge (see 5.2.6.1 for details).

This leads us to ask whether other established strategies in physics are truly distinct, or if they are simply labels we attach to different points in a continuum of possible solutions due to our inability to probe the entire solution space. In the latter case, this, coupled with the human desire to create identifiable patterns, might cause us

---

1 It is important to note that the topology of the fitness landscape critically depends on the choice of parametrization or representation of $\vec{u} \in S$.

2 Note, if the identified bridge is narrow in high-dimensional space the random sampling methods will be likely to fail to identify it.
5.2 Remote optimization of an ultra-cold atoms experiment

Figure 5.5: a) Real-time remote scheme for connecting experts and citizen scientists with the laboratory. The respective remote clients (RedCRAB for the experts, Alice Challenge game client for the citizen scientists) send experimental parameters through an online cloud interface. These parameters are turned into experimental sequences and executed by the Alice control program. The number of atoms in the Bose-Einstein condensate ($N_{\text{BEC}}$) serves as a fitness value and is extracted from images of the atom cloud taken at the end of each sequence. The Alice control program closes the loop by sending the resulting $N_{\text{BEC}}$ back to the remote clients through the same cloud interface. b) Left panel: illustration of the experimental setup. The RedCRAB algorithm and the Alice Challenge players can control the magnetic field gradient depicted by the yellow shaded coils and the intensity of the two dipole beams drawn in red and blue. Right panel: screenshots of the Alice Challenge [270]. The game client features a spline editor for creating and shaping the experimental ramps.
to terminate our search before discovering the true global optimum. This premature termination of the search nicely illustrates the stopping problem [272; 273] considered in both computer and social sciences: determining criteria to stop searching when the best solution found so far is deemed “good enough”.

We now apply this methodology to the high-dimensional problem of experimental BEC production [274]. In our case an increased BEC atom number, $N_{\text{BEC}}$, will provide significantly improved initial conditions for subsequent quantum simulation experiments using optical lattices [275]. Although extensive optimization has been applied to the BEC creation problem over the past decade by employing global closed-loop optimization strategies using genetic algorithms [117; 118; 276; 277], little effort has been devoted to the characterization of the underlying landscape topology and thereby the fundamental difficulty level of the optimization problem. In the global landscape spanned by all possible controls, it is thus unknown if there is a convex optimization landscape with a single optimal strategy for BEC creation as opposed to individual distinct locally optimal strategies of varying quality (as illustrated in Figure 5.6a) or a plethora of (possibly) connected solutions (Figure 5.6b). Recent experiments [49] indicated a convex and thereby simple underlying landscape, however, this study did not explicitly optimize $N_{\text{BEC}}$ and operated within a severely restricted subspace.

In our experiment [278], we capture $^{87}\text{Rb}$ atoms in a trap made of two orthogonal, focused 1064 nanometer 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 (see Figure 5.5b for an illustration). We evaporatively cool the atoms past the phase transition to a BEC 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 $N_{\text{tot}}$ and $N_{\text{BEC}}$.

This setting allows for evaporative cooling in 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) [279]. Second, a single laser beam can be combined with a weak magnetic gradient to form a hybrid trap [280]. In both cases, the traps are initially loaded from a pure tight magnetic trap (see 5.2.6.2 for details). Conventionally, two types of geometrically differing loading schemes are pursued: loading into a large volume trap that exhibits a nearly spatially mode-matched type of loading from the initial magnetic trap into the final trap configuration [274], or a small volume trap with only a small spatial overlap with the initial trap. The latter leads to a “dimple” type loading [281] in which a smaller but colder atom cloud is produced. 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 (BEC creation strategies): a small volume, “narrow” crossed dipole trap (NCDT), a large volume, “wide” counterpart (WCDT) and similarly a hybrid (HT) and “wide” hybrid trap (WHT).

We first optimize the system by applying a simple standard experimental approach (see 5.2.6.2 for details): 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, and we find that the HT is the best performing strategy. This hints at the landscape topology sketched in Figure 5.6a. Further systematic studies would then proceed to scans of two or more parameters simultaneously. However, allowing for scans of combined parameters enables prohibitively many different 2D parameter scans. Therefore, we proceed by applying the SCH derived from the Quantum Moves investigations.

Both the low-yield NCDT configuration and the WCDT are types of crossed dipole traps but with different effective volumes and thus represent ideal candidates for first exploration. However, a simple linear interpolation of all the available parameters between the NCDT and the WCDT fails to locate a bridge. Treating the effective trap volume as an independent second parameter realizes an extended 2D-interpolation and

![Figure 5.6:](image)

(a) Illustration of the apparent global landscape topology for BEC 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 [282] representation of the landscape showing the variety of different trap configurations that are accessible in our experiment [283]. The plot contains data of the four main configurations which were scanned and optimized by 1D and 2D parameter scans. For more details, see text.
leads to the emergence of a bridge 5.2.6.2. In this case, the change of the trap depth induced by changing the trap volume has to be counterbalanced by a quadratic increase of the laser intensities involved. Thus, changing to a different representation (i.e. a particular combination of parameters) efficiently encapsulating the underlying physics yields a bridge and disproves the initial assumption that each strategy was distinct as illustrated in Figure 5.6a. To illustrate this data, we created a dimensionality-reduced visualization [282] of the parameters scans (Figure 5.6c). 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 (see 5.2.6.2, not shown in Figure 5.6c). These scans identify a novel optimum away from the four initially defined experience-based trap configurations. This demonstrates that the HT, our initial candidate for a global optimum, is not even a local optimum when appropriate parameter sets are investigated. One is therefore inclined to view the topology of the landscape as closer to what is depicted in Figure 5.6b, where the four conventional strategies are now connected with bridges and at least one other higher-yield solution exists in the full landscape.

Having established that the global optimum must be found by employing unconventional strategies, we switch to the main topic of the paper: a remotely controlled strategy employing closed-loop optimization performed by experts implementing a state-of-the-art optimization algorithm and citizen scientists operating through the Alice Challenge game interface. As detailed below, our particular implementation allows for a quantitative assessment of the citizen scientist search behavior, but only a qualitative assessment of their absolute performance. As a result, the search behavior of the two parties can only be compared qualitatively.

5.2.4 RedCRAB optimization

As mentioned above, closed-loop optimization has been explored extensively for BEC creation using random, global methods [49; 117; 118; 276; 277] and is also routinely used to tailor radio-frequency fields to control nuclear spins or shape ultra-short laser pulses to influence molecular dynamics (see [185] and references therein). In our remote-expert collaboration, we employ the dCRAB [103] algorithm which is a basis-adaptive variant of the CRAB algorithm [95]. The main idea of both algorithms is to perform local landscape explorations, using control fields consisting of a truncated expansion in a suitable random basis. This approach makes optimization tractable by limiting
the number of optimization parameters and has, at the same time, the advantage of obtaining information of the underlying landscape topology. It has been shown that the unconstrained dCRAB algorithm converges to the global maximum of an underlying convex landscape with probability one [103]. That is, despite working in a truncated space, iterative random function basis changes allow the exploration of enough different directions in the functional space to escape traps induced by the reduced explored dimensionality [116; 193]. CRAB was introduced for the theoretical optimization of complex systems in which traditional optimal control theory could not be applied [129]. In closed-loop, CRAB was applied to optimize the superfluid to Mott insulator transition [127]. Very recently, dCRAB [103] was employed to realize autonomous calibration of single spin qubits [52] and to optimize atomic beam splitter sequences [131].

Unlike the closed-loop optimization performed in the past on other experiments in which the optimization libraries were installed directly in the lab control software, our Alice remote interface allows one to implement the dCRAB algorithm remotely (RedCRAB, see Figure 5.5a). This allows the team of optimization experts to easily adjust algorithmic parameters based on the quality of preliminary optimization runs and thus exploit its full potential. Moreover, algorithmic improvements to the control suite can easily be transferred to future experiments. As in the case of the IBM Quantum Experience, we believe that this will enhance the efficiency of experimentation and lower the barrier for even wider adoption of automated optimization in different quantum science and technology aspects, from fundamental science experiments to technological and industrial developments.

After initialization, the RedCRAB works unsupervised and controls the intensity ramps of the two dipole trap beams, the ramp duration $T_{\text{ramp}}$, and a single parameter that represents the value of the magnetic field gradient during evaporation (see 5.2.6.3). A set of parameters is sent through the remote interface to the lab and realized in the experiment. The corresponding yield $N_{\text{BEC}}$ is fed back through the same interface, which closes the loop. To prevent the algorithm from becoming misled due to noise-induced outliers, each RedCRAB iteration step is the result of an adaptive averaging scheme which repeats a set of parameters up to four times depending on $N_{\text{BEC}}$ reaching certain threshold values.

As illustrated in Figure 5.7a, we achieve a new maximal solution in about one hundred iterations that exceeds the performance of the HT by more than 10%. The solution is novel in the sense that it can be seen as a type of CDT combined with the magnetic field gradient of the HT. The beam intensities are adjusted to lead to relatively similar trap depths as in the HT. However, especially in the beginning of the
5 Optimized BEC production

Figure 5.7: (a) and (b) Experts’ optimization with RedCRAB. (a) Single unsupervised optimization run. By applying an adaptive averaging scheme, the mean $N_{\text{BEC}}$ is plotted in blue as a function of RedCRAB iteration steps (see text for details). The red solid line denotes the current best $N_{\text{BEC}}$. Comparing to the level of the previously best HT configuration (black dashed line), $N_{\text{BEC}}$ was improved by 10%. (b) Histogram of relative $N_{\text{BEC}}$ changes compared to the current best solution for the RedCRAB optimization. (c)–(e) Citizen scientists’ optimization in the ATC and the ASC (f): (c) Round-based performance in the ATC. The lines show the cumulated best achieved $N_{\text{BEC}}$ for teams with three or more active players as a function of ATC iteration steps (see text for definition). Although human players had only a very limited number of tries (13 iterations), they still achieve relatively good optimization scores. Overall, all teams but one achieve $N_{\text{BEC}}$ above $10^6$. (d) Histogram of $N_{\text{BEC}}$ changes relative to the current best solution for the ATC. In contrast to the experts’ RedCRAB searches (cf. panel (b)), humans engage in many search attempts that lead to poor $N_{\text{BEC}}$. The red bar denotes all solutions which showed a relative change in $N_{\text{BEC}} > +0.9$. (e) The players’ adaptive search behavior as a function of the relative performance with respect to the team’s best $N_{\text{BEC}}$. A linear regression with a 95% confidence bound is shown in red and yields a correlation of $-0.37(4)$. The distance measure captures the difference between the player’s current and own previous solution. The $N_{\text{BEC}}$ measure captures a player’s performance relative to the team’s best. Both measures are normalized across all ATC iterations and teams. (f) Histogram for the achieved $N_{\text{BEC}}$ for all submitted solutions in the ASC. More than 73% of the submitted solutions successfully yielded a BEC.
evaporation process, the trap is relaxed much faster, leading to an overall shorter ramp. By applying SCH, a bridge connecting the HT to this novel solution could be identified (see 5.2.6.3). This illustrates that the RedCRAB algorithm is highly effective at both locating novel, non-trivial optimal solutions as well as providing topological information of the underlying landscape. It substantiates the appearance of a complex but much more connected landscape than initially anticipated.

5.2.5 Citizen science optimization

In our second approach to remote optimization, we involve citizen scientists by employing a gamified remote user interface that we call the Alice Challenge. We face the difficulty of turning the adjustment of laser and magnetic field ramps into an interactive, engaging game. Therefore, we developed a client using the cross-platform engine Unity and promoted it through our online community [www.scienceathome.org](http://www.scienceathome.org). As depicted in Figure 5.5b, the ramps are represented by three colored spline curves and are modified by adjustable control points. The total ramp duration, $T_{ramp}$, is fixed. After manipulating the splines, the player submits the solution, which is then executed on the experiment in the lab (see Figure 5.5a). The obtained $N_{\text{BEC}}$ provides performance feedback to the players and is used to rank players in a high score list. Players have the option to see, copy and adapt everyone’s previous solutions. This setup generates a collective search setting where players emulate a multi-agent genetic search algorithm. Note, that two different game modes were designed: the Alice Team Challenge (ATC) and the Alice Swarm Challenge (ASC), both of which will be introduced later.

Citizen scientists have shown that they can solve highly complex natural science challenges [150; 267–269]. However, data from previous projects suffer from the fact that they merely showed that humans solve the challenges but did not answer how a collective is able to balance local vs. global search while solving these complex problems. Social science studies in controlled lab settings have shown that individuals adapt their search based on performance feedback [284]. Specifically, if performance is improving, humans tend to make smaller changes (i.e. local search), while if performance is worsening humans tend to make larger changes (i.e. search with a global component). Therefore, experimental evidence suggests that human search strategies are neither purely local nor global [284; 285]. Furthermore, studies have also established the importance of social learning and how humans tend to copy the best or most frequent solutions [286–288], which facilitates an improved collective search performance. However, these laboratory-based studies have been constrained by the low dimensionality and artificial nature of the tasks to be solved. This raises concerns with respect to the external
validity of the results: are these general human problem-solving patterns or are they merely behaviors elicited by the artificial task environment? Finally, previous citizen science results were based on intuitive game interfaces such as the close resemblance to sloshing water in Quantum Moves. In contrast, the Alice Challenge is not based on any obvious intuition. It is therefore interesting to investigate if and how citizen scientists are able to efficiently balance local and global search when facing a real-world, rugged, non-intuitive landscape.

In order to address this question, we created a controlled setting: the ATC. Unfortunately, due to the structure of the remote participation sufficient data could not be gathered to quantitatively study both the initial search behavior and the convergence properties of the human players (see 5.2.6.4). Our previous work [150] demonstrated that the human contribution lay in roughly exploring the landscape and providing promising seeds for the subsequent, highly efficient numerical optimization. We therefore chose a design focusing on the initial explorative search of the players knowing that this would preclude any firm statements of the absolute performance of the players in terms of final atom number. Concretely, teams of five players each were formed, with every team member being allowed one submission in each of the 13 rounds (ATC iterations). After the five solutions from the active team were collected, they were run on the experiment and results provided to the players. All teams were provided initially with the same five low performing solutions and $T_{\text{ramp}}$ was fixed to 4 seconds. Each ATC iteration lasted about three minutes, and a 13 ATC iteration game lasted approximately one hour. This was chosen as the best balance between keeping teams motivated over the whole experiment with minimal dropouts and gathering enough data for analysis.

As illustrated in Figure 5.7c players showcase substantial, initial improvements across all game setups, even though the system and its response were completely unknown to them beforehand. This demonstrates that humans can indeed effectively search complex, non-intuitive solution spaces (see also Figure 5.14). In order to make sense of how citizen scientists do this, we asked some of the top players how they perceived their own gameplay. One of them explained that he tried to draw on his previous experience as a microwave engineer by applying a black-box optimization approach. Because he did not need a detailed understanding of the underlying principles of the search space, this suggests that humans might have domain generic search heuristics they rely on when solving such high dimensional problems.

Differences in the player setup and the accessible controls precludes a direct comparison of the absolute performance of the RedCRAB and the citizen scientists. However, Figures 5.7b and 5.7d demonstrate the distribution of the relative changes in $N_{\text{BEC}}$ and
clearly reveal how fundamentally different the respective search behaviors are. The local nature of the RedCRAB algorithm leads to incremental changes in either the positive or negative directions. When optimizing the system 80% of RedCRAB’s guesses correspond to changes of 20% or less in their current optimal value of $N_{\text{BEC}}$. In contrast, humans engage in many search attempts that lead to poor $N_{\text{BEC}}$. In that case, 60% of the solutions yield $N_{\text{BEC}}$ which differs by more than 20% compared to the current best.

To investigate this quantitatively, we further analyzed observations from players in the ATC ($N = 110$). Supporting previous lab studies [284; 285], our results show that players engaged in 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 the player was far behind the best solution, the player tended to engage in more substantial adjustments to their current solution (see Figure 5.7e and 5.2.6.9). Advancing previous studies, we were also able to identify that players engaged in the same type of adaptive search, when engaged in social learning; i.e. when copying a solution from someone else in the team, and subsequently manipulating it before submitting their modified solution 5.2.6.9. 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, since poorly performing individuals search more distantly, breaking free from exploitation boundaries, while individuals that are near the top perform exploitative, local search. This out-of-the-lab quantitative characterization of citizen science search behavior represents the main result of this paper.

Finally, in order to explore the absolute performance of the citizen scientists, we created an open swarm version of the game, the Alice Swarm Challenge (ASC). The client was free to download for anyone and the number of submitted solutions was unrestricted. Participants could copy and modify other solutions freely. As this setting was uncontrolled, general statements about the search behavior are not possible. In the ASC, we had roughly 500 citizen scientists spanning many countries and levels of education. The submitted solutions were queued, and an estimated process time was displayed. In this way, players could join, submit one or a set of solutions and come back at a later time to review the results. The game was open for participation for one week, 24 hours per day, with brief interruptions to resolve experimental problems. As an additional challenge, the game was restarted two to three times per day while changing $T_{\text{ramp}}$ as well as the suggested start solutions. In a total of 15 sessions, we covered a range of ramp times from 1.75 seconds to 8 seconds and all in all 7577 solutions were submitted. Without the restriction of game rounds, players were able to improve the solutions further. Figure 5.7f shows the distribution of the attained $N_{\text{BEC}}$ across all
sessions. For short ramp durations it became increasingly difficult to produce BECs with high \( N_{\text{BEC}} \). Nonetheless, the players could adapt to these changing conditions and produce optimized solutions.

The largest BEC was found for \( T_{\text{ramp}} = 4s \) and contained about \( 2.8 \times 10^6 \) atoms, which set a new record in our experiment. The solutions found by the players were qualitatively different from those found by numerical optimization. Where the RedCRAB algorithm was limited by only having control over the evaporation process and being able to apply only a single specific value for the magnetic field gradient, the players had full control over all ramps throughout the whole sequence of loading and evaporation. This was utilized to create a smoother transition from loading to evaporation. The magnetic field gradient during evaporation was initially kept at a constant value but relaxed towards the end.

In conclusion, with the advent of machine learning methods, focus on cutting edge algorithmic design is shifting to the subtle interplay between exploration and exploitation elements at which human inspiration is acknowledged to be important. In our opinion, the growing emphasis on human-computer interaction, as computer algorithms are integrated more deeply into the scientific research methodology, will challenge the clear divide between social and natural science. We see our work as an example of the growing usefulness of bridging this divide. Concretely, we have introduced a novel interface that allowed for the first remote closed-loop optimization of a BEC experiment, both with citizen scientists interacting through a gamified remote client and by connecting to numerical optimization experts. Both yielded solutions with improved performance compared to the previous best strategies. The obtained solutions were qualitatively different from well-known strategies conventionally pursued in the field. This hints at a possible continuum of efficient strategies for BEC creation. Although quantitative studies of player optimization performance were precluded by the design, it is striking that with regards to overall performance, the players seemed to be able to compete with the RedCRAB algorithm and also exhibited the ability to adapt to changes in the constraints (duration) and conditions (experimental drifts). The controlled design of the ATC yielded quantitative insight into the collective adaptive search performed by the players. This points toward a future in which the massive amounts of data on human problem solving from online citizen science games could be used as a resource for investigations of many ambitious questions in social science.
Acknowledgements

We thank J. Arlt and C. Weidner for useful comments on the manuscript. Additionally, we thank J. Arlt for valuable experimental input. SM gratefully acknowledges the support of the DFG via a Heisenberg fellowship and the TWITTER grant. The Aarhus team thanks the Danish Quantum Innovation Center, ERC, and the Carlsberg, John Templeton, and Lundbeck Foundations for financial support and National Instruments for development and testing support. Finally, we thank all Alice Challenge players for participating in this study.

5.2.6 Supporting Information

5.2.6.1 Theoretical optimization of single particle transport

In the following we briefly review the theoretical framework for gamified investigation of single atom transport in a controllable potential at the quantum speed limit (QSL) [150]. The framework is the citizen science game Quantum Moves [289] and in particular the specific level Bring Home Water (BHW). Here a graphically illustrated wavefunction of an atom in one dimension (|ψ⟩) must be collected from a static Gaussian shaped potential well (optical tweezer) and subsequently transported into the ground state within a designated target area. To realize this, the player dynamically adjusts the depth and position of a transport tweezer. The fraction of the state in the target state (the motional groundstate, |ψ_T⟩) is given by the fidelity 0 ≤ F ≤ 1 defined as F = |⟨ψ_T|ψ⟩|^2. The player must also reach this state as quickly as possible (promoted by the introduction of a time penalty in the game). Due to constraints in available resources, computer-based optimization at the QSL exhibits exponentially growing complexity [84; 150]. Solutions to this particular challenge are valuable, for example for the realization of a large scale quantum computer based on ultracold atoms in optical lattices [290] and optical tweezers [42; 291; 292].

The focus in quantum optimal control has been primarily on developing tailored local optimization algorithms like Krotov, GRAPE and CRAB [44; 87; 178; 293]. The former two methods are very efficient, since they exploit the structure of the Schrödinger equation, whereas CRAB is universally applicable since it can use a gradient-free method to reach the optimum, and furthermore it has the attractive feature of operating in a reduced sized basis. Recently, gradient-based optimization in a reduced basis has also been exploited in the GOAT and GROUP algorithms [294? ]. All these local methods are typically turned into global optimizers by restarting over a wide range of initial seeds until they give sufficiently good results. In alternative efforts, global
Figure 5.8: Visualization of the construction of a bridge between the two clans of solutions of the BHW game. The process time was set to $T = 0.19$ (for a definition of units see [150]), which is below the estimated QSL 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 representation, the dimensionality has been reduced to two. Therefore, distances on the two horizontal axes should not be seen absolute and axes labeling was 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 GOLSS 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 GOLSS scheme by moving along the local optima of the underlying landscape towards an extremum in the superlandscape.

Search methods such as Differential Evolution, CMA-ES, and reinforcement learning have been applied directly to quantum control [84; 295; 296] and very recently the local and global methods have been combined [? ].

In Ref. [150] it was shown that the optimization of player solutions from the BHW challenge outperforms such purely numerical approaches for transport durations close to the QSL. The best results were found by optimizing the player solutions using the Krotov algorithm in a hybrid Computer-Human Optimization (CHOP). In order to
compare the different solutions obtained with CHOP, a distance measure was introduced. A clustering analysis revealed that solutions fall into two distinct clusters denoted as “clans”. The solutions forming a clan all follow a similar strategy, to which one can assign a physical interpretation. One of the clans exploits dynamics reminiscent of quantum tunneling, while the members of the other clan use a classically inspired shoveling strategy [150].

Here, we ask if these clans really represent physically distinct strategies in the sense that no mixed-strategy, high-yield solutions exist. Proving that a given solution is locally optimal and thus truly distinct involves extensive numerical work in either random sampling combined with methods like principal component analysis or systematic reconstruction of the full Hessian in the surrounding high-dimensional space [271]. Given the high dimensionality of the problem, an exhaustive exploration of the whole space is impossible to realize with a reasonable amount of time and resources. Instead, we investigated the topology of the landscape spanned by linear interpolation between the individual controls of representatives of the two clans. Given the interpolation parameter $\alpha \in [0,1]$, the interpolated control is defined as

$$\vec{u}_{\text{int}}(\alpha) = \alpha \vec{u}_1 + (1 - \alpha) \vec{u}_2,$$  \hspace{1cm} (5.1)

where $\vec{u}_1$ and $\vec{u}_2$ are the controls of the two solutions interpolated between. Figure 5.8a depicts a 2D visualization of the landscape corresponding to the interpolated solutions and local random perturbations to these using the t-SNE algorithm [282] as applied in Ref. [150]. 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 good, nearly optimal solutions in the underlying optimization landscape.

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 figure 5.8b.

In Ref. [150], we introduce a distance map $D_{ij}$ for two solutions $i$ and $j$, which compares the overlap between two corresponding wavefunctions $|\psi_i(x,t)\rangle$ and $|\psi_j(x,t)\rangle$ at each time step $t$ for a given total transport time $T$:

$$D_{ij} = \frac{1}{T} \int_0^T \langle f_{ij} | f_{ij} \rangle dt,$$  \hspace{1cm} (5.2)

where $|f_{ij}(x,t)\rangle = |\psi_i(x,t)\rangle - |\psi_j(x,t)\rangle$ is the difference between the wave functions at
each position $x$.

In terms of this distance metric, the displacement of each numerically optimized solution from the initial seed is relatively small: the optimization of the points in figure 5.8a leads to nearly vertical lines in the visualized landscape. This implies that as long as the right region of the landscape is explored, very close to the non-perfect trial solution lies a better 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 figure 5.8b. Thus, the landscape is locally very rugged, but rich in optima.

Given the density of locally optimal points, we now define the superlandscape as the approximately smooth envelope function spanned by the optimal points. If $O$ is a local optimization algorithm then this landscape is the composition $\tilde{J} = J[O[\tilde{u}(t)]]$. Figure 5.8d illustrates a simple, generic superlandscape. The underlying landscape consists of a dense collection of individual peaks with smoothly varying heights.

Physically, we interpret the sharpness of the peaks and the density of optima in terms of the population along the different instantaneous eigenvalues of the problem: Due to the time-energy uncertainty relation, rapid transfer in the vicinity of the QSL can only be achieved by significant excitations, which ensure rapid relative phase evolution of the individual wave function components [170]. Towards the end of the process, the population needs to be refocused into the ground state. Any minor perturbation to a path at some instant will in general lead to a decreased fidelity. However, it can (nearly) be compensated with another carefully chosen perturbation at another time, leading to many closely spaced locally optimal solutions.

If the superlandscape can be evaluated with sufficient speed (using efficient local optimization), then we propose to perform completely deterministic global optimization using a method that we call Gradient-Optimization and Local Superlandscape Search (GOLSS). Such a hybrid local-local optimization scheme has been recently proposed in the context of trap-free landscapes in the presence of noise [65]. In contrast, GOLSS entirely eliminates the random steps normally present in global optimization and could therefore potentially offer significant speedup compared to existing methods. (See figure 5.8d for an illustration.)

Here, we use a Nelder-Mead type search [97] combined with Krotov optimization to implement GOLSS. 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 figure 5.8c. 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, which is an improvement of nearly two orders of magnitude in terms of infidelity over the best player optimized solution (CHOP) yielding $F = 0.929$. It also represents an improvement of
the previously obtained numerical estimates of the QSL for both the tunneling and shoveling clans, which were at $T=0.25$ and $T=0.20$ respectively. 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 1D-line scan identifies a bridge is an illustration of a deeper underlying principle in numerical optimization: whereas nearly all of the many possible search directions yield poor behavior (illustrated by the blue and red points in figure 5.8a), once a good heuristic encapsulating the essence of the problem is determined, low dimensional search is sufficient [297]. Previously, in Ref. [150], we constructed these search spaces explicitly using parametrizations that emerged from data analysis of large amounts of numerical solutions. Recently, a similar approach was applied to extract low-dimensional search spaces for spin-chain dynamics using ML-generated data [84]. The search along convex linear combinations of existing solutions introduced here may provide a computationally inexpensive methodology to identify promising search directions in the multi-dimensional landscape. In addition, we believe that the concept of superlandscape and local search within it will be a useful metaheuristic for finding high-quality solutions for quantum optimal control problems.

5.2.6.2 Experimental details – Parameter scans

Each trap configuration that is presented in the main text is loaded from a pre-cooled Rb atom cloud prepared in the $|F = 2, m_F = 2\rangle$ state and trapped in a magnetic quadrupole trap. At this stage, we typically have $5 \cdot 10^8$ atoms at a temperature of $\approx 30\mu K$. The experiments of the Alice Challenge start from this point. The CDT consists of two perpendicular beams which overlap in the horizontal plane. They have $1/e^2$ waists of $45\mu m$ (beam A) and $85\mu m$ (beam B), respectively. The longitudinal focus position $x_{\text{focus}}$ of beam A can be adjusted, thereby changing its effective waist at the crossing point of the beams. This beam is used to realize HT and WHT. The beams are placed with a vertical offset of around $90\mu m$ below the centre of the magnetic trap. An offset magnetic field $B_{\text{off}}$ in that direction can be used to tune this distance.

In our experiment, the control of the time-dependent light and magnetic fields during the loading and evaporation is initially limited to eight parameters. Later, for the remote-controlled experiments we will relax this restricted representation to allow for the full high dimensional quasi-continuous control only restricted by hardware limitations.

For the parameter scans, the intensity ramps $I(t)$ of the dipole trap beams are described by a function inspired by a simple model of evaporative cooling based on
Optimized BEC production

scaling laws [298]

\[ \frac{I(t)}{I_i} = \left(1 + \frac{t}{\tau}\right)^{-\beta}. \]

(5.3)

Here, \( I_i \) is the initial intensity, whereas \( \tau \) and \( \beta \) influence the shape of the ramp. The duration of the ramp is fixed by defining the ratio of initial and final intensity \( I_i/I_f \) for a given \( \tau \) and \( \beta \). For simplicity, the intensity ratio, as well as \( \tau \) and \( \beta \) are chosen to be the same for the two beams. \( I_i \), however, is an independent parameter.

For the loading process from the magnetic trap into the final trap configuration, the dipole trap beams are regulated to their individual \( I_i \) and the magnetic field gradient is lowered in three linear ramps from 130 G/cm initially to a final value \( B'_f \), which is retained throughout the evaporation. In total, this leads to eight individual optimization parameters.

The result of 1D parameter scans is shown exemplary in figure 5.9a for the case of the HT. The scans clearly reveal a peak-like structure with a set of 1D-optimal solutions. For the conventional trap configurations (NCDT, WCDT, HT, WHT) we obtain the 1D-optimized values \( N_{BEC} = (0.53(9), 1.07(5), 1.8(2), 1.1(4)) \cdot 10^6 \) [299]. The resulting duration \( T_{\text{ramp}} \) of the whole evaporation (counting from the beginning of the loading process) differed for the individual traps and reached \( T_{\text{ramp}} = (2.66, 2.97, 5.56, 6.60) \) s, respectively.

To investigate the topology of the landscape we search for interconnecting bridges by simultaneously scanning several parameters. Both the low-yield NCDT configuration and the WCDT are types of crossed dipole traps but with different effective volumes dictated by \( x_{\text{focus}} \). A simple linear interpolation of all the available parameters between the NCDT and the WCDT (cf. equation 5.1) fails to locate a bridge as illustrated in the inset of figure 5.9b. This is consistent with the BHW case treated above in which the bridge did not appear until local optimization was performed on the interpolated seeds. Since local optimization is very time consuming in the experimental case, we instead try to extend the search space slightly beyond the simple 1D case.

Treating \( x_{\text{focus}} \) independently and introducing a second interpolation parameter realizes an extended 2D-interpolation, which leads to the emergence of a bridge as shown in figure 5.9b. The necessity of a 2D scan can be interpreted in terms of the distinction Simon and Newell [297] make between problem space (the subjective search space) and the objective task environment (physical subprocesses): the chosen parametrization in terms of different laser beam and magnetic field settings versus the time dependent trap depths and shapes during the loading and evaporation processes. In this case, the change of the trap depth induced by changing the trap volume has
Figure 5.9: (a) 1D parameter scans used to find locally optimized parameters displayed in a 2D tSNE representation exemplarily shown for the hybrid trap. In this case, there is no indication for a connection to any of the other “conventional” trap configurations. (b) 2D-interpolation between NCDT and WCDT showing a connecting bridge between the two. Along the first dimension, the interpolation parameter $\alpha_{\text{focus}}$ is varied influencing $x_{\text{focus}}$. The remaining parameters are scanned synchronously along the second dimension described by the interpolation parameter $\alpha_{\text{rest}}$. The solid red line is a simple theoretical estimate of where to find the bridge within the given parameter space. The inset shows a diagonal cut through the landscape and illustrates, a simple linear interpolation with all parameters fails to find a bridge.
to be counterbalanced by a quadratic increase of the laser intensities involved. The solid red line in figure 5.9b marks the position, which yields the same trap depth at the beginning of the evaporation stage.

Thus, changing to a different representation (i.e. a certain combination of parameters) efficiently encapsulating the underlying physics yields a bridge and disproves the local character of the solution strategies involved. We stress that although the bridge was located in a seemingly simple 2D scan, the heuristics introduced for the BHW case of identifying the multi-dimensional search direction between established strategies was crucial.

Comparing all other pairs of conventional strategies, the choice of parameter combinations for extended interpolation is much harder to motivate physically and simple 1D- and 2D-interpolation fail to locate bridges.

![Figure 5.10: Parameter scan in three dimensions from the NCDT to the HT. The normalized interpolation parameters are denoted $\alpha_1, \alpha_2, \alpha_3$. Each frame represents a scan point in the third dimension ($\alpha_3$). In frame $\alpha_3 = 0.75$, a trap configuration with $N_{\text{BEC}}$ larger than for the HT is found. This shows that the HT is not a local optimum which is in contrast to the initial indications (see main text).](image)

As described in the main text, parameter scans were not only performed in 1D or 2D. In figure 5.10 the extension of the scan space in the third dimension is presented for parameter sets linearly interpolated between the NCDT and HT. An optimum that moves from frame to frame is revealed and a parameter set is found yielding a slightly higher $N_{\text{BEC}}$ than in the HT is found. The scan disproves the local character of the HT which was implied by the 1D parameter scans.

5.2.6.3 Experimental details – Remote optimization

For this investigation the search space is restricted to the domain between HT and NCDT by fixing the effective volume of our CDT. Likewise, the vertical offset magnetic field is fixed to a value compensating the residual background magnetic fields and is not changed during optimization. For both the remote optimization with RedCRAB
and the Alice Challenge, non-optimized configurations were chosen as starting points of the optimization runs. The RedCRAB starting point was close to the HT configuration. In the Alice Challenge at each start of a new round, the high score list was emptied and filled with low quality solutions yielding typically $N_{\text{BEC}} \approx (1 - 2) \cdot 10^5$ atoms.

In contrast to the previous parametrization of the shape of the laser ramps, RedCRAB (based on the dCRAB algorithm [103]) focuses on a finite set of relevant basis functions that make up a sufficiently good ramp. Here, each of the ramps is composed of a Fourier basis up to the 5th harmonics in units of $2\pi/T_{\text{ramp}}$, where $T_{\text{ramp}}$ is the total ramp duration. $T_{\text{ramp}}$ itself as well as $B_f'$ during evaporation are chosen to be subject to optimization. The loading procedure of a certain trap configuration is the same predefined sequence described for the parameter scans above. To overcome shot-to-shot fluctuations and thus resulting in an optimization driven and influenced by noise, an adaptive averaging scheme is applied with a stepwise increasing number of averages for higher yields in $N_{\text{BEC}}$. Outliers to high $N_{\text{BEC}}$ are in this way re-evaluated. However, we still keep the number of time-consuming evaluations low at early stages of the optimization which decreases the overall convergence time.

The adaptive averaging scheme is implemented the following way: The current best $N_{\text{BEC}}$ is denoted by $N_{\text{rec}}$. Based on this value, a set of threshold values (thv) are defined via \{thv1,thv2,thv3,thv4\} = \{0.9N_{\text{rec}}, 0.96N_{\text{rec}}, 0.99N_{\text{rec}}, 1.01N_{\text{rec}}\}. Say the experimental apparatus will return the value $N_{\text{trial}}$ for a newly evaluated set of pulses. This value will have to go through the cascade of threshold values \{thv1,thv2,thv3,thv4\} before being chosen as the new current record $N_{\text{rec}}$. At first, $N_{\text{trial}}$ is compared to thv1. Only if it exceeds thv1, the same pulse will be re-evaluated yielding $N_{\text{trial}} = 0.5(N_{\text{trial}}^{-1} + N_{\text{trial}}^i)$. As a second step, $N_{\text{trial}}$ is compared to thv2. Again, only if it exceeds thv2, the very same pulse is re-evaluated for a third time. This procedure is repeated as long as $N_{\text{trial}}$ succeeds to jump over the respective thvn value until eventually $N_{\text{rec}}$ will be updated to $N_{\text{trial}}$ after 5 successful re-evaluations.

The optimized solution that was found with RedCRAB is a mixture between NCDT and HT. $B_f'$ resembles closely the one of HT, however 25% less intensity is used in beam A. This is partially compensated by adding beam B. This results in a trap depth which is about 15% lower compared to the HT at the beginning of the evaporation. The lowering rate of the trap is comparable in the first part of the evaporation and drops at the end below the one of the HT. At the same time the calculated geometric mean of the trap frequencies ($\bar{\omega}$) is higher. In both cases, the evaporation ends at a similar trap depth and similar $\bar{\omega}$. The total ramp duration is with $T_{\text{ramp}} = 4.92$ s shorter than the one of HT.

In the process of reaching this solution, RedCRAB identified of the order of 10
intermediate improved solutions (see steps of the current best $N_{\text{BEC}}$ depicted as red solid line in main text figure 5.7a). If the underlying landscape is sufficiently smooth, one might expect to be able to locate a bridge between the standard strategies and the novel optimum by a linear or possibly non-linear combination of these intermediate solutions. As illustrated in figure 5.11a, in which step-wise linear interpolation between these intermediate solutions is performed, this is nearly but not exactly the case. There are small intervals of decreasing yield. However, a direct linear interpolation between the HT and the novel optimum yields a monotonically increasing bridge (see figure 5.11b). Thus, the non-monotonicity of the stepwise interpolation is not likely to be caused by ruggedness of the underlying landscape. Rather, a more natural explanation would be that since the local simplex-optimization component of dCRAB is not purely gradient-based it does not necessarily have its axes oriented along the maximal slope and will have a tendency to find a slightly wiggly path towards the optimum. This increases the chance that experimental noise will occasionally cause the algorithm to find false search directions from which it is slowly recovering in the following iterations. Figure 5.11c gives a graphical visualization of this phenomenon.

**Figure 5.11:** (a) Stepwise, linear interpolation from the HT (first vertical dashed line) through the starting point of RedCRAB optimization (second vertical dashed line) to its optimum (last vertical dashed line). That is, the ramps between each optimization step (remaining vertical dashed lines) are interpolated. $\alpha'$ is a generalized parameter incorporating the individual interpolations. (b) Direct linear interpolation between the ramps for the HT and the found optimum of RedCRAB optimization (figure 5.7a in main text). $\alpha$ is the interpolation parameter. The error bars represent the standard error for five repetitions. (c) A possible way to explain the different results of (a) and (b) on an exemplified optimization landscape spanned by parameters $\alpha_1$ and $\alpha_2$. In the presence of noise, an optimizer like the RedCRAB algorithm could follow the non-ideal route towards the optimum marked by the orange triangles with error bars. However, following that trace by piece-wise interpolation (blue triangles, corresponding to data in (a)), renders the topology of the landscape visible and results in a non-monotonic trace. Direct linear interpolation (red circles, corresponding to data in (b)), in this case, exhibits continually increasing yields.
5.2.6.4 Design considerations of the Alice Challenge

Around the same time as the development of the remote interface for establishing a connection with RedCRAB, a simplified remote client with a graphical user interface was developed. It allows one to control the dipole beam intensities and the magnetic field gradient of the experiment via piecewise defined functions. Tests with an undergraduate student and a collaborator situated in the UK were successful. Both were allowed to optimize evaporation sequences via this client independently. Improved solutions were found which lead to the idea of gamifying the task of optimizing the evaporation process and give “non-experts” real-time access to our experiment.

The development of the Alice Challenge began. In the design phase, it was very unclear if large groups of citizen scientists could be recruited given the non-intuitive and relatively low level of gamification compared to previous games. To make recruitment more plausible we decided to run the challenge in a well-defined and relatively short time interval to ensure that we would have a fairly high number of simultaneous users at all times. A prototype version of the game was developed and presented at the National Instruments NIWeek 2016 in Austin, Texas. The event was used for a test-run of the game interface and to acquire a broad potential user base for the actual Alice Challenge.

Given the cycle time of the experiment, the relatively short duration of the Alice Challenge put severe restrictions on the total amount of data that could be collected. Due to the restricted amount of data available, we had to make two crucial choices. First, we had to decide between the controlled setting of teams optimizing in parallel or the full uncontrolled free access of every player to all previous optimization results. The former would allow for systematic investigation of the initial search behavior at the cost of obtaining little or no information about the convergence properties of the amateur players because each team would not have enough iterations to converge.

Since participants in the ATC are not confined to a lab setting nor paid, dropouts can create a major obstacle for such team-based experiments. In abstract terms, there is a $(1 - E)^N$ probability that all team members will finish all rounds, where $E$ is the dropout probability and $N$ is the number of people in the team. The bigger the $N$ and the more rounds, the bigger the dropout. Fewer and larger teams would also lead to fewer independent optimization runs and would therefore severely limit the statistical power of the quantitative characterization of player search behavior. This led us to settle on teams of five. Likewise, we settled on a relatively short amount of 13 rounds. Every team member was allowed one submission in each of the 13 rounds via the remote interface (see figure 5.5b in the main text). We explicitly instructed
participants that they would be working together in a team, but that collaboration was possible only through the visibility of team members’ solutions and scores. After the five solutions from the active team were collected, they were run on the experiment and results provided to the players. Each round lasted about 180 seconds. Therefore, a 13 round game lasted approximately one hour in total. This prioritization is in line with the findings from our previous citizen science work: the strength of the algorithms is to optimize a given seed systematically, whereas the contribution of the players was to provide good seeds by a more global rough search of the landscape.

To further increase engagement of the participants, we emphasized the importance of finishing the game as well as increasing the resilience of the team experience by replacing dropouts with bots. Following a recruitment campaign based on snow-balling, 142 participants from around the world committed to taking part in the experiment. Participants selected up to 10 one-hour slots during the week of the challenge. Once the recruitment campaign was over, they were randomly assigned to teams, while maximizing the number of complete teams. In incomplete teams, the slots of the missing players were taken by computational agents who would simply reshuffle existing solutions. One game session had to be excluded from the dataset, because the experiment setup had drifted significantly and thus the evaluation function had been corrupted. Overall, due to no-shows, dropouts, as well as the corrupted session, we analyzed observations from 110 players out of the original 142. All participants gave explicit consent to participate in the study, which was approved by an IRB at Aarhus University.

We are aware that, whereas this design optimizes statistical power to investigate the initial explorative behavior of the players, it unfortunately did not leave much room for investigating how well the players performed in absolute terms. To slightly compensate for this, we also inserted a short, uncontrolled “Swarm Challenge” in which players could copy and optimize freely. We do not have the statistical basis to make general statements about these results and we therefore do not want to place too high emphasis on the obtained total atom numbers or the relative merit on the final convergence properties of players versus experts.

5.2.6.5 Experimental details – The Alice Challenge

As described in the main text, the players control the loading sequence as well as the evaporation process through the game interface. In order to account for the high initial and low final parameter values of laser beam intensities and magnetic field gradients, the displayed ramps are represented on a logarithmic and normalized scale. In the Alice Swarm Challenge submitted solutions are placed in a waiting queue. Depending on
the length of the queue, an estimated process time is displayed. This allows players to join, submit a single or multiple solutions and come back at a later time to review the achieved score. All results are placed on a high score list and the players have the possibility to investigate and copy corresponding solutions completely or in parts. This facilitates reproducing working solutions and encourages the players to improve them further.

5.2.6.6 Analysis of solutions – The Alice Swarm Challenge

Due to the problem representation, player solutions feature, in general, a much smoother transition from loading to evaporation than the RedCRAB solutions or the parameter scans with a fixed loading sequence. Knowing the levitation gradient of Rb in $|F = 2, m_F = 2\rangle$, one can estimate when the loading of a given trap configuration is finished. In these terms, the loading in the case of the best performing player solutions happens within about 1s. This is about twice as fast compared to the standard loading sequence described above. Afterwards, the magnetic field gradient is lowered only very slowly and remains nearly constant just below the levitation gradient. This value is about 70% higher compared to the HT or best performing RedCRAB solution. Only in the last second of the sequence is it relaxed to a value similar as in the HT. The intensities of the dipole beams are lowered after 1s which is another indicator for the transition from loading to evaporation. Compared to the HT, extremely low intensities are reached at the end of the sequence and it seems that current experimental conditions, such as beam overlap and alignment, are optimally employed.

As explained in the main text, the Alice Swarm Challenge was restarted multiple times, thereby varying $T_{\text{ramp}}$. Investigating the quality of solutions as a function of pre-set $T_{\text{ramp}}$, we found that for ramps below 3s, $N_{\text{BEC}}$ decreases drastically and no solutions were found for $T_{\text{ramp}} < 1.75s$. The shape of the ramps for long and short durations are considerably different. For instance, the initial beam intensities are much higher for the short-duration player solution. This leads to a stronger confining trap accommodating the shorter ramp duration.

As a concrete example of the difference between obtained solutions, we compare the duration-robustness of the RedCRAB solution and player solutions obtained at short and long durations ($T_{\text{ramp}} = 1.75s$ and 4s, respectively) by stretching and compressing the solutions in time. Figure 5.12 demonstrates that solutions obtained at different durations exhibit different behavior. The features that are optimal for short ramps seem to be suboptimal for longer durations. Additionally, the scan reveals that the long-duration player solution is more robust than the RedCRAB solution against reductions
Figure 5.12: Sweep of the ramp duration, $T_{\text{ramp}}$, for different optimum solutions. The ramp shapes yielded through the RedCRAB optimization and from two Alice Swarm Challenge sessions were scanned as a function of total ramp duration and $N_{\text{BEC}}$ measured. As Alice Swarm Challenge solutions, the ramps resulting in the largest $N_{\text{BEC}}$ and the solution for the shortest set ramp duration were chosen. The data points are obtained by averaging over five repetitions, where the error bars represent the standard deviation. The big squared marks denote $T_{\text{ramp}}$ during optimization. Note, that the RedCRAB algorithm’s control was restricted compared to the one of the players. For details, see text.

5.2.6.7 Individual and collective problem-solving – The Alice Team Challenge

The fact that natural science challenges [150; 266–269; 300] can be solved efficiently by the general public is of interest to cognitive and social scientists as a source of insight into the general process of human problem-solving. However, data from these projects suffer from the fact that they were not gathered with particular social and cognitive science research questions in mind. Here, we set out a new kind of social science approach, where we investigate how a collective of citizens is able to balance local vs. global search in a real-world setting. Such insight is important for future designs of
large-scale citizen science investigations of natural science problems and advance general understanding of the process of individual and collective problem-solving.

On an individual level, experiments in the lab have shown that individuals adapt their search based on performance feedback [284] and their search strategies are thus not merely local, nor global [284; 285]. Specifically, if performance is improving, humans tend to make smaller changes (i.e. local search), while if performance is worsening humans tend to make larger changes (i.e. search with a global component). Studies in the field of cultural evolution have established that some of the most efficient social learning strategies are to copy the best or most frequent solutions and copy when the situation is uncertain [286–288]. Due to these social learning strategies, collective search is acknowledged to be effective at boosting the efficiency of a search process [301].

However, potentially constrained by the low dimensionality of the tasks to be solved, with few exceptions [302] prior studies on human problem solving have primarily focused on simplified situations where individuals have the option to either copy another solution or not [303]. As argued in [304], research should not only study option selection, but option generation where participants are not constrained by relatively few options, but are allowed to integrate and transform individual and social search information. This enables analysis of how much individuals are influenced, rather than merely if they are influenced.

Another key characteristic of previous studies is that they have primarily relied on a particular type of students at particular universities in a lab setting [305]. Furthermore, this work has relied on these participants solving artificially designed tasks. Researchers have usually designed the problem task to be solved, and therefore also the nature of the landscape to search and what constitutes a good solution (see e.g. [253; 284–287; 302; 306]). Our experimental framework points towards a possible solution to these challenges, since we investigate how citizen scientists, engaged in a real world, high-dimensional and mathematically well-defined problem, adapt their search strategies to performance feedback and inspiration from solutions of other players. Finally, while individuals’ ability to do adaptive search may make them uniquely suitable for navigating rugged fitness landscapes as evidenced in previous work [150; 307; 308], we do not know how this adaptive search mechanism plays out in a collective search environment.

5.2.6.8 Experimental treatments – The Alice Team Challenge

We randomly allocated participants into two different experimental conditions, in order to further investigate the collective search process. In the treatment condition, information regarding how often a certain solution was copied from the previous round
by team-members was available to participants, while in the control such information was not available. We conjectured that by presenting participants with this information, we could test the occurrence of explicit metacognitive social learning strategies [309]. More specifically, we hypothesized that because participants could see how many of their team’s solutions had been generated by social learning (as opposed to individual search not involving copying) they could compensate, among others, for under-reliance on social learning and copy more in a given round. In this way we take the first step to study if a collective of human searchers can function as an adaptive global search algorithm, gradually changing their recombination intensity according to their performance as well as the meta-information received. Figure 5.13c demonstrate the results. Participants in the condition that were exposed to meta-information about how often a solution had been copied in their team, engaged in more social learning than participants in the condition where this information was not available. This shows that it is possible, via a simple manipulation, to nudge the human players into relying more on social learning strategies, specifically “copy the best” (see also table 5.1). Considering that previous work [306; 310] argues that human solvers rely too little on social learning, and thus that an increase in relying on social learning strategies is desirable, this is a promising result. In the following we provide an analysis of the aggregated search behavior, in order to establish if and how a collective of humans are able to solve such a high-dimensional problem.

5.2.6.9 Analysis of individual and collective search – The Alice Team Challenge

Almost all teams and individuals managed to submit relatively good solutions above the 1 million threshold (see figure 5.7c in main text and figure 5.14). Similar to Ref. [302], we find that 53% of all moves were individual, 41% involved some form of social learning and the remaining 5% were random, i.e. randomly recombining existing solutions (see table 5.1). As outlined in the main text, we test and are able to show that players adapt their search based on the performance feedback they receive (see figure 5.7e where feedback and distance measures are ranked across all players and rounds). To further support this analysis, we normalize all scores in the following modelling efforts. We analyze the data using generalized linear mixed models with a Gaussian (or Binary with a logit function, where appropriate) error structure, and we control for individual variance by allowing for a random subject effect. This approach allows us to estimate a generalized model of adaptive search where individual heterogeneity and the repeated nature of the measures are taken into account. Models are constructed by forward
5.2 Remote optimization of an ultra-cold atoms experiment

![Figure 5.13:](image)

Figure 5.13: (a) and (b) How much players edit their own solution compared with relative performance in the previous round. (a) only includes submissions that did not involve any kind of copying, whereas (b) shows the data that involved copying another solution. In both cases, the solutions of players that performed well relative to the team are changed less than players who did not perform well. Both the distance and score measures are ranked within each round with the current team-best score as a reference point. A 95% confidence interval is shown. The distance measure is based on distance from the players own previous solution (a) and on distance from copied solution (b). (c) Participants in the condition where the meta-information of seeing the number of copied solutions in their team in the previous round was available (red bars), used the “copy the best” social learning strategy more than participants in the condition where this information was not visible (blue bars). A two-sample t-test showed the two to be statistically different (p<0.0001). (d) Variation of exploration in time. The measure for exploration was derived relative to the entire solution space covered by the players in the ATC. By computing distances from any two solutions submitted in the challenge we have obtained an average distance step of ∼ 6.13.

inclusion and reported effects are within a 95% confidence interval.

We model the distance from the players’ latest solutions as a function of the feedback players received in the form of a score. In this way we track how players responded to information about the underlying landscape. We find players are more likely to make minor changes to their solutions when they achieve a performance comparable to the team-best versus more significant changes when their performance is low comparative to team best feedback (CI: (-1.61: -1.21), p<0.0001). This supports former findings on individual adaptive search [284; 285]. Additionally, players tend to become more conservative as the rounds progress (CI: (-0.12:-0.06), p<0.0001), see also figure 5.13d).
Our collective setup also allows an investigation of if and how players engage in collective adaptive search, i.e. if they not only modify their own solutions, but also actively copy the solutions of other players (social learning). First, we establish that whether players engage in social learning or not depends on previous performance, where low performance leads to a higher likelihood of engaging in social learning (CI: (-0.92: -0.41), p<0.0001), see also [253; 287; 288; 302]. We also studied social learning in terms of which peer solutions players copy. This behavior follows a similar adaptive mechanism: players will tend to copy more dissimilar solutions, provided they had just experienced low performance (CI:(-1.35:-0.36), p<0.0001). Conversely, with low score differences, players are more likely to copy solutions similar to their own.

Furthermore, advancing previous studies, our setup allows us to investigate how players manipulate solutions after engaging in social learning (i.e. copying another solution), but before submitting their solution. As in pure individual search, players behave adaptively in this social situation: if they performed better than the previous team-best, their submitted solution will tend to stay closer to the copied solution (CI: (-2.33: -1.34), p<0.0001). Conversely, if their past score is below the same benchmark, their submitted solution will tend to drift further away. Interestingly, a comparison of the rank-based slopes in the graphs illustrating individual and social search shows that the adaptive effect performance induces in subsequent search, appears stronger for individual search than search that involved social learning (see figure 5.13a and Fig. 5.13b for details).

Overall, we created a novel, online gamified interface connecting a real-time physics experiment to citizen scientists. The setup provided a unique opportunity to both measure how much a collective of players change their solutions but also have an external measurement of the quality of the solution in the solution space. This enables one going beyond merely claiming human superiority [150; 267] and study how human problem solvers are efficient at balancing the trade-off between global and local search.

We show how individuals search adaptively depending on their own former performance, thus supporting lab-based studies based on artificial, low-dimensional problems [284; 285], while simultaneously expanding this adaptive mechanism to the realm of social learning. Even though the nature of the adaptive search mechanism is the same for both individual and social search, we find exploratory evidence for social search inducing less conservativeness for high performers. Finally, our innovative experimental game setup allows a genetic algorithm inspired opportunity to recombine and manipulate existing solutions, going beyond a simple imitation option in each round that simpler setups were constrained by [301; 306]. When searching for an optimal solution in a complex, high-dimensional problem space, our exploratory investigation shows that
humans don’t indiscriminately copy other solutions. They often only copy part of the solution and then further transform the copied solution in an adaptive manner. The fact that these individual and social adaptive search mechanisms systematically depend on the individual searchers’ relative performance creates a diverse mixture of search within a collective, shaping the collective balance of local vs. global search and when the collective stops searching.

5.2.6.10 Variables used in the analysis of individual and social search – The Alice Team Challenge

Feedback is defined as the ratio between the individual’s previous score and the best team score recorded so far. For individual adaptive search, a similar analysis was conducted using different benchmarks for feedback (i.e. either individual best or second to last submission). These models yielded qualitatively similar results. In the social adaptive learning situation only the social team-best score was considered to be a relevant benchmark. This operationalization follows previous studies that show individuals benchmark their performance against the best performance so far [284; 285].

The search distance is given by the Euclidean distance between consecutively submitted solutions, in the case of individual adaptive search. For social adaptive search, the similarity variable refers to the distance between the copied solution and the solution submitted in the previous time step by the player. The shorter the distance between two solutions, the more similar they are. In figure 5.7e of the main text, figure 5.13a and figure 5.13b feedback and distance measures are rank-based across all players and rounds, while scores are normalized in all modelling efforts. For the reported analysis, the feedback scores are normalized by dividing the individual’s current score with the team best so far in the game, leading to corresponding numbers between 0 and 1. Results are robust to varying modelling assumptions, such as standardizing the data, controlling for individual heterogeneity within round or taking round (time) as a fixed effect.

5.2.6.11 Characteristics of team-participants – The Alice Team Challenge

At the end of 13 rounds, participants were redirected and asked to fill in a brief survey. This allowed us to collect a number of demographic variables about the players. We had a response rate of 80% (89 respondents). The majority of our players were male (69%), with an average age of 30.1 (st.d = 10.12). With respect to education, 66% of the participants had obtained a higher education degree, and a little over half (53%) had physics as a subject in their education, after high school. Respondents were from
17 different countries, the majority being from Europe or North America.

![Histogram of individual highest achieved scores in the ATC. Overall, despite a very restricted number of tries, human players achieve relatively good scores. This histogram illustrates that only a handful of players have very low scores, while most players achieve scores above the 1 million threshold.](image)

**Figure 5.14:** Histogram of individual highest achieved scores in the ATC. Overall, despite a very restricted number of tries, human players achieve relatively good scores. This histogram illustrates that only a handful of players have very low scores, while most players achieve scores above the 1 million threshold.

<table>
<thead>
<tr>
<th>Overall search strategies</th>
<th>Social learning strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual search</td>
<td>Copy the best</td>
</tr>
<tr>
<td>53.4%</td>
<td>57.6%</td>
</tr>
<tr>
<td>Social learning</td>
<td>Other copying behavior</td>
</tr>
<tr>
<td>41.4%</td>
<td>42.4%</td>
</tr>
<tr>
<td>Shuffle</td>
<td></td>
</tr>
<tr>
<td>5.1%</td>
<td></td>
</tr>
</tbody>
</table>

**Table 5.1:** Search strategies in the ATC

The table shows how often players engaged in various types of search moves. Individual search refers to moves that did not involve any form of copying. Social learning strategies refer to moves that involved copying others. Copy the best refers to copying the entire solution (all three lines) that at the time was the best. Other copying behavior refers to copying any other solution or only partially copying the best (e.g. one or two of the three available lines). If a player “shuffled” he received a random combination of existing solutions, i.e. each line could be from different solutions.
6 Optimal control of a spin qubit in NV color center in diamond

6.1 Introduction into NV centers

Among the different possible platforms for quantum applications (see Chapter 1), color centers in diamond have attracted lots of attention over the last years. In particular, they might serve as a highly interesting candidate for quantum simulation and sensing with all the advantages resulting from a solid-state device with good prospect of miniaturization. Moreover, some color centers show highly promising properties at room temperature. Setting up the experimental apparatus is of relative ease because initialization and read out can be done optically and the actual manipulation is carried out via microwave (MV) manipulation [311; 312].

There is a whole set of defects in diamond and impurities that might be exploited for the aforementioned purposes. Specifically, the nitrogen vacancy (NV) center is frequently employed due to its remarkable optical and spin properties at room temperature and nitrogen is the most common impurity [313]. Hence, NV centers can be found in almost every diamond. However, the samples used in the lab are artificial diamonds, where one possible production process of them is via Chemical Vapor Deposition [314]. During that process the diamond is formed from a plasma at temperatures of above 700°C. Nitrogen is either added to the plasma or being implanted later on [147].

The NV center is made up by a nearest-neighbor pair of a nitrogen atom (substituting a carbon atom) and a lattice vacancy. A schematic visualization is shown in Fig. 6.1 where its location in the otherwise very regular diamond crystal structure (composed by two intersecting face-centered cubic lattices) is depicted. The axis of the NV is defined by the alignment of the two constituents and coincides with the [1 1 1] direction of the diamond lattice.

In this work we focus on the manipulation of the electronic hull: therefore all important properties follow from the electronic structure of the NV. The two charge states, negative NV− and neutral NV0 are known while there is still some ambiguity how the respective
number of electrons emerges [316]. Nevertheless, it is possible to suppress the NV\(^0\) state and work with the NV\(^-\) state only which for the sake of clearness of notation, for the rest of the thesis, be denoted by NV.

Its electronic structure is that of a three level system. There is a spin \(S=1\) triplet ground state \(^3A\) and a corresponding triplet excited state \(^3E\) as well as metastable intermediate singlet states, see Fig. 6.2. In absence of an external magnetic field, the \(|m_s=0⟩\) ground state level is separated from the (degenerate) \(|m_s=±1⟩\) states by a splitting of 2.87 GHz. For the excited states this splitting is lower and amounts to 1.4 GHz. Using green laser light, the NV can be transferred to its excited states. This process is mostly spin preserving. There are two possible decay paths from the \(|m_s=±1⟩\) excited states: either directly back to the \(|m_s=±1⟩\) ground states where red light is emitted or via some intermediate metastable states back to the \(|m_s=0⟩\) ground state. On application of an external magnetic field (along the axis of the NV), the \(|m_s=+1⟩\) and \(|m_s=−1⟩\) ground states get separated as described by the Zeeman effect. At about 1027 G, the \(|m_s=−1⟩\) and \(|m_s=0⟩\) states become comparable in energy resulting in a strong interaction (spin polarization). These two states also form the computational basis for controlled qubit manipulations as reported in the publication in Sec. 6.2. For a treatment of the control theoretical properties of single qubit operations see also Chapters 2.4.2, 3.1, 3.2, 8 and references therein.
Figure 6.2: Level structure of the NV. Excitation (green) and decay (red) of population from a triplet ground state $^3A$ to a triplet excited state $^3E$. Considering the ground state, the $|m_s = 0\rangle$ level is separated from the $|m_s = \pm 1\rangle$ levels by a zero field splitting of 2.87 GHz. Taken from [317].
6.2 Autonomous Calibration of Single Spin Qubit Operations

This section is a reprint of F. Frank, T. Unden, J. Zoller, Ressa S. Said, T. Calarco, S. Montangero, B. Naydenov, and F. Jelezko, *Autonomous calibration of single spin qubit operations*, npj Quantum Information 3, 48 (2017), doi:10.1038/s41534-017-0049-8, licenced under a Creative Commons Attribution 4.0 International License which can be found at https://creativecommons.org/licenses/by/4.0/. The author of this thesis was responsible for providing the numerical simulations and analysis as well as the optimal control algorithm.

6.2.1 Abstract

Fully autonomous precise control of qubits is crucial for quantum information processing, quantum communication, and quantum sensing applications. It requires minimal human intervention on the ability to model, to predict and to anticipate the quantum dynamics, as well as to precisely control and calibrate single qubit operations. Here, we demonstrate single qubit autonomous calibrations via closed-loop optimisations of electron spin quantum operations in diamond. The operations are examined by quantum state and process tomographic measurements at room temperature, and their performances against systematic errors are iteratively rectified by an optimal pulse engineering algorithm. We achieve an autonomous calibrated fidelity up to 1.00 on a time scale of minutes for a spin population inversion and up to 0.98 on a time scale of hours for a single qubit $\pi/2$-rotation within the experimental error of 2%. These results manifest a full potential for versatile quantum technologies.

6.2.2 Introduction

The ability to precisely control and calibrate single qubit operations in solids is a key element for reliable and scalable high-performance quantum technologies, for instance quantum-enhanced sensors and metrological devices. It is also the backbone of many quantum information processing tasks, which paves the way for the future realisation of quantum computation and communication. Together with efficient quantum system characterisations and dynamical predictions [229; 318], where human intervention is minimised, autonomous calibration of a single spin qubit is necessary for the realisation of such advanced quantum technologies. We report here experimental demonstrations of the autonomous calibration of a single spin qubit in diamond using closed-loop
6.2 Autonomous Calibration of Single Spin Qubit Operations

Our spin qubit implementation is a single nitrogen-vacancy (NV) colour centre in diamond. It provides a suitable platform for a precise qubit manipulation to be realised [319; 320]. Its remarkable features, such as optical initialisation and readout, and the ability to be manipulated by microwave fields at room temperature, make this physical system extremely attractive for many quantum technologies [321]. We have witnessed a vast array of demonstrations of the NV centres showing a great potential for future technologies, ranging from sub pico-Tesla magnetometry [322], electric field and temperature sensing [323; 324], to probing molecular dynamics [325], and single-cell magnetic imaging [326]. Furthermore, intertwinements between quantum information and metrology using NV centre based systems yield novel and effective techniques towards the realisation of high-performance technologies, e.g. applying quantum error correction [327] and phase estimation [328] to improve magnetic field sensitivity. One way to reach such technology is to apply the closed-loop optimisation method for auto-calibrating the controls required to drive the system in the presence of experimental limitations and noise. Closed-loop optimal control has been already applied to quantum information processing [127; 185]. However, to date no realisation of such autonomous calibration in room-temperature solids has been reported. Here, we apply a technique derived from optimal control theory, namely the dressed chopped random basis (DCRAB) algorithm [95; 103], to perform real-time closed-loop optimisations of two fundamental single qubit operations, a spin-1/2 population inversion and a $\pi/2$-rotation gate, against frequency detuning. The algorithm is adapted for use in the NV centre based experiment and directly embedded in the experimental apparatus, allowing the autonomous spin qubit calibrations to be fully performed as illustrated by Fig. 6.3. In contrast to a simple feedback circuit where usually only a single degree of freedom is adjusted at a time based on continuous measurement, our framework considers the whole quantum dynamics and is capable of addressing multiple degrees of freedom at every instance in time.

Optimal control methods have been applied to several quantum information processing tasks with NV centres [66; 113; 329–332], affirming their necessity and significance for quantum technology. However, the previously reported experiments [66; 113; 329–332], utilise open-loop optimisation techniques where the optimisation is performed before the actual experiment by separate computer simulations. The technique requires system–environment coupling information as detailed as possible to provide a robust solution. In contrast, the closed-loop technique requires no explicit system–environment information. Hence, it is utmost practical for the realisation of versatile quantum devices. One significant feature of the DCRAB algorithm is that it makes such closed-loop optimisation viable since the only quantity required from the experiment is a single
Figure 6.3: Schematics of the experiment. The closed-loop optimisation of electron spin qubit operations interfaced to a nitrogen-vacancy (NV) centre in diamond (nitrogen atom in yellow) (a) is performed on a homebuilt confocal microscope (c). The spin state is initialised and readout optically, and microwave pulses are applied to manipulate the state and can be used to create gates. To start an optimisation process, a guess pulse is applied to the sample through a microwave antenna and the figure of merit is evaluated by state tomography on the spin state. This fidelity estimate is then fed to an DCRAB based algorithm and a new test-pulse is generated. The spin trajectory (blue arrow) corresponding to an optimised pulse is shown on the Bloch sphere (b). As sketched, these steps are iteratively repeated until a previously defined fidelity is reached.

The figure of merit (e.g. state or gate fidelity). No further information, such as a gradient or a Hessian, is necessary. Moreover, recent theoretical work [115; 116], points out that the relevant number of degrees of freedom in the control is rather small for few qubit systems. A reasonable number of degrees of freedom can be addressed through a suitable parametrisation [49; 99; 100]. The DCRAB algorithm makes use of this foundation. It shapes high accuracy pulses with few iterations (or superiteration that is required for avoiding local optimisation traps), and maintains the robustness against noise and errors potentially occurred at any stage of experiments [103]. The key ideas are to expand the pulses in a reasonable function basis and subsequent basis function changes to avoid local traps. We provide more detailed discussions of the algorithm
and its implementations in the methods section and in Sec. A in the supplementary material.

6.2.3 Results

High fidelity population inversion via closed-loop control. In this experiment we search for an optimal microwave pulse to transfer the NV spin state $|m_s = 0\rangle$ to $|m_s = -1\rangle$ with high fidelity. Therefore, we performed a state tomography after applying a parametrised microwave pulse. Subsequent measurement of the state transfer fidelity is then used as a figure of merit. To estimate the particular fidelity $F_{cl}^{ex}$ we utilised quantum state tomography (see Sec. B and Sec. C of the supplementary information).

For convention, the subscripts of the fidelity $F_j^i$ indicate either experimental ($ex$) or theoretical ($th$) data. The superscripts ($cl$) and ($ol$) refer to closed-loop and open-loop control, respectively. In Fig. 6.4 we show the results of our state transfer optimisation. In part (a), we first identify via an open-loop simulation the expected performance of the employed optimisation method in presence of limited transfer time $T$ and static detuning $\Delta$. For relative process times $T/T_\pi$ exceeding $\sim 1.5$, our simulations identify robust solutions even up to a relative detuning of about $\Delta_\Omega = 10$. Here the maximal Rabi frequency is determined by $\Omega = \frac{T_\pi}{2T}$, which is connected to the maximal amplitude of the applied microwave pulse. We studied transfer times in the regime $T/T_\pi > 1$ because the maximum speed of a quantum system evolution is bound in general by the quantum speed limit [139; 166]. Numerical simulations support this fact when a rectangular microwave pulse shape is assumed (see supp. material section G). In part (b), we show our experimental results achieved via closed loop optimisation which support these results for small detuning, when $T/T_\pi = 1.5$. For comparison, we show the corresponding crosssection of the numerical results of part (a) in (c). Each optimisation, performed for a certain gate time and a certain detuning, bases on a DCRAB algorithm with 6 superiterations. Exemplarily, we show in Fig. 6.4 (d,e) the full closed-loop optimisation process in the case of no detuning, and when a small detuning is applied. The blue curve shows the currently best found solution, while the red line is an internal algorithmic figure of merit quantity (for details see Supp. Mat. Section A). The necessary time for achieving optimal fidelities depends on the accuracy of the tomography measurements ($F_{cl}^{ex}$) and on the initial optimisation parameters. In our case, a randomly chosen configuration (start-simplex) of the search algorithm was used. We achieve the maximal fidelity of 1.00 with an accuracy of $10^{-2}$ on a reliable timescale of about 2000 seconds in the case of no detuning and on a timescale of about 100 seconds with detuning. A faster optimisation with an off-resonant pulse is here
possible due to a fortunate choice of the initial search-configuration (optimal solution was already part of start-simplex). It is interesting to note, that all optimised pulses of the open-loop simulation were not able to beat the results of our closed-loop strategy in the case of moderate detuning. The best fidelity achieved via open-loop techniques \(F_{\text{ol}}\) is marked in Fig. 6.4 (b) and is only on the order of 0.6.

**Auto-calibration of a single qubit \(\pi/2\)-rotation.** A \(\pi/2\)-rotation is the basic block of generating coherent quantum processes like quantum metrology and quantum computing. To show the capabilities of our concept, we subsequently optimised the quantum gate

\[
G = e^{-i\frac{\pi}{2}S_x} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix},
\]

which is a typical experimental implementation of a Hadamard gate. Here, \(S_x\) identifies the spin-1/2 \(x\)-operator. In accordance to the previous experiment, we first performed the experiment with no detuning and second, when a detuning of 8.125 MHz (relative detuning \(\frac{\Delta}{g} = 0.7\) was applied. The results are shown in Fig. 6.5. To quantify the performance of the evaluated microwave pulse with respect to the defined quantum gate \(G\), quantum process tomography was used (for details see Supp. Mat. Section D and E). We observe after 99 evaluations a fidelity of 0.99 ± 0.01 and after 58 evaluations a fidelity of 0.98 ± 0.02 when the static microwave detuning was applied. The initial fidelity of the guess pulse was in both cases about 0.50. Compared to the previous experiment, the time needed for optimal results is about 4 times longer due to additional measurements, which are necessary for complete process tomography.

### 6.2.4 Discussion

Our experimental results demonstrate that the closed-loop feedback control overcomes static and unknown system errors to achieve the high fidelity autonomous calibration of single quantum gates that is necessary for future quantum technologies with room temperature solids. Our approach of the closed-loop optimisation uses minimal control resources and experimental knowledge that are accessible for users. The total time, required for autonomous calibration, is mainly determined by the duration of quantum tomography measurement and not by the optimisation algorithm. Hence, a significant speed up in the total time of calibration and its fidelity precision may potentially be achieved by employing fast and simplified tomography methods, for instance randomised benchmarking [333]. In addition to the autonomous calibration, our demonstrated
Figure 6.4: Optimal state transfer in presence of frequency detuning and limited time resources. (a) shows the resulting fidelity ($F_{th}^{ol}$) from the parameter study of the optimized state transfer as a function of relative process time $\frac{T}{T_\pi}$ and rel. detuning $\frac{\Delta}{\Omega}$. $T$ is the time duration of the applied pulse and $T_\pi$ is the time necessary for a population inversion when a constant Rabi frequency of $\Omega$ is applied. $\Delta$ is the detuning of the microwave frequency. In (b) we show the results of the closed-loop experimental optimisation ($F_{ex}^{cl}$, red points) for different rel. detuning and a fixed transfer time of $\frac{T}{T_\pi} = 1.5$. In addition, the level of best achieved fidelities from experimental evaluation of open-loop pulses is indicated (by black dashed line). A cross-section of the theoretical results from (a) at $\frac{T}{T_\pi} = 1.5$ is shown in (c). (d) shows the evaluated fidelities ($F_{ex}^{cl}$) when the microwave pulse is applied on resonance and in (e) when a relative detuning of 0.2 is applied. The blue lines show the last, highest fidelity achieved.
Figure 6.5: Optimisation of a single qubit $\pi/2$-rotation when the control pulse is applied on resonance with the NV transition (a) and when a detuning on the order of the Rabi frequency is artificially applied (b). For each scenario we show the measured fidelity $F_{cl}^{\text{ex}}$ with increasing number of function evaluations and the experimentally obtained process matrix for the real (left) and the imaginary part (right) in case of the highest evaluated fidelity.

closed-loop optimisation features stabilisation mechanism against experimental drifts, for instance due to fluctuations of the magnetic field strength and the resulting frequency detuning. Our procedure presented in this letter is not limited for application to single-qubit operations only. Further experimental implementations towards multi-pulse and multi-qubit gate autonomous calibrations are in principle feasible using our closed-loop optimisation method. Two qubit gates are often experimentally implemented with multiple pulses. In this case, each pulse could be parametrised individually or the overall pulse sequence can be parametrised by a proper basis function choice. The complexity will in general increase due to a more complex quantum process tomography or due to a larger parameter space. But the protocol itself, evaluation of a pulse sequence gate fidelity followed by a systematic pulse sequence update, can readily be extended as the algorithm does not require any a-priori system knowledge and is therefore flexible in its
6.2.5 Methods

The two-level quantum system considered in this work is composed of the NV ground spin states $|m_s = 0\rangle$ and $|m_s = -1\rangle$. Electron spin initialisation and readout are performed on a home built confocal microscopy setup at room temperature. To perform quantum operations on the NV spin with high fidelity, the microwave field source is controlled by an arbitrary waveform generator (AWG, Keysight M8195A), with a timing resolution of 65 GS/s and an amplitude resolution of 8 bit. In combination, we used a 50 W amplifier with a frequency bandwidth of about 4 GHz. The microwave field was created with a copper wire close to the NV. By controlling the amplitude and the phase of the microwave, we are able to rotate the system spin around arbitrary axis on the Bloch sphere. The time dependent control Hamiltonian is given by

$$H_c = 2\pi \Omega \left( X(t) \hat{S}_x + Y(t) \hat{S}_y \right),$$

while the system Hamiltonian in the rotating frame is $H_0 = 2\pi \Delta \hat{S}_z$, where $\hat{S}_x$, $\hat{S}_y$ and $\hat{S}_z$ are the spin operator of a two-level system. The functions $X(t)$ and $Y(t)$ define the corresponding microwave pulse, where $\Omega$ is the Rabi frequency and $\Delta$ is the microwave frequency detuning. In our experiments the following conditions were fulfilled,

$$X(t) + Y(t) \in [-1 1], \quad \Omega \in [0 10] \text{ (in MHz).}$$

The two functions $X(t)$ and $Y(t)$ are optimized simultaneously using the dCRAB algorithm, where both start off from an initial guess (randomly chosen start-simplex). At the core of the algorithm there is the use of superiterations which are composed again of an iterative update scheme that maps the time dynamics of all control input variables to a parameter variational problem via a suitable basis expansion. A brief introduction along with more general explanations can be found in SI/A. In this work, $X(t)$ and $Y(t)$ were expanded in (sub-) iteration denoted by $j$ of each superiteration denoted by $k$ as:

$$X(t)^{*,k} = a^{j,k}_{n=1} \sin(\omega^{j,k}_{n=1} t) + b^{j,k}_{n=1} \cos(\omega^{j,k}_{n=1} t) + X(t)^{*,k-1}.$$  

(6.4)
The two frequencies $\omega_{n=1}^{j,k} \in [0.5, 4.5]$ and $\Omega_{n=1}^{j,k} \in [0.5, 4.5]$ are randomly chosen at the beginning of each superiteration $k$ and only updated after all the containing $j = 1, \ldots, N_s$ (sub-) iterations were proceeded. This is where, in our case, the four coefficients $a_{n=1}^{j,k}, b_{n=1}^{j,k}, c_{n=1}^{j,k}, d_{n=1}^{j,k}$ are optimized using a direct search method (here Nelder-Mead, see Supp. Mat. Section A). When $j = N_s$ and thus a (sub-) iteration terminates, a basis update is done by updating (re-drawing of random numbers) $\omega_{n=1}^{j,k}$ and $\Omega_{n=1}^{j,k}$ and hence facilitating further improvement (see also SI/Fig. 2). The subsequent super-iteration ($k + 1$) search bases on the previous superiteration by calculating $X(t)/Y(t)^{*k-1}$ with the optimized (denoted by *) coefficients $a_{n=1}^{*k}, b_{n=1}^{*k}, c_{n=1}^{*k}, d_{n=1}^{*k}$ from the previous $k$-th superiteration. The algorithm terminates once the maximum number of superiterations set is reached.

**Data availability**

The authors declare that the main data supporting the finding of this study are available within the article and its Supplementary Information files. Additional data can be provided if necessary.

**Acknowledgements**

We acknowledge financial supports from European Union FET Projects DIADEMS, RYSQ, and SIQS, German Research Foundation (DFG) through the SFB/TRR21 projects, German Federal Ministry of Education and Research (BMBF) via the Q.Com projects, and the Center for Integrated Quantum Science and Technology (IQST). We thank S. Zaiser, P. Neumann, and J. Wrachtrup in Stuttgart University for their valuable support and discussions. S.M. gratefully acknowledges the support of the DFG via a Heisenberg fellowship.

**Author contributions**

F.F., T.U., and J.Z. contributed equally to this work. F.F., T.U., and B.N. carried out the experiment and analysed the data. J.Z. and R.S.S. provided theoretical supports and performed numerical simulations and analyses. F.F., T.U., J.Z., and R.S.S wrote the manuscript with feedback from all authors. T.C., S.M., B.N. and F.J. supervised and managed the project.
Competing interests
The authors declare that they have no competing interests.

6.2.6 Supplementary Material

6.2.6.1 A. Dressed CRAB Algorithm

The chopped random-basis algorithm (CRAB) is originally proposed to provide one simple yet powerful numerical optimisation method for controlling complex quantum many-body dynamics, where the gradient of the figure of merit to be optimised is hard to obtain and the controlling apparatus is realistically bandwidth-limited [95; 96]. Recently, an extended version of the algorithm, namely the dressed CRAB (DCRAB), has been developed to tackle possible local minima or false traps in the optimisation landscape [103]. The standard version of CRAB algorithm has experimentally been implemented and adapted in various physical systems, for instance in nitrogen-vacancy (NV) centre in diamond for precise spin controls beyond rotating-wave-approximation (RWA) [113], in ultra-cold atoms in optical lattice for controlling quantum phase-transition from a superfluid to a Mott insulator [127], in a non-classical state of Bose-Einstein condensate (BEC) [128], in superconducting qubits [224; 334], and in a hybrid system of BECs and cold atoms in optical lattice [129]. A theoretical relation to quantum speed limit in many-body quantum system has also been studied [173].

Here, we elaborate the DCRAB algorithms adapted for our current work on the real-time close-loop optimisation of NV electron spin control. To begin with, we provide in Fig. 6.6 a sketch of the open-loop and closed-loop optimisation methods. The closed-loop method offers users the ability to obtain an optimised control pulse for manipulating the system without having a full access to the complete information about it.

Given an unoptimised microwave pulse as a guess pulse $\Gamma_0(t)$, that controls the NV electron spin ground state, the DCRAB algorithm iteratively finds an updating pulse $g^{j,k}_n(t)$, which gives an optimal pulse $\Gamma(t)^{j,k} = \Gamma_0(t) \times g^{j,k}_n(t)$, such that a certain figure of merit $F^{j,k} = F^{j,k}[\Gamma(t)^{j,k}]$ is maximised (or minimised, if one uses $1 - F^{j,k}$), for a predetermined and fixed duration of time $t = T$. Here, we use $j$ to denote the update which can be either function evaluation or iteration, and $k$ to represent super-iteration. In a direct search method introduced later in this section, the figure of merit does not necessarily improve over a number of increasing function evaluations, nevertheless, it is not decreasing. The typical examples of figure of merit are state-to-state or state

\[\text{We distinguish between these quantities since one iteration may require multiple function evaluations.}\]
Figure 6.6: *Open-loop* and *closed-loop* optimisation methods. An optimised or a guess pulse $\Gamma_0(t)$, is fed into an optimal control algorithm (as gold, and in our work this is DCRAB algorithm). For each update (denoted by $j$, and can be either *function evaluation* or *iteration*), the *open-loop* method (red) uses a system dynamics simulator, e.g. a personal computer (PC) or a cluster of PCs, to numerically provide a single optimised figure of merit $F^j$. The simulator requires a set of information obtained indirectly from the actual physical system: a controllable Hamiltonian $H$, a system-environment interaction model $L$, and a description of noises or errors $E$. If the numerical $F^j$ reaches a certain target value, the process is terminated and the final optimised pulse $\Gamma(t)$, is passed to the actual physical process. In case the set of information from the actual physical system is not available or can only be partially obtained, one can resort to the *closed-loop* optimisation method (blue). The actual physical system acts as a simulator which outputs the actual $F^j$. If the target $F^j$ is satisfied, the desired process is directly started. In principle, both methods can be used complementarily as a *hybrid* method.

Fidelity, quantum gate fidelity, and entropy of entanglement [96]. Following Ref. [103], one may expand the updating pulse $g_{c}^{j,k}(t)$, as

$$g_{c}^{j,k}(t) = \frac{1}{\lambda(t)} \sum_{n=0}^{N} a_{n}^{j,k} \sin \omega_{n}^{j,k} t + b_{n}^{j,k} \cos \omega_{n}^{j,k} t.$$  (6.6)

where $N$ is commonly termed as a number of DCRAB *frequency components*, and $\omega_{n}^{j,k} = 2\pi(n + r)/T$ is a *randomised frequency* due to a random number $|r| < 0.5$ [113]. The function $\lambda(t)$ is a predetermined function limiting the update pulse to be zero at $t = 0$.
6.2 Autonomous Calibration of Single Spin Qubit Operations

and $t = T$. It allows for a smooth ramping at the beginning and at the end of the pulse. This limiting function is flexibly chosen and its parameters depend on some experimental factors, e.g. the resolution of the controlling apparatus (see Ref. [113] for further details). For each super-iteration (after a certain number of function evaluation $j = J$), DCRAB performs re-optimisations by transforming the updating pulse $g_c^{j,k-1} \rightarrow g_c^{j,k-1} + g_c^{j,k}$, where the parameters $\{a^{j,k}_{n}, b^{j,k}_{n}\}$ remain the same, while the randomised frequencies are updated $ω^{j,k-1}_{n} \rightarrow ω^{j,k}_{n}$. From Eq. 6.6, one finds that the optimisation problem is reduced to a search problem: “Find a set of parameters $\{a^{j,k}_{n}, b^{j,k}_{n}\}$, which maximises $F^{j,k}$, given that the DCRAB frequency components does not change at every function evaluations or iterations (i.e. $\{a^{j,k}_{n}, b^{j,k}_{n}, ω^{j,k}_{n}\} \rightarrow \{a^{j+1,k}_{n}, b^{j+1,k}_{n}, ω^{j,k}_{n}\}$).”

In principle, any direct search algorithm can be employed to solve this problem, such as Nelder-Mead or simplex algorithm [97], Powell algorithm [335], and pattern search algorithm [336]. Here, we employ the Nelder-Mead (NM) algorithm since it has shown its strength in optimizing various unconstrained minimization problems [337; 338], and has been implemented and efficiently tested in many code libraries, e.g. PYTHON [339] or MATLAB [340]. We encode the DCRAB method with the NM algorithm in PYTHON, and directly embed it on a PC interfaced with the microwave control apparatuses, for instance an arbitrary wave generator (AWG) [341].

A numerical evidence of the DCRAB capabilities is depicted in Fig. 6.7 which shows two instances of a control landscape. Control landscapes display the quality (e.g. fidelity) of an optimisation as a function of the control variables [124–126]. Here, the control variables are the amplitude and phase of a sinusoidal curve. Due to the limited function space to be explored at a time, the algorithm might get trapped at a sup-optimal point in the landscape. Further sub-optimal points may be induced by restrictions unavoidably imposed by the lab environment such as limited bandwidth, control power limitations, and noise occurring at various stages throughout the experiment. By moving on to the next super-iteration and hence doing a basis change (see arrow in the figure), the trap is unlocked with probability one allowing the algorithm to proceed and converge to a global optimum [103].

6.2.6.2 B. Quantum State Tomography

In our experiments the quantum system of interest is defined by the NV centre’s ground state spin levels $|m_s = 0\rangle$ and $|m_s = -1\rangle$. An experimentally prepared density matrix $\rho$, is measured by performing two Rabi experiments each with orthogonal rotation-axis $x$ and $y$, and initial state $\rho$. The rotation axis is controlled by the phase of the resonant microwave control. In the case of the $x$-rotation [342], after measuring the
6 Optimal control of a spin qubit in NV color center in diamond

Figure 6.7: Control landscape of the single qubit gate synthesis at two stages of optimisation. The Nelder-Mead search at the k-th super-iteration (si) goes towards the marked local minimum or trap (the left pane). Starting at that trap where the bases of the pulse information has been gained (the magenta circle on the right pane), the subsequent search with the basis changes unlocks the local trap allowing the optimisation to proceed further downhill. The dashed black triangle shows the possible span of the Nelder-Mead simplex.

population $P(|0\rangle,t)$ in the state $|m_s = 0\rangle$, we observe that

$$P(|0\rangle,t) = \frac{d + a}{2} + \frac{d - a}{2} \cos(2\pi\omega t) - c \sin(2\pi\omega t). \tag{6.7}$$

In the case of the $y$-rotation, we have

$$P(|0\rangle,t) = \frac{d + a}{2} + \frac{d - a}{2} \cos(2\pi\omega t) + b \sin(2\pi\omega t). \tag{6.8}$$

Here, the applied microwave duration is denoted by $t$, and the initial density matrix is given by

$$\rho = \begin{pmatrix} \rho_{|\bar{0}\rangle\langle\bar{1}|} & \rho_{|\bar{1}\rangle\langle\bar{0}|} \\ \rho_{|\bar{0}\rangle\langle\bar{1}|} & \rho_{|\bar{0}\rangle\langle\bar{0}|} \end{pmatrix} = \begin{pmatrix} a & b - ic \\ b + ic & d \end{pmatrix}. \tag{6.9}$$

Therefore, we identify the entries of the initial density matrix $\rho$ by fitting the experimentally observed Rabi oscillations to the described model. To guarantee an allowed density matrix, we use quadratic maximum likelihood estimation to extract the density
matrix, which is parameterised accordingly by

\[ \rho(\xi, \nu) = e^{-i\nu S_y} e^{-i\xi S_x} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} e^{i\xi S_x} e^{i\nu S_y}. \]  

(6.10)

Here, \( S_x \) and \( S_y \) are the spin-1/2 operators. We identify the error of the estimation of a single entry of \( \rho \), by

\[ \sigma(\rho_{ij}) = \sqrt{\text{Var}(\rho_{ij})} = \frac{1}{4} \sqrt{\sum_{i,j} (\rho_{ij} - \rho_{ij}(\xi, \nu))^2}. \]  

(6.11)

### 6.2.6.3 C. Fidelity Estimation: Optimising State Transfer

The fidelity of the state transfer \(|m_s = 0\rangle \rightarrow |m_s = -1\rangle \langle |−1\rangle\), is estimated by the following tasks: (i) initialising the system in \(|m_s = 0\rangle \langle 0\rangle\), (ii) applying the gate, and (iii) performing quantum state tomography. The fidelity is given by

\[ F = \rho_{|−1\rangle⟨−1|} \pm \sigma(\rho_{|−1\rangle⟨−1|}). \]  

(6.12)

### 6.2.6.4 D. Fidelity Estimation: Optimising a Hadamard Gate

The fidelity of a gate is estimated by firstly initialising the system in one of the four states, \( \psi_1 = |0\rangle, \psi_2 = |−1\rangle, \psi_3 = \frac{|0\rangle - |−1\rangle}{\sqrt{2}} \) and \( \psi_4 = \frac{|0\rangle + |−1\rangle}{\sqrt{2}} \). It is then followed by applying the gate, and applying the inverse of the ideal gate. Lastly, the population according to the initial state \( \psi_i \), is measured by quantum state tomography. If the final state matches the initial state, then the actual gate reaches the ideal gate. After applying this process for the full set of initial states \( \psi_i \), the fidelity \( F \), is estimated by

\[ F = E(\rho_{|\psi_i\rangle⟨\psi_i|}) \pm E(\sigma(\rho_{|\psi_i\rangle⟨\psi_i|})), \]  

(6.13)

where \( E(\cdot) \) denotes the expectation value, and \( \sigma(\cdot) \) is the standard deviation. The single qubit rotations that are necessary for the initialisation are repetitively calibrated. The corresponding fidelity is above 99.8%.

### 6.2.6.5 E. Quantum Process Tomography

Firstly, the system is initialised in one of the four states \( \rho_0^i = \psi_i \psi_i^\dagger \). Secondly, the gate is applied, and the final state \( \rho_f^i \) is measured by quantum state tomography. We choose
the following complete set of quantum operators:

\[ e_1 = 1, \]  
\[ e_2 = \sigma_x, \]  
\[ e_3 = -i\sigma_y, \]  
\[ e_4 = \sigma_z, \]

where \( \sigma_{x/y/z} \) are the Pauli matrices. If we define

\[
M = \begin{pmatrix}
0 \cdot 1 & 0 \cdot 1 & 0 \cdot 1 & 1 \cdot 1 \\
1 \cdot 1 & 0 \cdot 1 & 0 \cdot 1 & 0 \cdot 1 \\
0.5 \cdot 1 & -0.5i \cdot 1 & 0.5i \cdot 1 & 0.5 \cdot 1 \\
0.5 \cdot 1 & -0.5 \cdot 1 & -0.5 \cdot 1 & 0.5 \cdot 1, \\
\end{pmatrix}
\]  

(6.18)

and

\[
\beta = \begin{pmatrix}
1 & e_1 \\
e_1 & -1 \\
\end{pmatrix},
\]  

(6.19)

the process matrix is then defined by [15]

\[
\chi = \beta \begin{pmatrix}
M^{-1}\rho_1^f & M^{-1}\rho_2^f \\
M^{-1}\rho_3^f & M^{-1}\rho_4^f \\
\end{pmatrix} \beta.
\]  

(6.20)

6.2.6.6 F. Realistic Parameters Analysis

We perform numerical simulations on the optimised population inversion for a certain range of predetermined detuning \( \omega \), and control time \( T \), to demonstrate the robustness of the DCRAB algorithm. The simulated detuning is in the unit of Rabi frequency \( \omega_r \), and we use \( T_\pi \), to denote the control time of the theoretical rectangular \( \pi \)-pulse with no detuning. The result of the simulations is presented in Fig. 6.4 (a) in the main
6.2 Autonomous Calibration of Single Spin Qubit Operations

The Hamiltonian used in the simulations is

\[ H = \frac{2\omega}{\omega_r} \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} + \Gamma_x(t) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \Gamma_y(t) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \] (6.21)

The control amplitudes $\Gamma_x(t)$ and $\Gamma_y(t)$, are constrained by a condition of $|\Gamma_x(t) + \Gamma_y(t)| \leq 1$, $\forall t$, which corresponds to the voltage output limit of the AWG used in the experiment.

In performing the simulations, the algorithm and its parameters (e.g. the total number of function evaluations and super-iterations, the update parameters, the algorithm stopping criteria, and the pulses bandwidth) are taken to be the same as those used for generating the real pulse in the closed-loop optimisation experiment. Each of the fidelities, shown in Fig. 6.4 (a) in the main manuscript, are the average fidelities over 20 independent runs of initialisation and optimisation.

Finally, the set of pulses obtained from the simulation for $T = 1.5 \times T_\pi$, are directly applied to the experiment without any further closed-loop optimisation, and yields a state inversion fidelity of less than 0.6. This significant drop in fidelity is due to the experimental imperfections, the unknown system parameters, and the possible further constraints which are always present at any realistic setup. On the contrary, the closed-loop control optimisations take all the above-mentioned factors into account, and hence provide far better results.

6.2.6.7 G. Theoretical performance of a rectangular pulse for state inversion

Simulated fidelities in the case of a rectangular pulse with a constant Rabi amplitude $\Omega = \frac{1}{2T_\pi}$, detuning $\Delta$ and pulse duration $T$ is shown in Fig. 6.8.
Figure 6.8: State transfer fidelity in the case of a rectangular pulse for different pulse durations and frequency detunings.
7 Central Moment Dynamics

7.1 Introduction into approximate quantum simulation methods

All objects, macro- as well as microscopic, obey the laws of quantum mechanics. The only difference being, that the emerging quantum effects in the macro world are so weakly pronounced that they are hardly measurable. Therefore in the limit of large masses, high energies and large spatial scales (compared to Planck’s constant), quantum mechanics converges or reduces to the laws of classical mechanics whose founding fathers are Newton, Lagrange, Hamilton, Jacobi, Euler, and the Bernoullis [343].

Classical physics allows for a much more extensive description in terms of number of constituents $N_{co}$ and their interactions. On one hand this favors the classical description, especially when the thermodynamic limit ($N_{co} \to \infty$) shall be approximated [344]. On the other hand, going from objects on the scale of tens of nanometer (such as proteins and other multi molecular structures) down to the atomic ångström scale, quantum effects, such as tunneling, zero point energy, interference, state specific dynamics, and decoherence, become so pronounced that the classical description fails to grasp the essential physics, even qualitatively [345].

Precisely at this boundary between classical and quantum physics is the area of application of a wide variety of problems, especially in chemistry and biology [346–348]. Semi-classical and quantum-classical methods extending classical molecular dynamics have been proposed [349–352]. Originally designed to study spectral properties of many body quantum systems, density functional theory has also been extended to allow for calculations in real-time and real-space [353]. A minimalistic extension of the purely classical treatment is to add the width and correlation to the set of dynamical variables and approximate the potential energy operator via a second order Taylor expansion. This method has been named the thawed Gaussian approximation [354] and further simplified by the frozen Gaussian approximation [355]. Two more sophisticated approaches are being discussed in the next section, as they form a motivation for our work.
Then, we introduce a novel approximate framework to reduce the computational complexity of the study of quantum dynamics of high-dimensional continuous variables systems. Explicitly time-dependent systems are treatable as a viable connection to optimal control theory. The level of approximation is easily tunable and no restrictions are being posed on the physical processes.

7.2 Quantized Hamilton Dynamics and Quantal Cumulant Dynamics

One approach to capture quantum phenomena in many degrees of freedom is the Quantized Hamilton Dynamics (QHD) method [356–361]. The core idea is to propagate the expectation values

$$\langle x^k p^l \rangle_s = \frac{1}{2} \langle x^k p^l + p^l x^k \rangle,$$  \hspace{1cm} (7.1)

instead of the wave functions. The dynamics is then generated by the Heiseberg’s equation of motion

$$\frac{d}{dt} \langle O \rangle = \frac{i}{\hbar} \langle [H, O] \rangle + \langle \frac{\partial O}{\partial t} \rangle.$$  \hspace{1cm} (7.2)

The Hamiltonian of a particle with mass $m$ subject to a potential $V(x)$ reads

$$H = \frac{p^2}{2m} + V(x).$$  \hspace{1cm} (7.3)

In order to apply the method, two major approximations are necessary. One of them is summarized in the literature: “While arbitrary functions of position and momentum operators can be treated as dynamical variables of the QHD framework, the current formulation is restricted to polynomial moments of position and momentum operators” [357]. Therefore in general the potential energy operator needs to be Taylor expanded. This might either be done around the center of mass $x_0 = \mu = \langle x \rangle$ (moving frame) or around some fixed value $x_0 = \text{constant}$ (fixed frame). The Taylor expansion up to $p$-th order is:

$$V(x) \approx V(x_0) + \sum_{j=1}^{p} \frac{V^{(j)}(x_0)}{j!} (x - x_0)^j.$$  \hspace{1cm} (7.4)
The second approximation addresses the fact that apart from some often simplified models such as the harmonic oscillator and spin systems, where the dynamics can be completely described by a finite set of operators, in general, the emerging hierarchy of equations from the Heisenberg’s equation of motion are infinite. In the QHD framework, in order to terminate these equations, the expectation value of higher order operators is expressed in terms of products of expectation values of lower ones. This idea is borrowed from the many-body theory [362–365] and has been shown explicitly for up to fifth order moments of the general form $\langle abcd \rangle$ [359] which turns out to be rather lengthy. To give a flavor of this closure scheme, we paradigmatically show the approximation for third order moments:

$$\langle abc \rangle \approx \langle ab \rangle \langle c \rangle + \langle ac \rangle \langle b \rangle + \langle bc \rangle \langle a \rangle - 2 \langle a \rangle \langle b \rangle \langle c \rangle.$$  \hspace{1cm} (7.5)

Up to now, the QHD approach has been extended up to fourth order, which means that moment variables from Eq. (7.1) with indices \( \{k,l\} | k,l \in \{0,1,2,3,4\}, k + l \leq 4 \) were considered. As a final remark, we mention that the QHD framework has been slightly modified in a way to include the potential energy operator and its derivatives in the set of QHD variables [361].

Some years later, the Quantal cumulant dynamics (QCD) method has been introduced [366–370]. The two main ingredients of the QCD are a cumulant expansion technique and the use of a position shift operator acting on the potential operator. The advantage of the latter is that for some analytical forms of potentials, a Taylor expansion can be avoided. Cumulants and moments are both statistical quantities which are tightly related: The cumulant-generating function is simply defined as the logarithm of the moment-generating function. As a consequence, cumulants can be expressed in terms of moments and vice versa. One beauty of the cumulants is that for the normal distribution only the first two of them are nonzero.

The derivation of the QCD framework starts off from the Hamiltonian from Eq. (7.3). Before rewriting the potential part, a position shift operator is defined as

$$f(x + a) = e^{a(\partial/\partial x)} f(x) =: D_x(a) f(x)$$  \hspace{1cm} (7.6)

Making use of the shift operator $D_x$, the potential energy operator can be re-expressed as

$$V(x) = D_\mu(x - \langle x \rangle) V(\mu)|_{\mu = \langle x \rangle}$$  \hspace{1cm} (7.7)
Introducing a fluctuation operator of an observable $O$, $\delta O = (O - \langle O \rangle)$, the potential energy becomes:

$$\langle V(x) \rangle = \langle \exp(\delta x \frac{\partial}{\partial \mu}) \rangle V(\mu) = \sum_{j=0}^{\infty} \frac{1}{j!} \langle \delta x^j \rangle \frac{\partial^j V(\mu)}{\partial x^j}.$$  \hspace{1cm} (7.8)

After some straightforward calculations one can identify the cumulants $\lambda_j$, yielding:

$$\langle V(x) \rangle = \exp \left( \sum_{j=1}^{\infty} \frac{\lambda_j}{j!} \frac{\partial^j}{\partial \mu^j} \right) V(\mu)$$  \hspace{1cm} (7.9)

An approximation enters via the unavoidable truncation of the summation, which is often done for $j = 2, 4$. Consequently, higher order cumulants are neglected. The dynamics of the cumulants is again obtained from Eq. (7.2). Taking into account only up to second order cumulants, the second order total energy is given by

$$E_2 = \frac{\eta^2}{2} + \frac{\lambda_{0,2}}{2} + \exp \left( \frac{\lambda_{2,0}}{2} \frac{\partial^2}{\partial \mu^2} \right) V(\mu),$$  \hspace{1cm} (7.10)

where $\eta = \langle p \rangle$ and $\lambda_{k,l} = \frac{\partial^k}{\partial x^k} \frac{\partial^l}{\partial p^l} \langle \exp(cx) \exp(dp) \rangle_s$. One can show that $E_2$ is conserved, i.e. $\dot{E}_2 = 0$ \forall $t$, within in the QCD framework. Notice, however, that this does of course not imply that the energy is conserved nor it being equivalent to $E_2$. Furthermore, the QCD framework conserves the quantity $\gamma = \lambda_{2,0} \lambda_{0,2} - \lambda_{1,1}^2$ which allows to eliminate one of the three dynamical variables. But, once again, in general, physical processes do not necessarily conserve $\gamma$. The only thing that can be said is that $\lambda_{2,0} > 0$, $\lambda_{0,2} > 0$, and that the Heisenberg’s uncertainty principle must be fulfilled: $\lambda_{2,0} \lambda_{0,2} \geq \frac{1}{4}$.

### 7.3 Central Moment Dynamics

We build our central moment dynamics framework in analogy to the QHD approach, but use a completely different closure scheme. Moreover, our central moment framework is designed in a flexible way, allowing to set arbitrary orders without additional analytical calculations. Furthermore, in contrast to QHD as well as QCD, we explicitly do not assume the Heisenberg uncertainty to be minimal (least quantal principle). By keeping the approximation order flexible and hence allowing for a variable number of degrees of freedom tracked, we show that the resulting error becomes minimal when the effective dimensionality of the problem is covered. Finally, we explicitly apply our moment framework to time-dependent problems (i.e. non static parameters in the Hamiltonian), which has not been demonstrated yet, neither in QHD nor in QCD. The basic idea is to
represent quantum mechanical states using central moments of conjugate operators. For translational dynamics in continuous space, these are the position \( x \) and momentum \( p \) operators. Consequently, the \( k/l \)-th central moment can be expressed as the expectation value of the symmetrized product of the position and momentum fluctuation operators

\[
m^{[k,l]} = \frac{1}{2} \langle \{ (\hat{x} - \mu)^k, (\hat{p} - \eta)^l \} \rangle,
\]

where \( \mu = \langle \psi | \hat{x} | \psi \rangle \) and \( \eta = \langle \psi | \hat{p} | \psi \rangle \) are the center of mass in position and momentum space, respectively. The moment dynamics is derived based on Heisenberg's equation of motion (see Eq. (7.2)). Putting things together, the system of differential equations describing the moment dynamics can be written as

\[
\dot{m}^{[k,l]} = \frac{i}{\hbar} \left( \frac{p^2}{2m_0} + V(x), \frac{1}{2} \{ (x - \mu)^k, (p - \eta)^l \} \right) + \frac{1}{2} \left( \frac{\partial \{ (x - \mu)^k, (p - \eta)^l \}}{\partial t} \right).
\]

In order to map the right-hand side of Eq. (7.39) onto moments \( m^{[k',l']} \) \( k', l' \in \{1, 2, \ldots, k_{\text{max}}\} \), we have to have the potential in polynomial form. If that is not the case, the potential might either be Taylor expanded around the center of mass \( x_0 = \mu \) (moving frame) or around some fixed value \( x_0 = \text{constant} \) (fixed frame) [358], be treated fully analytically as 'natural variables' [361] or be transformed via a position shift operator and its expectation value being cumulant expanded [366]. Here, we go with the former ansatz, but use a rather distinct way to close the system of differential equations. Therefore, the extension of the potential reads

\[
V(x) \approx V(x_0) + \sum_{j=1}^{p} \frac{V^{(j)}(x_0)}{j!} (x - x_0)^j.
\]

Plugging the expanded potential \( V(x) \) into Eq. (7.39), a more evolved expression occurs whose size and closure properties depend on the highest tracked polynomial \( p \) of the potential. Here, we like to stress two facts: First, all the derivations and also the subsequent implementation can be automatized. As we will see in the following, this step is not runtime critical as the calculations are performed before the actual time evolution. Second, in order to obtain the system of differential equations for the moments

\[
\dot{m}^{[k,l]} = \mathcal{F}(m^{[k',l']}), \quad k' + l' \leq k_{\text{max}} + (d - 2),
\]
occurring commutators of the type \([\hat{x} - \mu]^k, (\hat{p} - \eta)^l] = [\tilde{x}^k, \tilde{p}^l]\) can be recursively broken down into lower orders anti-commutators that make up the moments (see Eq. (7.39)) in the following way

\[
[\tilde{x}^k, \tilde{p}^l] = - \sum_{q' = 1}^{u_o} \frac{(-1)^{q'}}{(2q' - 1)! (l - 2q' + 1)! (k - 2q' + 1)!} \times \frac{\hbar^{2q' - 1}}{2} \{\tilde{x}^{k-2q'+1}, \tilde{p}^{l-2q'+1}\} \\
- \sum_{q' = 1}^{u_e} \frac{(-1)^{q'}}{(2q'!) (l - 2q')! (k - 2q')!} \times \frac{\hbar^{2q'}}{2} [\tilde{x}^{k-2q'}, \tilde{p}^{l-2q'}]
\]

(7.15)

Here the upper summation borders are respectively:

\[
u_o = \begin{cases} \frac{\min(k, l) + 1}{2} & \text{if } \min(k, l) \text{ is odd} \\ \frac{\min(k, l)}{2} & \text{if } \min(k, l) \text{ is even} \end{cases}
\]

(7.16)

\[
u_e = \begin{cases} \frac{\min(k, l) - 1}{2} & \text{if } \min(k, l) \text{ is odd} \\ \frac{\min(k, l)}{2} - 1 & \text{if } \min(k, l) \text{ is even} \end{cases}
\]

(7.17)

The proof for this recursive relation can be found in the next Section 7.4. Using the formula from Eq. (7.15), the linear relation \(\mathcal{F}\) of the time derivative of all moments as a function of other moments (Eq. (7.40)), is achieved. It can be easily seen, that for any potential other than the harmonic one, that is for \(d \geq 3\), the system of differential equations forms an infinite hierarchical system. Therefore, a closure must be applied. This can either be done as presented in the publication (see Section 7.5), or according to a more evolved idea presented in appendix 7.6.1.
7.4 Proof of recursion formula

We proof the recursion formula from Eq. (7.15). For better readability, the tildes are omitted. First, we motivate how to come up with an appropriate ansatz. Second, we show that is indeed correct. Therefore, we begin with the definition of the commutator:

\[
[x^k, p^l] = x^k p^l - p^l x^k
\]  

(7.18)

As sketched, the idea will be to subsequently bring one x to the right. Hence, we get:

\[
= x^{k-1} p^l x + x^{k-1} [x, p^l] - p^l x^k
\]  

(7.19)

Once again, one x will be brought to the other side:

\[
= x^{k-2} p^l x^2 + x^{k-2} [x, p^l] x + x^{k-1} [x, p^l] - p^l x^k.
\]  

(7.20)

Repeating this k-times and applying \([x, p^l] = ilp^{l-1}\) for each step, brings us to:

\[
= il \sum_{j=1}^{k} \left( x^{k-j} x^{j-1} p^{l-1} - x^{k-j} [x^{j-1}, p^{l-1}] \right)
\]  

(7.21)

\[
= ilk x^{k-1} p^{l-1} - il \sum_{j=1}^{k} x^{k-j} [x^{j-1}, p^{l-1}]
\]  

(7.22)

These preconsiderations led us to the following ansatz (non centralized version):

\[
[x^k, p^l] = - \min(k,l) \sum_{q=1}^{\min(k,l)} \frac{(-i)^q k!}{q!(l-q)!(k-q)!} x^{k-q} p^{l-q}.
\]  

(7.23)

The proof will be carried out via induction. We start with Eq. (7.22) and plug in the recursion formula:

\[
[x^k, p^l] = ilk x^{k-1} p^{l-1} - il \sum_{j=1}^{k} x^{k-j} [x^{j-1}, p^{l-1}]
\]

\[
= ilk x^{k-1} p^{l-1} - il \sum_{j=1}^{k} x^{k-j} \left( - \sum_{q=1}^{\min(j-1,l-1)} \frac{(-i)^q (l-1)!(j-1)!}{q!(l-1-q)!(j-1-q)!} x^{j-1-q} p^{l-1-q} \right)
\]  

(7.24)
Some minor regrouping yields:

\[
= ilkx^{k-1}p^{l-1} - \sum_{j=1}^{k} \sum_{q=1}^{\min(j-1,l-1)} \frac{(-i)^{q+1}l!x^{k-(q+1)}p^{l-(q+1)}}{q!(l-(q+1))!} \cdot \frac{(j-1)!}{(j-(q+1))!} \quad (7.25)
\]

Next, we will swap the dependencies of the running indices in the double sum as sketched in Figure 7.1:

![Figure 7.1: Change of running indices from gray to blue.](image)

Hence, we get:

\[
= ilkx^{k-1}p^{l-1} - \sum_{j=1}^{k-1} \frac{(-i)^{q+1}l!x^{k-(q+1)}p^{l-(q+1)}}{q!(l-(q+1))!} \sum_{j=q+1}^{k} \frac{(j-1)!}{(j-(q+1))!} \quad (7.26)
\]

We will focus on the last sum and do a shift in the running index \( \tilde{j} := j - 1 \):

\[
\sum_{j=q+1}^{k} \frac{(j-1)!}{(j-(q+1))!} = \sum_{\tilde{j}=q}^{k-1} \frac{\tilde{j}!}{(\tilde{j}-q)!} \quad (7.27)
\]

What we obtain is the definition of the so called falling power, \( x^m = x(x-1)(x-2)...(x-(m-1)) \), which we will be using:

\[
x^m = \sum_{j=q}^{k-1} j^q \quad (7.28)
\]
Next, we apply the fundamental theorem of finite calculus [371]

\[
\sum_{j=q}^{k-1} \tilde{j}^q = \sum_{q} x^q \delta x \tag{7.29}
\]

and thus being able to discrete anti-derivative:

\[
= \left[ \frac{x^{q+1}}{q+1} \right]^k_q = \frac{1}{q+1} \left( \frac{k^{q+1} - q^{q+1}}{q!} \right) \tag{7.30}
\]

\[
= \frac{k!}{(q+1)(k-(q+1))!} \tag{7.31}
\]

Going back to Equation (7.26) and merging the \( q! \) yields:

\[
= i k x^{k-1} p^{l-1} - \sum_{q=1}^{k-1} \frac{(-i)^{q+1} l! k! x^{k-(q+1)} p^{l-(q+1)}}{(q+1)!(l-(q+1))!(k-(q+1))!} \tag{7.32}
\]

Substituting \( \tilde{q} := q + 1 \) results

\[
= i k x^{k-1} p^{l-1} - \sum_{\tilde{q}=2}^{k} \frac{(-i)^{\tilde{q}} l! k! x^{k-\tilde{q}} p^{l-\tilde{q}}}{\tilde{q}!(l-\tilde{q})!(k-\tilde{q})!} \tag{7.33}
\]

Finally, we see that the \( \tilde{q} = 1 \) summand is just the term in front of the sum. Hence, after fusing we arrive at:

\[
= \sum_{q=1}^{k} \frac{(-i)^{q} l! k!}{q!(l-q)!(k-q)!} x^{k-q} p^{l-q} \tag{7.34}
\]

which is just what we wanted to show.

What remains is to symmetrize the recursion formula (Eq. (7.23)) to make it applicable for the introduced purpose. Thereby we will be using the fact that \([x^k.p^l] = 0.5 \cdot [x^k.p^l] - 0.5 \cdot ([x^k.p^l]^+)\]. By doing so, we immediately see that the odd summands
(q=odd) yield the anticommutators and the even summands yield the commutators:

$$[x^k, p^l] = - \sum_{q'=1}^{u_o} \frac{(-1)^{q'} i^{l} k!}{(2q' - 1)! (l - 2q' + 1)! (k - 2q' + 1)!} \frac{\hbar^{2q'-1}}{2}$$

representing the odd summands

$$\{x^{k-2q'+1}, p^{l-2q'+1}\} +$$

$$\left(\frac{(-1)^{q'}}{(2q')! (l - 2q')! (k - 2q')!} \frac{\hbar^{2q'}}{2} [x^{k-2q'}, p^{l-2q'}]\right)$$

representing the even summands

Thereby the upper borders are respectively:

$$u_o = \begin{cases} \frac{\text{min}(k,l)+1}{2} & \text{if min}(k,l) \text{ is odd} \\ \frac{\text{min}(k,l)}{2} & \text{if min}(k,l) \text{ is even} \end{cases}$$

(7.37)

$$u_e = \begin{cases} \frac{\text{min}(k,l)-1}{2} & \text{if min}(k,l) \text{ is odd} \\ \frac{\text{min}(k,l)}{2} - 1^* & \text{if min}(k,l) \text{ is even} \end{cases}$$

(7.38)

(*) The minus one comes from the fact, that these commutators always yield zero.
7.5 Central moment dynamics of high-dimensional continuous variables quantum systems

The following article is in preparation. Zoller, J. and Silvi, P. and Montangero, S. (2018). “Central moment dynamics of high-dimensional continuous variables quantum systems.” The author of this thesis was responsible for the analytical calculations and the numerical simulations.

7.5.1 Abstract

We introduce a numerical method based on central moments to efficiently solve the time-dependent Schrödinger equation of high-dimensional quantum systems. The efficiency gain with respect to standard space discretization is exponential with the dimensionality of the system under study. We show that the central moments method introduced here can be merged with quantum optimal control, enabling the systematic control of high-dimensional continuous quantum processes. Finally, our quantum optimal control analysis corroborates the existence of a reduced effective dimension of continuous quantum systems.

7.5.2 Introduction

Solving the multi-dimensional Schrödinger equation for continuous systems is as challenging as fundamental to understand many ubiquitous processes in nature, ranging from chemical reactions and quantum information protocols to materials and quantum optical systems and their interface [48; 372; 373]. To efficiently attack this problem it is unavoidable to perform a massive reduction of degrees of freedom, focusing on few relevant ones while trying to keep most possible information on the system. In chemistry, for example, focusing on the classical degrees of freedom only, it is possible to study phenomena ranging from molecular vibrations and rotations [374; 375] to the description of reactions including proton tunneling [376]: methods on the field of classical molecular dynamics are routinely applied to study processes of the order of thousands up to billions of particles [345; 377; 378]. Several methods extending classical dynamics, e.g. by taking into account more degrees of freedom, quantum effects, and semi-classical approaches have been proposed [349–352]. For time-independent systems, more refined successful strategies have been proposed, from the thawed Gaussian [354] and the frozen Gaussian approximations [355], to the Quantized Hamilton Dynamics (QHD), where the Heisenberg equation of motion is applied to derive dynamical equations.
for different powers (and cross terms) of canonical conjugate variables [356–361]; and the Quantal Cumulant Dynamics (QCD) [366–370]. Finally, an alternative successful route to describe many-body quantum systems comes from the class of tensor network methods [379–383], whose application in quantum chemistry problems is currently increasing [384]. However, despite the optimal representation of quantum correlations provided by this class of methods, they suffer from very large local dimensions $d$ of each single interacting body, such as those typically needed for interacting particles in three dimensions. Being able to describe each constituent with a minimal number of local degrees of freedom would result in a potentially game-changing speed-up of such applications.

In this letter, we present a practical approach to describe the time-evolution of high-dimensional (single- and few-body) quantum systems with minimal resources. We capture the real-time dynamics of the central moments, in the canonical coordinates of the instantaneous wave function, and show that only a very limited number of them is needed to be evolved in order to obtain fairly precise results. The approximation in this CentrAl momeNts DYnamics (CANDY) scheme is controlled via the highest order of moments tracked, allowing for an estimation of the error introduced. In particular, we show how to extract the relevant degrees of freedom, and test the accuracy of our approach for one- and two-dimensional problems in comparison to space discretization, while highlighting the reduction in numerical complexity. Moreover, we show that the gain of the central moment approach increases exponentially with the system dimensionality. Finally, we discuss how to merge this approach with optimal control, adopting as a prototypical problem, the optimization of the transfer of a single particle wave function across a two-well potential. Optimal control is nowadays routinely used to control quantum dynamics both in theoretical analysis and experiments [21; 66; 175; 385; 386]. Here, we employ the dCRAB optimal control algorithm, based on an apriori truncation of the dimension of the space of the control functions to a finite value $N_f$ to enhance the algorithm effectiveness. Differently from other successful approaches [87; 92], the dCRAB has the advantage of also being capable to detect the effective number of degrees of freedom in the system: indeed it was shown that for discrete systems of Hilbert space dimension $N$, its overall accuracy scales exponentially with $N_f/N$ [103; 115]. Here we use optimal control not only to obtain an optimal transformation, but also as a diagnostic tool to estimate, for a continuous system, the effective number of independent degrees of freedom required (or effective dimension $N'$), for a given finite precision $\varepsilon$. Indeed, we base our approach on the reasonable physical intuition that in presence of a finite error $\varepsilon$ in the numerical simulations, it exists an effective cutoff of the independent number of variables needed to
describe the system, from $d^D$ to some lower value $N$. In general, the problem of quantum dynamics cast in Heisenberg picture requires one to solve an infinite system of hierarchic differential equations for the moments of position $x$ and momentum $p$. However, we observe that for a fixed finite error $\varepsilon > 0$, only $N$ of them have physical significance. Then, as we will show, it is possible to effectively close the infinite system of hierarchic differential equations, enhancing the efficiency of the system simulation and keeping the error limited. Viceversa, the fact that typical continuous systems are efficiently controllable indicates that their effective dimension $N$ is relatively small [77; 129]. We show that this is indeed the case, and that the error of the simulation becomes minimal when the number of tracked central moments $N_m$ is larger than the effective system dimension $N$, estimated via the dCRAB performance [103; 115].

![Figure 7.2: Log-plot of absolute value of computed (k,l) moments of a wave function $|\psi(t')\rangle$ for a typical instance $t'$ for a double-well tunneling process in $D = 1$.](image)

In the QHD method, the infinite hierarchy of coupled differential equations is truncated, decomposing higher order terms in terms of linear and quadratic terms [356–361]. In the QCD method the approximation enters via the highest order cumulant taken into account in the cumulant expansion [366–370]. The CANDY approach introduced here instead, does not assume the Heisenberg uncertainty to be minimal, and it allows to set arbitrary orders of the expansion without the need of additional analytical calculations. This enables an easy and efficient variation of the number of degrees of freedom tracked, which, in turn, allows to estimate the resulting error of the simulation.
7.5.3 CANDY method

Hereafter we describe the algorithm for a one-dimensional single-body point particle system, while generalization to higher dimensional problems being straightforward. The central moment \( m^{[k,l]} \) of the position \( \hat{x} \) and momentum \( \hat{p} \) operators, representing the quantum mechanical state, can be expressed as the expectation value of the symmetrized product of the position and momentum fluctuation operators \( m^{[k,l]} = \frac{1}{2} \langle \{ (\hat{x} - \mu)^k, (\hat{p} - \eta)^l \} \rangle \equiv \frac{1}{2} \langle \{ \hat{x}^k, \hat{p}^l \} \rangle \), where \( \mu = \langle \hat{x} \rangle \) and \( \eta = \langle \hat{p} \rangle \) are the center of mass in position and momentum space, respectively and \( \{ A,B \} = AB + BA \) is the anti-commutator. The real-time dynamics of the moments is described by the Heisenberg’s equation of motion

\[
\frac{d}{dt} \langle \hat{O} \rangle = i \hbar \langle [ \hat{H}, \hat{O} ] \rangle + \langle \partial \hat{O} / \partial t \rangle,
\]

where we consider a particle-in-a-potential Hamiltonian of the form \( \hat{H} = \hat{p}^2 / 2m + V(\hat{x}) \). The system of differential equations describing the moment dynamics thus reads

\[
\frac{d}{dt} m^{[k,l]} = \frac{i}{\hbar} \langle \{ \hat{p}^2 / 2m_0 + V(x,t), \frac{1}{2} \{ \hat{x}^k, \hat{p}^l \} \} \rangle + \frac{1}{2} \langle \partial \{ \hat{x}^k, \hat{p}^l \} / \partial t \rangle.
\]

In order to expand the right-hand side of Eq. (7.39) into moments \( m^{[k',l']} \) \( k',l' \in \{1,2...,k_{\text{max}}\} \), we assume the potential \( V(x,t) \) to be a finite-order \( w \) polynomial in \( x \); equivalently, we consider a finite number \( w \) of Taylor orders around the system’s center of mass \( \mu(t) \) or around a fixed value constant [358]. The results of the expansion of Eq. (7.39) are presented in Supplementary Material (SM), resulting in a system of coupled equations of motion for \( m^{[k,l]} \), where only a finite number of moments appear in each equation. Specifically, the system can be formally cast as

\[
\frac{d}{dt} m^{[k,l]} = \sum_{k',l'} I_{kl}^{k'l'}(t) m^{[k',l']}
\]

where the sum runs on indices \( k' \) and \( l' \) such that \( k' + l' \leq k + l + w - 2 \). The relatively simple form of Eq.(7.40) follows from the fact that the commutators of the form \( [\hat{x}^k, \hat{p}^l] \) appearing in Eq. (7.39) can be recast as lower order anti-commutators via Heisenberg relations (see SM). The Eq.(7.40) highlights the very well known fact that only harmonic systems (\( w = 2 \)) are closed, while for \( w \geq 3 \), the system of differential equations is formally infinite. However, here we assume that it is sufficient to explicitly track in real-time only a finite subset of moments \( m^{[k,l]}(t) \), specifically those moments that satisfy \( k + l \leq k_{\text{max}} \) for a chosen \( k_{\text{max}} \). In order to approximately close the system of equations for \( N_m \) tracked moments, we need to provide instantaneous estimates for all moments \( m^{[k,l]} \) with \( k_{\text{max}} < k + l \leq k_{\text{max}} + (w - 2) \), or extrapolated moments, as a function of the tracked moments themselves. Here we follow a simple but effective strategy to
extrapolate the missing moments, which follows from the observations of the strong
typical regularity of the higher moments themselves, as highlighted in Fig. 7.2. Indeed,
at any time \( t \), the instantaneous moments for high enough \( k,l \) are well-approximated
by an exponential law \( m^{k,l}(t) \propto \beta^k(t)\gamma^l(t) \). Physically, this reflects the fact that the
low-order moments encode most of the relevant information needed to describe the
dynamics. In conclusion, we extrapolate the missing moments by fitting an exponential
scaling of the tracked moments, via \( \beta \) and \( \gamma \), at any instant \( t \). More details of this
strategy are presented in the SM.

7.5.4 Results

As a prototypical scenario to test our approach, we focus on a two-dimensional quadruple-
well potential and compare the CANDY method versus the exact propagation of the
Schrödinger equation on a discretized real-space grid using the split-operator method.
This problem is relevant for many different platforms for quantum technologies, such
as in the Josephson tunneling [387; 388], or the transport of neutral atoms in optical
lattices [90; 389] and trapped ions [390–392]. We describe the potential landscape via
\( V^{2D}(x,y,t) = \alpha_0 + \alpha_x x^2 + \alpha_y y^2 + \alpha_4 x^4 + \alpha_y y^4 + u_x(t)x + u_y(t)y \) and set \( \hbar = 1 \).
The two time-dependent functions \( u_x(t) \) and \( u_y(t) \) are the control functions to be optimized to
transfer a Gaussian wave-packet across two opposite wells (see Fig. 7.3).

We optimize the state evolution, \( u_x(t) \) and \( u_y(t) \) for \( t \in [0,T] \) using the dCRAB
algorithm [95; 96; 103] and propagate the Schrödinger equation numerically (more
details in SM). Once the optimal control sequences are obtained, we use them to run the
CANDY simulation for the same process and compare the results. The resulting time-
evolution for the first-order spatial moments \( (\mu_x, \mu_y) \) along with the second moments
at \( t = 1 \) (central error cross) are plotted in Fig. 7.3 for the central moments scheme
(red) and the split-operator trajectories (white dashed/ blue), displaying excellent
agreement (plots displaying the accuracy \( \Delta \mu_x(t) \) and \( \Delta \mu_y(t) \) are reported in the SM).

A quantitative comparison of the two methods is reported in Fig. 7.4, where we plot
the relative energy-based error measure \( \Delta E_{\text{rel}} = \frac{1}{T} \int_0^T \frac{|E_{\text{mom}}(t) - E_{\text{ex}}(t)|}{|E_{\text{ex}}(t)|} \, dt \). Here, \( E_{\text{ex}}(t) \)
is the numerically exact energy calculated from the state evolution \( |\psi(t)\rangle \) based on
Schrödinger equation and \( E_{\text{mom}} \) is computed via the central moments propagation. In
Fig. 7.4 we report \( \Delta E_{\text{rel}} \) as a function of the total number of propagated central moments
\( N_m(k_{\text{max}}) \) (red line). The relative error is of the order of 1% for already \( N_m \sim 400 \)
(tracked moments up to \( k_{\text{max}} = 5 \)), while to reach a similar precision with the space
discretization solving exactly the Schrödinger equation, about \( N_{\text{Schr}} = 4 \cdot 10^4 \) points
are necessary, two orders of magnitude more. More precise results can be in principle
Figure 7.3: Comparison of first moments ($\mu_x$, $\mu_y$) for moment scheme (red) and exact dynamics (dashed white) along with second moments (red and blue bars) for $t = 0.425$ and $\alpha_x^2 = \alpha_y^2 = -3, \alpha_x^4 = \alpha_y^4 = 0.5, T = 2.5$ and $k_{\text{max}} = 8$. The contourplot visualize the underlying potential for $u_x = u_y = 0$.

achieved improving the fitting scheme of the extrapolated moments and adopting high precision simulations to control numerical instabilities that might arise due to the exponentially increasing values of the higher moments.

The computational cost for the two methods scales with the number of variables ($N_m \propto (0.5 \times k_{\text{max}}^2)^D$ and $N_{\text{Schr}} = 2 \times N_x^D$), which then play a fundamental role in the comparison. Since both problems are first-order differential linear systems, the (inverse) speedup $\kappa$ can be estimated by the ratio $\kappa = N_m / N_{\text{Schr}} \propto 0.5 \times \left( \frac{0.5 \times k_{\text{max}}^2}{N_x} \right)^D$: as $\left( \frac{0.5 \times k_{\text{max}}^2}{N_x} \right) \ll 1$ exponential speed-ups in $D$ are to be expected. Any algorithm exploited to speed up the numerical integration (e.g. exploiting the sparse matrix structure of the system of equations or parallelization) can be used for both methods. The parameter $k_{\text{max}}$ ultimately controls the accuracy of the simulation, as well as its complexity as $N_m(k_{\text{max}}) = 0.5 \times (k_{\text{max}} + 1) \times (k_{\text{max}} + 2) - 1$ for the one-dimensional case. In the two-dimensional case, the total number of tracked central moments is given by the condition $k_x + l_x \leq k_{\text{max}}$ and $k_y + l_y \leq k_{\text{max}}$ and $N_m(k_{\text{max}}) = (0.5 \times (k_{\text{max}} + 1) \times (k_{\text{max}} + 2))^2 - 1$.

Typical numbers $N_m$ are reported for $k_{\text{max}} \in \{2,3,\ldots,18\}$ in the SM.
7.5 Central moment dynamics of continuous variables quantum systems

We now argue how it is possible to obtain such accurate results with a drastically reduced number of effective degrees of freedom, which in turn leads to the identification of the effective dimension $N^\ast$ of a quantum dynamical process. We then turn our attention to the number of degrees of freedom in the control field that are necessary to control the system up to a given precision $\varepsilon$. As argued in Refs. [95; 96; 103; 115; 116], this is tightly connected to the number of independent degrees of freedom of the system, that might be drastically less than the natural degrees of freedom of the system $N = d^D$. We perform such an analysis for a one-dimensional transfer of a particle in a two-well potential. The goal is to tailor a control function $u(t)$ in such a way that the state transfer is enabled with high fidelity $F$ after a total time $T$ (further details are provided in SM). The results are displayed in Fig. 7.5 (lower panel), where we report the best infidelity $I^* = (1 - F^*)$ obtained as a function of the number of control basis functions $N_f$: it can be infered that $N^\ast \approx 20-27$. This finding agrees very well with the results of Fig. 7.5 (upper panel) where we show the final error of the central moment simulation as a function of $N_m$: for $N_m \approx 20-27$ we reach the minimum signalling that these moments are sufficient to grasp the major features the dynamical process.

**Figure 7.4:** Red: $\Delta E_{rel}$ versus number of moments tracked $N = N_m$ for 2D tunneling process. Blue: $\Delta E_{rel}$ versus number of spatial discretization points $N = N_{Schr}$ for ordinary wave function representation. Note that in order to reach the same $10^{-2}$ error level, two orders of magnitudes more variables have to be taken into account in the ordinary wave function representation.
7.5.6 Discussion

We introduced a novel tool to perform efficient approximate simulations of high-dimensional quantum systems. The efficiency gain is exponential in the dimensionality with respect to the standard grid discretization. We have shown how this approach also offers insights into the effective dimension of the system, that is, the number of independent degrees of freedom of a given problem at fixed precision \( \varepsilon \). This estimate is in agreement with an independent estimate based on quantum optimal control theory. We have also applied dCRAB optimal control directly on the central moments framework to enhance the two-well process and achieved results of similar quality (see SM). Future research will be dedicated to identify and characterize the limitations of this novel approach and the physical foundations of their appearance, e.g., understanding if and when it is not possible to reduce substantially the number of moments to be tracked: we foresee that, similarly as area laws explain the efficiency of tensor network methods, some general laws might be recognized for compressibility in momentum space.

Figure 7.5: Top panel: \( \Delta E_{\text{rel}} \) versus number of moments tracked \( (N_m) \). Lower panel: Optimized infidelities versus range of basis functions subject to search \( (N_f) \). Note that \( \Delta E_{\text{rel}} \) becomes small at around \( N_m \approx N_f \approx 27 \) where also the optimization converges (to \( I^* \approx 10^{-3} \)). As opposed to Fig.3, the analysis is carried out in one dimension \( (D=1) \).
This novel approach can be applied to describe a vast range of different experiments such as ultra cold atoms, trapped ions or neutral atoms [393; 394]. An intuitive argument to justify our expectations that in general $\mathcal{N}$ will not be too high, is that all these systems operate in a very narrow confinement spanned by, on the one hand, the maximal spatial extension and the maximal energy input (e.g. laser power limits) and, on the other hand, the limited experimental ability to resolve the state of the system, and finally the finite total time. These limitations ultimately introduce both infrared and ultraviolet cutoffs dictated by the experimental conditions, reflected by the effective number of independent degrees of freedom, drastically reducing the system’s effective dimension $\mathcal{N}$. Further developments in the direction of merging this approach with tensor network methods might results in a powerful method to describe efficiently high-dimensional many-body quantum systems.

Acknowledgements

The authors acknowledge support from the EU via UQUAM project and the ETN QUSCO, and the Baden-Württemberg Stiftung via Eliteprogramm for Postdocs. S.M. gratefully acknowledges the support of the DFG via a Heisenberg fellowship and the TWITTER project.

7.5.7 Supplementary Material

7.5.7.1 Derivation of moment differential equation

Starting off from Eq. (7.39) for $\dot{m}^{[k,l]}$ in the main text, we consider the kinetic part, the potential part and the partial time derivative separately. The former reads:

$$\dot{m}^{[k,l]}_{\text{kin}} = \frac{i}{4\hbar m_0} \langle [p^2, \{(x - \mu)^k, (p - \eta)^l\}] \rangle .$$ (7.41)

After some calculations, we arrive at:

$$\dot{m}^{[k,l]}_{\text{kin}} = \frac{k}{m_0} m^{[k-1,l+1]} + \frac{k\eta}{m_0} m^{[k-1,l]}$$ (7.42)\[= \frac{k(k-1)\hbar}{4m_0} i[(x - \mu)^{k-2}, (p - \eta)^l]. \tag{7.43}

The part with the partial time derivative

$$\dot{m}^{[k,l]}_{\text{ptd}} = \frac{1}{2} \left\langle \frac{\partial \{(x - \mu)^k, (p - \eta)^l\}}{\partial t} \right\rangle$$ (7.44)
yields
\[
\dot{m}_{\text{ptd}}^{[k,l]} = -k \frac{\eta}{m_0} m^{[k-1,l]} + \left( l(\alpha_1 + 2\alpha_2 \mu + 3\alpha_3 \langle x^2 \rangle + 4\alpha_4 \langle x^3 \rangle + 5\alpha_5 \langle x^4 \rangle + 6\alpha_6 \langle x^5 \rangle) m^{[k,l-1]} \right)
\] (7.45)

Note that the brackets \( \langle x^k \rangle \) can be centralized using:
\[
\langle x^k \rangle = \sum_{j=0}^{k} \binom{k}{j} \mu^j \langle (x - \mu)^{k-j} \rangle
\] (7.46)

(Further note, that the \( j = k - 1 \) summand always reduces to zero.)

The potential part
\[
\dot{m}_{\text{pot}}^{[k,l]} = \frac{i}{2\hbar} \langle [V(x), \{(x - \mu)^k, (p - \eta)^l\}] \rangle
\] (7.47)

might be expressed as:
\[
\dot{m}_{\text{pot}}^{[k,l]} = \frac{i}{2\hbar} \langle \{(x - \mu)^k, [V(x), (p - \eta)^l]\} \rangle
\] (7.48)

Hence for up to six-th order potentials we get:
\[
\dot{m}_{\text{pot}}^{[k,l]} = \frac{i}{2\hbar} \langle \alpha_6 \{(x - \mu)^k, [x^6, (p - \eta)^l]\} \rangle
\] (7.49)
\[
+ \alpha_5 \{(x - \mu)^k, [x^5, (p - \eta)^l]\} \rangle
\] (7.50)
\[
+ \alpha_4 \{(x - \mu)^k, [x^4, (p - \eta)^l]\} \rangle
\] (7.51)
\[
+ \alpha_3 \{(x - \mu)^k, [x^3, (p - \eta)^l]\} \rangle
\] (7.52)
\[
+ \alpha_2 \{(x - \mu)^k, [x^2, (p - \eta)^l]\} \rangle
\] (7.53)
\[
+ \alpha_1 \{(x - \mu)^k, [x^1, (p - \eta)^l]\} \rangle
\] (7.54)
Next we express the commutators in terms of anti-commutators using the recursion relation being discussed in the next paragraph.

\[
\hat{m}_{\text{pot}}^{[k,l]} = \frac{i}{2\hbar} < \alpha_6 \{ (x - \mu)^k, 3i\hbar \{ x^5, (p - \eta)^{l-1} \} \} + 5i(\ell - 1)(\ell - 2)\hbar^3 \{ x^3, (p - \eta)^{l-3} \} \]

(7.55)

\[
+ 3i(\ell - 1)(\ell - 2)(\ell - 3)(\ell - 4)\hbar^5 \{ x, (p - \eta)^{l-5} \} \]

(7.56)

\[
+ \alpha_5 \{ (x - \mu)^k, 2.5i\hbar \{ x^4, (p - \eta)^{l-1} \} \}
+ 2.5i(\ell - 1)(\ell - 2)\hbar^3 \{ x^2, (p - \eta)^{l-3} \}
+ 1i(\ell - 1)(\ell - 2)(\ell - 3)(\ell - 4)\hbar^5 (p - \eta)^{l-5} \]

(7.57)

\[
+ \alpha_4 \{ (x - \mu)^k, 2i\hbar \{ x^3, (p - \eta)^{l-1} \} \} + i(\ell - 1)(\ell - 2)\hbar^3 \{ x, (p - \eta)^{l-3} \} \]

(7.58)

\[
+ \alpha_3 \{ (x - \mu)^k, 1.5i\hbar \{ x^2, (p - \eta)^{l-1} \} \} + 0.5i(\ell - 1)(\ell - 2)\hbar^3 (p - \eta)^{l-3} \]

(7.59)

\[
+ \alpha_2 \{ (x - \mu)^k, i\hbar \{ x, (p - \eta)^{l-1} \} \}
+ \alpha_1 \{ (x - \mu)^k, i\hbar (p - \eta)^{l-1} \} \]

(7.60)

Unfolding each of the six terms separately, we start off with the \( \alpha_1 \) term:

\[
- \frac{\alpha_1}{2} l \langle \{ (x - \mu)^k, (p - \eta)^{l-1} \} \rangle \]

(7.65)

The \( \alpha_2 \) term reads:

\[
- \frac{\alpha_2}{2} l < 2 \{ (x - \mu)^{k+1}, (p - \eta)^{l-1} \} + 2\mu \{ (x - \mu)^k, (p - \eta)^{l-1} \}
- i(\ell - 1)\hbar [(x - \mu)^k, (p - \eta)^{l-2}] > \]

(7.66)

The \( \alpha_3 \) term reads:

\[
- \frac{\alpha_3}{4} l < 6 \{ (x - \mu)^{k+2}, (p - \eta)^{l-1} \} + 12\mu \{ (x - \mu)^{k+1}, (p - \eta)^{l-1} \}
+ 6\mu^2 \{ (x - \mu)^k, (p - \eta)^{l-1} \} - 2(\ell - 1)(\ell - 2)\hbar^2 \{ (x - \mu)^k, (p - \eta)^{l-3} \}
- 6i(\ell - 1)\hbar [(x - \mu)^{k+1}, (p - \eta)^{l-2}] - 6\mu i(\ell - 1)\hbar [(x - \mu)^k, (p - \eta)^{l-2}] > \]

(7.67)
The $\alpha_4$ term reads:

$$-\frac{\alpha_4}{2} l < 4 \{(x - \mu)^{k+3}, (p - \eta)^{l-1}\} + 12 \mu \{(x - \mu)^{k+2}, (p - \eta)^{l-1}\}$$

$$+ 12 \mu^2 \{(x - \mu)^{k+1}, (p - \eta)^{l-1}\} + 4 \mu^3 \{(x - \mu)^k, (p - \eta)^l\}$$

$$- 4(l - 1)(l - 2) \hbar^2 \{(x - \mu)^{k+1}, (p - \eta)^{l-3}\} - 4 \mu(l - 1)(l - 2) \hbar^2 \{(x - \mu)^k, (p - \eta)^{l-3}\}$$

$$- 6i(l - 1) \hbar[(x - \mu)^{k+2}, (p - \eta)^{l-2}] - 12 \mu i(l - 1) \hbar[(x - \mu)^{k+1}, (p - \eta)^{l-2}]$$

$$- 6 \mu^2 i(l - 1) \hbar[(x - \mu)^k, (p - \eta)^{l-2}] + i(l - 1)(l - 2)(l - 3) \hbar^3 [(x - \mu)^k, (p - \eta)^{l-4}]$$

The $\alpha_5$ term reads:

$$-\frac{\alpha_5}{4} l < 10 \{(x - \mu)^{k+4}, (p - \eta)^{l-1}\} + 40 \mu \{(x - \mu)^{k+3}, (p - \eta)^{l-1}\}$$

$$- 20(l - 1)(l - 2) \hbar^2 \{(x - \mu)^{k+2}, (p - \eta)^{l-3}\} + 60 \mu^2 \{(x - \mu)^{k+2}, (p - \eta)^{l-1}\}$$

$$- 40 \mu(l - 1)(l - 2) \hbar^2 \{(x - \mu)^{k+1}, (p - \eta)^{l-3}\} + 40 \mu^3 \{(x - \mu)^{k+1}, (p - \eta)^{l-1}\}$$

$$+ 2(l - 1)(l - 2)(l - 3)(l - 4) \hbar^4 \{(x - \mu)^k, (p - \eta)^{l-5}\}$$

$$- 20 \mu^2(l - 1)(l - 2) \hbar^2 \{(x - \mu)^k, (p - \eta)^{l-3}\}$$

$$+ 10 \mu^4 \{(x - \mu)^k, (p - \eta)^{l-1}\} - 20i(l - 1) \hbar[(x - \mu)^{k+3}, (p - \eta)^{l-2}]$$

$$- 60i \mu(l - 1) \hbar[(x - \mu)^{k+2}, (p - \eta)^{l-2}] + 10i(l - 1)(l - 2)(l - 3) \hbar^3 [(x - \mu)^{k+1}, (p - \eta)^{l-4}]$$

$$- 60i \mu^2(l - 1) \hbar[(x - \mu)^{k+1}, (p - \eta)^{l-2}] + 10i \mu(l - 1)(l - 2)(l - 3) \hbar^3 [(x - \mu)^k, (p - \eta)^{l-4}]$$

$$- 20i \mu^3(l - 1) \hbar[(x - \mu)^k, (p - \eta)^{l-2}]$$
The $\alpha_6$ term reads:

\begin{align}
+\alpha_6 l ( -6m_{[k+5,l-1]} + 30\mu m_{k+4,l-1} + 20(l-1)(l-2)\hbar^2 m_{[k+3,l-3]} & \quad (7.70) \\
- 60\mu^2 m_{[k+3,l-1]} + 60\mu \hbar^2 (l-1)(l-2)m_{[k+2,l-3]} - 60\mu^3 m_{[k+2,l-1]} & \\
- 6(l-1)(l-2)(l-3)(l-4)\hbar^4 m_{[k+1,l-5]} + 60\mu^2 (l-1)(l-2)\hbar^2 m_{[k+1,l-3]} & \\
- 30\mu^4 m_{[k+1,l-1]} - 6\mu (l-1)(l-2)(l-3)(l-4)\hbar^4 m_{[k,l-5]} & \\
+ 20\mu^3 (l-1)(l-2)\hbar^2 m_{[k,l-3]} - 6\mu^5 m_{[k,l-1]} & \\
+ 7.5i(l-1)\hbar [(x-\mu)_{k+4},(p-\eta)_{l-2}] & \\
+ 30i\mu (l-1)\hbar [(x-\mu)_{k+3},(p-\eta)_{l-2}] & \\
- 7.5i(l-1)(l-2)(l-3)\hbar^3 [(x-\mu)_{k+2},(p-\eta)_{l-4}] & \\
+ 45i\mu^2 (l-1)\hbar [(x-\mu)_{k+2},(p-\eta)_{l-2}] & \\
- 15i\mu (l-1)(l-2)(l-3)\hbar^3 [(x-\mu)_{k+1},(p-\eta)_{l-4}] & \\
+ 30i\mu^3 (l-1)\hbar [(x-\mu)_{k+1},(p-\eta)_{l-2}] & \\
- 0.5i(l-1)(l-2)(l-3)(l-4)(l-5)\hbar^5 [(x-\mu)_k,(p-\eta)_{l-6}] & \\
+ 7.5i\mu^2 (l-1)(l-2)(l-3)\hbar^3 [(x-\mu)_k,(p-\eta)_{l-4}] & \\
- 7.5i\mu^4 (l-1)\hbar [(x-\mu)_k,(p-\eta)_{l-2}] & 
\end{align}

Note, that the control input $\alpha_1 = u$ does not enter the dynamics directly. However, we must simulate $\mu$ and $\eta$ as well and at this point $\alpha_1$ comes into play:

\begin{align}
\dot{\mu} = \frac{1}{\hbar m_0} \eta & \quad (7.71) \\
\dot{\eta} = -\langle \frac{\partial V}{\partial x} \rangle = -\alpha_1 - 2\alpha_2 \mu - 3\alpha_3 \langle x^2 \rangle - 4\alpha_4 \langle x^3 \rangle - 5\alpha_5 \langle x^4 \rangle - 6\alpha_6 \langle x^5 \rangle & \quad (7.72)
\end{align}

As shown by Eq. (7.46) the $\langle x^n \rangle$ terms can be easily centralized.

In order to facilitate to grasp the idea, we consider an example where the basic parameters are set to be $d = 4$, $k_{max} = 8$. Fig. 7.6 displays a cartoon of the involved moments. The moments of order up to $k_{max}$ are denoted by the blue dots and will be propagated. For every instance in time, further moments outside of this set of moments, have to be approximated. These are marked by the green dots. This approximation is based on an exponential fit of tracked moments close to the boundary (marked by orange line). Hence, the extrapolated moments are not propagated in time but evaluated instantaneously. Finally, we like to remark that once $k_{max}$ is very large ($k_{max} \approx (d-2) \times N_t$), the lowest moments are not affected by any error as it travels towards the origin at
(d−2) orders per timestep $\Delta t$ and only arrives there after the time-evolution has finished.

To come up with the system of differential equations for the anti-commutators, occurring commutators have to be reformulated in terms of these, too. This can provably be done using the following recursion relation:

$$\begin{align*}
\left[\hat{x}^k, \hat{p}^l\right] &= -\sum_{q'=1}^{u_o} \frac{(-1)^{q'}}{(2q' - 1)! (l - 2q' + 1)!(k - 2q' + 1)!} \times \\
&\quad \frac{\hbar^{2q'-1}}{2} \left\{ \hat{x}^{k-2q'+1}, \hat{p}^{l-2q'+1} \right\} \\
&\quad - \sum_{q'=1}^{u_e} \frac{(-1)^{q'}}{(2q')! (l - 2q')!(k - 2q')!} \times \\
&\quad \frac{\hbar^{2q'}}{2} \left[ \hat{x}^{k-2q'}, \hat{p}^{l-2q'} \right]
\end{align*}$$

(7.73)

Here the upper summation borders are respectively:

$$u_o = \begin{cases} 
\frac{\min(k,l)+1}{2} & \text{if } \min(k,l) \text{ is odd} \\
\frac{\min(k,l)}{2} & \text{if } \min(k,l) \text{ is even}
\end{cases}$$

(7.74)

$$u_e = \begin{cases} 
\frac{\min(k,l)-1}{2} & \text{if } \min(k,l) \text{ is odd} \\
\frac{\min(k,l)}{2} - 1 & \text{if } \min(k,l) \text{ is even}
\end{cases}$$

(7.75)

Finally, we like to add a remark on the energy error. It can be expressed in terms of moments in (for $w = 4$, $D = 1$) the following way:

$$E = 0.5(m^{[0,2]} + \eta^2) + \alpha_1 \mu + \alpha_2 (m^{[2,0]} + \mu^2)$$

(7.76)

$$+ \alpha_4 (m^{[4,0]} + 4\mu m^{[3,0]} + 6\mu^2 m^{[2,0]} + \mu^4)$$

(7.77)

For $k_{\text{max}} = \{2,(3)\}$, the moments $m^{[4,0]}(m^{[3,0]})$ were taken into account based on the instantenous extrapolated ones since $m^{[4,0]}(m^{[3,0]})$ are not explicity propagated.

7.5.7.2 One dimensional optimized state transfer in two-well potential

The visualization in Fig. 7.2 and the optimal control analysis (depicted in Fig. 7.5 lower panel) are based on a state transfer across the potential of the form
7.5 Central moment dynamics of continuous variables quantum systems

Figure 7.6: Example of moment differential equation scheme for $k_{\text{max}}=8$ and $d=4$. The blue dots represent the to be propagated moments and the green ones the to be extrapolated moments. The orange line separates these two sets and is defined as $k+l = k_{\text{max}}$. The dashed arrows indicate which moments contribute to the time derivative of the exemplary selected moment $m_{[k=3,l=5]}$. For this particular moment, the to be extrapolated moments $m_{[k=5,l=4]}^e$ and $m_{[k=6,l=4]}^e$ contribute (among nine others). They are, in turn, obtained from a fit of tracked moments close to do boundary, yielding a closed system of differential equations.

$$V(x,t) = V_0(x) + u(t) \times x$$
$$= \alpha_0 + \alpha_2 \times x^2 + \alpha_4 \times x^4 + u(t) \times x.$$  \hspace{1cm} \text{(7.78)}

The $\alpha_i$ are set to be $\alpha_0 = 4.5, \alpha_2 = -3, \alpha_4 = 0.5$. The control function $u(t)$ is optimized using the dCRAB algorithm [95; 96; 103]. Initial state $|\psi_0\rangle$ and target state $|\phi\rangle$ are defined the following way:
\[ |\psi_0\rangle = |\Psi(t = 0)\rangle = \frac{1}{\sqrt{2}} (|\zeta_0\rangle + |\zeta_1\rangle) \]  
\[ |\phi\rangle = |\Psi_{\text{target}}(t = T)\rangle = \frac{1}{\sqrt{2}} (|\zeta_0\rangle - |\zeta_1\rangle). \]  

Here, \( |\zeta_0\rangle \) and \( |\zeta_1\rangle \) are ground- and first excited state of \( H_0 = p^2/2m + V_0(x) \). Using the definition of Eqs. (7.79), (7.80), we find that \( |\psi_0\rangle \) \( (|\phi\rangle) \) has almost all probability mass on the left (right) well.

The core idea of the dCRAB algorithm is to expand the control field in a suitable function basis

\[ u(t) = \sum_{i=1}^{N_b} c_i f_i(t) \]  

and subsequently re-initiate the basis functions \( f_i \) by randomly drawing them from a finite set \( \Omega \). The cardinality of it being \( N_f = |\Omega| \). If the function basis is chosen to be orthonormal, \( N_f \) independent degrees of freedom can be adjusted by the pulse. Fig. 1 displays the representation of an intermediate state \( |\psi(t')\rangle \) from an optimized (propagating the system using \( u^*(t) \)) evolution from \( |\psi_0\rangle \) to \( |\phi\rangle \). The parameters are: \( t' = 0.425, T = 1.5, |u(t)| \leq 50 \forall t; \) and \( \alpha_0 = 4.5, \alpha_2 = -3, \alpha_4 = 0.5 \). For all optimized pulses we are using the Fourier basis.

The results from Fig. 7.5 (lower panel) were obtained from optimizations setting various cardinalities \( N_f \) and allowing up to 100 super-iterations (for details see [103]) with 40 function calls each. The number of basis functions \( N_b \) to be visited at each super-iteration was set to be \( N_b = 2 \).

### 7.5.7.3 Propagation of Schrödinger Equation as a benchmark

To benchmark the results from the moment scheme, the state was propagated using the Schrödinger Equation via the split operator method. The core idea is to take the time evolution operator \( U(t, t + \Delta t) = e^{-i/\hbar H(t)\Delta t} \), which extends for the continous space problem (setting \( \hbar = 1 \)) to

\[ U(t, t + \Delta t) = e^{i/4\left(\frac{\Delta^2}{\alpha^2}\right)\Delta t} \cdot e^{-iV(x,t)\Delta t} \cdot e^{i/4\left(\frac{\Delta^2}{\beta^2}\right)\Delta t}, \]  

and use (discrete) Fourier transforms to alternate between position (\( x \)) and momentum.
(p) space. That way only diagonal operators have to be applied. To do the numerics, position and momentum space has to be discretized, while defining one of the two, also fixes the other. We used equidistant spatial discretization $x_i \in \{-x_{\text{max}}, -x_{\text{max}} + \Delta x, \ldots, x_{\text{max}}\}$, where $\Delta x = 2x_{\text{max}}/N_{\text{Schr}}$. The two-dimensional space representation (D=2) extends straightforwardly.

### 7.5.7.4 Inverse optimal control application on moments

As addressed in the discussion, we also did the test to directly work with the moment framework as the to be optimized variables via dCRAB. The goal set for the algorithm was to match the moments of the goal wave function ($|\phi\rangle$) possibly close. The resulting optimized pulse $u^*(t)$ was then evaluated along the lines of the previous two sections. As a result, the best obtained infidelity has been $I^* = 0.016$, which is of the same order as the error $\Delta E_{\text{rel}}(N_m)$ for the moments. Again, the number of moments involved in this optimization is significantly lower compared to $N_{\text{Schr}}$, yielding a large speed-up.

### 7.5.7.5 Number of moments as a function of chop-off level $k_{\text{max}}$ and dimension D
Table 7.1: Number of moments $N_m$ to be propagated as a function of chop-off level $k_{max}$ and space dimension $D$.

<table>
<thead>
<tr>
<th>$k_{max}$</th>
<th>$N_m(D=1)$</th>
<th>$N_m(D=2)$</th>
<th>$N_m(D=3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5</td>
<td>35</td>
<td>215</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>99</td>
<td>999</td>
</tr>
<tr>
<td>4</td>
<td>14</td>
<td>224</td>
<td>3374</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>440</td>
<td>9260</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>783</td>
<td>21951</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
<td>1295</td>
<td>46655</td>
</tr>
<tr>
<td>8</td>
<td>44</td>
<td>2024</td>
<td>91124</td>
</tr>
<tr>
<td>9</td>
<td>54</td>
<td>3024</td>
<td>166374</td>
</tr>
<tr>
<td>10</td>
<td>65</td>
<td>4355</td>
<td>287495</td>
</tr>
<tr>
<td>11</td>
<td>77</td>
<td>6083</td>
<td>474551</td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td>8280</td>
<td>753570</td>
</tr>
<tr>
<td>13</td>
<td>104</td>
<td>11024</td>
<td>1157624</td>
</tr>
<tr>
<td>14</td>
<td>119</td>
<td>14399</td>
<td>1727999</td>
</tr>
<tr>
<td>15</td>
<td>135</td>
<td>18495</td>
<td>2515455</td>
</tr>
<tr>
<td>16</td>
<td>152</td>
<td>23408</td>
<td>3581576</td>
</tr>
<tr>
<td>17</td>
<td>170</td>
<td>29240</td>
<td>5000210</td>
</tr>
<tr>
<td>18</td>
<td>189</td>
<td>36099</td>
<td>6858999</td>
</tr>
</tbody>
</table>

7.5.7.6 Difference Moment trajectory versus exact Schrödinger trajectory from Fig. 7.3

The differences between the moment trajectories $\mu_x^{\text{mom}}(t)$, $\mu_y^{\text{mom}}(t)$ and the respective exact Schrödinger trajectories $\mu_x^{\text{Schr}}(t)$, $\mu_y^{\text{Schr}}(t)$ (as displayed in Fig. 7.3 in the main text) are denoted by

\[
\Delta \mu_x(t) = \mu_x^{\text{mom}}(t) - \mu_x^{\text{Schr}}(t) 
\]

\[
\Delta \mu_y(t) = \mu_y^{\text{mom}}(t) - \mu_y^{\text{Schr}}(t). 
\]
The two quantities are plotted in Fig. 7.7.

\[ \Delta \mu_x, \Delta \mu_y \]

\[ \text{Time } t \]

**Figure 7.7:** Differences between \( \mu_{\text{mom}}^x(t) \), \( \mu_{\text{mom}}^y(t) \) and the respective exact Schrödinger trajectories \( \mu_{\text{Schr}}^x(t) \), \( \mu_{\text{Schr}}^y(t) \) as an insight on Fig. 7.3 from the main text. \( \Delta \mu_x(t) \) (\( \Delta \mu_y(t) \)) is shown on the full (dashed) line.
7.6 Appendix

7.6.1 Appendix A: Alternative closure scheme

The closure can be done, as shown in the publication, via the exponential fitting of the scaling of lower order moments. The instantaneously evaluated higher moments shall be denoted by \( m_{e_{k',l'}} \), \( k' + l' > k_{\text{max}} \). In essence, we evaluate \( m_e \) using a nonlinear function \( g(m) \):

\[
m_e = g(m)
\]  

(7.85)

Looking at the precise implementation of the moment framework, it turns out that the system of differential equations is linear (which is expected as the Schrödinger Equation, where its being derived from, is linear, too):

\[
\dot{m}_{n \times 1} = A_{n \times n}m_{n \times 1} + B_{n \times p}m_{p \times 1}^e
\]

(7.86)

The superscripts indicate the dimensions of the vectors and matrices. Here, \( n \) is the number of moments tracked and \( p \) is the resulting number of to be extrapolated moments. Here we will show the avenue about how one could obtain a different closure scheme. The main idea is to not only evaluate \( m_e \) as in Eq. (7.85), but dynamically simulate the \( m_e \), too. Therefore we apply the time derivative of Eq. (7.85), using the chain rule:

\[
\dot{m}_e = \frac{\partial g}{\partial m} \frac{\partial m}{\partial \dot{t}}
\]

(7.87)

Consequently, we have to solve

\[
\dot{m} = Am + Bm_e
\]

(7.88)

\[
\dot{m}_e = J(m) \dot{m}.
\]

(7.89)

Since \( g(m) \) from Eq. (7.85) is an approximation, so is \( J(m) \). Without loss of generality, we can say that the \( \tilde{J} \) becomes exact, once we add some, a priori unknown, correction \( C \). That way, Eq. (7.89) becomes:

\[
\dot{m}_e = \tilde{J}(m) \dot{m} = (J(m) + C) \dot{m}
\]

(7.90)
After plugging in Eq. (7.88), we arrive at

\[ \dot{m}_e = (J(m) + C)(Am + Bm_e). \]  

(7.91)

What remains, is to obtain \( C \). Therefore two routes are possible: \( C \) could either be estimated if a trial trajectory including the exact moments is available. As a second option, which we will elaborate on a bit more, \( C \) could be designed in the context of the entire system of differential equations (Eq. (7.88) and (7.91)). Therefore, we rewrite it in a block form:

\[
\begin{pmatrix}
\dot{m} \\
\dot{m}_e
\end{pmatrix}
= 
\begin{pmatrix}
A & B \\
(J(m) + C)A & (J(m) + C)B
\end{pmatrix}
\begin{pmatrix}
m \\
m_e
\end{pmatrix}.
\]

(7.92)

The dimensions of the blocks are \( n \times n \) (upper-left), \( n \times p \) (upper-right), \( p \times n \) (lower-left), and \( p \times p \) (lower-right). In the case that \( C \) could be perfectly determined, the matrix in Eq. (1.47) would resemble the one of a bigger matrix \( \bar{A}^{[n+p,n+p]} \). (This comes from the fact, that \( A^{[\infty,\infty]} \) exactly describes the dynamics). Looking at the eigenspectrum of the matrix \( \bar{A} \), an interesting observation can be made: As depicted in Fig. 7.8, all the eigenvalues \( \lambda_i \) come in sets of four. So for every \( \lambda_i \), there exists further eigenvalues \( \lambda_i^* \), \( -\lambda_i \), and \( -\lambda_i^* \). The symmetry with respect to the imaginary part can be explained by the fact that all unitary dynamics has to be periodical (under the mild assumption that all transition frequencies are rational numbers [160]). However, the real part symmetry can not be easily understood although it is always present. One idea might be to use this fact as a design principle to come up with an ansatz for the correction matrix \( C \).
Figure 7.8: Eigenspectrum of matrix $\bar{A}$. Its double symmetric property could lead to an approximation ansatz.
7.6 Appendix

7.6.2 Appendix B: Energy fluctuation in terms of moments

Besides the energy $E$ itself, its fluctuation $\Delta E$ can be expressed in terms of moments, too. For the case $D = 1$, we start off from its definition in bra-ket notation:

$$\Delta E = \sqrt{\langle \psi(t)|H(t)^2|\psi(t)\rangle - \langle \psi(t)|H(t)|\psi(t)\rangle^2} \quad (7.93)$$

Plugging in the Hamiltonian for potentials up to $d = 4$ and recalling the definition of the moments from Eq. (7.11), after some shuffling we arrive at:

$$\Delta E = \left( \frac{1}{4} (m^{[0,4]} + 4\eta m^{[0,3]} + 6\eta^2 m^{[0,2]} + \eta^4) ight.
\quad + \alpha_4^2 (m^{[8,0]} + 8\mu m^{[7,0]} + 28\mu^2 m^{[6,0]} + 56\mu^3 m^{[5,0]} + 70\mu^4 m^{[4,0]} + 56\mu^5 m^{[3,0]} + 28\mu^6 m^{[2,0]} + \mu^8)
\quad + 2\alpha_2 \alpha_4 (m^{[6,0]} + 6\mu m^{[5,0]} + 15\mu^2 m^{[4,0]} + 20\mu^3 m^{[3,0]} + 15\mu^4 m^{[2,0]} + \mu^6)
\quad + 2\alpha_1 \alpha_4 (m^{[5,0]} + 5\mu m^{[4,0]} + 10\mu^2 m^{[3,0]} + 10\mu^3 m^{[2,0]} + \mu^5)
\quad + \alpha_4^2 (m^{[4,0]} + 4\mu m^{[3,0]} + 6\mu^2 m^{[2,0]} + \mu^4)
\quad + 2\alpha_1 \alpha_2 (m^{[3,0]} + 3\mu m^{[2,0]} + \mu^3)
\quad + \alpha_1^2 (m^{[2,0]} + \mu^2)
\quad + \frac{\alpha_1}{2} (2m^{[1,2]} + 4\eta m^{[1,1]} + 2\mu m^{[0,2]} + 2\mu^2 )
\quad + \frac{\alpha_2}{2} (2m^{[2,2]} + 4\mu m^{[1,2]} + 4\eta m^{[2,1]} + 8\mu \eta m^{[1,1]} + 2\eta^2 m^{[2,0]} + 2\mu^2 m^{[0,2]} + 2\mu^2 \eta^2)
\quad + \frac{\alpha_4}{2} (2m^{[4,2]} + 8\mu m^{[3,2]} + 4\mu m^{[4,1]} + 12\mu^2 m^{[2,2]} + 16\mu \eta m^{[3,1]} + 2\eta^2 m^{[4,0]} + 8\mu^3 m^{[1,2]} + 24\mu^2 \eta m^{[2,1]} + 8\mu^2 \eta^2 m^{[2,0]} + 12\mu^2 \eta^2 m^{[2,0]} + 2\mu^4 m^{[0,2]} + 16\mu^3 \eta m^{[1,1]} + 2\mu^4 \eta^2)
\quad - \left( \frac{1}{2} (m^{[0,2]} + \eta^2) + \alpha_1 + \alpha_2 (m^{[2,0]} + \mu^2)
\quad + \alpha_4 (m^{[4,0]} + 4\mu m^{[3,0]} + 6\mu^2 m^{[2,0]} + \mu^4) \right)^{1/2} \quad (7.94)$$
8 Robust optimal control of single qubit operations

8.1 Abstract

A number of experimental realizations of quantum information protocols based on qubits either do not allow for precise enough measurement of the energy transitions or short time-scale drifts require the protocol to be robust for a distribution of detunings. In this scenario, carefully tailored pulses, which allow for the desired transformation regardless of drifts or precise knowledge of the system Hamiltonian parameters are highly demanded. In this section we show how gate fidelities above 99\% can be obtained when the variability of single qubit detunings are large, even exceeding the magnitudes of the control inputs. To achieve this goal we use a hybrid method combining two well established tools from optimal control theory: Direct Multiple Shooting and the CRAB algorithm.

8.2 Introduction

Any physical implementation of a quantum computing platform is prone to errors. Some errors, such as spin flips, can be cured using error correction protocols and dynamical decoupling techniques [395; 396]. However, there are nevertheless errors that cannot be addressed using generic, well established, correction or suppression methods. One kind of these errors are larger drifts in the involved energy levels that arise when using Zeeman splitting of degenerate states, which are for example used in solid state implementations [113; 397]. The external magnetic field might vary in strength and direction and nearby nuclear spins can add distortions, too. Furthermore atmospheric pressure and temperature fluctuations leads to non-stable conditions in most setups. In order to still run quantum computational algorithms on the system, robust pulses may be employed. Robustness against distortions on the control inputs have been studied for quantum transport [398] and furthermore also the presence of leakage when
Robust optimal control of single qubit operations

Synthesizing quantum gates for Josephson charge qubits can be cured [218; 399–401]. High fidelities can be reached since the desired quantum dynamics is obtained for a range of physical configurations, hence reaching the target almost independent of the occurring error. Another step in this direction has been recently demonstrated [66] theoretical and experimentally for NV centers in diamond. Starting off from the results of that paper, we present an optimization method that is capable of optimizing for larger detuning variations.

Another platform where robust pulses are highly needed is in the field of NMR: In [87; 402–404] the design of robust pulses with respect to frequency detunings and amplitude variations is discussed. In this context, broadband excitation shall be achieved while being able to compensate for $B_1$ field inhomogeneities (for a brief introduction into the system see also Sec. 2.4.2). Finding robust pulses yielding high precision, is a challenging task. We tackle this by using a direct sequential multiple shooting method [405], parameterizing the pulses according to the CRAB algorithm [95; 96]. This scheme is particularly suitable up to medium dimensional problems where it shows fast convergence to an optimum because it can address all relevant feature of the dynamics directly. Furthermore, it is fully parallelizeable and hence can yield full speed-ups on clusters.

This section is organized as follows: We define the control goal and system Hamiltonian in Sec. 8.3 and show that the system is controllable in Sec. 8.4. Solving the control problem we introduce our proposed algorithm in Sec. 8.5 and present the results for two different distributions of detunings in Sec. 8.6. We conclude our work in Sec. 8.7.

### 8.3 Problem definition

We consider the problem of a creating a $\pi$ gate on a detuned spin-$\frac{1}{2}$ system. However, any other gate can be synthesized, too. The target gate reads

$$U_g = e^{-\frac{i}{4}\pi \sigma_x}. \tag{8.1}$$

The Pauli operators are denoted by $\sigma_x$, $\sigma_y$ and $\sigma_z$. The system Hamiltonian can be written as

$$H^{(i)} = H_d^{(i)} + u_x(t)H_{x,x}^{(i)} + u_y(t)H_{x,y}^{(i)} \quad i \in \{1,2,...,N\} \tag{8.2}$$

$$= \delta_i \sigma_z + u_x(t)\eta \sigma_x + u_y(t)\eta \sigma_y, \tag{8.3}$$
where the subscript $d$ represents the drift Hamiltonian, e.g. non-controllable dynamics, and $c$ stands for the control Hamiltonian. The parameter $\delta_i$ brings in the uncertainty of experimental realizations of the system for which we are searching robust control pulses $u_x(t)$ and $u_y(t)$ with amplitudes $\eta$ that achieve the gate in Eq. (8.1), in fixed total time $T$. Thereby a finite number $N$ of different detunings will be regarded for the optimization. However, as we show in the results section, the finite number of detunings considered is sufficient to find high fidelity pulses for a continuous range of detunings.

Owing to experimental constraints, we assume that the controls are bound according to

$$|u_x(t) + u_y(t)| \leq b_u \forall t \in [0,T].$$

(8.4)

Without loss of generality, we can set $b_u = 1$ since we can map any other value via a canonical transformation between energy and time. When the number of different systems $H^{(i)}$ to be optimized for is small as well as the ratio $\frac{\max_i |\delta_i|}{\eta}$, the problem of finding high fidelity pulses becomes straightforward as rectangular pulses as an outcome from analytical calculations already lead to satisfying fidelities (see also Sec. 6.2.6.7). In turn, we are considering a broad distribution of detunings $\delta_i$ in a way that their range even exceeds $\eta$. Therefore advanced optimal control strategies must be considered when aiming for high precision control in reasonable total transfer time.

### 8.4 Controllability analysis

Before turning towards the optimal control strategy, we answer the question on whether different drift Hamiltonians allow for joint controllability, that is, are there sufficiently many Lie group generators to achieve $N$ times the su(2) algebra, which is the minimum requirement for full gate control [406]?

Our system consists of $N$ different replica of qubits that do not interacting with each other. However, their dynamics is still coupled via the shared control fields. For our study, we choose to define as the basis elements of the Lie algebra the direct sum over all subsystems of the various parts in the Hamiltonian

$$A := \bigoplus_{i=1}^{N} H_{d}^{(i)} \quad B := \bigoplus_{i=1}^{N} H_{c,x}^{(i)} \quad C := \bigoplus_{i=1}^{N} H_{c,y}^{(i)}$$

(8.5)
where $\oplus$ is the direct sum. In matrix form these are represented as

$$A = \begin{pmatrix} \delta_1 \sigma_z & 0 & \ldots & 0 \\ 0 & \delta_2 \sigma_z & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \delta_n \sigma_z \end{pmatrix}, B = \begin{pmatrix} \sigma_x & 0 & \ldots & 0 \\ 0 & \sigma_x & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_x \end{pmatrix}, C = \begin{pmatrix} \sigma_y & 0 & \ldots & 0 \\ 0 & \sigma_y & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_y \end{pmatrix}.$$  

Under the constraint that population may not transfer from one block to another, one can think of a joint Hamiltonian defined as $\tilde{H} = A + u_x(t) \cdot B + u_y(t) \cdot C$. What remains is to determine the number of generatable basis elements of the dynamical Lie algebra $g(N)$ as a function of $N$ different detunings via repeated application of Lie brackets on $A, B$ and $C$. In order to find this relation we start by considering the case $N=2$ and we will generalize from there on. When applying the commutators on all possible combinations of $A, B$ and $C$, we get for $N=2$:

$$D_1^{(N=2)} = [A, B] = 2i\delta_1 \begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_y \end{pmatrix},$$  

$$(8.7)$$

$$D_2^{(N=2)} = [A, C] = -2i\delta_1 \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x \end{pmatrix},$$  

$$(8.8)$$

$$D_3^{(N=2)} = [B, C] = 2i \begin{pmatrix} \sigma_z & 0 \\ 0 & \sigma_z \end{pmatrix}.$$  

$$(8.9)$$

We can see that these are, together with $A, B$ and $C$, the only linear independent elements generatable. For $N=3$, one further level of repeated commutators can be independently generated, yielding $A, B, C, D_1^{(N=3)}, D_2^{(N=3)}, \ldots, D_6^{(N=3)}$ as a basis for the Lie algebra. With the same geometrical argument, we can understand that $g(N)$ scales as $g(N) = 3 \cdot N$. Hence, the system remains fully controllable although one set of pulses has to generate all unitaries simultaneously. However, based on this finding, we can only guarantee to achieve arbitrary gates with unconstrained control, meaning no bounds on the control should be present and the total time can become inconveniently large. In order to ensure experimental feasibility, we will nevertheless apply constraints...
on the pulses and set the total time fixed. That way, the problem becomes non trivial. But, as we show in the next section, high precision pulses can still be found.

8.5 Optimal control strategy

We are using a direct multiple shooting as our control strategy, where the pulses are parametrized according to the CRAB scheme. The core idea is based on representing the state on a more coarse grid compared to the time grid chosen for actually propagating the system. Finally, the state on the intermediate coarse grid points together with final states and control coefficients form a bounded optimization problem which can be attacked with a standard interior point algorithm (a more detailed description is presented at the end of this section). We choose to work in Euler angles parametrization and hence the time evolution operator reads

\[
U(t) = \begin{pmatrix}
\cos\left(\frac{\theta(t)}{2}\right) \cdot e^{i\kappa(t)} & \sin\left(\frac{\theta(t)}{2}\right) \cdot e^{i\phi(t)} \\
\sin\left(\frac{\theta(t)}{2}\right) \cdot e^{i\phi(t)} & \cos\left(\frac{\theta(t)}{2}\right) \cdot e^{i\varepsilon(t)}
\end{pmatrix}.
\] (8.10)

Starting from the identity \( U(t = 0) = I \), we get the following initial values.

\[
\theta(t = 0) = 0, \quad \kappa(t = 0) = 0, \quad \varepsilon(t = 0) = 0, \quad \phi(t = 0) = \text{arbitrary} = \chi(t = 0).
\] (8.11)

When optimizing for the target gate from Eq. (8.1), we want the state representing variables to fulfill the following equations at final time \( T \):

\[
\cos\left(\frac{\theta(t = T)}{2}\right) = \pm \sqrt{2} \frac{\sqrt{2}}{2}, \quad \phi(t = T) - \kappa(t = T) = \begin{cases}
1.5\pi \pmod{2\pi}, & \text{in ‘+ case’}:\\
0.5\pi \pmod{2\pi}, & \text{in ‘- case’}:
\end{cases}
\] (8.12)

and

\[
\varepsilon(t = T) - \chi(t = T) = \phi(t = T) - \kappa(t = T) + \pi \pmod{2\pi}.
\] (8.13)

Note that there is one set of variables \((\theta, \kappa, \phi, \chi, \varepsilon)\) per replica.

As a next step, we calculate the time derivative of \( U(t) \) in the form of Eq. (8.10) using Schrödinger Equation. The outcome of this (full derivation shown in appendix in Sec. 8.8.1) is a coupled system of differential equations (omitting time dependencies for
better readability)

\[
\begin{align*}
\dot{\theta}_i &= -2 \left( \sin(\phi - \kappa)u_x - \cos(\phi - \kappa)u_y \right) \\
\dot{\kappa}_i &= \frac{1}{\cos(\frac{\theta}{2})} \left[ -\delta_i \cos(\frac{\theta}{2}) - \sin(\frac{\theta}{2}) \left( \cos(\phi - \kappa)u_x + \sin(\phi - \kappa)u_y \right) \right] \\
\dot{\phi}_i &= \frac{1}{\sin(\frac{\theta}{2})} \left[ \delta_i \sin(\frac{\theta}{2}) - \cos(\frac{\theta}{2}) \left( \cos(\phi - \kappa)u_x + \sin(\phi - \kappa)u_y \right) \right] \\
\dot{\chi}_i &= -\dot{\phi} \\
\dot{\varepsilon}_i &= -\dot{\kappa}
\end{align*}
\] (8.14-8.18)

\[i \in \{1, 2, \ldots, N\}.
\]

As we can see from Eq.(8.17, 8.18), we do not need to propagate the two phases \(\chi\) and \(\varepsilon\) as their dynamics follows the one of the other two phases, yielding effectively three real variables per qubit.

**Description of control strategy** The question is how to find the two pulses \(u_x(t)\) and \(u_y(t)\) that, given a total time \(T\), transfer the \(N\) qubits from their initial Euler angle representation given by Eq. (8.11) subject to the dynamical laws in Eq. (8.14) - (8.16) to fulfill the target gate in their respective representation given by Eq. (8.12, 8.13) while respecting the maximal excursions on the control pulses in Eq. (8.4). The short answer is: By mapping this problem onto a bounded statical optimization problem and attacking it via well established interior point solvers. In detail, we map the states \((\theta_i(t), \kappa_i(t), \phi_i(t))\) as well as the control pulses \((u_x(t), u_y(t))\). At first, let us focus on the states: the core of the employed direct multiple shooting method is to define a more coarse grid (illustrated by the gray vertical ticks in Fig. 8.1) than the actual time grid used for propagating the system (brown ticks, respectively). The states at all those intermediate transition points (assuming \(k\) in total) as well as the states at final time go into the transferred new static variables denoted by \(\vec{x}\). Hence, \(3 \cdot N \cdot (k + 1)\) (real) components of \(\vec{x}\) are required.
Second, we transfer the pulses into a static (e.g. non explicitly time dependent) variables as well. Therefore, we follow the idea of the CRAB algorithm [95; 96] to choose a chopped randomized function basis and embed its coefficients as further $2 \cdot N_c$ variables into the static search vector $\vec{x}$, resulting in a total dimension of $3 \cdot N \cdot (k + 1) + 2 \cdot N_c$. Up to this point we have set the stage for the variables subject to optimization. What remains is the definition of an objective function $f(\vec{x})$ as well as constraints on the optimization variables. As an objective function any distance metric from final time variables to the target values given by Eq. (8.12, 8.13) can be employed. It is preferable to choose a metric with convex property in the larger vicinity of an optimum in order...
to enhance convergence. Therefore, we choose \( f(\vec{x}) \) to be a sum of cosine values of all deviations from target values independently. That way, the modulus \( 2\pi \) condition is automatically met and the Taylor expansion around its extrema yields the favorable quadratic term as a leading order (the emerging constant is irrelevant for optimization). Finally, we have to include some equality constraints on the new state representing variables \( \vec{x} \) and we need to ensure that the pulses stay within the given boundaries. The state constraints follow from the construction principle that the whole system dynamics is simulated on sub-divisions (slices indicated by gray ticks), independently. In order to maintain physical trajectories all the way from \( t = 0..T \), we must assure the continuous condition such that the gaps in Fig. 8.1 are closed. Thus, for all transition points we put an equality constraint (as indicated by the green arrows) and only accept the solution if the sum of violations is orders of magnitudes lower than the goal precision. All the parts we have described until now will be handed over to an arbitrary constrained nonlinear programming solver. We choose to work with the \textit{fmincon} function under MATLAB(c).

### 8.6 Results

We first show the scenario of normal distributed detunings around zero in Sec. 8.6.1 and then provide a study of how the average fidelity scales with the width of the equidistantly chosen detunings in Sec. 8.6.2.

#### 8.6.1 Optimization of eight normal distributed detunings

To show the capabilities of this method, we optimize the aforementioned problem for \( N=8 \) detunings. The physical parameters are set to be:

- \( \eta = \frac{\pi}{1550} \)
- \( \delta_i \) chosen from normal distribution of standard deviation = \( 1.0 \cdot \eta \)
- Total transfer time \( T = 5600 \)
- Control bound \( b_u = 1 \)

The algorithmic parameter describing the number of coarse grid intermediate transition points is set to be \( k = 20 \). Furthermore we are using a 20 frequencies expansion
8.6 Results

per pulse:

\[ u_x(t) = \sum_{i=1}^{20} A_{x,i} \sin(\omega_{x,i} t) + B_{x,i} \cos(\omega_{x,i} t) \quad (8.20) \]

\[ u_y(t) = \sum_{i=1}^{20} A_{y,i} \sin(\omega_{y,i} t) + B_{y,i} \cos(\omega_{y,i} t) \quad (8.21) \]

where \( \omega_{x,i} \) and \( \omega_{y,i} \) are independently chosen randomly during the algorithm initialization according to:

\[ \omega_{x/y,i} = i + r_i, \quad r_i \text{ uniformly distributed random variable } \in [-0.5,0.5]. \quad (8.22) \]

Thus the effective total dimension for the static program equals \( 4 \cdot 20 \) for the CRAB coefficients plus \( 3 \cdot N \cdot (k + 1) = 504 \) resulting in 584 variables. We achieve an average fidelity of 99.13 percent. The resulting robust pulses are shown in the appendix in Sec. 8.8.2 in Fig. 8.5. For four detunings we show an exemplary state transfer from \( |0\rangle \) almost perfectly reaching the \( \frac{\pi}{2} \) rotated state. It can be seen that the optimized pulses induce many features in the dynamics resulting in a very unintuitive, but successful transfer.
8.6.2 Robust control of equidistant detunings of various ranges

Following a robust gradient-based optimization technique in a two-qubit system with NV-centers [407], we systematically study the best achievable average fidelities as a function of the detunings spreading range. For each range \( \delta_{\text{max}}/\eta = \{0.4, 0.45, 0.5, \ldots 1.4\} \) we optimized 10 equidistantly chosen detunings between \( -\delta_{\text{max}}/\eta \) and \( \delta_{\text{max}}/\eta \). As an average fidelity \( \bar{F} \) to be reported in Fig. 8.3, however, we evaluate the full average of the respective range

\[
\bar{F} = \frac{1}{2 \cdot \delta_{\text{max}}/\eta} \int_{-\delta_{\text{max}}/\eta}^{\delta_{\text{max}}/\eta} F(\delta) \, d\delta .
\]  

(8.23)
The quadratic fit $f(\delta, p_1, p_2, p_3) = p_1 \cdot (\delta/\eta)^2 + p_2 \cdot (\delta/\eta) + p_3$ in Fig. 8.3, yields the parameters $p_1 = -0.0219 \pm 0.0094$, $p_2 = 0.002685 \pm 0.014$, $p_3 = 0.9998 \pm 0.0042$. We can conclude, that the proposed optimization scheme is capable of finding a set of robust pulses that allows for high accuracy gate operations for a large range of detunings.

For $\delta_{\text{max}}/\eta = 0.5$ (turquoise), 0.85 (magenta) and 1.3 (green), the fidelity dependence $F(\delta_{\text{max}}/\eta)$ is explicitly shown in Fig. 8.4.

![Figure 8.3: Average Fidelity $\bar{F}$ as a function of the detuning range $\delta_{\text{max}}/\eta$. Note that high fidelity control can be reached even for large detunings. Dashed line shows quadratic fit which gives rise to an expected fidelity above 0.99 for detunings ranging up to $\pm 70\%$ of the driving field strength $\eta$.](image)

Closing this subsection, we like to add two remarks: First, all reported fidelities are based on highly accurate Schrödinger Equation standard integrators. Second, we stress the fact that this procedure, based on its construction principle, is highly parallelizable. At least to the degree of $k + 1$ threads a significant speed-up can be achieved to any problem solved this way. For the problem discussed here, even $(k + 1) \cdot N$ threads can go in parallel. In order to achieve compliance with the maximum control amplitudes allowed (see Eq. (8.4)), we chose to simply cut the pulse at the boundary whenever the algorithm demanded to put a larger excursion. Different strategies such as more smooth
8.7 Conclusion

We have shown that the problem of finding robust pulses for single qubit gate synthesis can be mapped onto a controllable system. Using Euler parametrization, we transformed the system into a boundary value problem, where the boundaries are not only to be met at the edges of the time interval but also at a number of intermediate times introduced to improve convergence and accuracy. The problem is then formulated as a constrained non-linear multivariable optimization problem, also referred as direct multiple shooting method, which is solved using a standard interior point algorithm. The outcome of this confinement e.g. via transfer functions of tanh type or penalty based enforcement are possible, too.
discrete system optimization is validated using high accuracy Schrödinger Equation propagation. Normally distributed and equidistantly distributed detunings over various ranges are optimized and the performance with respect to narrow sampled detunings is evaluated. As a result, average fidelities remain above 99% although detunings as large as up to ±70% of the control amplitudes are permitted.

8.8 Appendix

8.8.1 Derivation of qubit dynamics in angular variables

We start from the unitary representation from Eq. (8.10)

\[
U(t) = \begin{pmatrix}
\cos(\frac{\theta(t)}{2}) \cdot e^{i\kappa(t)} & \sin(\frac{\theta(t)}{2}) \cdot e^{i\chi(t)} \\
\sin(\frac{\theta(t)}{2}) \cdot e^{i\phi(t)} & \cos(\frac{\theta(t)}{2}) \cdot e^{i\varepsilon(t)}
\end{pmatrix}
\] (8.24)

and apply the Schrödinger operator Equation (setting \(\hbar = 1\):

\[
i \frac{d}{dt} U(t) = H(t)U(t),
\] (8.25)

where the Hamiltonian is the one from Eq.(8.3). We split the consideration into two parts: The two columns of \(U(t) = \begin{pmatrix} \vec{l} \\ \vec{r} \end{pmatrix}\) shall be derived independently. The left-hand side of Eq. (8.25) gives for \(\vec{l}\):

\[
i \frac{d}{dt} \vec{l} = \begin{pmatrix}
-\sin(\frac{\theta}{2}) \hat{\phi} e^{i\kappa} + \cos(\frac{\theta}{2}) e^{i\chi} i\kappa \\
\cos(\frac{\theta}{2}) \hat{\phi} e^{i\phi} + \sin(\frac{\theta}{2}) e^{i\phi} i\phi
\end{pmatrix}
\] (8.26)

The right-hand side gives:

\[
H(t)\vec{l} = \begin{pmatrix}
\delta_i \cos(\frac{\theta}{2}) e^{i\kappa} + \eta \sin(\frac{\theta}{2}) e^{i\phi}(u_x - iu_y) \\
\eta \cos(\frac{\theta}{2}) e^{i\kappa}(u_x + iu_y) - \delta_i \sin(\frac{\theta}{2}) e^{i\phi}
\end{pmatrix}
\] (8.27)
Putting these together, we arrive at (first line: upper, second line: lower entries respectively):

\[ -i \sin \left( \frac{\theta}{2} \right) \dot{\theta} - \cos \left( \frac{\theta}{2} \right) \dot{\kappa} - \delta \cos \left( \theta \right) = \eta \sin \left( \frac{\theta}{2} \right) e^{i(\phi - \kappa)} (u_x - iu_y) \]  

(8.28)

\[ i \cos \left( \frac{\theta}{2} \right) \dot{\theta} - \sin \left( \frac{\theta}{2} \right) \dot{\phi} + \delta \sin \left( \theta \right) = \eta \cos \left( \frac{\theta}{2} \right) e^{i(\kappa - \phi)} (u_x + iu_y). \]  

(8.29)

Next, we apply Euler’s formula to the exponential functions and split real and imaginary parts yielding:

\[ \dot{\theta} = -2 \eta (\sin(\phi - \kappa)u_x - \cos(\phi - \kappa)u_y) \]  

(8.30)

\[ \dot{\kappa} = \frac{1}{\cos \left( \frac{\theta}{2} \right)} \left[-\delta \cos \left( \frac{\theta}{2} \right) - \eta \sin \left( \frac{\theta}{2} \right) (\cos(\phi - \kappa)u_x + \sin(\phi - \kappa)u_y) \right] \]  

(8.31)

\[ \dot{\phi} = \frac{1}{\sin \left( \frac{\theta}{2} \right)} \left[\delta \sin \left( \frac{\theta}{2} \right) - \eta \cos \left( \frac{\theta}{2} \right) (\cos(\phi - \kappa)u_x + \sin(\phi - \kappa)u_y) \right] \]  

(8.32)

Analogously, we get for \( \vec{r} \) for the left-hand side of Eq. (8.25):

\[
\frac{1}{dt} \vec{r} = \begin{pmatrix}
\cos \left( \frac{\theta}{2} \right) & \frac{\delta}{2} e^{i\chi} + \sin \left( \frac{\theta}{2} \right) e^{ixi} \\
- \sin \left( \frac{\theta}{2} \right) & \frac{\delta}{2} e^{iz} + \cos \left( \frac{\theta}{2} \right) e^{iz}i
\end{pmatrix}
\]

(8.33)

And for the right-hand side, respectively:

\[
H(t)\vec{r} = \begin{pmatrix}
\delta \sin \left( \frac{\theta}{2} \right) e^{i\chi} + \eta \cos \left( \frac{\theta}{2} \right) e^{iz} (u_x - iu_y) \\
\eta \sin \left( \frac{\theta}{2} \right) e^{i\chi} (u_x - iu_y) - \delta \cos \left( \frac{\theta}{2} \right) e^{iz}
\end{pmatrix}
\]

(8.34)

Doing the exact same steps as before, we finally arrive at:

\[ \dot{\theta} = -2 \eta (\sin(\chi - \varepsilon)u_x + \cos(\chi - \varepsilon)u_y) \]  

(8.35)

\[ \dot{\chi} = \frac{1}{\sin \left( \frac{\theta}{2} \right)} \left[-\delta \sin \left( \frac{\theta}{2} \right) - \eta \cos \left( \frac{\theta}{2} \right) (\cos(\chi - \varepsilon)u_x - \sin(\chi - \varepsilon)u_y) \right] \]  

(8.36)

\[ \dot{\varepsilon} = \frac{1}{\cos \left( \frac{\theta}{2} \right)} \left[\delta \cos \left( \frac{\theta}{2} \right) - \eta \sin \left( \frac{\theta}{2} \right) (\cos(\chi - \varepsilon)u_x - \sin(\chi - \varepsilon)u_y) \right] \]  

(8.37)

From Eq. (8.30) = Eq. (8.35), we get that

\[ \varepsilon(t) - \chi(t) = \phi(t) - \kappa(t) + \pi \forall t. \]  

(8.38)
Using some trigonometric identities we can conclude:

\[
\sin(\chi - \varepsilon) = \sin(\phi - \kappa) \tag{8.39}
\]
\[
\cos(\chi - \varepsilon) = -\cos(\phi - \kappa) \tag{8.40}
\]

From here on, it is easy to see, that

\[
\dot{\chi} = -\dot{\phi} \tag{8.41}
\]
\[
\dot{\varepsilon} = -\dot{\kappa} \tag{8.42}
\]

holds for all times.

### 8.8.2 Optimized pulse shapes

Here, we show the optimized pulses from Section 8.6.1.

**Figure 8.5:** Optimized pulses over time. Blue: \(u_x(t) \in [-0.7143, 0.7143]\). Red: \(u_y(t) \in [-0.2857, 0.2857]\)
9 Conclusions

In this thesis we studied quantum optimal control theory from an algorithmic perspective but also applied it in theory and experiment. On the theoretical side, we applied open-loop control on a time-dependent Rabi Hamiltonian which serves as a model for the parametric dynamical Casimir effect. We computed time-dependent control functions that represent the qubit frequency such that photon generation in the cavity could be enhanced by two orders of magnitude with respect to an unoptimized on-off strategy. We also demonstrated the robustness of the achieved controls, paving the way for future experimental implementations of the system. Furthermore, we introduced a hybrid algorithm, combining two well established concepts from optimal control theory – direct multiple shooting and the chopped random-basis (CRAB) algorithm – which is capable of generating robust pulses with respect to large detuning variations. This feature is highly demanded in a broad class of different systems such as nitrogen-vacancy (NV) centers and nuclear magnetic resonance (NMR).

Quantum optimal control theory and complex system analysis also served as the core motivation for the Central Moment Dynamics framework. Based on the expectation value of the symmetrized product of powers of canonical conjugate variables, we show how to efficiently find an approximate solution of the time-dependent Schrödinger equation for high-dimensional continuous variables systems. The introduced framework can be merged with quantum optimal control theory enabling systematic control and we could confirm the existence of a reduced effective dimension of continuous quantum systems.

Besides these theoretical studies, we directly applied closed-loop optimal control in the lab for two different setups. Shaping laser pulses and magnetic field strengths, we enhanced the evaporative cooling process to yield a large BEC in terms of atom numbers. For a spin qubit represented by a single NV color center in diamond, we demonstrated autonomous calibrations via closed-loop optimal control. The control enters through a microwave field, whose amplitude, frequency and phase is tunable.

In order to facilitate the experimental control process, we have written a software suite named Remote dCRAB optimization (RedCRAB): it is capable of remotely interfacing with arbitrary quantum experiments and to optimize pulse shapes, enabling a broad
range of control goals. For example state transfers, gate synthesis, optimization of some observables, or entanglement maximization are possible. Various constraints can be respected and multiple features have been implemented to meet the needs for laboratory control.

In conclusion, we have studied various physical systems in theory and experiment and we have developed a hybrid quantum control algorithm as well as a novel approach to simulate continuous variables systems. Finally, we developed a software suite we envisage having the potential to make a crucial contribution towards the realization of game-changing quantum technologies.
A Quantum Optimal Control Software Projects

During the last two decades, many software projects to simulate quantum dynamics and to control quantum devices have been started. An excellent overview is provided on the quantiki website [408]. Some of them do have a feature to optimally control quantum systems – either as an option or as a main purpose. We briefly introduce some of the frameworks in the following. A matlab-based toolset for gradient based optimal control has been introduced under the terms dynamical optimisation platform (DYNAMO) [409]. As a result, a highly recommended paper emerged [43]. Focusing on and specializing in spin dynamics, a simulation library especially for the NMR community has been developed under the term Spinach [410]. Special emphasis is devoted to the optimal control part where multiple options are available [411]. The HamPath package [412] is capable of solving optimal control problems via indirect methods but also to study Hamiltonian systems. The focus is on mathematical control theory while quantum physics is one of the targeted applications. The Machine-Learner Online Optimization Package (M-LOOP) [413] is a closed-loop control software that can automatically optimize the parameters of a scientific experiment or computer controlled system. At the core of the optimization, a Gaussian process is used to develop a statistical model of the relationship between the parameters and its controls with regard to the respective figure of merit. M-LOOP has also been applied to the production of a BEC [49]. The QDYN python package [414] provides a collection of utilities to augment the Fortran QDYN library [415] for quantum dynamics and control developed at Christiane Koch’s group.

A.1 Quantum Toolbox in Python (QuTiP)

One of the most professionally implemented and best maintained quantum software suites is QuTiP [416]. Therefore, several hundreds citing works have been recorded since its launch in 2011 [417; 418]. Its focus lies on the simulation of the dynamics
of open quantum systems. In addition to the standard Lindblad and Monte Carlo Solvers, QuTiP includes routines for Bloch-Redfield evolution, periodic systems using the Floquet formalism, and stochastic solvers. It is very flexible in use as multiple customized quantum objects such as states and operators can be defined and manipulated. Moreover, it takes care of some side tasks such as parallelization, efficient propagation and plotting. Furthermore, QuTiP comes with an optimal control package providing GRAPE and CRAB as possible algorithms. The latter algorithm has been implemented in a cooperation between Aberystwyth and Ulm universities.

A.2 RedCRAB

Over the course of the last four years, we wrote the Remote dCRAB optimization (RedCRAB) software suite. It is capable of iteratively solving various open and closed-loop (quantum) optimal control problems\(^1\). The peculiarity of RedCRAB is that it remotely interfaces laboratories or system simulations with our RedCRAB servers based at the university of Ulm via the web. The foremost design principle is that this way also non control experts can directly apply optimal control to their respective system in an easy to use and permanently accessible manner. Consequently, arbitrarily many scientists can use its capabilities in tailoring dynamical pulse sequences (i.e. time dependent degrees of freedom to be adjusted) for a wide spectrum of physical systems. Applications are therefore not limited to quantum information protocols, as needed for example for metrology and sensing, but also cover different branches of physics, like energy science, AMO physics, and condensed matter. To achieve experimental feasibility, a series of features such as subsequent reevaluation steps, incorporation of standard deviation resulting from measurement and fault tolerance were added to the bare optimization algorithm. Besides experimental applications, theorists are benefiting as well as they can test and optimize various systems in the open loop optimization branch of the software. RedCRAB has already been successfully employed in synthesizing a single qubit gate in the NV center (Chapter 6) and enhancing the evaporative cooling ramps needed to achieve a BEC (Chapter 5).

The flow of functionality of RedCRAB is as follows: at first key figures about the optimal control problem (e.g. how many pulses are required, what amplitude limits have to be respected, process duration, ...) are defined in a configuration file. After the connection to a RedCRAB server has been established, customized trial control pulses

\(^1\) For an introduction into the field and our employed dCRAB algorithm, see Chapter 2 and references therein.
are sent for evaluation. The outcome from the experiment (or simulation) in terms of a quality measure (figure of merit) is fed back to allow the algorithm to converge. The number of iterations (=feedback loops, function evaluations) required depends on the number of pulses and free parameters subject to search, and the complexity of the problem (as well as possibly other factors beyond our control such as statistical disturbances and drifts of the experiment and in case an initial guess based on random components is employed, its quality). All data generated are automatically recorded and connection losses can be automatically resumed.

Some of the most important features\(^1\) of RedCRAB are summarized in Fig. A.1. The major options regarding the actual shaping of the pulses are visualized in Fig. A.2: an initial guess and amplitude limits can be considered as well as the magnitude of variations of the to be probed pulses can be adjusted. The total transfer time \(T\) might also be subject to search and the maximum bandwidth of the pulses can be limited. Global shape functions such as smooth ramping up and down of the pulses (which might be required for some experimental apparatuses) at \(t = 0\) and \(t = T\) can be factored in. Furthermore, multiple pulses can be optimized simultaneously.

Besides options for pulse shaping, a multitude of other functionality is available. We are going to describe one of the features in more detail in the following. In the case of relative large uncertainty on the figure of merit evaluation of the experiment (e.g. for the BEC setup optimization in Chapter 5, fluctuations up to 15% emerged), multiple re-evaluation steps (of the same (set of) pulses) should be considered. In order to save function evaluations, a sophisticated \(\mu - \sigma\) criterion can therefore be employed. The central idea is to consider (an estimate on) the standard deviation (denoted by \(\sigma\)) as a second feedback quantity (besides the (estimate on) the actual figure of merit \(f\) (also denoted by \(\mu\))). Then we first calculate the probability for a newly evaluated control pulse to be better than the current record \(P(f_{\text{new}} > f_{\text{rec}})\) under some assumptions (see below). Second, based on user defined threshold steps, possible further re-evaluations of the same trial pulse is being done subsequently until the pulse either suffices all of the assigned surpass probabilities or it is being rejected as a new record (example pending). To clarify the working principle, we begin by explaining how to calculate the probability \(P(f_{\text{new}} > f_{\text{rec}})\). Therefore also see Fig. A.3. Assuming that both random variables \(f_{\text{new}}\) and \(f_{\text{rec}}\) are normally distributed with \(\mu = f\) and \(\sigma = \text{Std}(f)\), the probability \(P(f_{\text{new}} > f_{\text{rec}})\) is calculated via an auxiliary random variable \(z\) which is scaled to be

\(^1\) More details can be found in the RedCRAB documentation and on the project website being launched soon
Figure A.1: Overview of the most important specifications of the current version (v. 1.0.4) of the RedCRAB program suite.
standard normally distributed ($\mathcal{N}(0,1)$):

$$
\mu_z = f_{\text{new}} - f_{\text{rec}} \quad \text{(A.1)}
$$

$$
\sigma_z = \sqrt{\sigma^2_{\text{new}} + \sigma^2_{\text{rec}}} \quad \text{(A.2)}
$$

$$
\Rightarrow \frac{\mu_z}{\sigma_z} = z \sim \mathcal{N}(0,1) \quad \text{(A.3)}
$$

Employing the normal cumulative distribution function $\Phi$ (see Fig. A.4), we can calculate the probability as

$$
P(f_{\text{new}} > f_{\text{rec}}) = \Phi(z). \quad \text{(A.4)}
$$

Due to the symmetry of the probability density function $\rho(x)$ when $x \sim \mathcal{N}(0,1)$, we can
Figure A.3: Schematic drawing of the two assumed probability density functions $\rho(f_{rec})$ (blue) and $\rho(f_{new})$ (red). The goal is to calculate $P(f_{new} > f_{rec})$. (Note that the green shaded region is not the actual probability we are looking for, but gives an intuition.)

Figure A.4: Plot of cumulative distribution function of the normal distribution.
conclude that anytime $f_{\text{new}} > f_{\text{rec}}$, the calculated probability $P(f_{\text{new}} > f_{\text{rec}}) > 0.5$ and smaller than 0.5 for the opposite case. Once a new set of $f_{\text{new}}$ and $\sigma_{\text{new}}$ is transferred and the probability $P(f_{\text{new}} > f_{\text{rec}})$ is calculated, re-evaluation rules based on the probability can be defined by the user. To illustrate that, we give an example of possible re-evaluation steps in the following table.

<table>
<thead>
<tr>
<th>re-evaluation step #</th>
<th>threshold P</th>
</tr>
</thead>
<tbody>
<tr>
<td>step 1 → 2nd evaluation</td>
<td>0.33</td>
</tr>
<tr>
<td>step 2 → 3rd evaluation</td>
<td>0.5</td>
</tr>
<tr>
<td>step 3 → 4th evaluation</td>
<td>0.501</td>
</tr>
<tr>
<td>step 4 → 5th evaluation</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Table A.1: Configurable thresholds for re-evaluations

In this example, a second re-evaluation is only requested once the calculated probability $P(f_{\text{new}} > f_{\text{rec}})$ exceeds 0.33. A third one only if the probability exceeds 0.5 and so forth. Only if after the 5th evaluation the (combined) probability exceeds 0.51, the new pulse will be assigned as the new record.

Note that $f_{\text{new}}$ and $f_{\text{rec}}$ in Eq. (A.1) are always the mean of all re-evaluation (up to what is available) and similarly for the $\sigma$ in Eq. (A.2); they are calculated as

$$\sigma_{\text{new/record}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sigma_i.$$

Finally, a brief summary of the most important modules of RedCRAB, which is implemented in python, is depicted in Fig. A.5.

Possible future developments of RedCRAB are the implementation of a graphical user interface (GUI) on the client side. Usability can still be enhanced when configuration file creation as well as the actual operation can be done also graphically. More choices for the underlying optimization algorithm such as genetic algorithms and differential evolution might be implemented. Moreover, different function basis to be selected – currently the options are Fourier basis, Tschebyscheff polynomials and Gaussian bell curves – can be easily included. However, the most important input comes from (already ongoing) further experimental applications and their requirements.
RedCRAB – module overview: v 1.0.4 → scheduled

Redcrabmain.py
- Checks for new incoming opti jobs
- Initiates their execution (threaded)
- Monitors their status

Redcrab.py
- Performs optimizations
- Manages file exchange with user home folder on server

Redcrabclient.py
- Central program for client
- Creates client logs
- Handles communication with server and other client modules

remote_module.py
- Paramiko based file exchange

Opti_instance.py
- Handles communication with server and other client modules

Config file maker

Logging_module.py

ssh via paramiko

Client’s home folder on server

Local folder on client’s pc

Figure A.5: Overview of the major modules of the RedCRAB software suite.
# List of Figures

2.1 Open-loop versus closed-loop control – a sketch .................. 9
2.2 Unlocking false traps for the dCRAB algorithm – an example .... 18
2.3 Sketch of potential $V(x)$ and $|\Psi(x)|^2$ trajectories for h.o. shifting problem 23
2.4 Results for optimal transport of a single particle in a harmonic well .. 27
2.5 Optimal Gaussian Focusing - problem definition ..................... 36
2.6 Results for optimal Gaussian focusing .............................. 38
2.7 Results for state transfer of greedy one step approach ............... 48

3.1 Visualization of QSL trajectories .................................. 59
3.2 Visualization of control landscapes ................................. 63

4.1 Sketch of Rabi Hamiltonian as being the central model for the DCE .. 68
4.2 DCE paper – Results for on-off protocol ............................ 76
4.3 DCE paper – Results for bang-bang strategy and dCRAB algorithm .. 77
4.4 DCE paper – Summary of results achieved .......................... 78
4.5 DCE paper – Robustness analysis ................................. 80
4.6 DCE paper – Working principle of bang-bang strategy ............... 81

5.1 Cooling steps towards Bose-Einstein condensation .................. 84
5.2 Employed BEC trap geometry of setup ............................. 85
5.3 Scheme for RedCRAB remote optimal control loop .................. 86
5.4 BEC lab setup in Aarhus ............................................ 87
5.5 BEC paper – Overview of studied remote control options .......... 91
5.6 BEC paper – Control landscape topologies and scanned landscape .. 93
5.7 BEC paper – Relative and absolute BEC atom number results ....... 96
5.8 BEC paper SI – Visualization of bridgeability of solutions of BHW game 102
5.9 BEC paper SI – Demonstration of 1&2D parameter scans and found bridge107
5.10 BEC paper SI – Demonstration of 3D parameter scans from NCDT to HT108
5.11 BEC paper SI – Landscape exploration from HT to RedCRAB optimum 110
5.12 BEC paper SI – Ramp duration sweep for different OC strategies .... 114
5.13 BEC paper SI – Analysis of players’ behavior ........................ 117
5.14 BEC paper SI – Distribution of highest achieved results in team challenge 120
6.1 NV crystal structure ........................................... 122
6.2 NV level structure .............................................. 123
6.3 NV paper – Schematics of the experiment ................ 126
6.4 NV paper – Results for state transfer ...................... 129
6.5 NV paper – Results for gate optimization ............... 130
6.6 NV paper SM – Open and closed-loop optimization methods 134
6.7 NV paper SM – Unlocking false traps, see Fig. 2.2 .... 136
6.8 NV paper SM – Rectangular pulse state transfer fidelity 140
7.1 Proof of recursion formula: double sum indices change 148
7.2 CANDY paper – Log-log plot of central moments ........ 153
7.3 CANDY paper – Lower moments phasespace trajectories 156
7.4 CANDY paper – Results for method in 2D ............... 157
7.5 CANDY paper – Results for method in 1D vs. optimal control analysis 158
7.6 CANDY paper SM – Example for dependencies of CANDY equations 165
7.7 CANDY paper SM – In depth plot for Fig. 7.3 .......... 169
7.8 Alternative closure scheme: eigenspectrum of moment matrix $\bar{A}$ 172
8.1 Working principle of Direct Multiple Shooting method .... 181
8.2 Bloch sphere visualization of resulting optimized trajectories 184
8.3 Equidistant detunings results: average fidelity $\bar{F}$ vs. detuning range $\delta_{max}/\eta$ 185
8.4 In depth plot for Fig. 8.3 ...................................... 186
8.5 Paradigmatic plots of optimized pulse shapes ............ 189
A.1 RedCRAB – Specifications overview ..................... 196
A.2 RedCRAB – Pulse shaping features ....................... 197
A.3 Working principle of $\mu$ - $\sigma$ criterion .................. 198
A.4 Cumulative distribution function for normal distribution 198
A.5 RedCRAB – module overview .............................. 200
List of Tables

5.1 Distribution of search strategies in the ATC . . . . . . . . . . . . . . . . 120
7.1 Number of moments $N_m$ tracked as a function of $k_{max}$ and D . . . . . . . 168
A.1 $\mu$ - $\sigma$ criterion: example of thresholds for re-evaluations . . . . . . . . . . 199
Bibliography


[58] K. Graichen. METHODEN DER OPTIMIERUNG UND OPTIMALEN STEUERUNG.


2017. Note, the displayed data points stem from an older set of measurements, where relative $N_{\text{BEC}}$ between the trap configurations were not represented correctly. Therefore, the labelling of the color scale was omitted.


[299] The error bounds represent the standard deviation and are based on benchmarks taken randomly during one month’s period. However, all our experiments presented in the following were conducted over the span of six months and long term drifts of the experimental apparatus were present. RedCRAB optimization and the parameter scans spanned over five months, the Alice challenge and experimental investigation of specific solutions over one month. Individual trap configurations were therefore subjected to an overall drift of up to 14%. Despite this, the relative yield of different strategies has been investigated with low noise within a short time frame, and over longer time scales by relating to benchmark sequences.


[302] M. Derex, R. Feron, B. Godelle, and M. Raymond. Social learning and the


[315] F. Frank (2016), with kind approval.


[414] https://github.com/goerz/qdynpylib,.

[415] https://www.qdyn-library.net,.


List of publications

Teile dieser Dissertation wurden bereits in folgenden Fachartikeln veröffentlicht:


F. Frank, T. Unden, J. Zoller, Ressa S. Said, T. Calarco, S. Montangero, B. Naydenov, and F. Jelezko, *Autonomous calibration of single spin qubit operations*, npj Quantum Information 3, 48 (2017), doi:10.1038/s41534-017-0049-8, licensed under a Creative Commons Attribution 4.0 International License which can be found at [https://creativecommons.org/licenses/by/4.0/](https://creativecommons.org/licenses/by/4.0/).

Danksagung

Die Danksagung wurde aus Gründen des Datenschutzes entfernt.
Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, sowie wörtlich oder inhaltlich übernommene Stellen als solche kenntlich gemacht habe und die Satzung der Universität Ulm zur Sicherung guter wissenschaftlicher Praxis beachtet habe.

Ulm, den .................................................. ..............................................................

Jonathan Emanuel Zoller