Machine Learning based Optimization for the Preparation of Rubidium Rydberg Atoms and Characterization of the Detection

Valerie Mauth

Masterarbeit in Physik
angefertigt im Institut für Angewandte Physik

vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
Bonn

July 2025

I hereby declare that this thesis was formulated by myself and that no sources or tools other than those cited were used.

Bonn, ...
Date Signature

1. Supervisor: Prof. Dr. Sebastian Hofferberth
2. Supervisor: Prof. Dr. Daqing Wang

Contents

CHAPTER 1

Introduction

The idea that quantum computers could solve computational task exponentially faster than a classical computer was proposed by Richard Feynman in the early 1980s [1]. His work laid the foundation for a rapidly growing research field, which explores how quantum phenomena can be used to achieve significant performance improvement over classical technologies [2]. Quantum computing and quantum technology in general have a wide range of applications, including secure communication [3], high-precision sensing [4], machine learning [5], chemistry [6], and many more [7, 8]. A variety of quantum hardware platforms exist, ranging from trapped ions [9], photonic systems [10] and neutral atoms [11, 12] to superconducting circuits [2]. To profit from the strength of different platforms while mitigating their limitations, hybrid systems that combine two or more quantum platforms have been proposed [13]. The goal is to create implementations that can store, process and transfer quantum information [14].

The Hybrid Quantum Optics (HQO) experiment in the Nonlinear Quantum Optics research group of Sebastian Hofferberth investigates a novel hybrid system between an electromechanical oscillator and Rubidium Rydberg atoms.

Rydberg atoms, which are highly excited atoms with long coherence times [15], are promising candidates for hybrid systems. They possess strong electric dipole transitions over a large range of the electromagnetic spectrum [12], with optical transitions between ground and Rydberg states and transitions in the microwave regime between adjacent Rydberg states [15].

Electromechanical oscillators are often used in hybrid systems, since they have high quality factors of up to 10^8 [16, 17], long coherence times [13], and couple to microwaves [16, 18]. These properties make them suitable for microwave-to-optics transduction [19], quantum acousto-dynamics [16] and fundamental research of the boundary of classical and quantum systems [20].

The first goal of the HQO experiment is to cool down one vibrational mode of a high-overtone bulk acoustic resonator (HBAR) to its quantum mechanical ground state via interaction with Rubidium Rydberg atoms. By bringing the Rydberg atoms close to the HBAR, which will be placed on an atom chip, it will be possible to exchange microwave photons with the Rydberg atoms via dipole interactions. The coupling between HBAR and Rydberg atoms will enable the optical control and readout of the HBAR.

The experiment has to be conducted in a cryogenic environment to reduce the initial thermal occupation of the HBAR. To prevent laser induced heating and to improve the vacuum in the cryogenic region, the atom loading is performed in a separated vacuum chamber. This requires the transport of the atom cloud from the loading chamber to the science chamber, hosting the HBAR. In the scope of this

thesis the cold atom preparation sequence, including loading and transport of the atom cloud, was optimized. Furthermore, the Rydberg ion detection setup was characterized to determine the detection efficiency. At the time of this thesis the cryostat was not yet implemented into the setup, meaning that only measurements at room temperature were possible.

For the optimization of the preparation sequence a machine learning online optimization method, based on the python package M-LOOP [21], was employed. Machine learning (ML), as a subfield of artificial intelligence [22], is a continuously growing field with already significant impact across a wide range of application, including industry [23, 24], medicine [25, 26] and science [27–29]. It is based on the development of algorithms that learn, purely from data, about patterns within complex systems, with the goal to predict the system's future behavior and to enable informed decision-making [22]. ML algorithms have already become a common tool in physics for processing large datasets, for example in particle physics [30] or cosmology [31], where they are widely used for classification and regression tasks [29]. Beyond applying ML algorithms after a complete dataset was generated, online machine learning methods are increasingly being used to optimize ultra-cold-atom experiments in real time [21, 32, 33]. The advantage of online machine learning is that the algorithm learns about the experimental system during the process of generating new data, enabling it to make informed decisions about which experimental parameters to test next, thereby significantly improving the efficiency and speed of the optimization process [21]. The algorithm builds an internal model of the system which is continuously improving by receiving real time feedback on its predictions, despite having no prior knowledge of the experiment [21, 32, 34]. This makes machine learning optimization well suited for the optimization of complex experimental setups with unknown perturbations [33], sometimes even leading to unexpected optimization results with significant improvement [34].

In this thesis, the machine learning online optimization routine was implemented into the HQO experiment and applied to the optimization of the magneto-optical trap (MOT), as a proof of principle application, and the magnetic transport, for optimizing a complex system with a large optimization parameter space.

The structure of this thesis is as follows. At first the experimental setup and sequence for the preparation of the Rubidium Rydberg atoms is introduced in Chapter 2. Chapter 3 provides an overview of the general principles of machine learning online optimization and its implementation in the M-LOOP Python package. Next, the integration of the machine learning optimization cycle into the HQO experiment is discussed in Chapter 4. Furthermore, this chapter outlines the specific realization of the MOT and magnetic transport optimization, and discusses the results. Chapter 5 explains the ion detection setup in more detail and discusses the characterization measurements for this setup. Finally, Chapter 6 provides a summary of the optimization and characterization conducted in this thesis for the Rubidium Rydberg preparation and detection in the HQO experiment.

CHAPTER 2

Experimental Setup and Sequence

This chapter provides an introduction to the experimental sequence and setup in the HQO experiment. First, a brief overview is presented. Next the different experimental steps composed of initial cold atom preparation (see Section 2.1), magnetic transport (see Section 2.2) and Rydberg excitation and detection (see Section 2.3) are discussed in more detail. Additionally, the experiment control system, which controls the different experimental steps, is introduced in Section 2.4.

The full experimental setup required to realize the desired hybrid system is shown in Fig. 2.1. It is composed of four parts: the MOT vacuum chamber for the initial cold atom preparation; the science

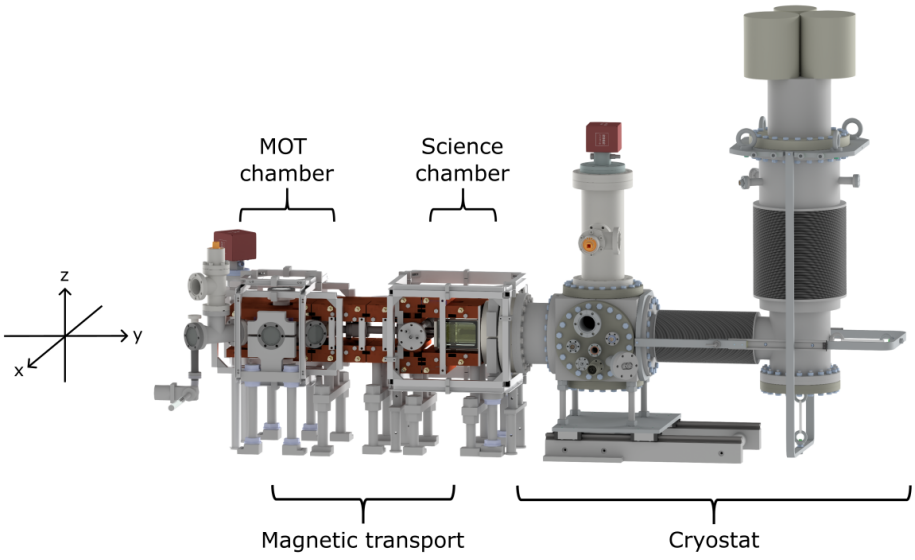


Figure 2.1: CAD drawing of the full experimental setup. It consists of the MOT chamber, where the Rubidium atoms are initially prepared, the science chamber, where the actual experiment takes place and the magnetic transport connecting both chambers. The cryostat will enable experiments in a cryogenic environment in the future. At the time of this thesis, experiments could only be conducted in the setup shown without the cryostat. This is referred to as the 'room temperature setup'. The coordinate system on the left describes the laboratory system which will be used as reference system throughout this thesis.

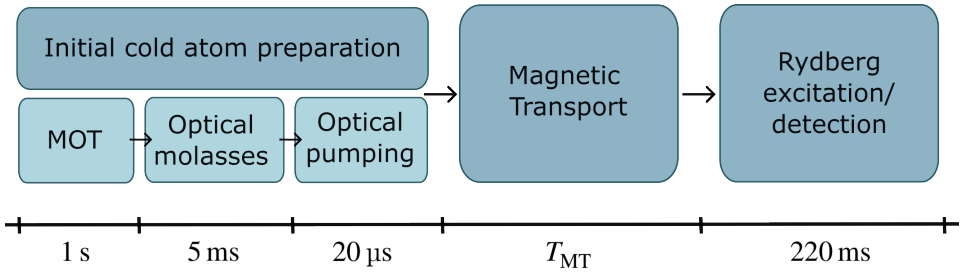


Figure 2.2: The experimental sequence for experiments in the room temperature part of the setup. The atoms are initially prepared in the MOT chamber, where they are at first trapped and cooled down in a magneto-optical trap. After this an optical molasses phase is performed for cooling to sup-Doppler temperatures. Next the atoms are optically pumped to a low field seeking state, to be able to trap them in a magnetic field. After this they are magnetically transported to the science chamber where the Rydberg excitation and detection takes place. The x-axis shows the duration of the steps, T_{MT} describes the magnetic transport time which is either chosen as 1.5 s or 2 s.

chamber, which is an ultra-high vacuum chamber hosting the atom chip; a magnetic transport connecting both vacuum chambers; and the cryostat. The cryostat will be attached to the science chamber to enable experiments in a cryogenic environment. However, the cryostat was not yet included in the setup at the time of this thesis. Therefore, all optimizations and characterizations were conducted in the room temperature part of the setup. The two chamber design was chosen to prevent laser induced heating in the cryogenic region during the initial cold atom preparation. The Rubidium atoms, specifically the isotope ^{87}Rb, are loaded from Rubidium background gas in the MOT chamber and then transported to the science chamber, where the actual experiment will take place. Due to this loading process, the pressure in the MOT chamber is higher than required for the science chamber. Using a differential pumping tube between both vacuum chambers separates the MOT chamber with relatively high pressure of around $\sim 10^{-9}$ mbar from the science chamber at ultra-high vacuum of around $\sim 10^{-10}$ mbar. The Rubidium background gas is produced by heating up a broken Rubidium ampulla connected to the MOT chamber. For more information about the Rubidium atom source in this experiment see Johanna Popp's Master's thesis [35]. Minimizing the heating effects in the science chamber is also why a magnetic transport is used to connect both chambers instead of an optical transport, which could be realized, for example, with a movable optical lattice [36].

The experimental sequence in the room temperature setup is summarized in Fig. 2.2 and can be divided into three parts. At first the atoms are trapped in the MOT chamber and initially cooled down in a magneto-optical trap and subsequent optical molasses phase. To prepare the atoms for the magnetic transport they also need to be optically pumped into a low field seeking state. The initial cold atom preparation in the MOT chamber is explained in more detail in Section 2.1. In the next step the atoms are magnetically transported to the science chamber where the Rydberg excitation and detection take place. The working principle of the magnetic transport and the realization in the experiment are discussed in Section 2.2. The Rydberg excitation and detection sequence can be found in more detail in Section 2.3.

Furthermore, an absorption imaging system is built around the MOT chamber and the science chamber. For the MOT chamber the absorption imaging is conducted along the x-axis, for the science chamber it is conducted along the y-axis. The orientation of the axes are defined by the coordinate system in Fig. 2.1, which will be used as reference throughout the thesis. Absorption imaging can be used to characterize the properties of a trapped ultracold atom cloud by illuminating the atoms with a resonant and collimated

laser beam and detecting the transmitted light. In the experiment this is realized by taking three images. The first one is the shadow image, where the atoms are trapped in the chamber and illuminated with the imaging light. After this a laser image without any atoms in the beam path is taken. This captures the undisturbed laser beam. The third image is a background image with the imaging light turned off and no trapped atoms. From this measurement it is possible to determine the 2D column density of the atom cloud. By integrating over the column density the atom number can be extracted [37]. If the density distribution of the atom cloud can be assumed to be Gaussian, one can extract the cloud width σ (FWHM) and the center position of the cloud (along the axes of the imaging plane) by fitting Gaussian distributions along both axes [37].

The absorption images are taken after a fixed time of flight (TOF). The time of flight is a free evolution time for the atoms, during which the trapping potential is turned off. If not stated otherwise, the time of flight is set to 10 ms here. This ensures that the magnetic fields are fully turned off when doing the absorption imaging measurement and the imaged atom cloud can be described by a Gaussian distribution. This is important for the subsequent analysis of the absorption images. The expansion of the atom cloud for different TOF depends on the temperature of the cloud. By measuring the cloud width σ at different times t after free expansion, it is possible to extract the temperature T of the cloud. This is done by fitting [38]

$$\sigma(t) = \sqrt{\frac{k_B T}{m} t^2 + \sigma_0} \tag{2.1}$$

to the measured cloud widths. Here, k_B is the Boltzmann constant, σ_0 is the initial cloud width, and m the mass of the atoms. This is called time of flight (TOF) measurement. For a more detailed and mathematical description of the absorption imaging method and the TOF measurement see e.g. [37–39]. The absorption imaging setup in the HQO experiment was characterized and discussed in Julia Gamper's master thesis [40]. As most of the laser setups required for the other experimental steps have been documented in previous Bachelor's and Master's theses[1], it will not be explained in much detail in this thesis.

[1] see Julia Gamper's Bachelor's and Master's thesis [40, 41], Samuel Germer's Bachelor's thesis [42], Johanna Popp's Master's thesis [35] and my Bachelor's thesis [43]

2.1 Initial Cold Atom Preparation

The initial cold atom preparation sequence is used to trap and initially cool the Rubidium atoms in the MOT chamber before transporting them to the science chamber. The setup of the MOT chamber is shown in Fig. 2.3. The chamber is a custom-made vacuum chamber (from VACOM [44]) in octagon form. It has ten viewports, six for aligning the MOT beams through the chamber (see Section 2.1.1), one for the connection to the magnetic transport, two for being able to do absorption imaging and another one for fluorescence imaging. Fluorescence imaging and absorption imaging are methods used for characterizing the ultra-cold atom cloud in the MOT chamber. The fluorescence imaging captures the fluorescence of the atoms trapped in the MOT with a camera placed along the fluorescence imaging axis, as shown in Fig. 2.3. This section will explain the step required to prepare the atoms for the magnetic transport. Section 2.1.1 explains the principle of a magneto-optical trap, next the optical molasses phase will be explained in Section 2.1.2 and the first magnetic trap in the MOT chamber for the magnetic transport is discussed in Section 2.1.3.

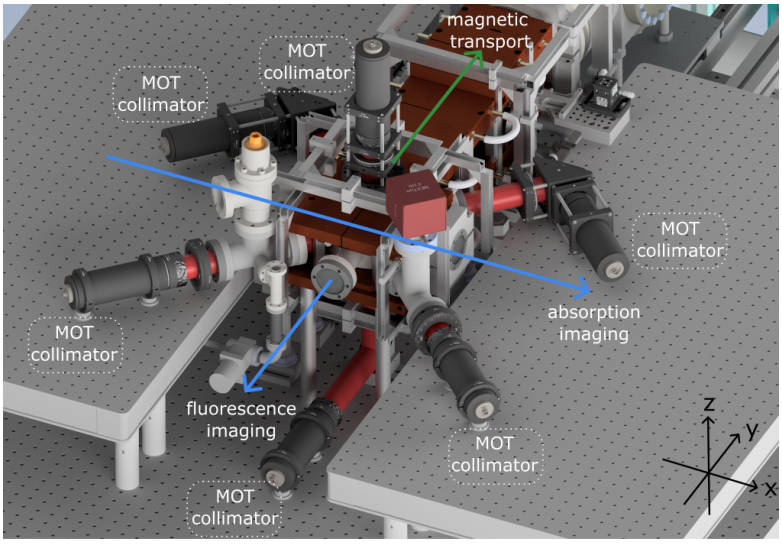


Figure 2.3: CAD drawing of the MOT chamber. The MOT collimators are used for aligning the six MOT beams through the chamber. They are placed in front of the respective viewports. The blue arrows indicate the axes used for absorption imaging and fluorescence imaging. The green arrow shows the direction of the magnetic transport to the science chamber.

2.1.1 Magneto-Optical Trap

The first cooling and trapping stage is a magneto-optical trap (MOT) [46]. The principle of a MOT is based on laser cooling in a non-uniform magnetic field creating spatially dependent forces. Laser cooling results from the conservation of energy and momentum in a scattering process. When an atom is moving with velocity $\vec{v}$ in one direction it can be slowed down with a counterpropagating laser beam with wavevector $\vec{k}$. In the rest frame of the moving atom, the photon's frequency is shifted to higher frequencies due to the Doppler effect. In order to get absorbed, the frequency ω_L of the light must be red detuned from the atomic transition [47]. The absorption of a counterpropagating, red-detuned

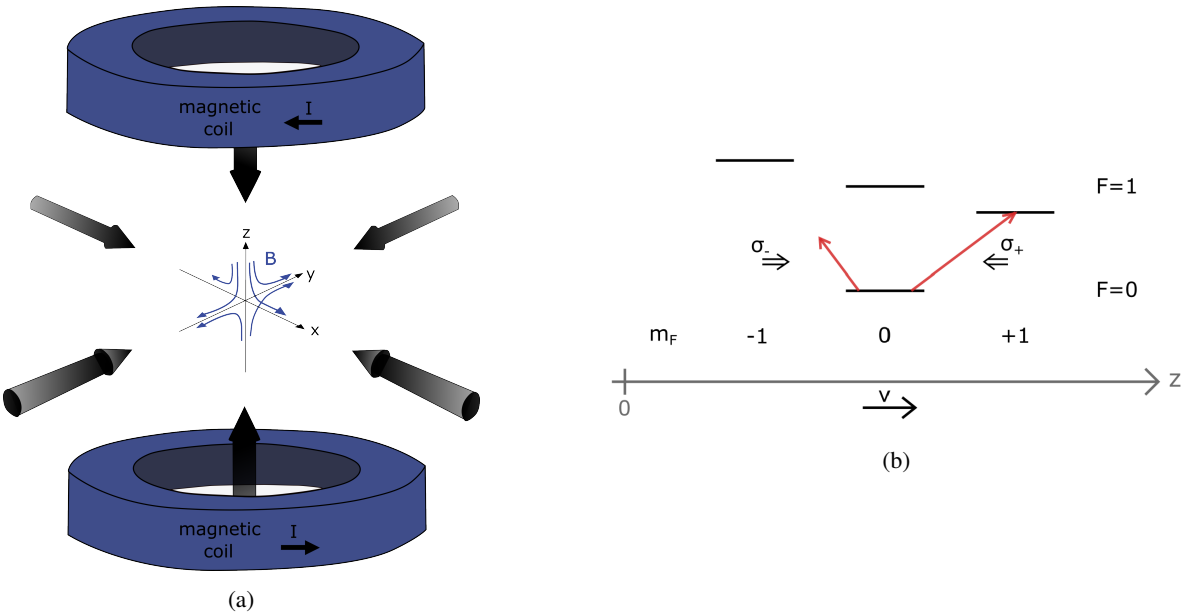


Figure 2.4: **(a)** Realization of the magneto-optical trap with magnetic coils in anti-Helmholtz configuration. The arrows indicate the three pairs of counterpropagating MOT beams. The coils produce a quadrupole magnetic field B. **(b)** Schematic of the energy level shift of the hyperfine m_F states in a magnetic field for an atom moving along the z-axis with velocity v. Counterpropagating σ_+-polarized light is shifted into resonance, while σ_--polarized light travelling in the same direction as the atom is shifted out of resonance. Both images are adapted from [45].

photon excites the atom to a higher lying energy state. Furthermore, the atom experiences a recoil in the direction of the incident photon. The subsequent spontaneous decay of the atom causes an isotropic emission of a photon. Therefore, the average transferred momentum during the spontaneous decay is zero, leading to an average scattering force in the opposite direction of the atoms initial motion [45]. To be able to cool down an atom cloud in one dimension, two counterpropagating red detuned laser beams are required. The difference between the scattering forces of both beams gives the resulting cooling force $\vec{F}_{\text{cooling}}$. For a two level system this can be approximated as [47]

$$\vec{F}_{\text{cooling}} \approx \hbar k^2 \frac{8\delta}{\Gamma} \frac{I/I_{\text{sat}}}{(1 + I/I_{\text{sat}} + (2\delta/\Gamma)^2)^2} \cdot \vec{v} = -\alpha \cdot \vec{v}, \tag{2.2}$$

where I is the laser intensity, I_{sat} the saturation intensity, Γ the natural linewidth and δ the detuning of the laser frequency to the atomic transition. If $\delta < 0$, this describes a damping force which can slow the atoms down. Aligning two counterpropagating red-detuned laser beams on all three cartesian coordinate axes enables the atoms to be slowed down and cooled in all three dimensions [45]. This is called optical molasses, and will be employed in Section 2.1.2.

A non-uniform magnetic field is required to not only cool the atoms but also confine them spatially. This can be realized, for example, with coils in anti-Helmholtz configuration. This is shown in Fig. 2.4(a). The magnetic field is created using two magnetic coils, placed at a distance equal to their radii, and carrying the same current I in opposite direction. Near the symmetry center of the coil pair, the magnetic field can be approximated as a quadrupole magnetic field. The field is zero at the center position and

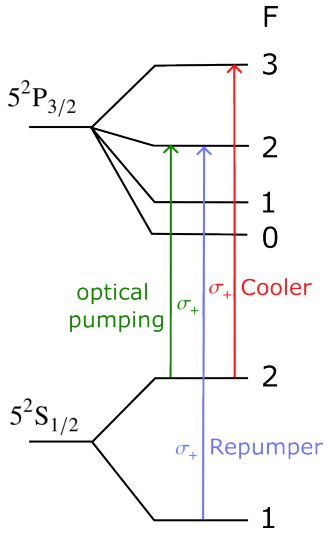


Figure 2.5: Energy level structure of ^{87}Rb showing the D_2 line [49]. The Cooler laser drives the σ_+ transition $|5S_{1/2}, F = 2, m_F\rangle \rightarrow |5P_{3/2}, F = 3, m_F + 1\rangle$ and the Repumper laser drives the transition $|5S_{1/2}, F = 1\rangle \rightarrow |5P_{3/2}, F = 2\rangle$. For trapping atoms in a magnetic field they have to be optically pumped into a low field seeking state. This is done with the optical pumping laser driving the transition $|5S_{1/2}, F = 2, m_F\rangle \rightarrow |5P_{3/2}, F = 2, m_F + 1\rangle$.

increases linearly with distance to the center [45]. The properties of the magnetic quadrupole field generated by anti-Helmholtz coil pairs will be discussed in more detail in Section 2.1.3. The magnetic field minimum must be aligned with the crossing point of all six laser beams. Due to the Zeeman effect [48], induced by the magnetic field, the hyperfine energy states will split up. The energy shift of the atomic levels depends on the magnitude of the magnetic field and the magnetic quantum number m_F. Due to the energy shift the absorption of a red detuned, counterpropagating σ_+ photon becomes more probable than the absorption of a σ_- photon propagating in the same direction as the atom. Fig. 2.4(b) visualizes that the σ_+-transition is shifted into resonance for this configuration. This causes an imbalance in the scattering forces. The atoms will be slowed down and pushed back to the center. Since the magnetic field amplitude changes linearly with the distance to the center, the Zeeman splitting depends on the position of the atom in the magnetic field. The total (approximated) force in one dimension (here e.g. in the z direction) is thus position z and velocity v dependent [47]

$$F_{\mathrm{MOT}} \approx -\alpha v - \alpha\beta z, \tag{2.3}$$

where $\beta = \frac{dB}{dz}\frac{g_F \mu_B}{\hbar k}$, with μ_B being the Bohr magneton and g_F the Landé factor, and α is defined in Eq. (2.2). The realization of such an implementation is called Magneto-optical trap (MOT).

Besides the cooling effect in the MOT, there is also heating due to the recoil during the spontaneous emission and absorption. This leads to a random walk of the atom. The minimal temperature the atoms can reach in the MOT is defined by the Doppler limit [47]

$$T_D = \frac{\hbar\Gamma}{2k_B}, \tag{2.4}$$

with k_B being the Boltzmann constant.

To realize a magneto-optical trap of ^{87}Rb in the experiment, three main requirements have to be met. The laser beams have to be frequency stabilized to drive the chosen atomic transition in Rubidium. Also, the polarization of the light has to be set correctly, and a magnetic quadrupole field has to be generated. In Fig. 2.5 the level structure of ^{87}Rb is shown. The cooling process is realized by the so-called Cooler laser driving the σ_+ transition from $|5S_{1/2}, F = 2, m_F\rangle$ to $|5P_{3/2}, F = 3, m_F + 1\rangle$. After optically pumping all the atoms into $|5P_{3/2}, F = 3, m_F = 3\rangle$ the transition induced by the cooling laser from $|5S_{1/2}, F = 2, m_F = 2\rangle$ to $|5P_{3/2}, F = 3, m_F = 3\rangle$ describes a closed cooling cycle. However, since the laser is red detuned and the polarization may not be perfect, it can happen that some atoms are excited to $|5P_{3/2}, F = 2\rangle$ and can thus decay to $|5S_{1/2}, F = 1\rangle$. To bring these atoms back to the cooling cycle, a Repumper laser drives the transition $|5S_{1/2}, F = 1\rangle \longrightarrow |5P_{3/2}, F = 2\rangle$.

Cooler and Repumper laser both need to be split up into six individual beams that can be aligned along all three axes though the vacuum chamber. This is realized by the use of fiber beam splitters that guide the light to the MOT chamber. The polarization is adjusted to circular polarized light with the help of quarter-wave plates (QWP) mounted behind the fiber output. Both lasers are frequency stabilized relative to a so-called Master laser. The Master laser itself is frequency stabilized to an external Ultra Low Expansion (ULE) cavity, relative to a cross-over resonance of ^{85}Rb, by the Pound-Drever-Hall (PDH) method [50]. The frequency and phase stabilization of other lasers used in the experiment relative to the stable Master laser is realized by an Offset-Lock. This is done by overlapping the laser in question with the Master laser on a photodiode, creating a beat-note signal. The offset locking scheme is implemented using a high frequency divider (ADF4007 from ANALOG DEVICES [51]). For more information on the laser setup and realization of the frequency locking scheme see the Bachelor thesis of Julia Gamper [41] and my Bachelor's thesis [43]. The quadrupole magnetic field is generated by one magnetic coil placed on top and one coil placed at the bottom of the MOT vacuum chamber. These coils will be referred to as MOT coils. Both coils are mounted in a copper block which is water cooled. To reduce eddy currents in the coils, the copper mounts have a slit. The magnetic coil pair creates a quadrupole magnetic field with a magnetic field gradient of 1.93 G/cm/A. Furthermore, bias magnetic coils are mounted in a compensation cage around the chamber. These coils are used to compensate external magnetic fields and also provide a quantization axis, for example for absorption imaging. In addition to that the combination of the compensation coils makes it possible to move the minimum of the quadrupole field. With the given configuration an atom cloud temperature of $\sim 300\,\mu$K is reached, which is still above the Doppler temperature of 146 µK [49] for the cooling transition of ^{87}Rb.

For more detailed information on how the magneto-optical trap was set up see my Bachelor's thesis [43] and Julia Gamper's Master thesis [40]. The optimization by hand of the magneto-optical trap parameters was done in the Master's thesis of Julia Gamper. For testing the machine learning optimization algorithm on a simple optimization problem, the MOT optimization was repeated as a proof of principle. This will be discussed in Section 4.2.

2.1.2 Optical Molasses

As discussed in Section 2.1.1, the minimal temperature of the atom cloud in a MOT is limited by the Doppler temperature. Due to the hyperfine structure of the ^{87}Rb atoms it is possible to achieve Sub-Doppler cooling by doing an optical molasses after the MOT phase. In Section 2.1.1 it has already been discussed that an optical molasses can be realized by turning off the magnetic fields that were required for the MOT. In the MOT the imbalance in the scattering force is caused by the Doppler shift, making the counterpropagating σ_+ polarized light more probable to get absorbed. However, when the

Doppler shift becomes smaller, due to a reduced velocity of the atoms, this imbalance is not sufficient to cool down the atoms. This gives rise to the Doppler limit. When turning off the magnetic field while still illuminating the atom cloud with counterpropagating σ_+ and σ_- light, a new effect appears that causes an imbalance in the scattering force. Due to the overlap of both laser beams the atom sees a polarized light field with a polarization axis rotating around the propagation axis. This causes a motion-induced population difference which will in turn create an imbalance in the scattering force [52]. This leads to a cooling process below the Doppler temperature. Since the optical molasses had already been characterized by Julia Gamper [40] and further optimization is not subject of this thesis, a more in depth derivation is not discussed here. For a mathematical derivation of Sub-Doppler cooling with $\sigma_{+/-}$ light see [52].

Experimentally, the optical molasses after the MOT phase is realized by turning the MOT coils off. Furthermore, the Cooler and Repumper power are decreased, and the Cooler frequency has to be detuned further from the resonance transition. In this way the atoms are cooled down to $\sim 26\,\mu\text{K}$, what is below the Doppler temperature of $146\,\mu\text{K}$ [49].

2.1.3 Magnetic Trap

In the next step the atoms need to be prepared for being transported from the MOT chamber to the science chamber. The magnetic transport is realized by trapping the atoms in a magnetic potential and moving the potential along the transport axis. To be able to do so, an initial magnetic trap has to be generated in the MOT chamber as the first part of the magnetic transport. The general principle of magnetic trapping and the corresponding experimental steps will be explained here. The realization of the moving magnetic trap will be explained in Section 2.2.

An atom with magnetic dipole $\vec{\mu}$ can be trapped in a non-uniform magnetic field $\vec{B}$. The interaction energy of a dipole in the magnetic field is given by [45]

$$V = -\vec{\mu} \cdot \vec{B}. \tag{2.5}$$

If the magnetic moment is processing around the magnetic field faster than the magnetic field direction is changing, the magnetic moment can follow the magnetic field adiabatically [53]. The adiabatic approximation of the potential for an atom in the hyperfine state $|F, m_F\rangle$ is given by $V = \mu_B g_F m_F |\vec{B}|$ [54]. Atoms which are in a state with $g_F m_F > 0$ can be trapped by a magnetic field with a local minimum. These states are known as 'weak field seeking' states. Atoms with $g_F m_F < 0$ are anti-trapped by a magnetic field with a local minimum. They are so-called 'strong field seeking' states and could, in principle, experience trapping in a magnetic field maximum. According to Maxwell's law, however, there are no static magnetic fields with a local maximum. Therefore, only configurations that trap weak field-seeking states can be realized [53].

There are different configurations that can be used to generate a magnetic field with a local minimum. In the experiment such a magnetic field is created with two coils in anti-Helmholtz configuration. This is the same configuration as used for the MOT. The generated magnetic field has a minimum $|\vec{B}| = 0$ at the symmetry center of the two coils and, near the minimum, can be well approximated as a quadrupole field. In a coordinate system with the magnetic field minimum located at the origin, the magnetic field $\vec{B}$

and corresponding potential V are given by [55]

$$\vec{B}(x,y,z) \approx \frac{B'}{2}\begin{pmatrix}-x\\-y\\2z\end{pmatrix} \tag{2.6}$$

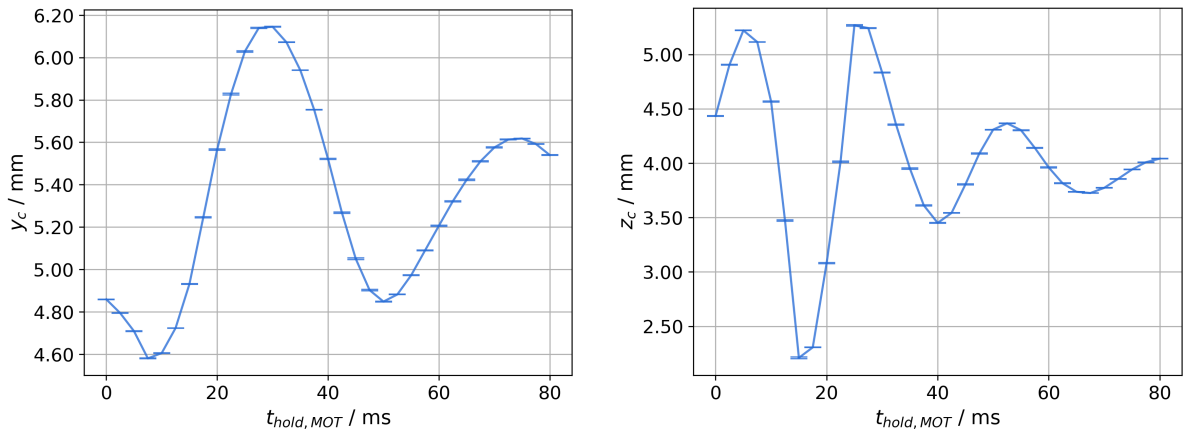
$$V(x,y,z) = \mu_\mathrm{B} g_\mathrm{F} m_\mathrm{F} \frac{|B'|}{2}\sqrt{x^2+y^2+4z^2}, \tag{2.7}$$

where B' is the magnetic field gradient along the strong z-axis. The enhanced magnetic field gradient along the z-axis follows from Maxwell's equation div $\vec{B} = 0$ [45]. The gradient B' depends on the coil geometry and the current I through the coils [55]. Since the coil geometry is fixed by the design of the experiment, the gradient and thus the trapping potential are modified during the experimental sequence by changing the current applied to the coils. In contrast to the quadrupole magnetic field in the MOT, a stronger magnetic field gradient is required for the magnetic trap to cancel out gravity induced forces [56]. Furthermore, the generated zero magnetic field minimum in the anti-Helmholtz configuration can induce Majorana spin flips. This describes the process in which the atoms transition from a trapping spin state ($g_\mathrm{F} m_\mathrm{F} > 0$) to an anti-trapped states ($g_\mathrm{F} m_\mathrm{F} < 0$) due to non-adiabatic perturbations at small magnetic fields. This will lead to atoms getting lost from the trap [53, 54].

To realize a magnetic trap in the experimental setup, the atoms have to be optically pumped to weak field seeking states. For ^{87}Rb the $|5S_{1/2}, F = 2, m_F = 2\rangle$ dark state is such a weak field seeking state. Even though the cooling happens on the $|5S_{1/2}, F = 2, m_F = 2\rangle \rightarrow |5P_{3/2}, F = 3, m_F = 3\rangle$ transition, it is not ensured that all atoms are in the desired dark state after the molasses phase. Therefore, an additional σ_+ polarized laser driving the $|5S_{1/2}, F = 2, m_F\rangle \rightarrow |5P_{3/2}, F = 2, m_F + 1\rangle$ transition, optically pumps the atoms to the dark state in which they can be trapped. Similar to the MOT, the Repumper laser (driving $|5S_{1/2}, F = 1\rangle \rightarrow |5P_{3/2}, F = 2\rangle$) is used for bringing atoms which have decayed to the state $|5S_{1/2}, F = 1\rangle$ back to the optical pumping transition. The optical pumping transition is also shown in the ^{87}Rb level scheme in Fig. 2.5. The optical pumping and Repumper laser are turned on for 20 µs after the optical molasses phase. In the next step, the magnetic trap in the MOT chamber is initialized. The MOT coils generate the magnetic field for trapping with a magnetic field gradient of 130 G/cm along the strong axis.

The process of turning the magnetic fields off after the MOT phase and turning them on again for the magnetic trap introduces some unwanted dynamics of the atom cloud. During the cooling process in the MOT the atoms are trapped and the trapping force in z-direction is counteracting gravity. Subsequently, the atoms fall down when the magnetic field and the MOT beams are turned off. After the optical molasses and the optical pumping the magnetic field is ramped up again. However, eddy currents in the magnetic coils prevent the magnetic field to build up instantaneously. It takes 13 ms for the magnetic field of the MOT coils to reach its maximum value when being ramped up from 0 G/cm to 130 G/cm and another 25 ms to stabilize (see Julia Gamper's Master's thesis for the characterization of the coils [40]). This gives a rather long time for the atom cloud to fall down and accelerate before being trapped again. The increase in kinetic and potential energy of the atoms lead to a sloshing of the atom cloud around the magnetic trap minimum and subsequently an increase in temperature of the atom cloud.

This can be verified by trapping the atoms in the magnetic trap and performing absorption imaging after different holding times in the trap. The experimental sequence for this measurement is the following. The atom cloud is first cooled and trapped in the MOT for 1 s, then the magnetic field is turned off for 5 ms to realize sub-Doppler cooling and after this the atoms are pumped to the weak field seeking



(a) (b)

Figure 2.6: Measurement of the cloud center positions y_c and z_c along the **(a)** y-axis and **(b)** z-axis for different holding times $t_{\text{hold, MOT}}$ in the magnetic trap generated by the MOT coils. The experimental sequence is composed of MOT phase, optical molasses phases, optical pumping and subsequent trapping in the magnetic trap. The datapoints are averaged over 20 measurements.

dark state for 20 µs. After this the magnetic trapping field is turned back on and absorption imaging is performed after different hold times of the atom cloud in the magnetic trap.

The center position of the atom cloud along the y- and z-axis can be extracted from these absorption images and is plotted against the trap holding time in Fig. 2.6. The center position is defined relative to an arbitrary origin in the capture area of the MOT absorption imaging camera. This is sufficient here, since the aim is to extract the relative oscillation amplitude of the center position. The amplitude of the oscillation describes the actual distance that was covered by the cloud position in the MOT, since the pixel size of the camera and the magnification of the imaging system were taken into account here. The atoms are sloshing with a maximal amplitude of ~ 1.5 mm in y-direction and with a maximal amplitude of ~ 3 mm in z-direction. As expected the sloshing in z-direction is higher than in y-direction due to the gravitational acceleration during the molasses and optical pumping time.

The increase in temperature can be observed by doing a time of flight measurement before and after turning on the magnetic trap for the first time. The expansion of the atom cloud is measured along both axes y and z of the imaging plane. Thus, two temperatures $T_{y,z}$ for both axes can be extracted, by fitting Eq. (2.1) to the data. For thermalized clouds the measured temperature values along both axes are approximately the same. However, when the cloud is not fully thermalized the temperatures along the two axes can differ. Since the atom cloud has only one temperature, the extracted temperatures along both axes can not always be understood as the actual temperature of the atom cloud. However, they still give information about the dynamic of the atom cloud and are a measure for the energy in the system. Therefore, always both extracted temperature values are stated here for the TOF measurements.

The results of the TOF measurement immediately after the optical pumping phase are shown in Fig. 2.7(a) and Fig. 2.7(b). The measurement yields temperatures of $T_y = (56.4 \pm 2.2)$ µK and $T_z = (56.9 \pm 2.2)$ µK, which lie within a 1σ range. The atom cloud temperature is below the Doppler temperature due to the Sub-Doppler cooling during the optical molasses. For the second measurement the magnetic field, with a magnetic field gradient of 130 G/cm, was turned on after the optical pumping phase. The atoms were trapped for 500 ms to ensure that the sloshing is sufficiently reduced before

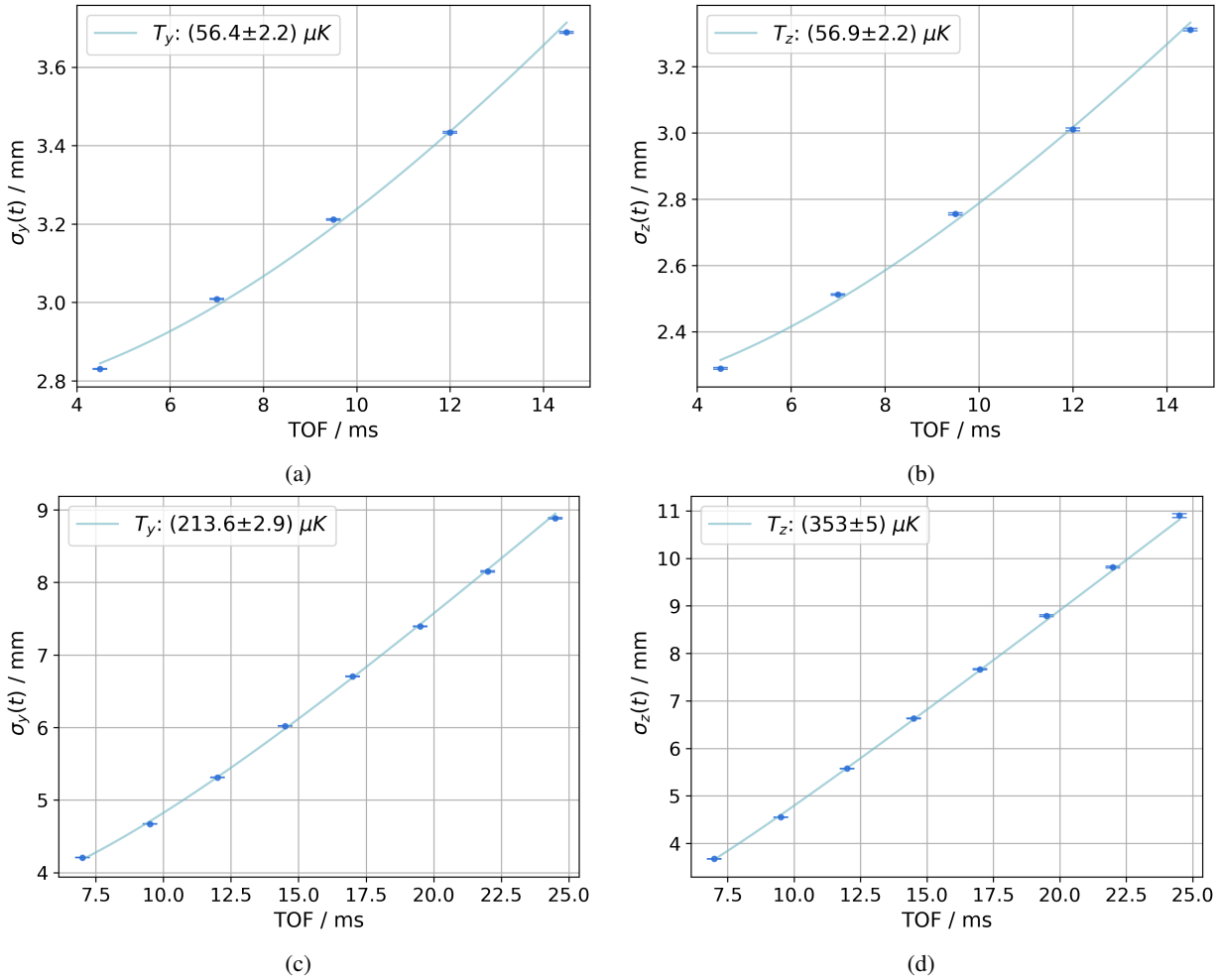


Figure 2.7: Time of flight measurement in the MOT chamber. $\sigma_{y,z}$ describe the atom cloud widths along the axes of the imaging plane. The temperatures are extracted by fitting Eq. (2.1) to the datapoints. The datapoints were averaged over 20 measurements. **(a)-(b)** Results for the TOF measurement after the optical pumping phase, without turning on the magnetic trap in the MOT chamber. **(c)-(d)** Results for the TOF measurement after 500 ms holding time of the atoms in the magnetic trap.

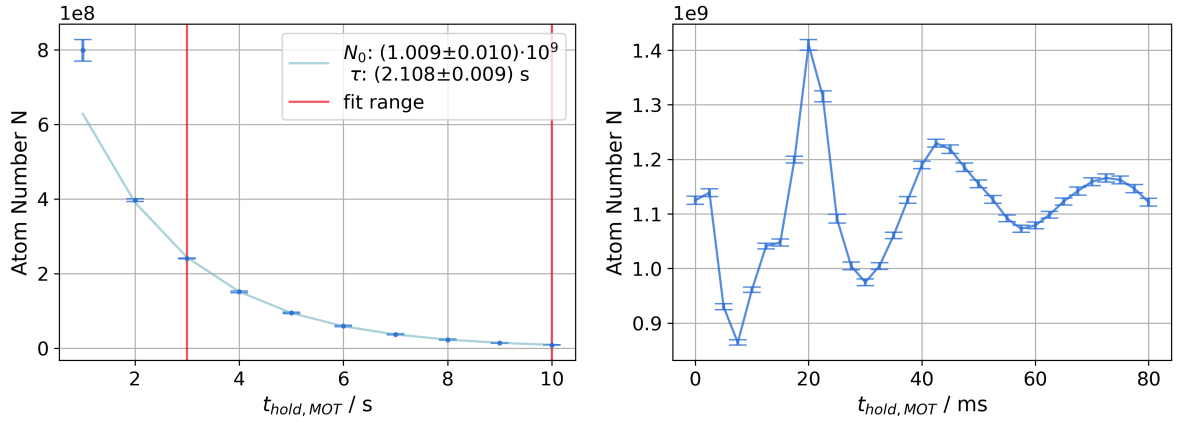
repeating the TOF measurement. The temperature extracted for the y-direction increased by almost a factor of ~ 3.8 and for the z-direction by a factor of ~ 6.3, leading to temperatures even above the Doppler temperature. The measurements show that the cloud gets heated up when the magnetic trap is turned on.

Another important characteristic of the atom cloud in the magnetic trap is their lifetime. Atom losses in the magnetic trap are mainly induced by collisions with background gas [57], and Majorana spin flips [58]. With an atom cloud temperature of the order of 200 µK, the contribution of Majorana spin flips to the atom loss mechanism can be neglected [58]. Therefore, the number of atoms trapped as a function of time can be described by an exponential decay [57]

$$N(t) = N_0 \cdot \exp\left(-t/\tau\right), \tag{2.8}$$

where τ is the trap lifetime. The lifetime can be experimentally measured by repeating the holding time scan over a larger scan range, extracting the atom number for each holding time $t_{\text{hold, SC}}$ and fitting the exponential function to the datapoints. This is shown in Fig. 2.8(a). The lifetime is measured to be $\tau = (2.108 \pm 0.009)$ s. The lifetime measurement starts at a trap holding time of 1 s, because the determination of the atom number with the Gaussian fits is not reliable for shorter holding times. The cloud is not fully thermalized at these points due to the added sloshing dynamics. For this reason, the cloud cannot be perfectly described by a Gaussian distribution. The measurement of the atom number for holding times between 0 ms and 80 ms can be seen in Fig. 2.8(b). The increasing and decreasing of the atom number is not expected and is probably the result of a poor fit, the measured atom number for such short trap holding times thus only a rough estimate of the real value. When taking into account the result of the fit parameter N_0 (see Fig. 2.8(a)), describing the number of atoms at $t_{\text{hold,MOT}} = 0$, and the rough estimates from this measurement one can approximate the number of atoms trapped in the first trapping stage as $(1.0 \pm 0.1) \cdot 10^9$ atoms. The stated error for the atom number takes into account the systematic uncertainty of the measurement. This value will be used as a reference for calculating the transport efficiency from the MOT chamber to the science chamber.

The oscillation of the atoms, the extracted temperature values and the number of trapped atoms describe the properties of the atom cloud at the beginning of the magnetic transport, since the magnetic trap generated by the MOT coils is the first trapping stage of the transport.



(a) (b)

Figure 2.8: Measurement of the atom number N in the MOT chamber for different holding times $t_{\text{hold, MOT}}$ in the magnetic trap generated by the MOT coils. The experimental sequence is composed of MOT phase, optical molasses phases, optical pumping and subsequent trapping in the magnetic trap. **(a)** Measurement for holding times between 1 s and 10 s. The datapoints are averaged over 6 measurements. The quantities N_0 and τ are extracted by fitting Eq. (2.8) to the datapoints. **(b)** Measurement for holding times between 0 ms and 80 ms. The datapoints are averaged over 20 measurements.

2.2 Magnetic Transport

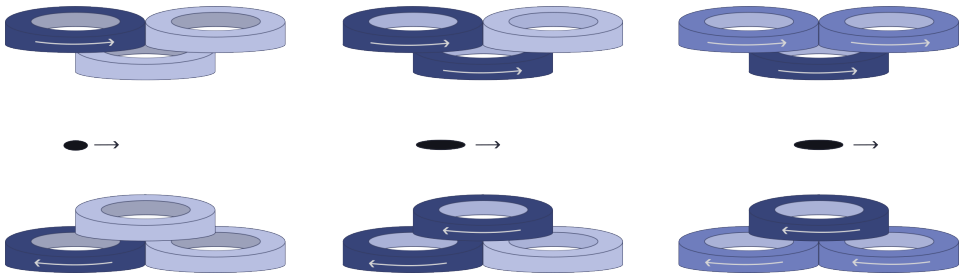


Figure 2.9: Illustration, adapted from [59], of the magnetic transport principle. The left most picture shows an atom cloud trapped in a magnetic field generated by current, indicated by the dark blue color, running through a single coil pair. By applying current to two partly overlapping coil pairs, shown in the middle picture, the center of the atom cloud is moved between the symmetry axes of both coils. The cloud is elongated along the transport axis. For the right most picture the trapping potential is generated from three adjacent coil pairs. The current running through the outer coil pairs is reduced compared to the center coil pair, illustrated by a lighter blue color.

After magnetically trapping the atoms they need to be transported over 450 mm from the MOT chamber to the science chamber. The design for the magnetic transport was done by Cedric Wind, based on the implementation of Greiner et al. [59]. The magnetic transport in the HQO experiment was set up by Cedric Wind and Johanna Popp [35]. This section will explain the basic principle of the magnetic transport and discusses the implementation into the experiment.

The idea behind the magnetic transport design is to move a magnetic trapping potential along the transport axis, thereby transferring the trapped atoms from one side to the other. The moving magnetic trapping potential can be realized by a chain of magnetic coils that partly overlap [59]. Three such partly overlapping coil pairs are illustrated in Fig. 2.9. As discussed in Section 2.1.3 applying current to only one coil pair will produce a magnetic field according to Eq. (2.7), with the position of the potential minimum located at the symmetry center of the coils. An atom cloud trapped in such a magnetic trap is illustrated in the leftmost picture of Fig. 2.9. When decreasing the current in the first coil pair and at the same time increasing the current in the adjacent coil pair the generated magnetic fields of both coil pairs overlap. The resulting magnetic field can be described again as a quadrupole field [59]. The minimum position of the respective trapping potential lies between the symmetry center of the first and second coil pair. The exact position depends on the ratio of the current carried by the two coils [55]. If the current in coil pair two is further increased and the current in coil pair one decreases to zero the potential minimum will be shifted to the symmetry center of the second coil pairs. In this way it is possible to iteratively displace the potential along the transport axis [59].

However, this would lead to varying magnetic field gradients along the weak $\vec{B}$-field axes during the transport, assuming the magnetic gradient B' along the strong axis is kept constant during the transfer. The magnetic field gradient along the transport direction would be weaker when being generated by two overlapping coils compared to the pure Antihelmholtz configuration, while the other weak axis would be increased accordingly [55]. This would lead to a flatter potential along the transport axis, thus producing a more elongated trapped atom cloud, what is shown in the center image of Fig. 2.9. The changing properties of the potential during the transport could cause the atoms to heat up. By adding a third coil pair it is possible to preserve the flattened magnetic field gradient along all axes during the transport [59]. Instead of using the magnetic field of a single coil pair to generate the trapping potential

at the symmetry center of the respective coil pair, the potential is generated by the overlapping magnetic field of three adjacent coil pairs. This is realized by applying suitable currents to both coil pairs sitting next to the central one [59], as it is indicated in the rightmost picture of Fig. 2.9. The resulting magnetic field around the trap center, which is displaced along the transport axis y in this way is given by [55]

$$\vec{B}(x,y,z) \approx B' \cdot \begin{pmatrix} -\frac{\alpha}{1+\alpha} \cdot x \\ -\frac{1}{1+\alpha} \cdot y \\ z \end{pmatrix}, \tag{2.9}$$

where the parameter α, describes the ratio of the magnetic field gradient in x-direction B'_x and y -direction B'_y.

The magnetic transport assembly in the HQO experiment is shown in Fig. 2.10. It consists of seven coil pairs placed next to each other which are partly overlapping. This enables the transport of atoms over a distance of 450 mm along the y-axis. The coordinate axis refers to the reference system introduced in Fig. 2.1. The first coil pair is the one that is also used to generate the magnetic field for the MOT and is therefore called MOT coil. Coil pair 5 is a double-stacked pair, consisting of 4 coils in total. This ensures that even though the coil pair is placed further from the transport axis, it produces approximately the same magnetic field gradient as the rest of the transport coils. All coils are water cooled to prevent them from heating up. The start point of the magnetic transport at $y = 0\,\text{mm}$ and the end point at $y = 450\,\text{mm}$ correspond to the symmetry center of the MOT coils and the SC trap coils respectively. At these positions it is not possible to preserve the flattened magnetic field potential since the magnetic field is generated by a single coil pair each. Therefore, the $\vec{B}$-field gradient along the x-axis is weaker while for the y-axis it is greater compared to the rest of the transport. However, due to the chosen coil configuration this can not be prevented here. Apart from this it is possible to remain the geometry of the trapping potential throughout the transport with the given transport design. The generated magnetic fields have a magnetic field gradient along the strong axis (z-axis) of $B' = B'_z = 130\,\text{G/cm}$. The magnetic field gradients in x- and y-direction are constrained to $B'_x \approx 95\,\text{G/cm}$ and $B'_y \approx 35\,\text{G/cm}$. This results from Eq. (2.9), with α fixed to ~ 2.71 by the setup. At $y = 0\,\text{mm}$ and $y = 450\,\text{mm}$ the magnetic field is given by Eq. (2.7) resulting in magnetic field gradients of $B'_x = B'_y = 65\,\text{G/cm}$.

The currents that are required to generate a quadrupole magnetic field at each position y_{V_0} along the transport axis can be seen in the upper part of Fig. 2.10. The position y_{V_0} refers here to the zero crossing of the magnetic field, what corresponds to the position of the potential minimum. The current traces were extracted from a magnetic field simulation that was written by Cedric Wind.

For the implementation in the experiment we want to define a time dependent transport curve $y_{V_0}(t)$. This requires a mapping of y_{V_0} to the corresponding currents as a function of time. This mapping is done by using the simulated relationship between y_{V_0} and the current traces, shown in Fig. 2.10. The mapping between the transport curve to the time dependent current traces is illustrated for an example trajectory $y_{V_0}(t)$ in Fig. 2.11. The $y_{V_0}(t)$ trajectory is used as input for the magnetic transport simulation, written by Cedric Wind. The simulation generates the corresponding current time profiles for every coil pair. Thus, the magnetic transport can be controlled and modified by the choice of transport curve. $y_{V_0}(t)$ must be a smooth curve that fulfills the boundary conditions $y_{V_0}(t = 0) = 0\,\text{mm}$ and $y_{V_0}(t = T_{\text{MT}}) = L$, where T_{MT} is the total transport duration and $L = 450\,\text{mm}$ is the length of the transport. The time dependent current traces generated by the magnetic transport simulation are saved in files, that can be read by the experiment control system. The experiment control will be introduced in Section 2.4. The power supplies, which are used for applying the current to the different transport coils, are remotely

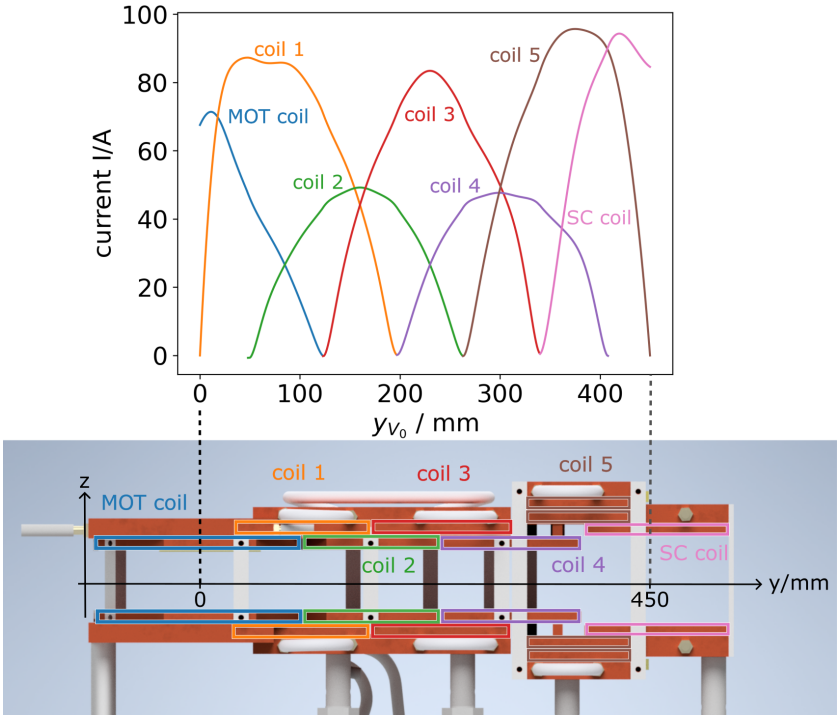


Figure 2.10: The inner assembly of the magnetic transport, consisting of seven partly overlapping coil pairs. The MOT coil is placed in the MOT chamber and the SC coil in the science chamber. The total transport length is $L = 450\,\text{mm}$. The plot above the magnetic transport shows the required current to generate a magnetic trapping potential at the respective position along the transport axis.

controlled by the computer control system. This is indicated in Fig. 2.11 by the arrow from experiment control system to PSU 1-4. In total four power supplies (from Delta Elektronika (2x type SM15-100 P218, 1x type SM66-AR-110, 1x SM70-90 P069)) are used. The power supplies are all connected to a home build IGBT (insulated-gate bipolar transistor) switching box, as depicted in Fig. 2.11, to change their connection to the coil pairs. In that way a single power supply can be used to power two coil pairs during the transport. Although current is going through maximally three coils at any given time, a fourth power supply is used to allow for smooth switching. The time of the switching is controlled by TTL signals, which are timed by the experiment control system. The start and stop point of the respective TTL signals is extracted from the current time profile of each coil pair. Seven out of the eight outputs of the switching box are connected to the seven transport coil pairs. The connection between switching boxes and coil pairs is also illustrated in Fig. 2.11. In this way it is possible to apply the correct current at the corresponding time to the magnetic coils to successfully transport the atoms from MOT chamber to the SC chamber.

When arriving in the science chamber, the atoms remain trapped in the last trap potential generated by the SC coil pair. After some holding time, one can either do absorption imaging to characterize the performance of the transport or the experimental sequence continuous with Rydberg excitation and subsequent detection. In the future, when the atom chip is built into the setup, the next step would be to transfer the atoms from the quadrupole trap into the Z-wire trap integrated on the atom chip. The simulated transfer process is discussed in Leon Sadowski's Master's thesis [60].

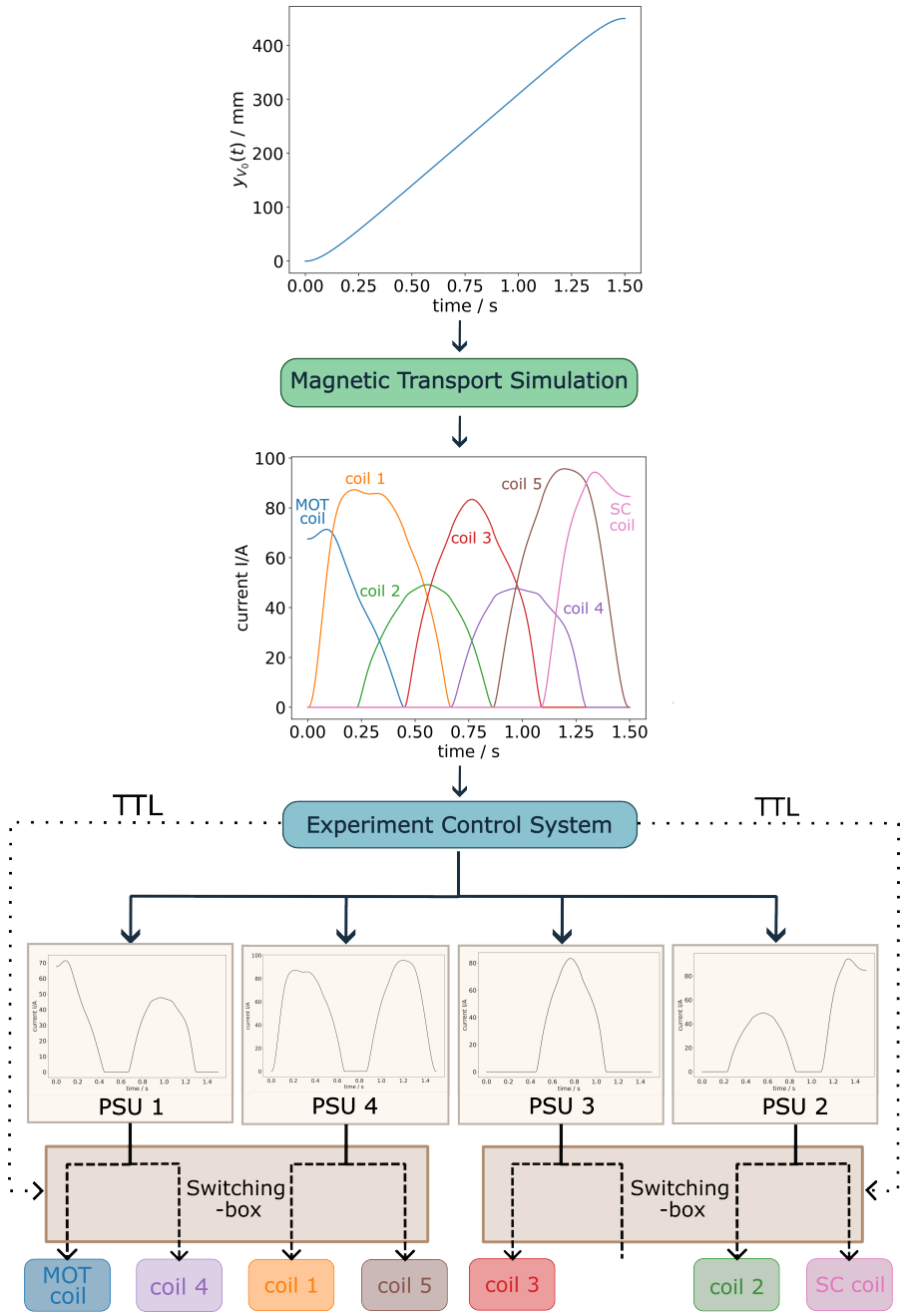


Figure 2.11: Implementation of the magnetic transport in the experiment. The time profile for the potential minimum position $y_{V_0}(t)$ is used as input for the magnetic transport simulation, written by Cedric Wind. This will generate current traces as a function of time for each coil pair. The experiment control system builds its internal sequence based on the generated current traces and forwards the current traces to the respective power supplies (PSU). The switching boxes control to which coil pair the output of the PSUs is applied to. The switching boxes themselves are controlled by TTL signals, which are timed by the experiment control system.

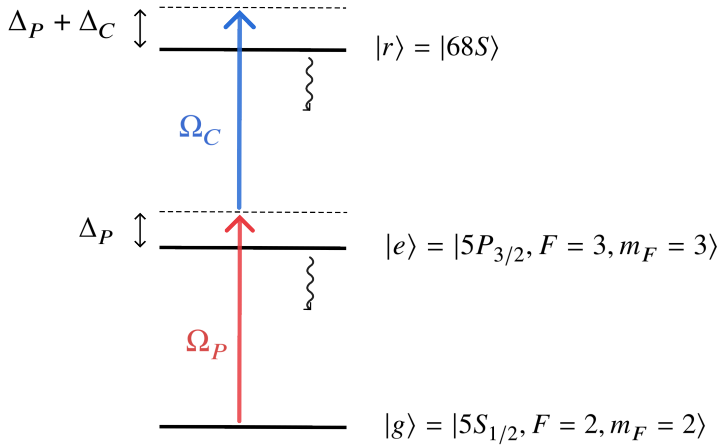


Figure 2.12: Off-resonant two-photon excitation scheme for exciting ^{87}Rb to the $|r\rangle = |68S, m_J = \frac{1}{2}\rangle$ Rydberg state. Ω_P and Ω_C describe the Rabi frequency for the probe and control laser and Δ_P and Δ_C are the detunings of the respective lasers.

2.3 Rydberg Excitation and Detection

In order to couple the Rubidium atoms to the electromechanical oscillator, they must be excited to a Rydberg state. Therefore, once the atoms have been transported to the science chamber, Rydberg excitation and subsequent detection takes place. The atoms are excited to a Rydberg state via two-photon excitation. For the detection of Rydberg atoms two methods are implemented in the experiment. The first one is based on measuring the few photon transmission spectroscopy with Single Photon Counter Modules (SPCMs). Additionally, the successful excitation to Rydberg states can be verified by ionizing the Rydberg atoms and subsequently detecting the Rydberg ions with a microchannel plate (MCP). The following section gives an overview of the Rydberg excitation scheme and the two detection methods. In particular, the setup for ionization of Rydberg atoms and ion detection is discussed here, whereas the characterization of the detection setup is presented in Chapter 5.

The Rubidium atoms can be excited to a Rydberg state via two-photon excitation. This is schematically shown in Fig. 2.12. A weak probe beam, with a wavelength of 780 nm, couples the ground state $|g\rangle = |5S_{1/2}, F = 2, m_F = 2\rangle$ to the intermediate state $|e\rangle = |5P_{3/2}, F = 3, m_F = 3\rangle$. Simultaneously, a strong 480 nm control beam drives the transition from the intermediate state to the Rydberg state $|r\rangle = |68S, m_J = \frac{1}{2}\rangle$. For the Rydberg states the hyperfine splitting is negligible and thus the $|J, m_J\rangle$ basis represents good quantum numbers [61]. The desired Rydberg state was chosen in this way, since a first generation chip was designed to feature a resonator with a frequency corresponding to the $68S \rightarrow 68P$ transition in ^{87}Rb [60]. *Weak* and *strong* beam refers to the laser power, which are chosen such that $\Omega_C \gg \Omega_P$, where Ω_C and Ω_P are the Rabi frequencies of control and probe laser respectively. The low power of the probe laser and the higher power of the control beam are required to achieve similar coupling strengths for the $|g\rangle \rightarrow |e\rangle$ transition and the $|e\rangle \rightarrow |r\rangle$ transition. This is due to the smaller dipole matrix element of the latter [61]. The power of the probe laser lies in the single photon regime and the control laser power is set to ~ 100 mW.

It is necessary to use SPCMs to detect the probe transmission signal after it has passed through the atom cloud, due to the low power of the probe beam. The probe transmission spectroscopy is measured

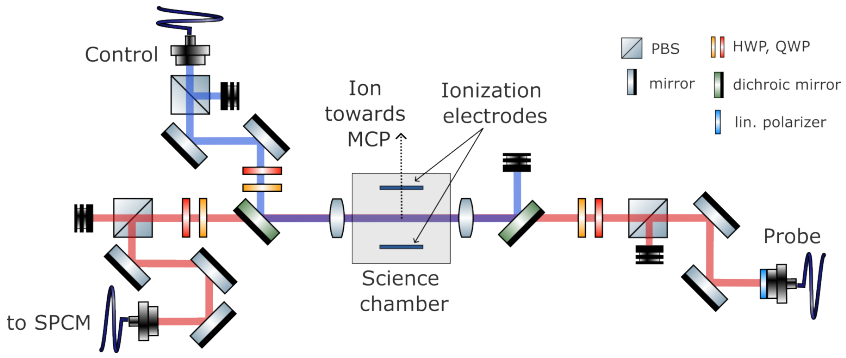


Figure 2.13: Optical setup for Rydberg excitation. Probe and Control laser are counterpropagating through the science chamber. QWP-plates and HWP-plates are used for circular polarizing the light. The position of the ionization electrodes are also schematically shown. The dotted arrow indicates the flight direction of the Rydberg ions after ionization.

by scanning the detuning Δ_P of the probe laser. With the control light turned off, the typical transmission valley of the probe light is observed. If control and probe light are simultaneously turned on, Rydberg excitation can take place. However, two cases can be distinguished in terms of Rydberg excitation. If both probe and control laser are on resonance the probe laser drives the $|g\rangle \rightarrow |e\rangle$ transition and the control laser couples the intermediate state to the Rydberg state. In this case one expects the medium to become transparent for the probe photons leading to a transmission peak inside the probe transmission valley around $\Delta_P = 0$. This is called electromagnetically induced transparency (EIT) and is the result of destructive interference of possible excitation paths. For more information about EIT see e.g. [62].

If both laser are detuned from resonance, off-resonant Rydberg excitation is possible. This can be realized if the detuning of probe and control compensate each other, such that $\Delta_P + \Delta_C \approx 0$. In this case an additional dip in the probe transmission signal at the probe detuning Δ_P can be observed. Both signals indicate that the intermediate state was coupled to the Rydberg state, and can therefore be used for Rydberg detection.

Fig. 2.13 shows the optical setup used for Rydberg excitation and the detection of the probe transmission. The frequency stabilized probe and control laser are guided by fibers to the science chamber. Both light beams are circularly polarized by passing through a half-wave plate (HWP) and a quarter-wave plate (QWP) before being focused into the chamber. The two aspheric lenses on both sides of the chamber were chosen such that the control beam has a waist of 50 µm at the center of the chamber and the probe laser has a waist of 9 µm. After the probe beam passed though the atom cloud the light is coupled to the SPCM (COUNT-250C-FC from Laser Components [63]) to detect the transmitted photons.

The counting of the single photon signals is handled by a Time Tagger (from Swabian Instruments [64]), which is a streaming time-to-digital converter. The dichroic mirrors are used for overlapping probe and control beam. They also prevent control light being coupled into the SPCM. With the given setup it is possible to measure the on resonance EIT peak in the transmission signal and observe an off resonant absorption dip in the probe transmission spectrum. The detection with SPCM is documented in the Master thesis of Julia Gamper [40] and will not be discussed here in more detail. The off resonant signal will be used in Section 5.2 to determine the ion detection efficiency.

The second implemented detection scheme is based on the ionization of Rydberg atoms and subsequent detection of created ions. The ionization scheme makes use of the fact that the Rydberg valence electron

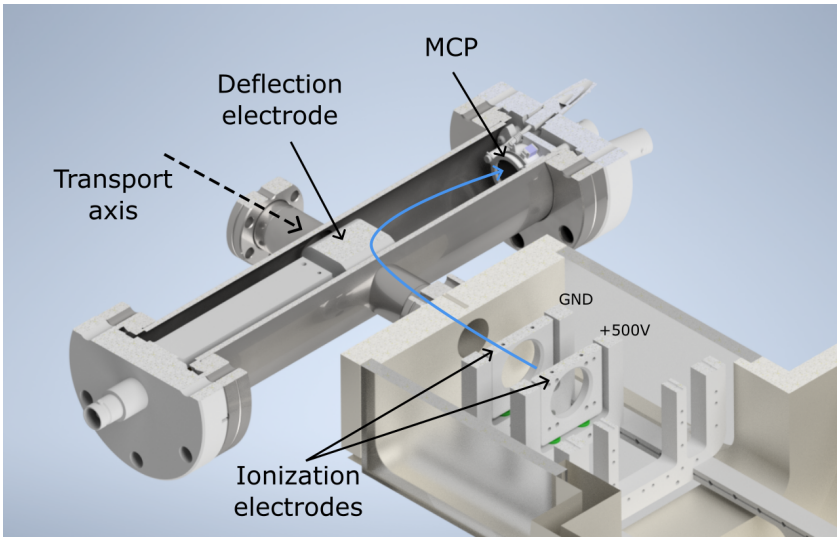


Figure 2.14: Inner assembly of the ionization and detection setup. The ionization electrodes are placed in the science chamber. One of the ionization electrodes is positively charged, while the other one is grounded. In this way the Rydberg atoms are ionized and subsequently pushed away. The positively charged deflection electrode guides the ions towards the MCP. The blue arrow indicates the ion trajectory.

is weakly bound to the core. Thus, it is possible to ionize the Rydberg atom by applying an external electric field [65]. The external field E modifies the Coulomb potential of the Rydberg valence electron. Classically the total potential can be approximated by the superposition of the Coulomb potential with the additional electric potential as $V(r) \approx -\frac{1}{r} - eE \cdot r$ [15]. The required electric field for ionizing Rydberg s-states derived from this approximation is given as [66]

$$E_I \approx 3.2 \cdot 10^8 \frac{\mathrm{V}}{\mathrm{cm}} \frac{1}{n^{*4}}, \tag{2.10}$$

where $n^* = n - \delta_{nlj}$ is the effective principle quantum number with n the principle quantum number of the Rydberg state and δ_{nlj} the corresponding quantum defect for a Rydberg state with quantum numbers n, l, j [15].

From Eq. (2.10) it is apparent that the required electric field for ionization is state dependent, enabling in principle state-selective field ionization. By linearly ramping the electric field, adjacent Rydberg states can be ionized at distinct times according to their ionization energies. This leads to time dependent ionization signals from which the Rydberg state distribution could be extracted [67]. The subsequent detection of the Rydberg ions is commonly done with MCPs [61]. The working principle of the MCP will be explained in more detail in Section 5.2, where the characterization of the ion detection setup is discussed.

In the experiment the ionization of the Rydberg atoms and the ion detection are realized with the setup shown in Fig. 2.14. Two electrodes are placed inside the science chamber along the axis of the magnetic transport. After the Rydberg excitation pulse, the ionization field is turned on, and one of the electrodes is supplied with up to 500 V. The other electrode remains grounded. This creates an electric field of up to 192 V/cm. The electric field required for ionizing the $|68S\rangle$ Rydberg state can be determined, with Eq. (2.10), to be ~ 18.1 V/cm. Therefore, it is possible to ionize the atoms excited to the desired Rydberg

state. After ionization the positively charged ions are pushed away from the positively charged electrode. They are accelerated towards the grounded electrode, which they can pass through an integrated hole. An additional deflection electrode is placed behind the second electrode along the same axis. By positively charging the deflection electrode the ions get deflected and accelerated towards the front plate of the MCP (F4655-11 from Hamamatsu Photonics K.K. [68]). The signals generated by the MCP are counted as well by the Time Tagger. The generation of the MCP output signals and the collection of the signals by the Time Tagger is explained in more detail in Section 5.2.

It should be noted that the described ionization setup is only implemented temporarily for the room temperature setup. Once the cryostat is integrated into the experiment, the ionization unit will be implemented according to the design of Leon Sadowski [60]. However, the deflection electrode and the detection of the ion signals with the MCP will remain unchanged. Therefore, understanding and characterizing the MCP signals in the room temperature setup will also be useful for the future setup at cryogenic temperatures.

The Rydberg excitation/detection sequence is mainly controlled by a high speed multi-channel arbitrary digital pulse pattern generator[2] (in the following called pulse generator). The pulse generator with a time resolution of 2 ns is suitable for generating the fast Rydberg excitation and detection pulses. The Rydberg excitation and detection sequence is illustrated in Fig. 2.15. After the magnetic transport the atoms are held for another 80 ms in the magnetic trap of the SC coil before the Rydberg excitation and detection sequence starts. This is indicated in Fig. 2.15 by the magnetic field gradient $B'_{\mathrm{SC\,coil}}$ of the SC coil set to 130 G/cm for $t < 0$. A single Rydberg excitation and detection sequence takes 100 µs and is structured as follows. Probe and Control light are turned on for 11 µs and 12 µs respectively. As soon as the probe pulse is turned off again, the ionization of the Rydberg atoms begins. The ionization and deflection control signals in Fig. 2.15 are trigger signals for the high voltage switch (AMXT-500-EF High Voltage Switch from CGC INSTRUMENTS [70]), which controls the voltage supply to the electrodes. As soon as the voltage switch is triggered, it switches the supply from ground to the desired voltage. The ionization and deflection electrodes are turned on for 42 µs. The SPCM counts are measured by the Time Tagger for 14 µs. This is indicated in Fig. 2.15 by the SPCM Time Tagger trigger set to high during this time. The measurement of the MCP signals with the Time Tagger begins shortly after the ionization start and stops shortly before the end of the ionization process. The reason for this is that cross-talk to the MCP was induced when applying the high voltages for ionization and deflection to the corresponding electrodes. Shortening the measurement time window could eliminate cross-talk from the data. The Rydberg excitation and detection with the SPCM and the MCP is repeated 1000 times in one experimental cycle. After the 1000 Rydberg pulses, which take a total 100 ms, the magnetic field of the SC coil pair is turned off. The pulse generator then waits 20 ms before repeating the 1000 excitation and detection pulses. These are background measurements, since no atoms are trapped anymore. This sequence enables the excitation of Rydberg atoms, and allows both Rydberg detection schemes to be employed simultaneously.

[2] developed by Felix Engel during his Bachelor's thesis [69]

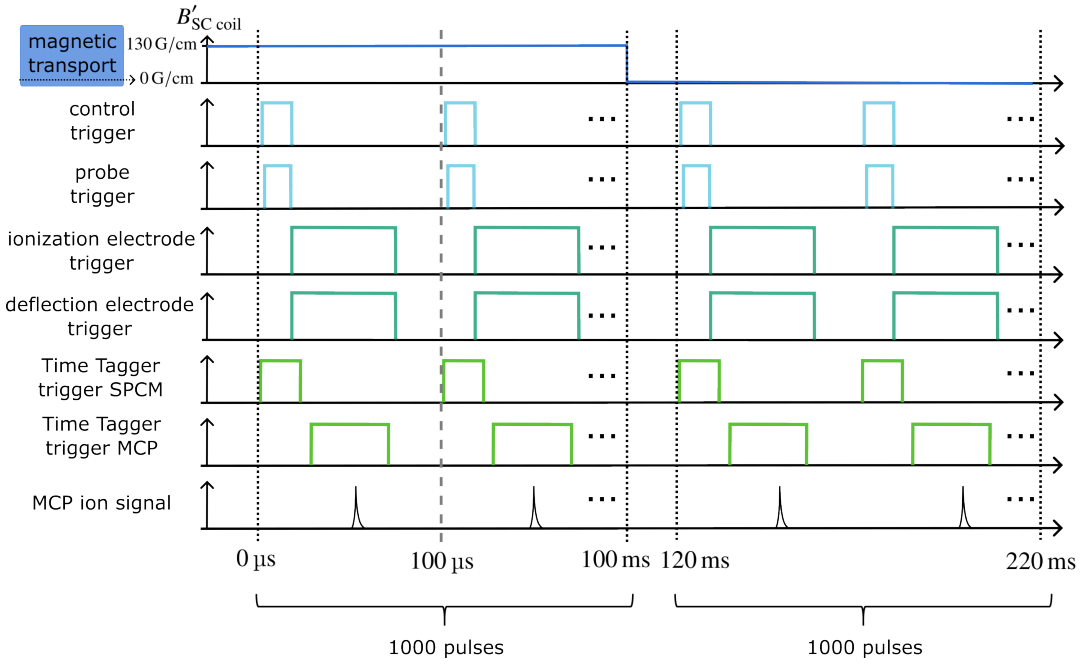


Figure 2.15: Schematic description of the Rydberg excitation and detection sequence. The time axis starts with the beginning of the sequence. The sequence is divided into two parts. During the first 100 ms 1000 Rydberg excitation and detection pulses are applied with the magnetic trap in the SC turned on. This is indicated by the magnetic field gradient B'_{SC} set to 130 G/cm. One pulse corresponds to one Rydberg excitation and detection pulse, which take in total 100 µs. After a short waiting time the 1000 pulses are repeated without atoms trapped, since the magnetic field is turned off. The pulses required for Rydberg excitation are colored in light blue, the pulses for the ionization and deflection process are colored in dark green and the measurement with the Time Tagger is indicated by the bright green pulses. The MCP ion signal describes the signal generated by the MCP.

2.4 Experiment Control System

The experimental cycle is fully controlled by the experiment control system. This is a software, written in C#, which can be found on GitHub[3]. The experiment control system generates a sequence that controls and times the output of an ADwin. The ADwin is a real-time-processor with digital and analog outputs. In this experiment, an ADwin-Pro II with a total of 16 analog channels and 32 digital channels is used. It has a time resolution of 20 µs [72]. The digital outputs can be set to 0 V (low) and 5 V (high) and are used as TTL trigger signals for different hardware devices. The TTL signals are used to control, for example, the acousto optical modulators (AOMs) in the optical setup (see Section 4.2), trigger the imaging camera or to switch the power supply output between different coils (see Fig. 2.11). The analog channels provide signals between 0 V and 10 V and are used to control, for example, the output of the power supplies for the magnetic coils or adjusting the RF power of the AOMs.

Controlling all hardware devices via a single experiment control interface ensures that their output signals are synchronized with the rest of the sequence. The machine learning optimization routine will run in parallel with the experimental cycle. For a successful implementation of the optimization routine, both routines must be coordinated with each other. The experimental cycle will be briefly explained here

[3] https://github.com/coldphysics

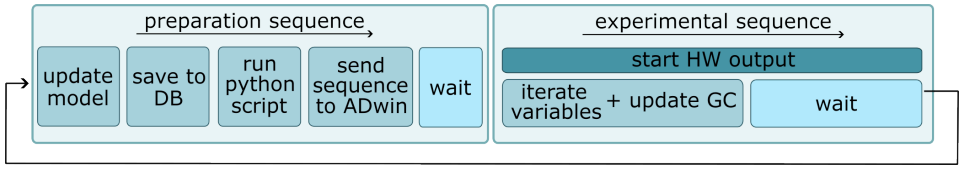


Figure 2.16: Cycle of the experiment control system, adapted from [71]. During the preparation sequence the experiment control system builds a new sequence model, that is saved to the database (DB). This sequence model will be forwarded to the ADwin, which triggers different hardware devices and controls the hardware output signals. The communication with the hardware devices is initialized via Python scripts. During the experimental sequence the hardware (HW) output is started, and the experiment control handles the iteration of computer controlled variables for the next model. A global counter (GC) is assigned to each sequence model.

to set the ground for the discussion about implementing the machine learning routine into the experiment in Chapter 4.

The output routine for each channel (of the digital and analog ADwin cards) is built up by a sequence of steps with variable duration. The step duration can be set with a variable defined in the user interface (UI) of the experiment control system. In addition, the step output of a digital channel can be defined as either high or low in the UI. For the analog channels there are different ways to program the output value in the experiment control system. It can be set to a constant value, linearly ramped to a certain value, or the output signal can be loaded from a csv or binary file. Furthermore, all the computer controlled variables can be iterated during the experiment. This is often used to determine the optimal value of a parameter in an experimental sequence when optimizing manually. A global counter GC is assigned to each experimental sequence. Every sequence model and corresponding result of the experiment, that means the absorption images taken, are saved under the respective global counter.

The experiment control cycle is shown in Fig. 2.16. It can be divided into two parts of fixed length: the preparation sequence and the hardware output sequence, that means the experimental sequence. The length of the latter is defined by the total ADwin sequence length. The preparation time is set by the user when defining the basic underlying model in the experimental control. During the preparation time an internal model is created based on the sequence that was implemented in the experiment control UI. This model is then saved to the MySQL database, which is used for saving the experimental data in this laboratory. Next, the communication with all the hardware devices, which are needed for the sequence (e.g the pulse generator, DDS board, . . .), is initialized via a Python script. A set of control values, defined in the experiment control interface, is transferred to the hardware to configure the state of the devices for the specific run.

The generated output sequence is then sent to the ADwin. The experimental sequence will be started when the fixed preparation time is over. The ADwin forwards all the output signals to the respective hardware devices and the experimental sequence starts. During this time, the experiment control handles potential iterations of variables and updates the global counter. It then waits for the experimental sequence to finish. After this, the internal model is updated again. Potential files, storing signal traces for specific sequence steps, are reloaded for this. The new model is saved again to the database under a new global counter and the cycle repeats itself. The full cycle duration can be kept constant by adjusting the waiting time at the end of the preparation sequence according to how long it took to create the new sequence model. The hardware is set to the last values of the sequence in the preparation phase. The cycle is programmed such that the MOT light and the MOT magnetic fields are turned on during this time, meaning that the MOT loading already starts in the preparation sequence.

CHAPTER 3

M-LOOP - a Machine Learning Online Optimization Package

A scientific experiment consists of a variety of parameters that influence the outcome of a measurement. With a growing parameter space the manual search for an optimum becomes more difficult and at some point nearly impossible. Applying machine learning algorithms for the optimization of a complex scientific system can help to find the optimum more efficiently. The implementation of the machine learning optimization into the HQO experiment is based on the open source machine learning online optimization package (M-LOOP) [21, 73]. It was designed with the purpose to enable machine learning based optimization in computer controlled scientific experiments and was first applied to the optimization of the production of Bose-Einstein condensates [21]. The M-LOOP package combines the idea of online optimization (OO) with online machine learning (OML) [21]. Section 3.1 gives a short introduction to online optimization and online machine learning. The combination of both in the M-LOOP package and the resulting optimization routine is discussed in Section 3.2. In Section 3.3 the optimization algorithms implemented into the M-LOOP package are presented.

3.1 Introduction to Machine Learning Online Optimization

Classical optimization algorithms aim to find the best solution for a mathematical problem. This often refers to finding the minimum or the maximum of an objective function [74]. These optimization problems can be described as either 'online' or 'offline'. Offline optimization refers to the case where a complete dataset is known from the beginning. On the basis of that the given object can be optimized. However, when the optimization is happening in real time, this means data is generated during the optimization process, the problem is described as 'online' [21, 75]. The principle of online optimization is often employed for the optimization of experimental control parameters X in ultra-cold-atom experiments [76]. The optimization routine is here based on first proposing new control parameters and then testing them in the experiment. In this way the performance of the parameters can be directly measured. This will then be used as feedback for the optimization algorithm to further improve the control parameters [77, 78]. The optimization problem is solved by applying classical optimization algorithms, for example, gradient based [76] or genetic algorithms [79].

The idea of online optimization is combined with online machine learning in the M-LOOP package, to make better decisions for new parameters X to test [21]. Machine learning algorithms aim to solve a

given task T by learning about its relevant properties from a training dataset $\mathcal{D}$. The algorithm builds a model, based on the available data, which is then used to make predictions about unseen data. The accuracy of the model is characterized by a so-called performance measure P, which is based on the deviation of the models predictions to the expected results. The optimization of the model's performance describes the learning process [80]. Machine learning algorithms are thus constructed based on the type of dataset, the task that should be solved and the learning process.

The tasks that can be solved by a machine learning algorithm include 'classification' and 'regression' [81]. For classification problems, the algorithms task is to assign the input data to specific and discrete categories. For example, this could be the classes 'True' or 'False' for a binary classification [80]. In the case of a regression problem, the task is to map unseen input variables to their corresponding output target. This task is solved by finding a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ that describes the mapping between input and output variable correctly. Based on the underlying model, the algorithm can make predictions for future inputs [80, 81]. Several other types of tasks are possible, making machine learning applicable to many different areas [81].

The way the dataset $\mathcal{D}$ is presented determines how the algorithm can learn and which tasks can be solved. The dataset can consist of a collection of input datapoints X with target values or labels C assigned to each input. In this case the learning process is described as 'supervised'. If the correctly assigned targets or labels are missing, and the dataset only contains the input datapoints X, 'unsupervised' learning algorithms have to be applied. A third learning category is 'reinforcement' learning, in which the dataset is generated through interactions with the environment. Every interaction and the gained experience, meaning e.g.the success of the action, will be used as a tuple in the dataset. The aforementioned learning categories only cover some of the types of learning in machine learning. However, they represent the most common ones and therefore additional ones are not discussed here [80].

Another distinction that can be made is whether the learning is happening online or offline. This is defined similarly to online/offline optimization. For offline learning, also called batch machine learning, a full set of data is provided at the beginning and will be used in the learning algorithm as training data to construct the model. In contrast to that, for online learning there is a stream of data instead of a fixed dataset in the beginning. The datapoints arrive one at a time and can only be seen once by the algorithm. For each new datapoint the previous model is updated and therefore improves over time. This allows making predictions already during the learning process and not only after having processed the full dataset [82, 83].

3.2 M-LOOP Optimization Routine

The combination of online optimization with online machine learning allows choosing new testing parameters X based on the prediction of the machine learner model. This can make the optimization process more efficient. The resulting machine learning online optimization (MLOO) algorithm is a supervised learning algorithm based on a dataset that is continuously updated during the optimization process [21]. A single datapoint in the dataset consists of a D-dimensional optimization parameter set X that is labeled with a corresponding cost value $C(X)$. The D optimization parameters can be any of the experiment's parameters that are computer controlled. This is necessary so that they can be changed automatically during the optimization process. The cost $C(X)$ is a scalar quantity that characterizes the performance of the experiment with respect to the chosen parameters. The cost function $C(X)$ in dependency of a parameter sets X is the object that should be modeled and minimized at the same time.

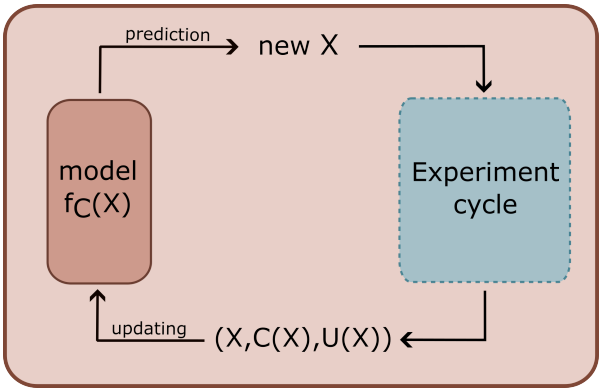


Figure 3.1: Illustration of the M-LOOP optimization routine. The machine learner builds a model f_C of the objective function based on the current dataset. The model is used to make a prediction about a new parameter set X that should be tested. The experiment cycle is then conducted with the updated parameters. A cost value $C(X)$ with associated uncertainty $U(X)$ is extracted from the measurement results. The growing dataset is used to update the model. The cycle repeats itself until some halting condition is met.

The task of the algorithm is to build an underlying model f_C of the objective function that maps each parameter set X to the respective cost value $C(X)$. This model can be used to make predictions about a new parameter set that is most likely to minimize the objective function. The dataset $\mathcal{D} = \{\mathcal{X}, (C(\mathcal{X}), \mathcal{U}(\mathcal{X}))\}$ used to build the model is updated during the optimization process. It consists of all tested parameter sets $\mathcal{X} = (X_1, \ldots, X_N)$, the respective cost values $C(\mathcal{X}) = (C(X_1), \ldots, C(X_N))$ and the uncertainties of the cost $\mathcal{U}(\mathcal{X}) = (U_1, \ldots, U_N)$ [73].

The optimization routine required to update the dataset and improve the prediction about the optimal parameter set is illustrated in Fig. 3.1. The algorithm first generates new testing parameters X based on the current model f_C. The parameter values are chosen within their maximal and minimal boundaries, which must be defined beforehand. In the next step the experiment is conducted with the new parameter values. At the end of the experiment cycle a cost value $C(X)$, with an uncertainty $U(X)$, is extracted based on the measurement results with the given parameter set. The cost value with associated uncertainty and the corresponding parameter set form a new datapoint $(X, C(X), U(X))$ in the dataset. The machine learning algorithm will learn from the updated dataset and improve its internal model f_C [73]. When deciding on the next best testing parameters, the algorithm considers two different types of parameter sets. These are parameter values that are most likely to minimize the current model and parameters with relatively high cost prediction uncertainty. Taking into account parameter ranges for which the model is rather uncertain prevents the algorithm from becoming trapped in local minima. The algorithm selects the parameter type in an alternating manner [21]. The experiment is performed again with the new set of parameters and the full cycle repeats itself until some halting condition is fulfilled. The halting condition can be defined by a maximal number of total optimization runs, a maximal number of optimization runs without finding better parameters, a target cost that has to be reached or a maximal duration of the total optimization routine. After this the optimization is finished and the best found parameters will be returned [73].

The advantage of online optimization over offline optimization is that the algorithm receives immediate feedback about its predictions. As the dataset grows over time, the model becomes more precise, allowing more accurate predictions. This means that the choice of the next testing parameters is an informed

decision, based on the goal of minimizing the cost. This can lead to a significant improvement in efficiency of the optimization process compared to a randomly chosen dataset, as would be the case with offline optimization [21].

3.3 Optimization Algorithms

Machine learning online optimization can be implemented using different algorithms. One of the most common ones are Gaussian processes and neural nets, which are also the ones that are used by the M-LOOP package. Both algorithms develop a model that describes the mapping between parameter set and observed cost. The generated model will then be used to find the next best parameters to test [73].

The first ten parameter sets in every optimization process are set by a training algorithm which can be either the Differential Evolution algorithm [84], the Nelder-Mead algorithm [85], or by choosing parameters randomly [73]. These initial training data are then used to train the machine learner. The Differential Evolution algorithm and the Nelder-Mead algorithm are classical optimization algorithms. Throughout this thesis the default Differential Evolution algorithm was used for training. Furthermore, the decision about which parameter set to test next is sometimes made by the training algorithm rather than the machine learning algorithm. This is, for example, the case when the experiment cycle is ready to try a new parameter set, but the learning process of the machine learner is not yet finished. To not waist any time and further generate training data in between, the experiment is then performed with parameter values chosen by the DE algorithm [73]. Using the DE algorithm alongside the machine learning optimization, also prevents the machine learner to get trapped in a local minimum.

The following section introduces the Differential Evolution algorithm and both machine learner algorithms.

3.3.0.1 Differential Evolution (DE)

Differential Evolution (DE) is a direct search algorithm that search for optimal parameters, that minimize the objective function $f_C(X)$, by iteratively updating a population of parameter sets [84]. The parameter population is defined as a parameter vector $\mathcal{X}_G = (X_{G,1}, X_{G,2}, \ldots, X_{G,N_\mathrm{P}})$ consisting of N_P D-dimensional parameter sets $X_{G,i}$. The dimension D of the parameter sets is defined by the number of optimization parameters. The first generation G of parameter sets is chosen randomly. N_P new parameter sets are generated by adding the difference between two randomly chosen parameters sets to a third parameter set of the current generation [84]:

$$V_l = X_{G,i} + F \cdot (X_{G,j} - X_{G,k}),\ l \in [1, N_\mathrm{P}] \tag{3.1}$$

where $i, j, k \in [1, N_\mathrm{P}]$ are random indices and $F \in [0, 2]$ is some constant weighting factor. The entries of a new parameter vector $X_{G+1,l}$ are defined by randomly choosing entries from $X_{G,l}$ and V_l to fill up the new vector. This is called crossover and is done to ensure higher variability in the new vector. All new parameter sets are then tested in the experiment. They are accepted if they produce a lower cost value than the corresponding parameter set of the previous generation, meaning $f_C(X_{G+1,l}) < f_C(X_{G,l})$ is fulfilled. If this is not the case the vector entry of the previous generation is used for the new parameter vector $\mathcal{X}_{G+1}$. This is repeated until some halting conditions is met, and the parameter set resulting in the lowest cost is returned.

Even though differential evolution is the fastest evolutionary algorithm [84], it still converges much slower than Gaussian processes or a neural net. Nevertheless, the machine learner can profit from new parameter sets that are chosen by the DE algorithm. Since the DE parameter sets are independent of the internal model built by the machine learner, they can help the machine learning algorithm to jump out of a local minimum.

3.3.0.2 Gaussian Process (GP)

One of the algorithms used to model the objective function is a Gaussian process. For this method, the continuous updating of the model as the dataset grows is based on Bayesian optimization. This describes a supervised learning method [86]. Bayesian optimization uses a probabilistic approach for modelling the objective function $f_C : \mathcal{M}_X \to \mathbb{R}$ over an infinite domain $\mathcal{M}_X$. This assumes that the objective function is randomly distributed and can thus be described as a stochastic process [87].

Gaussian processes are used to model these stochastic processes. They are the extensions of multivariate functions into the infinite domain [87]. This means that the function values at every point $X \in \mathcal{M}_X$ in the infinite domain of the function are Gaussian distributed. The mean vector in the multivariate Gaussian is replaced with a mean function $\mu : \mathcal{M}_X \to \mathbb{R}$, describing the expectation value of the function value at every point X. The covariance matrix will be replaced as well with a covariance function $K : \mathcal{M}_X \times \mathcal{M}_X \to \mathbb{R}$. The covariance function $K(X, X', H)$ describes the correlation between function values at X and X'. H is a set of D hyperparameters, for each optimization parameter in the parameter set X. The hyperparameters have to be fitted in ever optimization step by finding the maximum of the likelihood function for the given observations [21]. A Gaussian process $\mathcal{GP}$ of the objective function f_C is then defined as [87]

$$p(f_C) = \mathcal{GP}(f_C; \mu, K) \tag{3.2}$$

and describes the probability distribution of the objective function.

The solution to the regression task is found by starting with a prior Gaussian process $p(f_C)$. Since the dataset is empty at the beginning, this distribution is not based on the actual process and just serves as an arbitrary smooth starting distribution. The hyperparameters are set to some arbitrary initial values. The prior distribution describes the belief about the model before seeing data. The idea behind Bayesian optimization is to find a posterior distribution that updates the prior distribution based on observed data. This can be achieved by determining the likelihood function. The likelihood function describes the probability of observing the data for a given model. Multiplying the prior by the likelihood updates the prior distribution based on how well each possible model of the prior distribution can describe the observed data. This forms the posterior distribution. The posterior process will then serve as the prior process in the next step, when a new observation was made. It will then be updated again based on the new observations. Thus, Bayesian optimization is an inductive process, in which the posterior distribution becomes more precise with every newly observed datapoint, and is thus suitable for online machine learning [87].

The generation of the posterior process can be described more formally as follows. The starting point is the prior Gaussian process $p(f_C)$. The probability distribution of the function values $\Phi = f_C(X)$ is given as $p(\Phi|X)$. In the next step some observations $\mathcal{D} = \{X, C\}$[1] are made. To update the prior

[1] It is assumed that the observation has no noise here, to make the short introduction more clear. However, when applying the machine learning algorithm the noise of the observations are taken into account. For more information on Bayesian optimization with noisy observations see [87]

distribution with the new dataset, a posterior process $p(f_C|\mathcal{D})$ has to be evaluated. This is done by first applying Bayes' theorem to get the posterior distribution of the function values [86, 87]

$$p(\Phi|\mathcal{D}) = p(\Phi|(\mathcal{X}, C)) = \frac{p(C|\mathcal{X}, \Phi) \cdot p(\Phi|\mathcal{X})}{p(C|\mathcal{X})} \hat{=} \frac{\text{likelihood} \cdot \text{prior}}{\text{marginal likelihood}}, \tag{3.3}$$

where $p(C|\mathcal{X}, \Phi)$ is the likelihood function. The distribution $p(f_C|\mathcal{D})$ at an arbitrary location X, the so-called posterior predictive process, can be obtained by averaging over all possible function values Φ weighted by their posterior probability [86]. The posterior process is again a Gaussian process and describes the objective function distribution conditioned on the observed data. As this is only intended as a short introduction to Bayesian optimization, no further mathematical details are given here. For more information see [86] and [87].

The generation of the posterior process describes how an online model, based on Gaussian processes, can be built. However, to be able to efficiently find the minimum of the modelled cost function, the algorithm has to make smart choices about which parameter set X should be used for the next observation. In Bayesian optimization this is achieved by defining an acquisition function that returns for each potential observation point X a scale for their contribution to the optimization process [87]. In the M-LOOP package the score depends on two conditions. The first one is based on the possibility of some parameter set X to minimize the mean function $\mu_C(X|\mathcal{D}, H)$ of the posterior Gaussian process. This approach tries to directly find the minimum of the objective function. However, since the model is only built upon a small set of data, it is necessary to further improve it by making observations at positions where the model is still rather uncertain about its prediction. This is done by choosing parameter sets that maximize the variance $\sigma_C(X|\mathcal{D}, H)$ of the posterior Gaussian process. The acquisition function is designed such that by finding the minimum of the function both conditions are taken into account[2]. The influence of both conditions can be adapted by changing their weighting in the function. In the implementation of the M-LOOP package the weighting is continuously scanned, to realize a more robust decision-making process [21].

3.3.0.3 Neural Net (NN)

Neural networks (NN) are also a common method used in machine learning optimization [34, 88]. A neural network is composed of several layers that consist of multiple nodes, called artificial neurons. An illustration of an example neural net can be found in Fig. 3.2. All layers are connected via links between their nodes. The number of input data defines the number of neurons in the first layer. The subsequent layers are called hidden layers and are used to map the input data to activations of the nodes in the last layer, the output layer. Each node in the hidden layer is connected to the nodes of the previous layer. The activation $a_{i,l}$ of the neuron in layer l is generated via so-called activation functions $a_{i,l} = \sigma(\sum_j^{N_n} w_{ij} a_{j,l-1} + b_{i,l})$ that depend on the activation $a_{j,l-1}$ of all nodes in the previous layer weighted with the parameters w_j, and an additional parameter $b_{i,l}$ called bias. N_n describes the number of nodes in the previous layer. The weights $w_{i,j}$ control the contribution of the previous nodes to the activation of the neuron in question [80]. There are different options for the activation function which influence the behavior of the neural network. The activations of the neurons in the output layer describe the output values. In this way it is possible to map the input data to respective outputs via the neural net

[2] The acquisition function is defined as $A(X) = b\mu_C(X|\mathcal{D}, H) - (1-b)\sigma_C(X|\mathcal{D}, H)$, where b defines the weighting between both conditions [21].

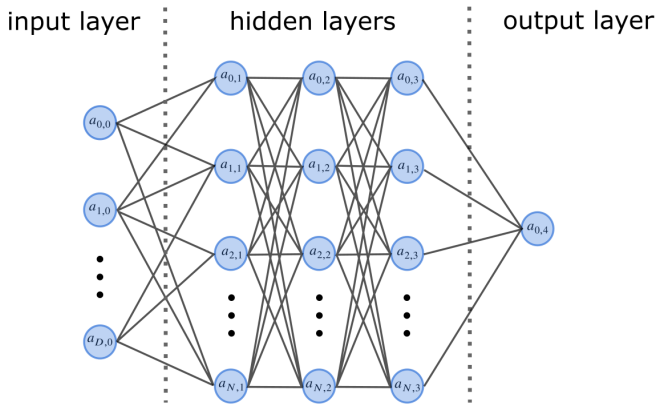


Figure 3.2: Illustration of an example neural network with three hidden layers. $a_{i,l}$ describes the activation of the i-th node in layer l. The input layer has D nodes, corresponding to D optimization parameters. Each hidden layer has N nodes. The output layer has one node, describing the output. For the implementation of the optimization algorithm the activation of the last node would correspond to the cost. In general, more output nodes are possible in a neural net. Image adapted from [89].

and thus build a model of the objective function f_C [80].

The learning process is based on adjusting the weights and biases for each neuron. This is done by calculating the loss function, which corresponds to the deviation of the expected output to the measured activation of the output nodes. Next, the gradient of the loss function with respect to the weights and biases is determined. The weights and biases are adjusted in the opposite direction of the gradient to minimize the loss. This is called gradient descent method. The process of adjusting the weights and biases with respect to the loss function gradient describes the learning process of the neural net [80].

There are different algorithms that can be applied to control the training of the neural net. In the M-LOOP implementation training of the network is achieved by applying the Adam optimizer [21]. This is an adaptive learning rate optimization algorithm that combines the stochastic gradient descent method with additional momentum based methods [81]. The training is repeated whenever a new datapoint is generated. The model is thus continuously updated and becomes more precise as the dataset grows. In the M-LOOP package three independent artificial neural networks (ANNs) are initialized and make up a stochastic artificial neural network (SANN). Each single neural net is built up of 5 hidden layers with 64 neurons for each layer. The activation function is given by the Gaussian error linear unit [34]. After training each ANN independently, the next step is to find the next best parameters for testing. This is done by applying the L-BFGS-B algorithm to find parameters that minimize the current model of the neural net [34]. L-BFGS-B uses the Quasi-Newton method. This is a second-order gradient method [81]. For more information on the optimization algorithms and deep learning in general, see e.g. [81]. To be more robust against getting trapped in a local minimum each of the three neural nets are trained independently, and thus each network returns a new parameter set for testing. Additionally, the training algorithm, meaning the Differential Evolution algorithm, provides a forth parameter set. All parameter sets are tested in the next step in the experiment, to efficiently explore the full parameter space. These observations are then used to train all three networks to continuously improve the model [34].

The model built by ANNs may not be as reliable as the ones built by a Gaussian Process. However, the advantage of using deep neural networks over using Gaussian processes is that the neural net learning

process is computationally faster for larger parameter spaces. The computational time for the fitting procedure with a Gaussian process grows with the cube of the number of training data, whereas for a neural network it scales linearly [34, 73]. For a larger number of parameters a larger training dataset is needed to capture all relevant information. Therefore, the use of artificial neural networks is beneficial when optimizing a larger parameter space [34].

CHAPTER 4

Machine Learning Online Optimization of the Experiment

The experimental sequence in the HQO experiment is fully computer controlled as explained in Section 2.4. This enables the implementation of the machine learning online optimization package M-LOOP [21] into the experiment control system for the optimization of the experimental sequence. The implementation of the machine learner cycle into the experiment cycle is explained in more detail in Section 4.1. Following its successful implementation, the MOT sequence was optimized and the performance of different optimization algorithms was compared. This is discussed in Section 4.2. Section 4.3 describes the machine learning optimization of the magnetic transport sequence. It discusses how different cost functions and parametrizations influence the outcome of the machine learning optimization.

4.1 Implementation into the HQO Computer Control

The machine learning cycle, explained in Section 3.2, and the experimental cycle, explained in Section 2.4, must be combined so that they can work in parallel, enabling the machine learning optimization of the experimental sequence. Apart from the fact that the parameters have to be computer controlled for the implementation, it must also be possible to update them during the optimization process. For a normal experiment run all parameters are set manually in advance in the experiment control interface. Possible iterators and their respective scan range are also defined beforehand. To enable continuous updating of the experimental parameters, an interfacing layer was introduced to which both the experimental and machine learning cycles have access. The interplay of machine learner cycle, experiment cycle and the connecting interfacing layer is illustrated in Fig. 4.1.

The interface layer fulfills two tasks. Firstly, it provides a channel for exchanging the new parameters X_{ML} found by the machine learner with the computer control, in order to test them in the next experimental cycle. Secondly, the results of the experiment run are fed back to the machine learner in the form of a cost value via the interfacing layer.

The exchange of parameter sets is realized by saving the new parametrization in files which are accessible for both machine learner and computer control. For the MOT optimization the optimization parameters X_{ML} correspond to parameters defined in the experiment sequence and can thus be directly saved to the files. For the magnetic transport the new parameter set X_{ML} must first be converted to current traces and respective TTL pulses for all coil pairs. This is done with the magnetic transport

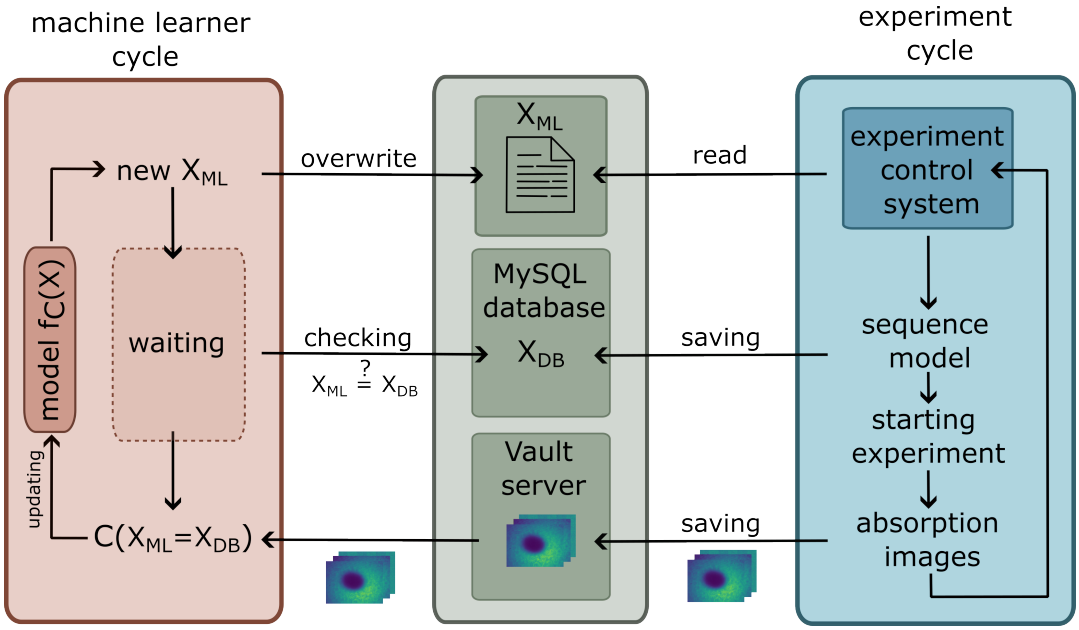


Figure 4.1: Schematic description for the implementation of the machine learning cycle into the experiment cycle. The green box combines both cycles by introducing an interfacing layer to which both cycle have access. The interfacing layer is composed of three units. The first one describes the parameter files with the new parameter values X_{ML} chosen by the machine learner. The experiment control system reloads them in each cycle. The MySQL database is used for saving the experimental sequence model. Furthermore, the machine learner accesses the database to check if the parameters of the last sequence model X_{DB} correspond to the last found parameter X_{ML} of the machine learner. The last unit is the vault server, to which the absorption images are saved. The absorption images are then loaded by the machine learner controller to extract a cost value C if the datasets match and thus $X_{\mathrm{ML}} = X_{\mathrm{DB}}$.

simulation, as discussed in Section 2.2. In total four files with the current traces for the power supplies 1 to 4 and additional seven files for the coil switching TTL signals are created for a new parameter set. These are the files that the computer control can load to update the sequence model. The chosen parametrization for both MOT optimization and magnetic transport optimization are discussed further in Section 4.2 and Section 4.3.

Once the machine learning controller has overwritten all the parameter files, it enters a waiting state. The experiment control system loads the files at the beginning of every cycle and updates the sequence model. Besides updating the sequence model, the computer control system also saves the model to the MySQL database alongside a corresponding global counter. A new measurement cycle then starts with the updated sequence model. Each experimental sequence during the optimization ends with taking absorption images either in the MOT chamber or in the science chamber. The absorption images can characterize the performance of the experimental run and are thus used for extracting a cost value $C(X)$. The definition of the respective cost function for the MOT optimization and the magnetic transport optimization is given in Section 4.2 and Section 4.3. After the absorption images are saved on the vault server, under the same global counter as the sequence model, the experimental cycle repeats itself. The database and the vault server can be accessed as well from the machine learner controller. While the machine learner cycle is in its waiting state, it continuously checks whether the value of the parameters

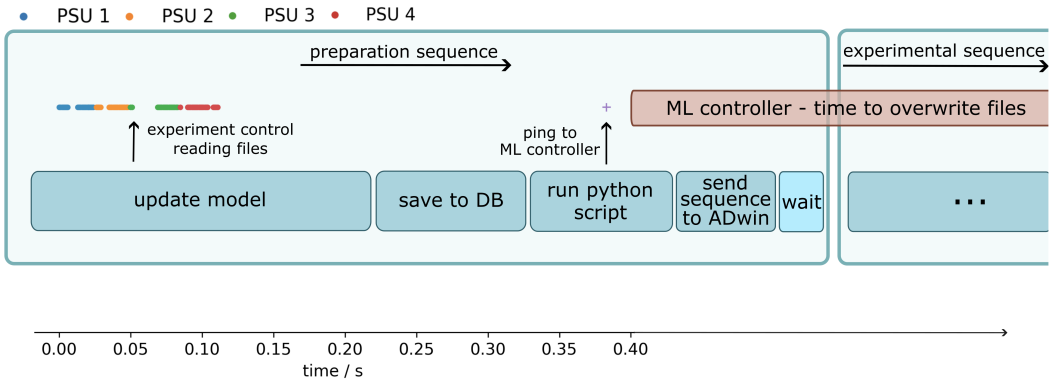
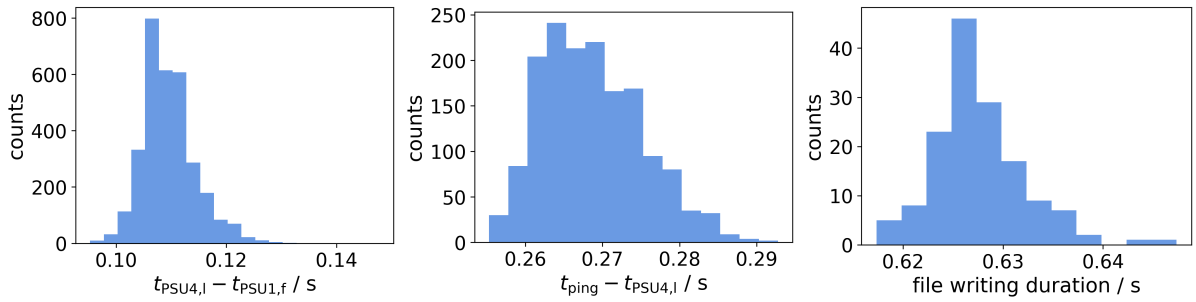


Figure 4.2: Timing between different operations during one magnetic transport optimization cycle. The blue boxes describe the experiment cycle, the red box operation of the machine learner controller. The colored points show times at which the experiment controller accessed the power supply files. PSU 1-4 stands for the files containing the current traces for the power supplies 1-4. The purple cross shows the point when the experiment control sends the ping to the machine learner controller. The time axes only corresponds to the colored datapoints that were measured during one specific experimental cycle.

X_{DB} in the latest sequence model that was saved to the database matches with the latest testing parameter set X_{ML} that was generated by the machine learner. In this way the machine learner can check if the new parameter set X_{ML} was already tested in the experiment. If this is the case it loads the absorption images and extracts the cost $C(X_{\mathrm{ML}} = X_{\mathrm{DB}})$ for the current parameter set. In this way the training dataset grows by one. The machine learner than updates the underlying cost model based on the new datapoint, chooses new parameters for testing, and the cycle restarts. The interfacing layer can therefore be described as a kind of buffer that can be accessed and updated by the machine learning controller and the experiment controller as required.

However, since both cycles operate independently of each other in principle, they must be synchronized to prevent them accessing the different units of the interfacing layer at the same time. This is realized by sending a ping from the experiment cycle to the machine learner cycle. The ping is triggered after the experiment control accessed the parameter files and build up the model for the next sequence. The timing of the ping with respect to the experiment cycle is illustrated in Fig. 4.2. The blue boxes describe the experiment cycle and the red box indicates operations of the machine learner cycle. The purple cross indicates the timing of the ping to the machine learner controller. After receiving the ping the machine learner controller is allowed to overwrite the files with the updated values for the next optimization run. It then goes into its waiting state where it continuously checks the database. This can happen in parallel to the experiment controller saving new sequences to the database, since the experiment control always creates a new entry in the database instead of overwriting an old one. The same holds for loading the absorption images from the Vault server after the condition $X_{\mathrm{ML}} = X_{\mathrm{DB}}$ is fulfilled. In this way both cycles can in principle work in parallel without disturbing the other, if the process of writing and reading the parameter files works fast enough.

For the MOT optimization this is the case since the parameter files only consist of single values, as explained in Section 4.2. For the magnetic transport optimization however this still led to permission errors when overwriting or reading the files containing the current traces for the four power supplies. The files storing the current traces have a larger size compared to the MOT optimization parameter files,



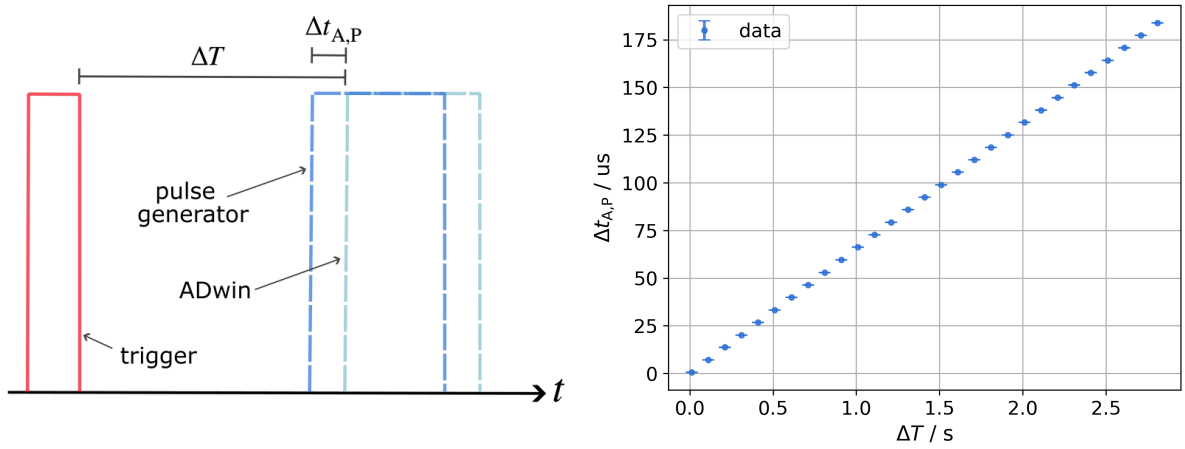
(a) (b) (c)

Figure 4.3: Analysis of the access times for the files storing the current traces for PSU 1-4. **(a)** Distribution of the time period during which the experiment control access one of the PSU files. The file access happens during the preparation sequence in the experiment cycle. $t_{PSU1,f}$ is the time of the first access of the PSU 1 file, $t_{PSU4,l}$ is the time of the last access of the PSU 4 file. **(b)** Distribution of the time interval between the point when the experiment control is done loading the PSU files and the time the machine learner controller receives the ping t_{ping}. The end point of the loading process from the experiment control is marked by the last access time to the PSU 4 file $t_{PSU4,l}$. **(c)** Distribution of the time it takes the machine learner controller to overwrite all four PSU files during the optimization problem.

which lead to enhanced writing and reading times. The timing of the different file access points was therefore analyzed in more detail for the magnetic transport optimization.

As discussed before, Fig. 4.2 shows the timing between different operations during one optimization cycle. Furthermore, the time points when the experiment control accessed the four power supply files (PSU 1-4) during one experimental cycle are indicated by the colored points. The access times were extracted by querying the time of the latest file access in a separate thread during the sequence. The time on the x-axis is defined relatively to the first access time of the PSU 1 file. The time axes only specifically match the colored datapoints. The length of the boxes beneath them is chosen arbitrarily since it was not measured how long the different operations in the preparation sequence take. The full preparation sequence however has a fixed duration of 1 s. The measurement show that the files are accessed several times during the time interval in which the internal computer control model is updated. The file for PSU 1 is always accessed first and the file for PSU 4 is accessed last.

The measurement of the access points during one experimental cycle was repeated several times. The time interval between the first access of the PSU 1 file $t_{PSU1,f}$ and last access of the PSU 4 file $t_{PSU4,l}$ describe the total duration during which the experiment control access one of the power supply files. The distribution of these time intervals $t_{PSU4,l} - t_{PSU1,f}$ can be seen in Fig. 4.3(a). The mean time differences between first and last access is (0.10965 ± 0.00009) s. Furthermore, the time between the point when the experiment control accesses the last time the PSU 4 file $t_{PSU4,l}$ and the time the machine learner controller receives the ping was measured for several runs (see Fig. 4.3(b)). The minimal time difference that was measured is 0.255 s. This shows that the ping was always received after the computer control finished reading all files. In a separate measurement the time required by the machine learner controller to overwrite the four files was determined (see Fig. 4.3(c)), with a maximal measured write duration of 0.645 s. In principle, this should enable the machine learner to overwrite the files while the experiment is running for ~ 2 s and allow the experiment control to read them during the preparation time of 1 s, without interference between the two processes. However, permission errors still occurred from time to time, making it difficult to reach the desired number of optimization runs. Since the measurements in



(a) (b)

Figure 4.4: **(a)** Schematic description for the analysis of the clock delay between pulse generator and ADwin. The red trigger signal is generated by an ADwin channel. It is used to trigger the pulse generator. Both ADwin and pulse generator are programmed to generate a TTL signal after a variable time of Δ_T from the trigger signal. The TTL signals are indicated by the blue dotted boxes. The time delay $\Delta_{A,P}$ between the pulse generator TTL signal and the ADwin TTL signal is measured. **(b)** Results for the measurement of $\Delta_{A,P}$ with Δ_T ranging from 0 s to 2.8 s.

Fig. 4.3 only capture the latest access time, but not the time the file is open, it could be that one of the controllers has the file open for a longer time. This could explain the observed access time problems. It was decided to not further analysis this and rather get around the problem by implementing an additional try except statement that prevents the program to shut down when running into a permission error. Due to the machine learner control and experiment controller not being perfectly timed, it can happen that an experiment run goes to waist. However, the dominating factor in terms of time delay between two optimization runs is the algorithm requiring a certain amount of time to build the new model. Thus, this additional delay does not have a major impact.

An additional modification of the experimental sequence implementation was made for the purpose of optimizing the magnetic transport. Instead of using the digital output of the ADwin for generating the magnetic coil switching TTLs (explained in Section 2.2), the pulse generator (which also generates the Rydberg excitation and detection pulses) is used. Even though the ADwin will be used for everyday experiment cycles, it has been found that the pulse generator is more suitable for optimization. The reason for this is that the digital ADwin sequence is programmed into the computer control interface before starting the cycle. For each channel an individual number of steps with a variable duration and corresponding high or low TTL can be set in the UI. The duration of the steps can be dynamically adapted by reading and writing corresponding csv files, as it is done in the MOT optimization cycle. However, the number of steps and the decision if the step corresponds to a high or low TTL is fixed when starting the sequence. With this implementation it was only possible to switch the current output between the coils for a fixed number of times, giving the machine learner less freedom for optimization. Therefore, it was decided to use the pulse generator, for which it is possible to dynamically adapt the number and duration of the TTL switching signals. The start of the pulse generator sequence is triggered by the computer control, but the output sequence itself is programmed by a Python script.

A potential issue with this implementation is that the coil switching TTLs and the power supply outputs are no longer controlled by the same device. Even though the pulse generator is triggered by the ADwin, this will lead to time mismatches as the internal clocks of the ADwin and the pulse generator are not synchronized. The time delay between ADwin and pulse generator was measured with the test sequence, schematically shown in Fig. 4.4(a). One output channel of the ADwin and one output channel of the pulse generator are programmed to generate a TTL signal after a variable delay time ΔT relative to a trigger signal. The trigger signal is programmed into the computer control sequence and generated by a different ADwin channel. ΔT is the time between the ADwin trigger signal and the ADwin TTL in question. The trigger is additionally used to start the output sequence of the pulse generator. The time delay ΔT is scanned in the computer control from 0 s to 2 s and the time difference $\Delta t_{\mathrm{A,P}}$ between ADwin signal and pulse generator signal are measured with an oscilloscope. The measured datapoints are plotted in Fig. 4.4(b). The time difference between the signals increases linearly with the delay time ΔT. This shows that the internal clock of the pulse generator is running faster than the one of the ADwin. The maximal duration of the magnetic transport T_{MT} that will be tried out for the optimization process is 2 s. This corresponds to a maximal time delay of 130 µs for the pulse generator TTLs during the sequence. Since the experimental sequence for the magnetic transport is programmed in a way that the TTL signals trigger the coil switching box 10 ms before the output sequence of the power supply starts, a time delay of up to 130 µs from the pulse generator TTLs will not affect the experimental results.

Furthermore, it was observed that the pulse generator sets all channel outputs to high during its reprogramming process, which takes place in the preparation phase of the experiment control cycle. This means that all coil switching TTLs are set to high. If both switching TTLs for one power supply are set to high, both connected coils are supplied with current. Since the MOT loading already begins within the preparation time, it is important to make sure that the current supplied by PSU 1 is only applied to the MOT coil during this time. Otherwise, the magnetic field gradient generated by the MOT coils would be lower than required. Implementing an AND gate controlled by one ADwin TTL could ensure that the TTL for coil 4, which is supplied by the same PSU as the MOT coil, is set to low during the preparation time (see Fig. 2.11 for reference of the coils). Building on this, the pulse generator is suitable for generating the coil switching TTL signals during the optimization routines. At the same time, the results generated during the optimization process using the pulse generator are easily transferable to an experimental sequence with the ADwin TTLs for everyday measurements.

4.2 MOT Optimization

The magneto-optical trap is the first trapping and cooling stage in the experiment. The number of atoms being trapped and the temperature of the atom cloud are influenced by various parameters. These parameters can be divided into two categories, the computer controlled parameters and the ones that can not be controlled by the computer control system. The latter comprises the correct alignment of the Cooler and Repumper laser beams and the polarization of the laser light. These optimizations have to be performed manually and can not be part of any machine learning online optimization routine. This was done when initially setting up the magneto-optical trap (see [40, 43]). The second category of computer controlled parameters includes the magnetic field gradient, the detuning of Cooler and Repumper laser and their powers. The number of computer controlled parameters influencing the MOT is manageable, which makes manual optimization feasible. Manual optimization refers to scanning two or three parameters iteratively during the experimental sequence. This kind of optimization was already done by Julia Gamper [40]. However, this optimization was repeated here using the implemented machine learning algorithm to demonstrate the principle by first applying the learner to an easy application. Furthermore, the performance of different optimization algorithms was tested. The implementation into the setup (see Section 4.2.1) and the results of the optimization (see Section 4.2.2) will be discussed in the following.

4.2.1 Parametrization and Cost Function

In order to use the implemented optimization cycle, summarized in Fig. 4.1, for the MOT optimization, one has to decide on suitable optimization parameters X and a cost function $C(X)$. The parametrization is discussed in Section 4.2.1.1 and the chosen cost function is described Section 4.2.1.2.

4.2.1.1 Parametrization

As stated above the optimization parameters of interest are the frequency and the power of the Cooler and Repumper laser and the magnetic field gradient. The dependence of the cooling and trapping force on the Cooler laser detuning and power, and on the magnetic field gradient, is evident in the definition of F_{MOT} in Eq. (2.3). Furthermore, the efficiency with which the atoms are pumped back into the cooling cycle depends on the power and detuning of the Repumper beam [45]. Since the applied magnetic field causes an energy splitting of the hyperfine states the detuning of the lasers and the magnetic field gradient are interdependent.

All parameters to be optimized can be changed by the experiment control system, which makes them feasible for online machine learning optimization. The magnetic field gradient is controlled by the current supplied to the coils. The output current of the magnetic coil power supplies can be programmed by applying an external voltage [90]. The control voltage is supplied by the analog outputs of the ADwin. The ADwin output is programmed by defining the desired magnetic field gradient $B'_{\mathrm{MOT\,coil}}$ (along the strong z-axis) in the Computer control system and converting it to the corresponding voltage signal. This makes it possible to adjust the magnetic field gradient in the MOT with the computer control system.

As described in Section 2.1.1 the Repumper and Cooler laser are locked relative to a Master laser via an Offset-Lock. The reference frequency for adjusting the lock point is provided by either voltage controlled oscillators (VCO) (ZX95-100-S+ from Mini-Circuits [91]) (for the Cooler laser) or by direct digital synthesizers (DDS) (AD9959 4 channel 500 MSPS DDS with 10-bit DACs from ANALOG

DEVICES [92]) (for the Repumper laser). The output frequency of the VCOs can be directly controlled by the analog control voltages from the ADwin channels. The desired frequency is set in the experiment control system and converted to the corresponding VCO input voltage. An additional Arduino provides a computer interface for the DDS board. The Arduino is programmed with a Python script in every cycle. As explained in Section 2.4, the execution of the Python script is controlled by a ping from the experiment control system. This script reads in the required variables, which are set in the computer control interface. Both frequency variables are defined as the detuning from the respective resonance transition.

The laser power is controlled by an acousto optical modulator (AOM), which is placed in front of the fiber that guides the light from the laser table to the experiment table. Only the first order, generated by the AOM, is coupled into the fiber. Turning the RF signal for the AOM on and off using a TTL signal turns the light on the experiment table on and off. The AOMs are therefore used as optical switches. The power in the first order, and thus the power of the light coupled out of the fiber, can be adjusted with the amplitude of the RF signal. The AOM RF power is controlled by the analog output voltages of the ADwin. No power calibration converting control voltage into laser power is implemented in the computer control system for the AOM. This is because the actual power at the MOT chamber fluctuates due to polarization and temperature fluctuations that affect the fiber coupling.

In conclusion, the laser frequency, the laser power and the magnetic field gradient, can be changed automatically via the computer control interface. The parametrization for the MOT optimization is thus straightforward to select, since it simply involves finding optimal constant values for the magnetic field gradient along the strong axis B'_{MOT}, the Cooler and Repumper detuning from resonance (Δ_{C}, Δ_{R}) and the control voltage for the AOM RF power for both lasers ($V_{\mathrm{AOM,C}}$, $V_{\mathrm{AOM,R}}$).

4.2.1.2 Cost function

In order to enable online optimization, it is necessary to define a cost function that quantifies the performance of the measurement. There are different characteristic properties of an atom cloud trapped in a MOT, for example, the temperature of the cloud, the atom number or the phase space density. It was decided to optimize with respect to the number of trapped atoms N. The atom number can be measured by performing absorption imaging after a MOT loading phase of 1 s and 10 ms of time of flight (TOF). The measurement is repeated three time for the same parameter set X. The cost function is defined as the scaled mean of the atom numbers over three measurements with the respective parameter set X

$$C(X) = -\overline{N(X)} \cdot 10^{-8}. \tag{4.1}$$

The factor -1 in the definition is necessary, since the machine learner aims to minimize the cost function, whereas the atom number should be maximized. The scaling with a factor of 10^{-8} is chosen so that the resulting cost lies in a range between 0 and -10. The uncertainty $U(X)$ of the cost is given by the standard error of the mean.

4.2.2 Optimization of MOT Parameters

The results of the machine learning optimization for the magneto-optical trap are discussed in this section. The optimization process was repeated three times using different optimizer: the two machine learning optimizers (Gaussian Process (GP) and Neural Net (NN)), and the classical search algorithm (Differential Evolution (DE)). All three optimization processes were started with the same first parameter set X_0 for the

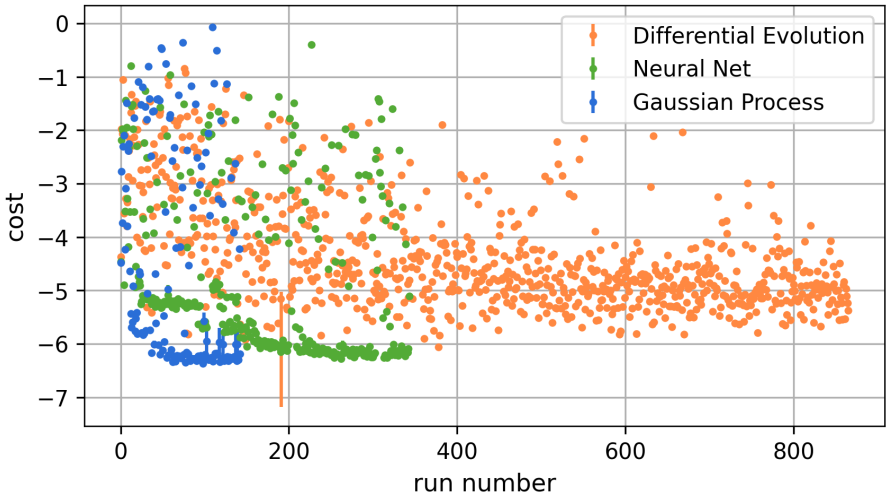
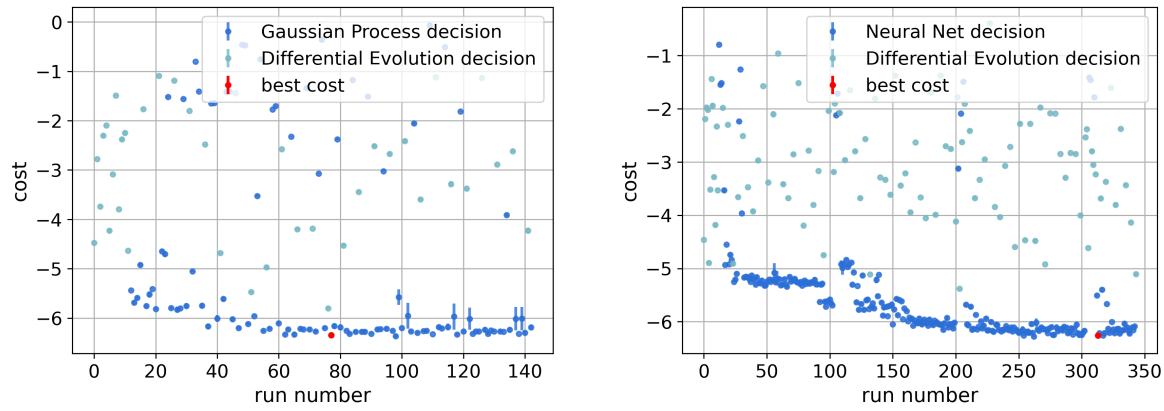


Figure 4.5: Cost curve for the machine learning based MOT optimization with different optimization algorithms. Each run number corresponds to a new parameter set that was tested. The performance of the optimization process was tested with the Gaussian Process, the Neural Net and the Differential Evolution algorithm. The large uncertainty for one run during the Differential Evolution optimization is due to a timing problem that occurred during one experimental cycle. The stopping condition was defined by a maximal duration of the total optimization routine. However, all optimizations were stopped manually beforehand, once a sufficient number of runs had been conducted.

purpose of comparison. The permitted parameter ranges are based on values that are known to produce a MOT. For the AOM RF power control voltage the boundaries are set to $[4\,\mathrm{V}, 10\,\mathrm{V}]$, whereas 10 V is the highest possible value. For the Cooler this corresponds to a power range of $[42\,\mathrm{mW}, 107\,\mathrm{mW}]$ and for the Repumper to a power range of $[1.2\,\mathrm{mW}, 3.0\,\mathrm{mW}]$ in each of the six MOT beams. The magnetic field gradient range is set to $[5\,\mathrm{G/cm}, 20\,\mathrm{G/cm}]$, the Repumper detuning range is set to $[-25\,\mathrm{MHz}, 25\,\mathrm{MHz}]$ and the Cooler detuning range is set to $[15\,\mathrm{MHz}, 35\,\mathrm{MHz}]$. The Cooler detuning describes a frequency shift to lower frequencies with respect to the resonant Cooler transition. For the Repumper detuning the positive values correspond to blue detuning and the negative values to red detuning from the resonant Repumper transition.

The results of the three optimization processes are shown in Fig. 4.5 and summarized in Table 4.1. Fig. 4.5 displays the measured cost values with respect to the run number. Every run number corresponds to a new parameter set that was tested. The optimization with the Neural Net converges in 250 runs to a cost value around a −6.3. The Gaussian Process is faster than this and converges after around 80 runs to a similar cost value. The DE algorithm on the other hand is clearly slower than the machine learning algorithms. Even after 800 runs it seems to stagnate around a cost value of -5 ± 1, and did not converge. An optimization run consisting of more than 800 runs could probably still lead to some improvements. However, this result already shows that both machine learning algorithms perform better than the classical search algorithm. Therefore, the DE algorithm is only used to generate the first training datasets for the machine learning algorithms.

Furthermore, for each optimization controlled by a machine learning algorithm, the DE algorithm is sometimes used in between to generate an additional parameter set for testing. This is important for the machine learning algorithms to not get lost in a local minimum, as discussed in Section 3.3. The



(a) (b)

Figure 4.6: Cost curve for the machine learning based MOT optimization with **(a)** the Gaussian Process as machine learner controller and **(b)** the Neural Net as machine learner controller. In each machine learning based optimization the Differential Evolution algorithm is used in between to generate some new parameter sets for testing. The decisions made by the machine learning algorithm are marked in dark blue and the decisions made by the Differential Evolution algorithm are marked in light blue. The red datapoint shows the run that resulted in the best cost value. The larger uncertainties for some of the runs are due to some unknown disturbance in the experiment that sometimes lead to a failed experimental cycle.

optimization runs shown in Fig. 4.5 with the GP and the NN were analyzed more closely in relation to this in Fig. 4.6(a) and Fig. 4.6(b). The plots show when the decision about the next parameter set was made by the machine learner and when by the DE algorithm. The first 10 parameter sets are always determined by the DE algorithm, enabling the machine learner to build its initial model. Furthermore, both plots show that the DE algorithm sometimes decides on the new parameter sets in between. Runs with a parameter set from the DE algorithm mainly resulted in higher cost values.

Table 4.1 summarizes the results of the three optimization runs. It lists the best found parameters during the different optimization runs and the cost measured with these parameters. The best run for the DE optimization resulted in a cost of -6 ± 1. This corresponds to the run with the large error bar in Fig. 4.5. The reason for the large standard error of the mean is that one out of the three measurements per parameter set yielded an atom number that was 1.6 times greater than the other two. One possible reason for this could be that the MOT loading time was longer than usual due to some timing problem of

Optimizer	$V_{\mathrm{AOM,C}} \hat{=} P_{\mathrm{C}}$	$V_{\mathrm{AOM,R}} \hat{=} P_{\mathrm{R}}$	B'_{MOT}	Δ_{C}	Δ_{R}	Cost
GP	$10\,\mathrm{V} \hat{=} (107 \pm 1)\,\mathrm{mW}$	$10\,\mathrm{V} \hat{=} (3.02 \pm 0.01)\,\mathrm{mW}$	$10.9\,\mathrm{Gcm}^{-1}$	29.4 MHz	12.8 MHz	-6.37 ± 0.04
NN	$10\,\mathrm{V} \hat{=} (107 \pm 1)\,\mathrm{mW}$	$9.4\,\mathrm{V} \hat{=} (2.84 \pm 0.01)\,\mathrm{mW}$	$10.9\,\mathrm{Gcm}^{-1}$	29.7 MHz	14.9 MHz	-6.280 ± 0.018
DE	$9.5\,\mathrm{V} \hat{=} (101 \pm 1)\,\mathrm{mW}$	$10\,\mathrm{V} \hat{=} (3.02 \pm 0.01)\,\mathrm{mW}$	$11.3\,\mathrm{Gcm}^{-1}$	28.6 MHz	10.9 MHz	-6.055 ± 0.020

Table 4.1: Best found MOT parameters for the optimization process with the Gaussian Process (GP), the Neural Net (NN), and the Differential Evolution (DE) algorithm. $V_{\mathrm{AOM,C/R}}$ is the control voltage for the RF amplitude of the Cooler and Repumper AOM. The respective laser power $P_{\mathrm{C/R}}$ for Cooler and Repumper was measured for one MOT beam. B'_{MOT} describes the magnetic field gradient for the MOT magnetic field and $\Delta_{C,R}$ is the Cooler/Repumper detuning.

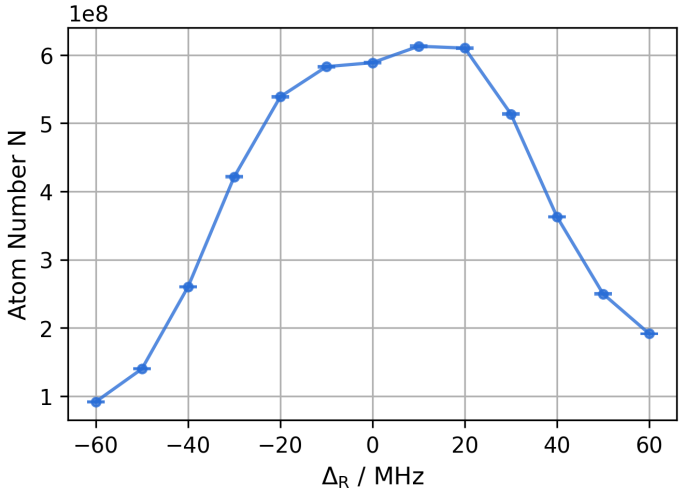


Figure 4.7: Atom number N measurement for different Repumper detunings Δ_R. Positive Δ_R values correspond to blue detuning and the negative values to red detuning from the resonant Repumper transition. For the measurement the MOT was loaded for 1 s. The datapoints are averaged over 17 measurement.

the computer control sequence. Therefore, the result does not seem reliable. In Table 4.1 the parameter set of the second best cost is listed for the DE optimization. The parameter values returned by the optimizer are given without uncertainties and are rounded to the second decimal place. Setting the values with higher accuracy in the computer control does not affect the outcome. The GP and NN optimization resulted in similar parameter sets. All of their parameter values differ by less than 6%, except for the Repumper detuning Δ_R. When doing a 1D scan of the Repumper detuning with the same experimental sequence used for the optimization (see Fig. 4.7), it becomes apparent that the optimal Repumper detuning is not significant. The parameter set found by the GP learner results in a slightly lower cost value. However, both NN cost value (-6.280 ± 0.018) and GP cost value (-6.37 ± 0.04) still lie within a 3σ-interval. Therefore, the resulting cost difference could be caused by the small deviations in the best found parameter values but also by fluctuations in the experiment. In summary, both the GP and NN learner were able to optimize the magneto-optical trap by identifying a parameter set within a minimum valley of the cost function. Even though the best parameter set tested by the DE algorithm results in a cost value of the same order, the optimizer did not converge within 800 runs, and thus does not seem suitable for optimizations in a larger parameter space. With the optimized parameter set the number of atoms trapped in the MOT is in the order of $6.3 \cdot 10^8$. This is similar to what was reached by manual optimization, as discussed in Julia Gamper's Master's thesis [40].

4.3 Magnetic Transport Optimization

The magnetic transport is an important part of the experiment since it connects the MOT chamber with the science chamber. However, the additional step of transporting the atoms from one chamber to the other can lead to atom loss and heating effects. Therefore, optimizing the magnetic transport is crucial for the experiment's overall performance. Absorption imaging in the MOT and science chamber can be used to characterize the cloud density distribution and temperature at the start and end of the transport. It is, however, not possible to image the cloud during the transport. The magnetic transport itself is thus some kind of black box, since only the start and end point can be characterized. This makes the manual optimization more difficult. Furthermore, finding a suitable parametrization for y_{V_0} is a complex task due to the large parameter space for the optimization problem. The magnetic transport is thus a suitable application for a machine learning optimization, that can possibly lead to improved and faster results.

This chapter characterizes at first the initial magnetic transport performance in Section 4.3.1. The following Section 4.3.2 discusses the chosen parametrization for the potential minimum trajectory and the cost function required for the machine learning optimization. In the next step the optimization results are analyzed in Section 4.3.3 with respect to the influence of the cost function (see Section 4.3.3.1), the frequency range (see Section 4.3.3.2) and the transport time (see Section 4.3.3.3).

4.3.1 Initial Magnetic Transport Characterization

The magnetic transport of the atoms can be controlled by modifying the trajectory of the trap potential minimum $y_{V_0}(t)$ along the transport axis. Therefore, the optimization task for the magnetic transport is to find a suitable curve for the potential minimum trajectory. Badr et al. [55], who worked with a similar magnetic transport implementation, discussed the performance of different time profiles for the trajectory of the magnetic trap minimum. In their experiment they are magnetically transporting sodium atoms over a length of 310 mm in 600 ms. Various parametrization were tested, including trajectories with constant velocity and trajectories with constant acceleration. Furthermore, they designed a parametrization with the error function as a foundation. They concluded that the transport curve based on the error function, which is defined as [55]

$$y_{V_0}(t, T_{\mathrm{TM}}) = \frac{L}{2} \cdot \left(1 - \mathrm{erf}\left[\log\left(\sqrt{\frac{T_{\mathrm{MT}} - t}{t}}\right)\right]\right) := y_{\mathrm{err},V_0}(t) \tag{4.2}$$

performed best out of the tested parametrizations. The trajectory is constructed such that all time derivatives are continuous, resulting in a smooth curve that reaches the end of the transport after a fixed time T_{MT} [55]. In this way it fulfills the boundary conditions $y_{V_0}(t = 0) = 0\,\mathrm{mm}$ and $y_{V_0}(t = T_{\mathrm{MT}}) = L$, with L being the length of the transport. The $y_{\mathrm{err},V_0}(t)$ trajectory with corresponding velocity and acceleration profiles, adapted for a magnetic transport length of $L = 450\,\mathrm{mm}$ and a total magnetic transport time $T_{\mathrm{MT}} = 1.5\,\mathrm{s}$ is plotted in Fig. 4.8. The advantage of y_{err,V_0} compared to a simple linear function is that the trapping potential accelerates smoothly at the beginning of the transport and smoothly decelerates to a velocity of zero at the transport end. The velocity is constant in the middle of the transport. The trajectory defined in Eq. (4.2) will be referred to as error function throughout this thesis.

To test if this trajectory also gives meaningful results for the implementation in our experiment, the magnetic transport was conducted with the transport trajectory shown in Fig. 4.8. The performance of the magnetic transport was analyzed by doing two characterization measurements, namely a

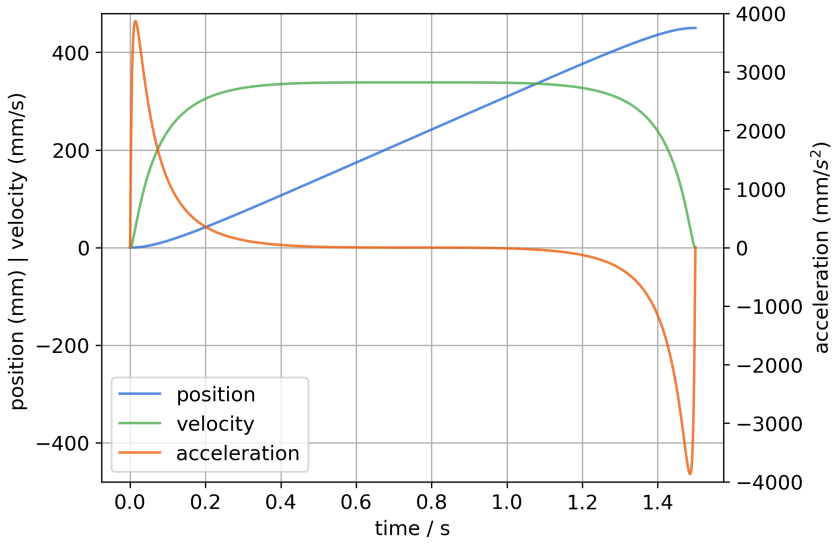
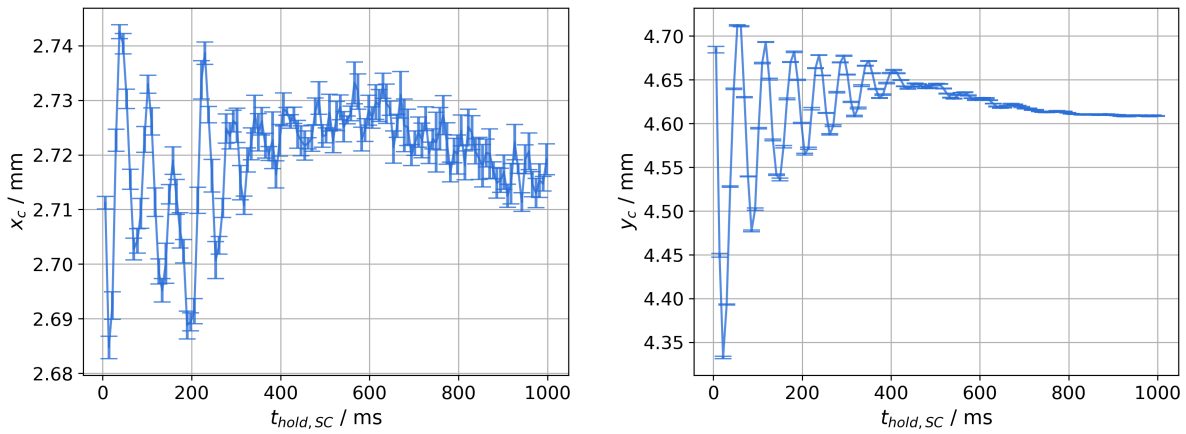


Figure 4.8: Position curve with respective velocity and acceleration profile for the potential minimum parametrization constructed by [55] and defined in Eq. (4.2). For the plot the transport length L is set to 450 mm and the total magnetic transport time is set to $T_{\mathrm{MT}} = 1.5$ s.

magnetic trap holding time measurement and a TOF measurement after the transport sequence. The experimental sequences of the characterization measurements will be explained in the following. For both measurements, the atoms are first prepared in the MOT chamber according to the steps discussed in Section 2.1 and then transported to the science chamber.

The first characterization measurement captures the dynamics of the atom cloud immediately after the transport. This is done by holding the atoms for different durations in the last trap of the transport, generated by the SC coils, and measuring the cloud center position. The center position is extracted by doing absorption imaging along the z-axis with a TOF of 10 ms. Similar to the measurement plotted in Fig. 2.6, the center position is defined relative to an arbitrary origin in the capture area of the SC absorption imaging camera. The distances correspond to actual distances covered by the cloud in the science chamber. The displacement of the atom cloud center position along the x- and y-axis can be seen in Fig. 4.9(a) and Fig. 4.9(b). The maximal displacement along the x-axis, meaning the distance between maximal and minimal center position, is around $S_{x,\mathrm{max}} \sim 0.05$ mm. The small sloshing amplitude could be induced by the changing magnetic field gradient along the x-axis at the end of the transport, as discussed in Section 2.2. The sloshing is stronger along the y-axis, with a maximal sloshing amplitude of around $S_{y,\mathrm{max}} \sim 0.38$ mm. This sloshing occurs when the magnetic trapping potential stops at the end of the transport, but the trapped atoms still have an increased velocity component along the transport axis. Since the trap is not moving anymore, the atoms can only move within the trapping potential and thus start sloshing around the trap center. The sloshing amplitude reduces over time. This is due to the atoms in the cloud oscillating with different frequencies, since the magnetic trap is not a harmonic trap [45].

It is planned that the atom chip, featuring the HBAR, will be mounted on a sample holder perpendicular to the transport axis. The planned setup in the science chamber region is shown in Fig. 4.10. After arriving in the last quadrupole trap, the atoms will be transferred to a Z-wire trap which is integrated on the atom chip. The aim is to bring the atoms with the quadrupole trap as close as possible to the surface of the atom chip for loading into the Z-wire trap. The Z-wire trap will then move the atoms even



(a) (b)

Figure 4.9: Displacement of the cloud center x_c and y_c along the **(a)** x-axis and **(b)** y-axis for different holding times $t_{\text{hold, SC}}$ in the SC trap after the magnetic transport with y_{err,V_0} as parametrization for the trajectory of the magnetic trap center. The images were taken after a TOF of 10 ms. The cloud width is extracted by a Gaussian fit to the summed optical density along the respective axes. The center position is defined relative to an arbitrary origin in the capture area of the imaging camera. The results are averaged over 20 measurements.

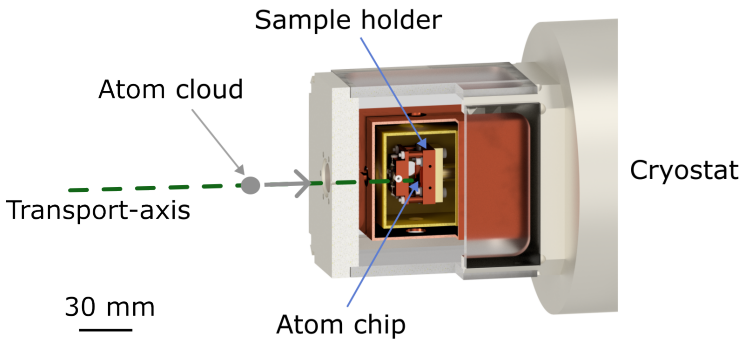
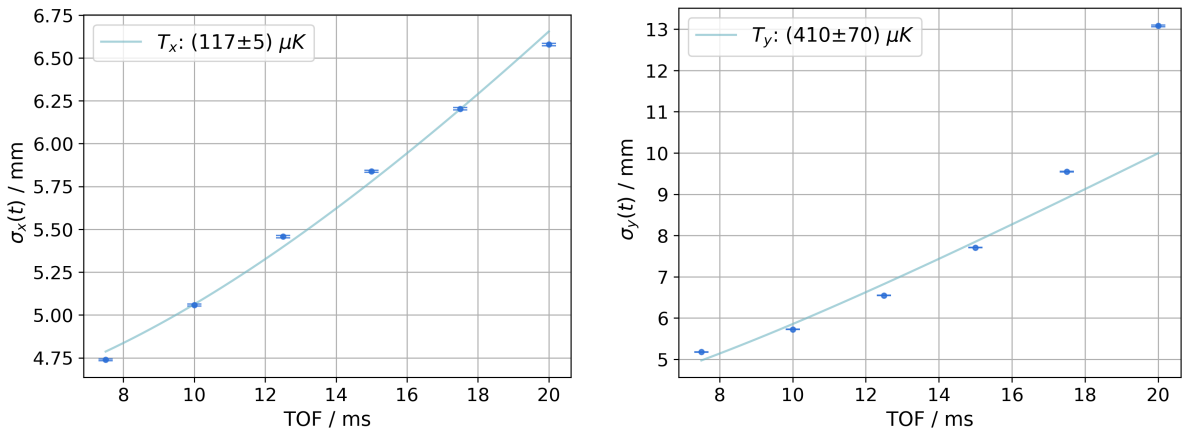


Figure 4.10: CAD drawing, by Leon Sadowski [60], illustrating the science chamber region with orientation of the atom chip and the transport axis of the magnetic transport. The atom chip is mounted on a sample holder perpendicular to the transport axis. The sample holder is connected to the cryostat.

closer to the chip. Thus, the sloshing of the atom cloud along the transport axis in the last quadrupole trap, could lead to the atoms hitting the chip surface resulting in atom loss and coating of the chip with Rubidium atoms. This coating may result in the buildup of surface charges, which would disturb the Rydberg atoms due to their high sensitivity to stray electric fields [60, 93]. Therefore, it is important to reduce the sloshing as far as possible to minimize the sloshing induced coating. In addition to that the sloshing causes the atom cloud to heat up.

The overall heating induced by the transport can be characterized with a time of flight measurement in the science chamber. To do so the atoms are transported to the science chamber and held in the last SC trap with $B' = 130\,\text{G/cm}$ for a fixed time of 450 ms. At this point the sloshing amplitude is already



(a) (b)

Figure 4.11: Time of flight (TOF) measurement in the science chamber after the magnetic transport with $y_{\mathrm{err},V_0}(t)$ as potential minimum parametrization. The measurement is performed after 450 ms holding time in the last transport trap with $B' = 130$ G cm and subsequent ramping down to $B' = 130$ G cm in 50 ms. **(a)** Measured cloud width σ_x along the x-axis. **(b)** Measured cloud width σ_y along the y-axis. The measurement was repeated 20 times. $T_{x,y}$ are extracted by fitting Eq. (2.1) to the data.

significantly reduced. This is important so that the cloud is approximately Gaussian distributed and it is thus possible to extract the cloud width from a Gaussian fit to the cloud density distribution. It is not waited for a longer time, since this would increase the measurement time and cause higher atom loss due to collisions with background gas. After the holding time the magnetic field is linearly ramped down to a magnetic field gradient of $B' = 60$ G/cm in 50 ms. In the next step the TOF measurement is performed. The reason for ramping down the magnetic field before performing absorption imaging is that the magnetic fields can not be turned off instantaneously (as discussed in Section 2.1.3). During the time the magnetic field drops off the residual magnetic field, due to eddy currents, induces random dynamics in the atom cloud. This prevents the atoms from expanding freely during the intended free expansion time in the TOF measurement and influences the imaging due to the inhomogeneous magnetic field. To minimize this effect, the magnetic field is decreased to 60 G/cm before turning it off completely for the TOF measurement. This reduces eddy current, making it easier to capture the actual dynamics of the atom cloud.

The results of the TOF measurement are shown in Fig. 4.11. Eq. (2.1) is used as fit function to extract fit values for T in x- and y-direction. The poor fit along the y-axis is a result of the sloshing induced dynamics and the fact that the atom cloud has not yet thermalized. Due to the additional velocity in y-direction, the velocity distribution can not be described by a Maxwell-Boltzmann distribution which is a prerequisite for extracting the temperature from the TOF measurement [38]. Furthermore, the temperatures in x- and y-direction differ, since the atom cloud is not fully thermalized. Thus, $T_x = (117 \pm 5)$ µK and $T_y = (410 \pm 70)$ µK, extracted from the fits, are not the actual temperatures of the atom cloud. However, both quantities are still a measure for the kinetic energy in the system and reflect the temperature that would be reached when waiting long enough for the cloud to fully thermalize. The fit results for $T_{x,y}$ show that the atom cloud gets heated up during the transport. The higher value for T_y is caused by the increased velocity component along the transport axis.

The last important measure for characterizing the performance of the magnetic transport with y_{err,V_0}

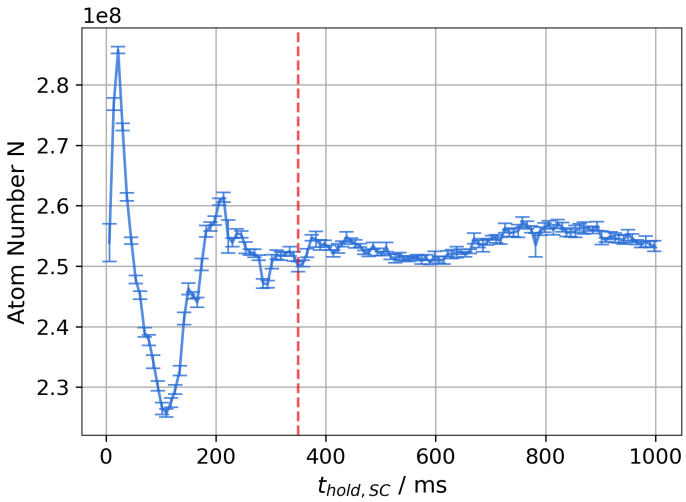


Figure 4.12: Atom number extracted from absorption images for different holding times $t_{\text{hold, SC}}$ in the SC trap after the magnetic transport with y_{err,V_0} as parametrization for the trajectory of the magnetic trap center. The datapoints show the results averaged over 20 measurements. The red dashed line is at 350 ms and indicates the time point where the atom number is extracted for the transport efficiency η_T.

as transport curve is the transport efficiency η_T describing the relative number of atoms reaching the end of the transport. Fig. 4.12 shows the measured atom number for different holding times in the SC trap. The sloshing measurement (from Fig. 4.9) was analyzed with respect to the atom number for this plot. The oscillation of the atom number for holding times between 0 ms and 300 ms is not expected and is probably due to the poor Gaussian fit when the sloshing amplitude is too large. Therefore, the extracted atom number values are only a rough estimate. After around 350 ms the atom number starts to stabilize. Therefore, the value for the atom number after a trap holding time of 350 ms in the last trap will be used as reference for the number of atoms arriving in the science chamber. The number of atoms ending up in the SC for this measurement is thus $(2.500 \pm 0.009) \cdot 10^8$. The error estimate does not include systematic errors. Therefore, the true uncertainty is much higher due to the aforementioned imaging difficulties. The magnetic transport was started with $(1.0 \pm 0.1) \cdot 10^9$ atoms being trapped in the MOT chamber (see Section 2.1.3). This results in a transfer efficiency of $\eta_T = (25.0 \pm 2.5)\%$ for the magnetic transport with the error function y_{err,V_0} and a total magnetic transport time of 1.5 s.

In conclusion, when the error function $y_{\text{err},V_0}(t)$ is used to parametrize the potential minimum trajectory, around 1/4 of the atoms initial trapped in the MOT are successfully transported to the SC. Background collisions alone however can not fully explain this atom loss. An atom cloud trapped in a static potential with a lifetime of (2.108 ± 0.009) s, would only lose around 50% of the initial atoms in 1.5 s. The stated lifetime is the MOT chamber lifetime, extracted from Fig. 2.8(a). The transport is mostly through a region outside the MOT Chamber, where we expect the pressure to be better. Therefore, the loss is overestimated here. This shows that there is still room for improvement with regard to the choice of the transport trajectory. Furthermore, using y_{err,V_0} as the transport trajectory leads to sloshing effects in the science chamber. These effects could potentially lead to coating of the atom chip with Rubidium atoms, as well as atom loss and heating of the cloud. This motivates further optimization of the trajectory with the machine learning online optimization package M-LOOP.

4.3.2 Transport Curve Parametrization and Cost Function

To apply the implemented optimization routine to the magnetic transport, a suitable parametrization and cost function with regard to the optimization task has to be defined. This section will introduce the definition of the parametrization in Section 4.3.2.1 and the definition of the cost function in Section 4.3.2.2.

4.3.2.1 Parametrization of the Transport Curve

The transport curve must be parametrized in such a way that a discrete number of optimization parameters can modify the trajectory while preserving the boundary conditions. This can be realized by choosing a fixed base function and adding variable modulation terms to the base function during the optimization process. Even though the error function $y_{\mathrm{err},V_0}(t)$ induces sloshing and heating effect in the transport, it is a reasonable choice for the base function of the parametrization. It is a smooth curve with continuous time derivatives that fulfills the given boundary conditions, and is thus vastly superior to a linear transport curve. Furthermore, already $\sim 25\%$ of the atoms reach the end of the transport with this curve, which is a good starting point for the optimization.

The full parametrization that is used for the optimization process is defined as follows

$$y_{V_0}(t, T_{\mathrm{MT}}, X) = y_{\mathrm{err},V_0}(t, T_{\mathrm{MT}}) + \sum_{k=k_0}^{k_{\max}} A_k \cdot \sin\left(\frac{\pi k}{T_{\mathrm{MT}}} t\right), \tag{4.3}$$

where the base function $y_{\mathrm{err},V_0}(t, T_{\mathrm{MT}})$ is modulated by a finite sine series. This parametrization is inspired by the chosen parametrization for the optimal quantum control search of a quantum brachistochrone trajectory by Manolo et al. [94]. The motivation for adding the finite sine series is that it generates new functions with periodic variations without introducing any discontinuities. Furthermore, choosing $k \in \mathbb{N}$ ensures that the boundary conditions $y_{V_0}(t = 0, T_{\mathrm{MT}}, X) = 0\,\mathrm{mm}$ and $y_{V_0}(t = T_{\mathrm{MT}}, T_{\mathrm{MT}}, X) = 450\,\mathrm{mm}$ are still met. The frequency range $f = \left\{\frac{\pi k}{T_{\mathrm{MT}}}\right\}$, with $\{k \in \mathbb{N} | k_0 \leq k \leq k_{\max}\}$ must be fixed before starting the optimization. The respective amplitudes $X = \{A_k\}$, which define the contribution of the different frequency components, are the free optimization parameters that are handed over to the machine learner. The number of optimization parameters is thus defined by the number of modulation terms.

Depending on the chosen amplitudes it can happen that Eq. (4.3) results in values smaller than 0 mm or greater than 450 mm during the transport time. However, the transport is restricted to a range of $[0\,\mathrm{mm}, 450\,\mathrm{mm}]$ along the transport axis by the assembly of the magnetic coils. Therefore, the additional restriction that every $y_{V_0}(t, T_{\mathrm{MT}}, X) > 450\,\mathrm{mm}$ will be set to 450 mm and every $y_{V_0}(t, T_{\mathrm{MT}}, X) < 0\,\mathrm{mm}$ will be set 0 mm is applied. This introduces points of discontinuity in the first and second derivative of the trajectory, which could lead to large slopes in the current traces. Since the response of the magnetic field to the current changes is not instantaneously (as discussed in more detail in Section 4.3.3.2), this will in most cases not be transferred to fast changes of the magnetic trapping potential. In the case where it does lead to a rapid movement of the trapping potential this will result in a worse performance of the transport. However, the machine learner learns about this due to the associated higher cost and will therefore avoid the corresponding parameter regions. It is therefore assumed that the chosen parametrization is suitable for the given problem.

4.3.2.2 Cost Function

Besides choosing a suitable parametrization, it is important to decide on a cost function that can capture the performance of the magnetic transport in one value. The goal is to increase the atom number transferred to the science chamber and at the same time decrease the sloshing and the transport induced heating of the atom cloud. The atom number $N(X)$ for each parameter set X can be extracted from one absorption image in the science chamber taken after the transport.

To measure the temperature or the sloshing, several absorption images, either at different TOF values or different holding times have to be taken (see Fig. 4.9 and Fig. 4.11). In principle this could be implemented into the machine learning cycle. However, the time per optimization cycle would be significantly extended by this. To reduce the time it takes the algorithm to find a minimum, it is preferable to define a cost value that can be extracted from a single absorption image. Therefore, the widths $\sigma_{x,y}(t_{\mathrm{TOF}})$ of the atom cloud, at a fixed TOF time, are used as further measures for the transport performance. The indices x and y refer to the axes of the imaging plane in the science chamber. This choice was made because the widths are influenced by the cloud temperature and by sloshing. Since $\sigma_{x,y}(t_{\mathrm{TOF}})$ scales with $\sqrt{T}$ [38], T being the cloud temperature, it can be used as a temperature measure, assuming a thermal distribution. Even if the cloud is not fully thermalized by the time of the imaging, the cloud width is still a measure for the velocity distribution in the atom cloud. A larger sloshing amplitude corresponds to an increased velocity component along the transport axis and would thus result in an elongated cloud. Therefore, it can also quantify the sloshing. The widths of the atom cloud thus seem to be suitable quantities for characterizing the transport with respect to the induced heating and sloshing.

The absorption imaging sequence to extract the atom number $N(X)$ and the cloud widths $\sigma_{x,y}(X)$ is conducted after 350 ms holding time in the SC trap. The chosen holding time is a trade of between waiting long enough such that the sloshing is reduced (see Fig. 4.12 and Fig. 4.9) and having a short enough measurement cycle to reduce the total optimization time. The free expansion time before the absorption imaging is set to $t_{\mathrm{TOF}} = 10\,\mathrm{ms}$. The individual cost factors $N(X)$, $\sigma_x(X)$ and $\sigma_y(X)$ are mapped to one single cost value $C(X)$ by taking the inverse of the cloud widths and multiplying it with the atom number

$$C(X) = -\left(\frac{N(X)}{N^0}\right)^a \cdot \left(\frac{\sigma_x^0}{\sigma_x(X)} \cdot \frac{\sigma_y^0}{\sigma_y(X)}\right)^b. \tag{4.4}$$

Since the extracted atom number $N(X)$ and cloud widths σ_x and σ_y are of different orders of magnitude, they are scaled with $N^0 = 1.9 \cdot 10^8$, $\sigma_x^0 = 5.1\,\mathrm{mm}$, and $\sigma_y^0 = 5.9\,\mathrm{mm}$ respectively. All cost factors are now of the order of one and can be multiplied together in the cost function. Since the machine learner is designed such that it tries to minimize the cost function, a lower cost value should correspond to a better performance of the experiment with the current parameter set. This is taken into account by the minus in the definition of $C(X)$. It ensures that a higher atom number and smaller cloud widths lead to a lower cost value. By introducing the weight factors a and b (as done in Eq. (4.4)) the contributions of the two cost factors to the total cost can be adjusted. Thus, it is possible to decide if the optimization should either happen with stronger focus on the atom number or with respect to the cloud widths. The influence of different weights on the optimization process will be discussed in Section 4.3.3.1.

Furthermore, an additional condition is applied to the cost function. Whenever the atom number is smaller than $3 \cdot 10^7$ the cost is set to zero. The reason for this is that the absorption imaging does not work reliably for atom clouds with an atom number much smaller than $3 \cdot 10^7$. This sometimes causes unreasonable results for atom number and cloud width from the Gaussian fit. To prevent giving the

machine learner wrong feedback, this condition was introduced. Additionally, a transport sequence resulting in an atom number smaller than $3 \cdot 10^7$ does not resemble as a good run for the purpose of the actual experiment, even if the cloud is really compressed. The experiment aims for at least an atom number of this order to achieve a good coupling between electromechanical oscillator and Rydberg atoms.

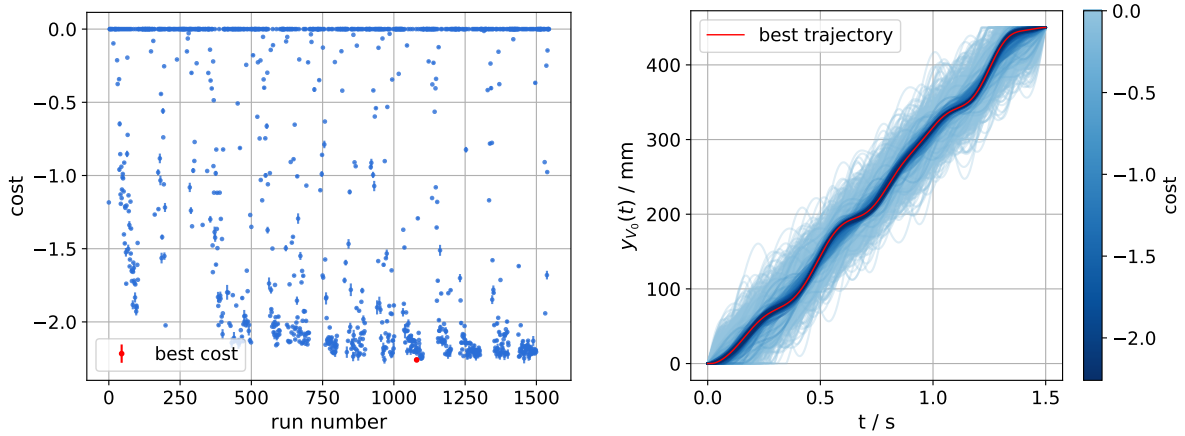
The measurement will be repeated three time for the same parameter set X, and the mean over all resulting cost values will be fed back to the machine learner. The uncertainty $U(X)$ of the cost is given by the standard error of the mean. The cost function can thus quantify the performance of the magnetic transport with respect to the atom number and the cloud width, which is a measure for the temperature and sloshing, by doing a single measurement at the end of the transport.

4.3.3 Characterization of the Optimization Process

This section discusses the results of the magnetic transport optimization with different configurations. It was decided to use the neural network as the machine learner controller for all optimizations instead of the Gaussian Process. Even though the Gaussian Process performed better for the MOT optimization (see Fig. 4.5), it quickly became apparent that for the magnetic transport the optimization with the Gaussian Process takes longer than with the neural net. This is expected due to the larger parameter space, as discussed in Section 3.3.0.3. The following optimizations are thus all conducted with the neural network as machine learner. If not stated otherwise the first parameter set to be tested during the optimization is chosen as $X_0 = \{A_k = 0, \forall k\}$. The parameter set X_0 describes the default error function y_{err,V_0}. In this way it is possible to immediately check if newly generated parameter sets perform better or worse than the pure base function, by comparing the cost values with the initial cost value of the run. As discussed in Section 3.2 there are different possible conditions that can define the stopping point of the optimization routine. It was decided to choose a maximal duration of 11 hours as the halting condition in order to ensure a predictable end point of the optimization. In the following the results of different optimization runs will be analyzed. At first the influence of different weights a and b in the cost function will be discussed in Section 4.3.3.1. Next, the modulation frequency range will be varied to check if this can further improve the results (see Section 4.3.3.2). These characterizations are all conducted for a magnetic transport parametrization with a total transport time of $T_{\mathrm{MT}} = 1.5\,\mathrm{s}$. In the last step it will be checked if a longer magnetic transport of $T_{\mathrm{MT}} = 2\,\mathrm{s}$ and the respective machine learning optimization lead to better results (see Section 4.3.3.3).

4.3.3.1 Influence of the Cost Function Weights

The cost function, defined in Eq. (4.4), has two free parameters that have to be set before starting the optimization run. These are the weights a and b, which define the contribution of the atom number $N(X)$ and the atom cloud widths $\sigma_{x,y}(X)$ to the cost value, respectively. Their impact on the performance of the optimization will be the focus of the following discussion. The total magnetic transport time is set to $T_{\mathrm{MT}} = 1.5\,\mathrm{s}$ and the modulation frequency range for the sine series is set to $f = \left\{\frac{k\pi}{1.5\,\mathrm{s}}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$. This results in 15 free parameters A_k which are the subject of optimization. The maximal and minimal boundaries of the optimization amplitudes A_k are chosen so that the allowed amplitude range linearly decreases with increasing frequency. The maximal absolute value of the amplitude for the lowest frequency component ($k = 1$) is set to 20 mm leading to boundaries defined as $[-20\,\mathrm{mm} + 0.7\,\mathrm{mm} \cdot k, 20\,\mathrm{mm} - 0.7\,\mathrm{mm} \cdot k]$ for each frequency component k. On the one hand this is



(a) (b)

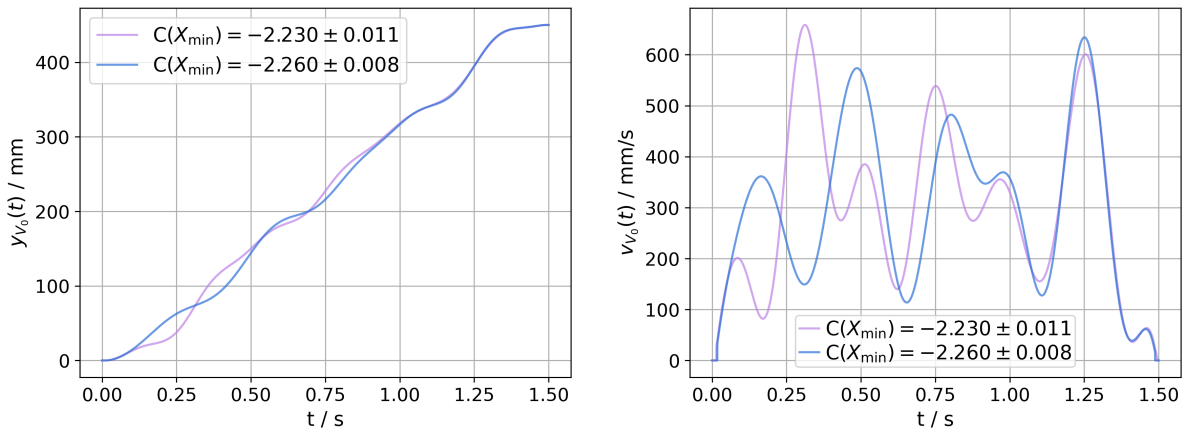
Figure 4.13: Optimization results for the 1.5 s magnetic transport with the parametrization defined in Eq. (4.3), the frequency range $f = \left\{\frac{k\pi}{1.5s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and the cost weight ratio $a : b = 1 : 2$. **(a)** Measured cost value for each run with a new parameter set. **(b)** Plot with all potential minimum trajectories that were tested during the optimization run.

done to decrease the possible parameter search space to accelerate the optimization process. On the other hand, higher frequency modulations with large amplitudes lead to large current slopes. Since the magnetic field can only follow current slopes up to a given range, which will be discussed in more detail in Section 4.3.3.2, the restriction to lower amplitude values for higher frequencies was chosen. Since the optimal parameters are not found at the border of the set boundaries, the boundary ranges were not altered any further.

The performance of the magnetic transport optimization was tested with the weight ratios $a : b = \{1 : 1, 1 : 2, 1 : 4\}$. At first the results for $1 : 2$ weights will be discussed. After this the results will be compared to the outcome of the optimization with the other weight factors.

Analysis of the optimization results with $a : b = 1 : 2$ Fig. 4.13(a) shows the optimization process for the weight ratio $a : b = 1 : 2$. The measured cost value for each run is plotted against the run number. Each run corresponds to a different tested parameter set X. The algorithm starts with a cost value of -1.1839 ± 0.0009 for the base function y_{err,V_0} and converges around a cost of -2.3. The reason for the relatively huge amount of datapoints lying exactly at zero, is due to the cost function condition that assigns every run with an atom number smaller than $3 \cdot 10^7$ to a cost value of zero. All trajectories that were tried during the optimization process are plotted in Fig. 4.13(b). The color of the trajectory corresponds to its respective cost. The red curve shows the trajectory that resulted in the best cost value. This trajectory corresponds to the cost value $C(X_{\mathrm{min}}) = -2.260 \pm 0.008$ marked in red in Fig. 4.13(a). The best found trajectory differs significantly from the base function y_{err,V_0} and improves the performance of the transport in terms of the cost value. This shows the effectiveness of the machine learning optimization with the given transport curve parametrization.

Moreover, the plot shows that a variety of parameters were tried out. The optimal trajectory found lies well within the range that can be reached with the set parameter bounds. This indicates that the parameter bounds were chosen appropriately, giving the learner the freedom to reach the optimal trajectory with



(a) (b)

Figure 4.14: Comparison of **(a)** best found trajectories $y_{V_0}(t)$ and **(b)** respective velocity profiles $v_{V_0}(t)$ of the potential minimum along the transport axis for two optimization runs with the same configurations. Both optimization runs for the 1.5 s magnetic transport were conducted with the parametrization defined in Eq. (4.3), the frequency range $f = \left\{\frac{k\pi}{1.5s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and the cost weight ratio $a : b = 1 : 2$. The blue curves correspond to the first optimization run in Fig. 4.13(a) and the purple curves correspond to the second optimization run in Fig. A.1(a) in the appendix.

the given parametrization. In some cases the cutoff condition, that sets every y value below or above the start or end point to 0 mm or 450 mm respectively, generates trajectories where the trapping potential is not accelerated and decelerated smoothly at the start and end point. One possible improvement of the parametrization could thus be to apply some kind of smoothing to these curves. However, since the algorithm seems to learn that these curves result in bad outcomes, this was not done here.

Furthermore, curves resulting in a good cost value in the order of −2.26, all have a similar shape. Curves with higher cost values, however, seem to clearly differ. This could mean that the solution found is near a global minimum. Alternatively, the algorithm may have become trapped in a local minimum and did not explore different trajectories that could still lead to better outcomes. To test if the solution is reproducible a second optimization run with the same configurations was started. The optimization process, in terms of measured cost for each run, is shown in Fig. A.1(a) in the appendix. The best trajectory that was found for the second optimization run is plotted together with the trajectory of the first run in Fig. 4.14(a). Their respective velocity profiles can be found in Fig. 4.14(b). The shapes of both curves are similar at the beginning and end. Between 0.1 s and 0.9 s the form differ slightly. For the trajectory of the first optimization run the trapping potential is faster in the beginning than for the trajectory of the second optimization, but is decelerated again at the point where the second one is further accelerated. However, it seems like the small differences in the first half of the transport do not affect the performance significantly, since the purple curve results in a minimal cost of $C(X_{\text{min}}) = -2.230 \pm 0.011$, which lies in a 3σ environment around the cost of the blue curve. Therefore, it seems like there are several minima in the parameter landscape with a very similar performance of the magnetic transport. For such a complex optimization problem it is to not surprising that it was not possible to find one specific global solution. Nevertheless, the comparison shows that the dynamic of the potential minimum at the start and end point is an important factor for the performance of the magnetic transport. Apparently it is beneficial for the transport if the potential minimum is not continuously decelerated at the end of the

transport, like it is the case for the error function y_{err,V_0}, but rather experiences some short acceleration before being fully decelerated to a velocity of zero. Furthermore, in contrast to the error function, the velocity is not kept constant during the majority of the transport time. The results show that the introduction of additional dynamics is beneficial for the successful transport of the atoms from the MOT chamber to the science chamber.

However, the position and velocity curves plotted in Fig. 4.14(b) and Fig. 4.14(a) are only theoretical time profiles for the motion of the trapping potential. The actual dynamic behavior of the trapping potential and the atoms trapped within can not be imaged along the transport. The trajectories probably differs from the theoretical ones due to the retarded response of the magnetic fields or other perturbations introduced by the hardware. For example, the differential pumping tube, with a diameter of 5 mm and a length of 10 cm is located between the two chambers. The atoms pass the differential pumping tube between 50 mm and 150 mm distance to the transport start point. This could possibly introduce further disturbance of the cloud, if the transport axis is not aligned with the axis of the tube, possibly leading to atom cutoff. However, the two optimization results shown in Fig. 4.14(a) differ slightly in the region of the tube but still produce similar cost values. Thus, it seems that the chosen trajectory through the tube does not have a major influence on the outcome here. Although it is difficult to understand the full dynamics of the transport process, machine learning can optimize the magnetic transport without knowing the entire system, as it indirectly learns about the system through the feedback it receives. Therefore, machine learning optimization is very suitable for the magnetic transport, since it exact dynamics can not be observed and are influenced by various factors that are not fully known.

The correlation between the atom number cost factor and the cloud widths cost factor for the first optimization run (blue curve in Fig. 4.14(a)) is further analyzed in Figs. 4.15(a) to 4.15(c). The scaled cost factors are plotted against each other, and the respective total cost is given by the colorbar. However, only the runs leading to a cost smaller than zero are plotted here. The reason for this is that the runs resulting in a cost of zero are those that result in an atom number below the cut-off value. In these cases, the Gaussian fit does not work reliably, resulting in unreasonable values for the cloud widths. X_{min} and X_{initial} describe the parameter set resulting in the best cost and the initial parameter set, meaning the error function y_{err,V_0}. Fig. 4.15(c) shows that σ_x and σ_y are clearly correlated with each other. Parameters that led to a decrease of the cloud width along the x-axis also decreased the cloud width along the y-axis. This is expected for an atom cloud in a magnetic trap that can be described as a Gaussian distribution. However, σ_x and σ_y are not perfectly correlated, meaning during the optimization process trajectories were tested that could e.g. improve the width in y-direction without strongly influencing the width in x-direction. This is mainly due to the transport introducing additional dynamics along the transport axis. This is reflected by the significantly greater sloshing along the transport axis compared to the x-axis. The best parameter set X_{min} leads to a relative decrease of 18.1% in σ_y and to a relative decrease of 10.7% in σ_x. This shows that the y-cloud width could be further reduced than the x-cloud width, resulting in a less elongated atom cloud at the end of the transport. The outliers in the plot, lying within the range $\sigma_x/\sigma_x^0 = [3, 4]$, mainly correspond to runs where the Gaussian fit did not properly work, but which were not caught by the atom number cutoff condition. However, it is important to notice that the plots are only displaying the cost factors for the parameters that were tested during this specific optimization process. Therefore, the described correlations do not necessarily correspond to overall correlations between the different factors. It is possible that different parameters leading to different cost factor relations were simply not tested during this optimization run.

Fig. 4.15(a) and Fig. 4.15(b) show the relation between the measured atom number with the respective widths. One can see that for the tested parametrizations the result of the atom number is almost

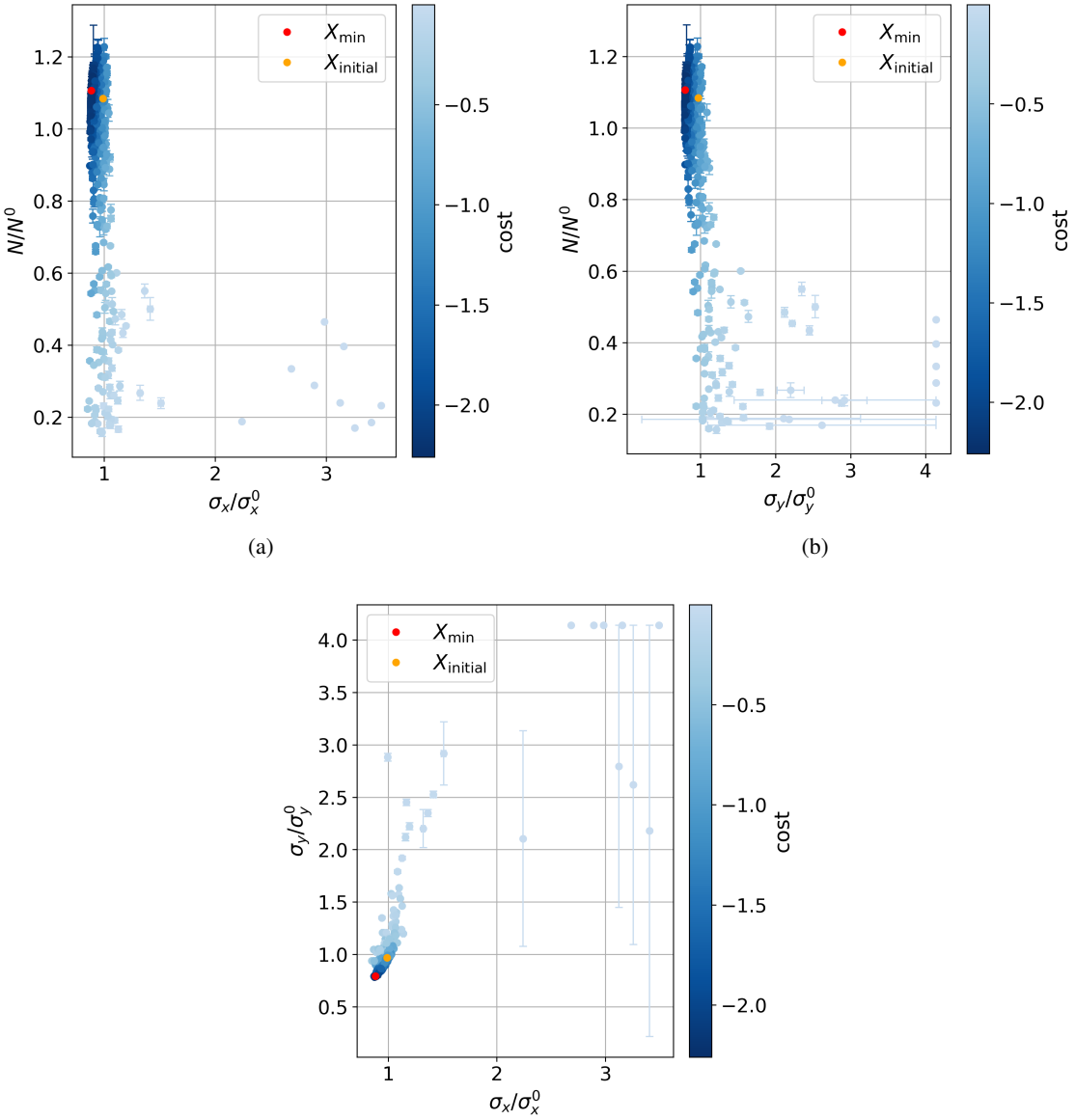


Figure 4.15: Relation between the three cost factors N/N_0, σ_x/σ_x^0, σ_y/σ_y^0 for the 1.5 s magnetic transport optimization run with the parametrization defined in Eq. (4.3), the frequency range $f = \left\{\frac{k\pi}{1.5s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and the cost weight ratio $a : b = 1 : 2$. The color bar describes the respective total cost. The cost factors are only plotted if the respective total cost value was smaller than zero.

uncorrelated to the measured x- and y-width for $N/N^0 > 0.9$. The outliers can again be attributed to a poor fit. However, for atom numbers below and above the X_{min} point the width slightly increases. This shows that X_{min} is actually located in some flat minimum for the cloud widths in the parameter space that was tested during this optimization run. Furthermore, it seems that there is a small anti-correlation between the atom number and σ_y for $N/N^0 < 0.9$. Magnetic transport realizations during which more

atom loss occurs, thus often also led to an elongated cloud along the transport axis. This does seem reasonable since strong and fast acceleration and deceleration can cause both atom loss and increases the momentum along the transport axis. The best found parameters lead to a relative atom number increase of 2.1%, compared to the error function. Although Fig. 4.15(a) and Fig. 4.15(b) indicate that there is more room for improvement in terms of the atom number, the best parameter set was chosen according to the cost factor weight ratio of 1 : 2, resulting in a larger improvement of the cloud widths.

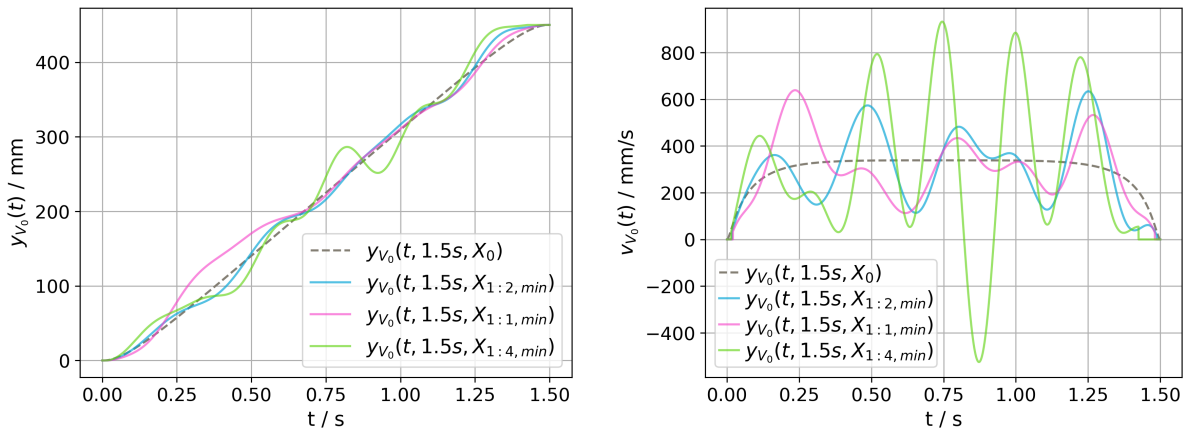
To check if the cost function fulfills the aim of reducing the sloshing and the temperature, by minimizing the cloud widths, the characterization measurements for the magnetic transport previously done with the error function as shown in Section 4.3.1 are repeated for the optimal trajectories found during the optimization runs. As discussed in Section 4.3.2, the performance of the magnetic transport is characterized by the atom number after 350 ms holding time in the last trap, the fit values for $T_{x,y}$ are determined with a TOF measurement after a total holding time of 500 ms in the last trap, and the maximal sloshing amplitude $S_{\mathrm{y,max}}$. The fit results $T_{x,y}$ are not the real cloud temperatures but still indicate the temperature that will be reached when the system thermalizes, as discussed in Section 4.3.1. The measurements are shown in Figs. 4.17(a) to 4.17(c) and the results are summarized in Table 4.2. With the best trajectory found by the optimization the sloshing amplitude could be reduced by almost a factor of two, from $S_{\mathrm{y,max}}(X_0) \sim 0.38\,\mathrm{mm}$ to $S_{\mathrm{y,max}}(X_{1:2,\mathrm{min}}) \sim 0.20\,\mathrm{mm}$. Furthermore, the sloshing is damped much more quickly than it is the case for the magnetic transport with the error function. The T_y values is also reduced by a factor of almost 2.2. The temperature fit result along the x-axis is the same within the uncertainties. The results show that by optimizing the magnetic transport with respect to the cloud widths and the atom number, it was possible to indirectly optimize the temperature of the atom cloud in the science chamber and decrease the sloshing of the cloud in the last transport trap.

Influence of different weight ratios $a : b$ The machine learning optimization of the magnetic transport was repeated with a cost function weight ratio $a : b$ of 1 : 1 and 1 : 4. The cost curves for both optimization runs are plotted in Fig. A.1(b) and Fig. A.1(c) in the appendix. The stopping condition for the optimization run was again defined as a maximal duration of 11 hours. The best parameter set for both optimizations was found in the final 350 from 1600 runs. However, the algorithm did most likely not converge at the end of the optimization time, since the measured cost does not asymptotically approach a cost value for the last optimization runs. A longer optimization time could still lead to better results.

Rather than repeating the same optimization for a duration of more than 11 hours, a second optimization process was started that builds on the results of the first. The results of the initial optimization can be incorporated by using the best found trajectory in the first optimization, $y_{V_0}(t, T_{\mathrm{MT}}, X_{\mathrm{min},1})$, as the new base function. The new parametrization for the second optimization process is given by

$$y_{V_0}(t, T_{\mathrm{MT}}, X) = y_{V_0}(t, T_{\mathrm{MT}}, X_{\mathrm{min},1}) + \sum_{k=k_0}^{k_{\mathrm{max}}} A_k \cdot \sin\left(\frac{\pi k}{T_{\mathrm{MT}}} t\right), \tag{4.5}$$

where $X_{\mathrm{min},1}$ is the best found parameter set of the respective initial optimization run. The frequency range is again defined by $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$. This parametrization biases the optimizer, and it might be more likely to become trapped in a local minimum. Despite this parametrization being slightly more biased and restricted compared to Eq. (4.3), it was tested in subsequent optimization runs to reduce the optimization time. For the cost factor ratio 1 : 1 the second optimization run with the new base function did not lead to any further improvement. The results from the first optimization with the original



(a) (b)

Figure 4.16: Comparison of the **(a)** best found potential minimum trajectories $y_{V_0}(t)$ and **(b)** respective velocity profiles $v_{V_0}(t)$ for the 1.5 s magnetic transport optimization runs with the frequency range $f = \left\{\frac{k\pi}{1.5s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and different cost factor weights $a : b = \{1 : 1, 1 : 2, 1 : 4\}$. $X_{a:b,\text{min}}$ is the best found parameter set for the optimization run with wight ratio $a : b$. X_0 is the initial parameter set, describing the base function y_{err,V_0}.

parametrization are therefore analyzed here in more detail. These results are referred to as $X_{1:1,\text{min}}$. For the optimization with the cost factor ratio of $1 : 4$ the results could be significantly improved compared to the first optimization. The cost curve for the second optimization run is plotted in Fig. A.1(d) in the appendix. The base function used for the second optimization process, which corresponds to the best trajectory found in the first optimization, is plotted in Fig. A.2(a) in the appendix. The algorithm also did most likely not converge for the second optimization, since the measured cost is still significantly improving for the last optimization runs. However, by analyzing the contribution of the different cost factors to the cost one can conclude that the parameters that decrease the cost primarily improve the cloud widths while simultaneously leading to a significant decrease in the atom number. The atom numbers almost reach the defined atom number threshold of $3 \cdot 10^7$ atoms. For this reason, the machine learning optimization was not repeated any further. The best found parameter set from the second optimization run is defined as $X_{1:4,\text{min}}$ and will be used for comparing the results of the different cost functions. The final trajectory results for all three machine learning optimizations with different cost ratios are plotted in Fig. 4.16(a) with their respective velocity curves in Fig. 4.16(b).

The characterization measurements, consisting of the sloshing measurement and the TOF measurement, are repeated for these trajectories. The measurements are shown in Figs. 4.17(d) to 4.17(i) and the extracted quantities for comparison are summarized in Table 4.2. With the parametrization $X_{1:1,\text{min}}$, 14% more atoms were transported to the science chamber compared to the realization with the $X_{1:2,\text{min}}$ parametrization. The resulting T_y value however is a factor of 1.2 greater and the maximal sloshing amplitude is almost about a factor of 2 larger. Nevertheless, compared to the initial parametrization y_{err,V_0} the temperature along the transport axis could still be decreased, even though the sloshing amplitude is approximately the same. However, the sloshing for $X_{1:1,\text{min}}$ is damped more strongly than for y_{err,V_0}. This was already observed for the parametrization with $X_{1:2,\text{min}}$. The different sloshing amplitudes can possibly be attributed to the slightly different shape of the curves at the end of the transport. For all parametrizations, the trapping potential accelerates again at the end of the transport, around $t = 1.1$ s,

X	η_T	$\frac{\sigma_x(X_0)-\sigma_x(X_{\min})}{\sigma_x(X_0)}$	$\frac{\sigma_y(X_0)-\sigma_y(X_{\min})}{\sigma_y(X_0)}$	$T_x(X_{\min})/\mu\mathrm{K}$	$T_y(X_{\min})/\mu\mathrm{K}$	$S_{y,\max}(X_{\min})/\mathrm{mm}$
$X_{1:2,\min}$	$(26.5 \pm 2.6)\%$	10.1%	18.1%	120.6 ± 2.0	189 ± 9	~ 0.20
$X_{1:1,\min}$	$(30 \pm 3)\%$	12.2%	5.3%	142.6 ± 2.4	228 ± 12	~ 0.39
$X_{1:4,\min}$	$(5.1 \pm 0.5)\%$	33.3%	34.1%	92 ± 4	126.8 ± 1.2	~ 0.09
$X^s_{1:4,\min}$	$(4.9 \pm 0.5)\%$	-	-	86 ± 4	121.5 ± 0.9	~ 0.10
X_0	$(25.0 \pm 2.5)\%$	–	–	117 ± 5	410 ± 70	~ 0.38

Table 4.2: Comparison of characterization measurements for the 1.5 s magnetic transport optimization runs with the frequency range $f = \left\{\frac{k\pi}{1.5s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and different cost factor weights $a : b = \{1:1, 1:2, 1:4\}$. $X_{a:b,\min}$ is the best found parameter set for the optimization run with wight ratio $a : b$, η_T is the transport efficiency, $T_{x,y}$ are the temperature fit results and $S_{y,\max}$ describes the maximal sloshing amplitude along the transport axis. $\frac{\sigma_{x,y}(X_0)-\sigma_{x,y}(X_{\min})}{\sigma_{x,y}(X_0)}$ describes the relative improvement of the cloud widths for the best found trajectories compared to the magnetic transport realization with the initial parameter set X_0, describing the base function y_{err,V_0}. $X^s_{1:4,\min}$ is the parametrization of the smoothed and shortened $X_{1:4,\min}$ trajectory.

before decelerating and stopping at the transport end. For $X_{1:2,\min}$ and $X_{1:4,\min}$ there is an additional small acceleration phase around $t = 1.4\,\mathrm{s}$. This results in an S-curve like shape of the potential minimum trajectory at the end of the transport in all cases. A flatter S-curve at the end, results in a smaller T_y fit result. One reason for this could be, that the trapping potential approaches the transport end more slowly. This decelerates the atoms earlier, resulting in a smaller velocity along the transport axis at their arrival in the science chamber. In this way, the $X_{1:4,\min}$ decreases the sloshing amplitude to $\sim 0.09\,\mathrm{mm}$. The $T_y = (126.8 \pm 1.2)\,\mu\mathrm{K}$ fit value, is reduced by a factor of ~ 3.2 compared to the error function y_{err,V_0}. Additionally, this parametrization results in a decreased $T_x = (92 \pm 4)\,\mu\mathrm{K}$.

However, the measured sloshing amplitude can not be directly compared to the results with the other trajectories. The respective velocity curve has a point of discontinuity at 1.43 s where the velocity jumps from 55 mm/s to 0 mm/s. This discontinuity in the velocity is a result of the cutoff condition at $y = 450\,\mathrm{mm}$. Thus, for the $y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:4,\min})$ parametrization the atom cloud already arrives at the end of the transport after 1.43 s. The trajectory thus features a small plateau at the end of the transport. During this time the sloshing is already damped, leading to a lower maximal sloshing amplitude at the start of the sloshing measurement. To capture the movement of the center position immediately after the atoms arrive in the science chamber the characterization measurements were repeated with a total magnetic transport time of 1.43 s. The results can be found in Figs. A.4(a) to A.4(c) in the appendix. It can be seen that the sudden stop of the transport leads to an increased sloshing amplitude of 0.65 mm, while the measured temperature stays the same since no additional heating effects are introduced.

Since the sloshing is still damped relatively quickly, it might be possible to prevent the first large sloshing amplitude by continuously decelerating the trap minimum until it reaches a velocity of zero. To test this, the trajectory is adapted by smoothening the end, by applying a quartic B-spline interpolation[1]. The condition $y \leq 450\,\mathrm{mm}$ is still taken into account for the smoothening. The total magnetic transport time is reduced to 1.45 s, which corresponds to the time when $y_{V_0} = 450\,\mathrm{mm}$ for the smoothed curve. The new trajectory is described with the parameter set $X^s_{1:4,\min}$. The position curve and the respective velocity curve are shown in Fig. A.2(b) in the appendix. The characterization measurements for this curve can be found in Figs. A.4(d) to A.4(f) in the appendix and are summarized in Table 4.2. With the

[1] For the smoothening the *make_interp_spline* function from the scipy.interpolate module in Python was used.

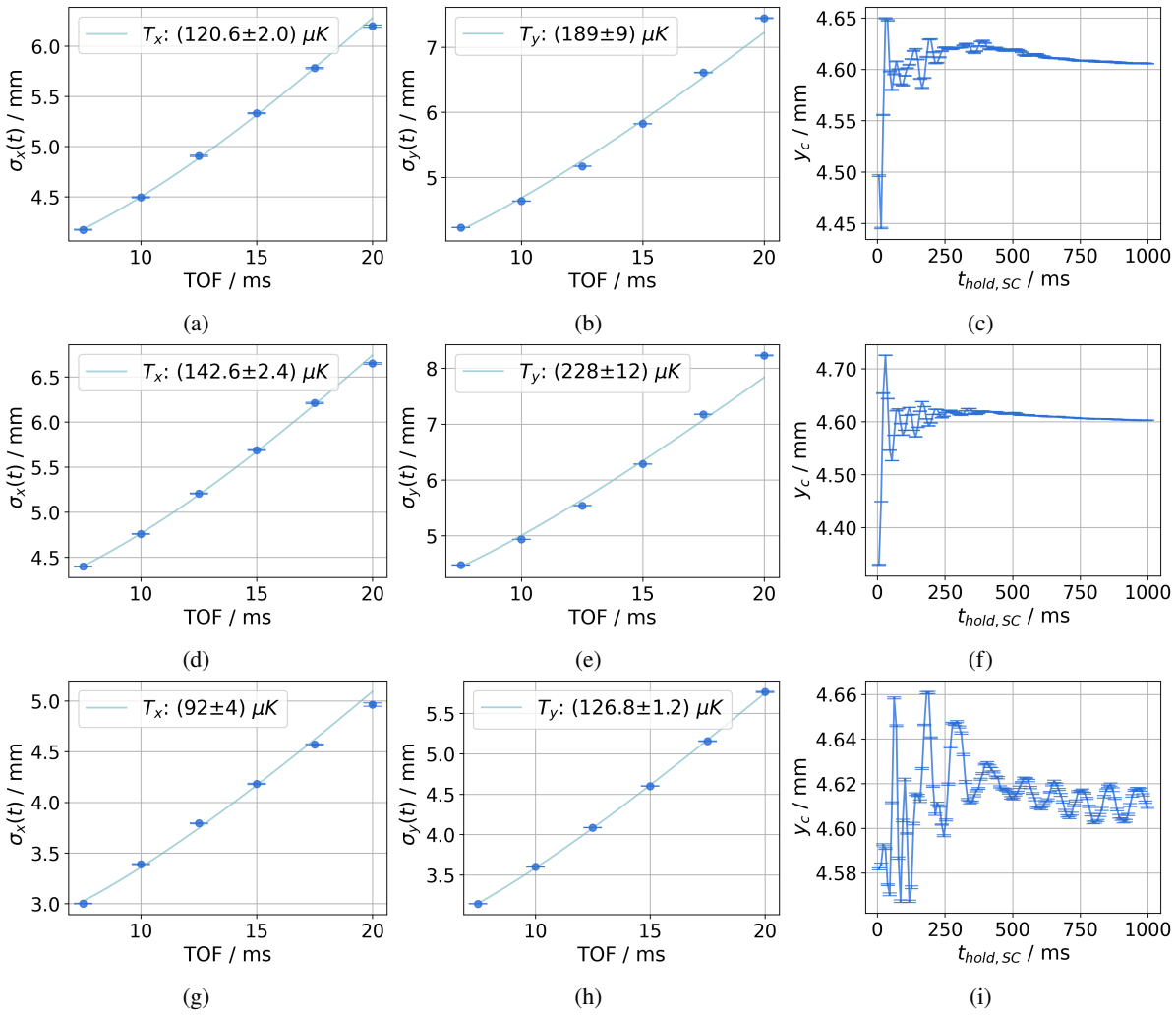


Figure 4.17: Characterization measurements for the 1.5 s magnetic transport with the best found trajectories from the optimization run with frequency range set to $f = \left\{\frac{k\pi}{1.5s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and different cost factor weights $a : b = \{1:1, 1:2, 1:4\}$. Results for **(a)-(c)** $X_{1:2,\text{min}}$, **(d)-(f)** $X_{1:1,\text{min}}$ **(g)-(i)** $X_{1:4,\text{min}}$ parametrization. The characterization measurements consist of a TOF measurement in the science chamber and the measurement of the displacement of the cloud center y_c along the transport axis for different holding times $t_{\text{hold, SC}}$ in the SC trap after the magnetic transport. $\sigma_{x,y}$ describes the cloud width along the x- and y-axis. The temperatures are extracted by fitting Eq. (2.1) to the data. The sloshing measurements are averaged over 10 loops and the TOF measurements are averaged over 20 loops.

smoothed curve at the end of the transport it is possible to reduce the sloshing to a maximal amplitude of $S_{y,\text{max}} \sim 0.10\,\text{mm}$. Furthermore, the temperatures are reduced even a bit further compared to the original $X_{1:4,\text{min}}$ parametrization. Thus, it was possible, with small post optimization by hand, to significantly reduce the sloshing and the temperature of the atom cloud at the science chamber with the best found trajectory of the optimization process with weight ratio $a : b = 1 : 4$. A small plateau at the end of the other two trajectories can also be observed. However, since this happens in the last 2 ms of the transport, the sloshing behavior can still be captured sufficiently, and the characterization measurements are thus

not repeated here for corrected trajectories.

Due to the complexity of the system, it is not possible to identify a clear effect-cause relationship between the transport curve and the resulting transport performance. One can conclude that slowly transporting the atoms at the end seems crucial for reducing heating effects. A flat S-curve like trajectory at the end points seems optimal here.

Atom Loss Analysis Apart from the fact that the $X_{1:4,\mathrm{min}}$ parameter set leads to less heating, it does cause a significant atom loss. Only $\sim 5\%$ of the atoms arrive at the end of the transport with the $X_{1:4,\mathrm{min}}$ parametrization. This atom loss also occurs similarly for the smoothed curve $X^{s}_{1:4,\mathrm{min}}$. One prominent difference between this trajectory and the other optimization results is that the magnetic trap potential goes backwards once at $t \sim 0.9\,\mathrm{s}$. The magnetic trap is strongly decelerated and accelerated again around this point. This could result in a built-up of the atom oscillation around the trap center, causing the atoms to gain enough energy to escape the trap.

To identify if the proposed loss mechanism is indeed present in the system, a simplified 1D Monte Carlo simulation of the magnetic transport was implemented. The movement of the atoms inside the moving trapping potential was simulated by solving Newtons equation of motion along the transport axis y in 1D [55]

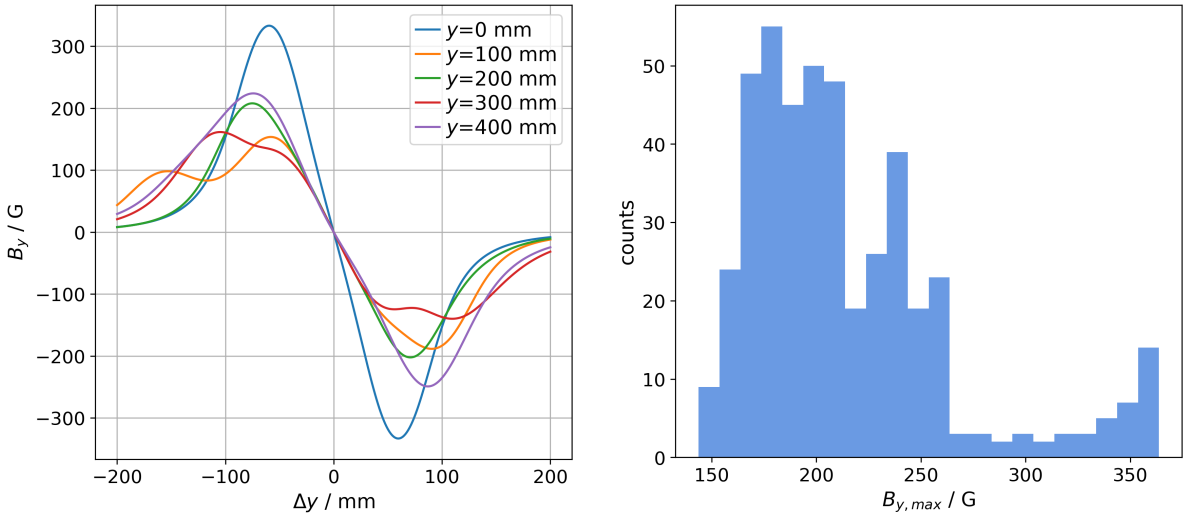
$$F_y = m \cdot \frac{d^2}{dt^2} y_A(t) = -\Big(\nabla V(y_A(t), t)\Big)_y - m \cdot \frac{d^2}{dt^2} y_{V_0}(t) \tag{4.6}$$

$$= -\,\mathrm{sgn}(y_A(t)) \cdot g_F m_F \mu_B B'_y - m \cdot \frac{d^2}{dt^2} y_{V_0}(t). \tag{4.7}$$

Here, $y_A(t)$ describes the atom position in the comoving frame where the origin is given by the position of the potential minimum, $y_{V_0}(t)$ is the potential minium trajectory during the transport, m is the mass of the Rubidium atoms and B'_y the magnetic gradient along the transport axis. The first term describes the trapping force and the second term is the additional acceleration force due to the moving potential minimum. The simulation is a classical simulation of the one-dimensional dynamics during the magnetic transport and does not take into account collisions between the atoms. It is not a full simulation of the dynamics of the atom cloud in the moving trapping potential, since only one dimension is simulated, collisions between the atoms are neglected and additional losses like Majorana spin flips are not included.

Furthermore, a constant magnetic field gradient of $B'_y = 35\,\mathrm{G/cm}$ is assumed for the simulation. This approximation only holds for small distances to the trap minimum. Due to the coil geometry the magnetic field will decrease again for larger distances. This can be seen in Fig. 4.18(a), where the magnetic fields at four different positions along the transport axis, extracted from the magnetic transport simulation written by Cedric Wind, are plotted. The different forms of the magnetic field are caused by small differences in the geometry of the transport coils. At the start and end point ($y = 0\,\mathrm{mm}$ and $y = 450\,\mathrm{mm}$) the magnetic field has a greater gradient of $B'_y = 65\,\mathrm{G/cm}$ since it is only generated by one coil pairs as discussed in Section 2.2. Due to the different coil geometry the maximal magnetic field that is generated during the transport varies along the transport axis. The distribution of the magnetic field maxima $B_{y,\mathrm{max}}$ of the magnetic fields generated along the transport axis is shown in Fig. 4.18(b).

Since the magnetic potential is proportional to the magnetic field strength, the maximal magnetic field is a measure for the trap depth of the magnetic potential during the transport. By solving the differential equation for an initial atom cloud sampled from a thermal distribution with a temperature corresponding



(a) (b)

Figure 4.18: **(a)** Form of the magnetic field B_y at different positions y along the transport axis. Δ_y describes the distance to the zero crossing of the magnetic field. Due to the different geometries of the coils generating the magnetic field, the curves have slightly different shapes. The results for the magnetic fields were extracted from the magnetic transport simulation written by Cedric Wind. **(b)** Distribution of the magnetic field maxima $B_{y,\max}$ of the magnetic fields generated along the transport axis y.

to the root mean square temperature $T \sim 290\,\mu\text{K}$ of the measured temperatures T_y and T_z in the first magnetic trap (see Figs. 2.7(c) and 2.7(d)), and with a constant magnetic field gradient of $B'_y = 35\,\text{G/cm}$, one can extract the maximal potential energy that each atom reaches in the potential.

The differential equation is solved for 4000 sampeld atoms with the magnetic transport parametrization $y_{V_0}(t) = y_{V_0}(t, 1.5s, X_{1:4,\min})$, for which the backward propagation occurs. From the simulation results one can extract the maximal distance $y_{A,\max}$ to the trap center reached for each simulated atom. By multiplying it with the magnetic field gradient $B_{y_{A,\max}} = y_{A,\max} \cdot B'_y$, one derives a measure for the maximal potential energy of each atom during the transport, in units of the magnetic field. $B_{y_{A,\max}}$ describes the maximal magnetic field that is required to trap the atom that reached a maximal oscillation amplitude of $y_{A,\max}$. Fig. 4.19 displays the distribution of $B_{y_{A,\max}}$ for all sampled atoms. The distribution of the magnetic field maxima in Fig. 4.18(b) shows that the maximal magnetic field strength generated during the transport is 360 G. Thus, one can conclude that the atoms which have a $B_{y_{A,\max}} > 360\,\text{G}$ would be lost during the transport, since they reach energies greater than the maximally generated potential trap depth during the whole transport. The trapping threshold of 360 G is indicated by the red dotted line in Fig. 4.19. It divides the $B_{y_{A,\max}}$ distribution into two distinct regions. Around 8% of the initialized atoms lie above the threshold and would thus be lost from the trap.

Fig. 4.20(a) and Fig. 4.20(b) show example trajectories for a randomly chosen atom from the distribution below the threshold and from above the threshold. Both results show that the atom is shifted back and forth at the point where the potential minimum is going backwards. However, only for the atom from the second region above the threshold the backward propagation causes an increase in the oscillation amplitude for it to be lost from the trap. This demonstrates that the assumption that additional atom loss is caused by backwards propagation is valid. However, it can not explain the amount of atoms loss

observed in the experiment, leading to a transport efficiency of only 5%. Even though it is expected that including all three dimensions and collisions between the atoms in the simulation would influence the observed dynamics, there still must be other more dominant atom loss mechanisms during the transport.

Furthermore, it was also analyzed if the atom loss due to the backward propagation introduces an effect similar to evaporative cooling, where all the fast atoms are removed from the ensemble. This could explain the reduced temperature and decrease in atom number for this parametrization. However, no indication for this could be observed when comparing the velocity distribution of the simulated cloud at the start and end point. The same simulation was also repeated for the base function y_{err,V_0} as the parametrization for the potential minimum. This did not lead to any atom loss with regard to the defined condition.

An additional possible atom loss source could be the differential pumping tube placed between the MOT chamber and the science chamber. The tube has a length of 100 mm and a diameter of 5 mm. If an atom cloud is not perfectly aligned with the tube's axis, it may be cut off at the edges. Since the atoms with higher energies can reach larger distances from the trap center, the cut off at the edge could also induce an evaporative cooling effect. However, the differential pumping tube is located between 50 mm and 100 mm from the starting point. In this region the $X_{1:4,\mathrm{min}}$ trajectory only slightly differs from the $X_{1:2,\mathrm{min}}$ trajectory, for which a much larger number of atoms arrive in the science chamber. Therefore, it can not be completely clarified which mechanism causes the atom loss for the $X_{1:4,\mathrm{min}}$ parametrization.

Additional things that could be tried out, for example, is stitching the initial part of $y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:2,\mathrm{min}})$ for $t < 1.1$ s together with the end of $y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:4,\mathrm{min}})$. In this way the backward propagation would be removed but the flattened S-curve would be preserved. Further manual modification could help to identify the reasons behind the different modulations chosen by the machine learner and the effect they have on the atom loss and temperature of the atom cloud.

In conclusion, it was possible to steer the optimization by using different cost factor weights a and b. For $a : b = 1 : 1$ a maximal transfer efficiency of 30% could be achieved. The lowest atom cloud temperature was measured for the $X_{1:4,\mathrm{min}}$ magnetic transport realization, but 80% of the atoms are lost compared to the error function. The result of the optimization run with the weight ratio of $a : b = 1 : 2$ seems to be a good trade-off between atom number, temperature and sloshing effects.

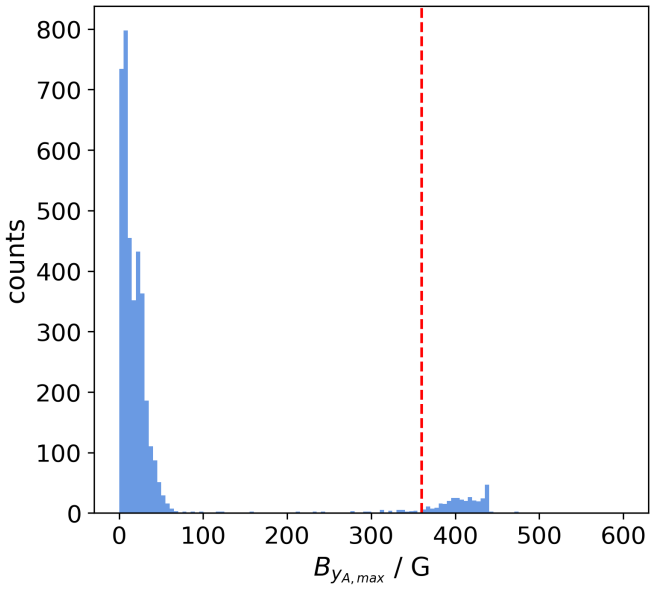
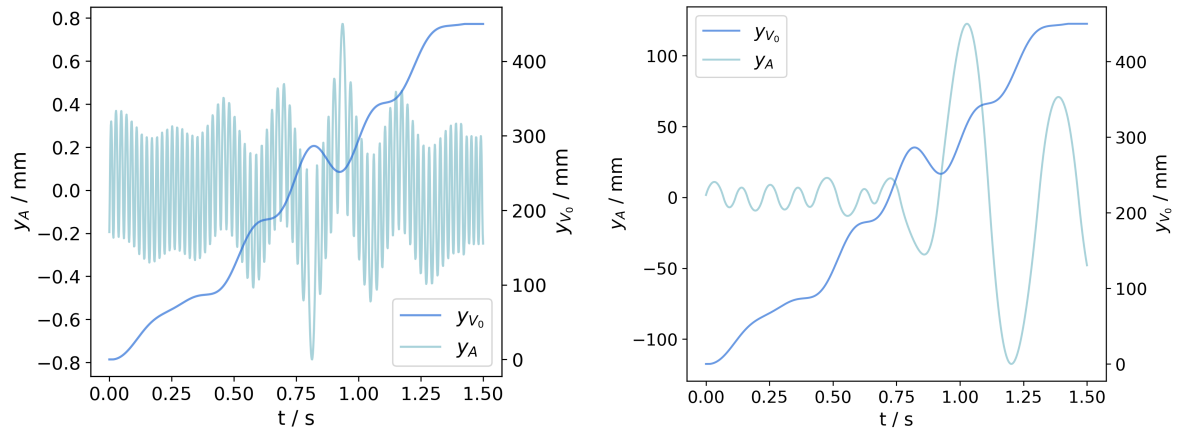


Figure 4.19: Distribution of the maximal magnetic fields required to trap the 4000 sampled atoms from the Monte Carlo simulation during the simulated magnetic transport. $B_{y_{A,\max}} = y_{A,\max} \cdot B'_y$, where $y_{A,\max}$ describes the maximal distance of an atom to the trap center during the transport. The distribution of $y_{A,\max}$ was extracted from a simplified 1D simulation of the magnetic transport with parametrization $y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:4,\min})$. $B'_y = 35\,\mathrm{G/cm}$ y_A is the magnetic field gradient along the transport axis, which was used for the simulation. The red dotted line indicates the trapping threshold and is located at $B_{y_{A,\max}} = 360\,\mathrm{G}$.



(a) (b)

Figure 4.20: Solution of Newtons equation of motion Eq. (4.7) for an atom trapped in a moving magnetic potential, which moves along the transport axis according to $y_{V_0} = y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:4,\min})$. $y_A(t)$ describes the distance between the position of the trapped atom and the potential minimum position and is thus the oscillation amplitude of the atom in the comoving frame. **(a)** Example trajectory for an atom which would stay trapped during the transport ($B_{y_{A,\max}} < 360\,\mathrm{G}$). **(b)** Example trajectory for an atom which would be lost during the transport ($B_{y_{A,\max}} > 360\,\mathrm{G}$).

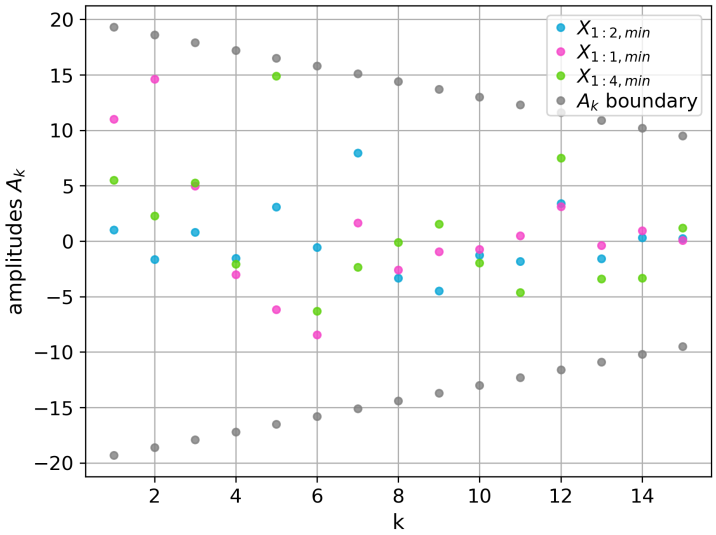


Figure 4.21: Contribution of the different frequency components to the optimization results with the weight ratios $a : b = \{1 : 1, 1 : 2, 1 : 4\}$ and frequency range $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$. The amplitudes A_k correspond to the amplitudes defined in Eq. (4.3) (for $X_{1:2,\text{min}}$ and $X_{1:1,\text{min}}$) or Eq. (4.5) (for $X_{1:4,\text{min}}$). The A_k boundaries describe the boundaries for each amplitude value during the optimization. The best found parameters are returned without uncertainties from the machine learner controller.

4.3.3.2 Influence of Modulation Frequency Ranges

In the next step it was analyzed if the modulation frequency range $f = \left\{\frac{k\pi}{1.5\,\text{s}}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ is a good choice or if changing the parameter space can give better results. The contribution of the different frequency components to the optimization results with the weight ratios $a : b = \{1 : 1, 1 : 2, 1 : 4\}$ and frequency range $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ are shown in Fig. 4.21. The defined boundaries for each amplitude are plotted in gray. For $X_{1:4,\text{min}}$ the A_5 amplitude is near the upper boundary. The rest of the amplitudes are well below the limits. Since the choice of the boundaries is always a trade of between optimization time and degree of optimization freedom, the set boundaries seem to be an acceptable choice. Furthermore, no clear trend between frequency k and respective amplitude is visible. Both even and odd amplitudes contribute, since the resulting curves show no point symmetry. For point symmetric curves the odd sine amplitudes would be negligible [94]. The highest three frequencies ($k = 13, 14, 15$) only make a minimal contribution in the $X_{1:1,\text{min}}$ and $X_{1:2,\text{min}}$ parametrization. This is a first indication, that choosing higher frequencies may not be beneficial.

Optimizations with different modulation frequency ranges were tested to see if this assumption is correct or if the results can be improved further. The magnetic transport optimization was thus repeated with a cost weighting of $a : b = 1 : 2$ and 20 optimization parameters for the frequency components $f = \left\{\frac{k\pi}{1.5\,\text{s}}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 20\}$. The boundaries are chosen again accordingly to $[-20\,\text{mm} + 0.7\,\text{mm} \cdot k, 20\,\text{mm} - 0.7\,\text{mm} \cdot k]$. The optimization results are shown in Figs. 4.22(a) and 4.22(b). The machine learning optimization was stopped again after 11 hours. A larger number of runs result into a cost of zero or near zero, compared to the optimizations with lower frequencies and a smaller parameter space. This can be explained by the fact that the parameter space is about five optimization parameters larger making it more difficult to find a reasonable parameter set. Additionally, the fast trajectory modulations may cause current modulations that are too rapid for the magnetic field

to follow. This will be analyzed at the end of this section. The trajectories producing the lowest cost values do not feature high frequency modulations. For the best found curve the last five amplitudes A_k have values below 0.5. Furthermore, the best trajectory only lead to a cost of $C = -1.67 \pm 0.05$. For the optimization with frequency range $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ the optimizer found an optimum after 11 hours with the lowest cost $C = 2.260 \pm 0.008$, and was therefore a better and faster optimization. One can conclude from this run, that a larger parameter space with higher frequencies requires an optimization run longer than 11 hours. At the same time, it is not clear if higher frequencies will contribute to further improvements.

To speed up the optimization it was decided to use the previously determined optimal trajectory taking into account up to 15 frequencies as a base function and let the machine learner only act on higher frequency contributions. Therefore, the parametrization from Eq. (4.5) is used again. The trajectory $y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:2,\mathrm{min}})$ is defined as the base function. The optimization is repeated once for the frequency range $f = \left\{\frac{k\pi}{1.5\,\mathrm{s}}\right\}$ with $\{k \in \mathbb{N} | 11 \leq k \leq 20\}$ and once with $\{k \in \mathbb{N} | 16 \leq k \leq 25\}$. This way it can be checked if either the larger parameter space of 20 parameters increases the difficulty for the optimizer or if the modulation with higher frequencies do not lead to good magnetic transport results. The optimization results, given in Figs. 4.22(c) to 4.22(f), show that the higher frequency ranges did not lead to any improvement of the initial cost. All trajectories with significant contribution from higher frequencies result in a cost around zero. The best found parametrization for both optimizations does not differ in shape from the base function $y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:2,\mathrm{min}})$. This shows that the magnetic transport does not work properly when the potential minimum trajectory features frequency components above $f = \frac{15\pi}{1.5\,\mathrm{s}}$.

When discussing possible modulation frequency ranges it is also important to characterize the magnetic field response to such modulations. Because of eddy currents, the magnetic field can not follow changes in the current that happen too fast. Modulations of the potential minimum trajectory with a certain frequency do not directly map to modulations of the current traces, and thus the magnetic field, of the same frequency. However, higher frequencies still cause faster changes of the coil current and thus correspond to higher current slopes in the current traces. As it is also not possible to map modulation frequencies directly onto respective current slopes, the generated current slopes from an actual optimization run with a desired frequency range were sampled. This gives a good estimation for the order of magnitude of the current slopes. This was done by calculating the current traces for each trajectory that was tested in the 1.5 s magnetic transport optimization run with $a : b = 1 : 2$ and $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$. For every coil pair the highest current slopes were extracted for each parametrization. The current trace of coil 1 and the SC coil feature the highest maximal current slopes. The sampled distribution of the maximal current slopes for these two coils is given in Fig. 4.23. The maximal current slopes lie in a broad range between 100 A/s to 14 000 A/s.

The response of the magnetic field to current slopes of this order was tested by measuring the magnetic response to a triangular current trace with varying slope. Coil 1 was used as the test coil pair because the highest slopes occurred for these current traces. The maximal and minimal current values are chosen such that the magnetic field is ramped from 0 G/cm to 130 G/cm and back to 0 G/cm. The current applied to coil 1 was measured with a current clamp (AC/DC Current Clamp CC-65 from Hantek [95]). The magnetic field generated by coil 1 was measured with a fluxgate magnetic sensor (Automative, Fully Integrated Fluxgate Magnetic Sensor DRV425-Q1 [96]), with a slew rate of 6.5 V/µs [96], placed above the coil. The current traces and the magnetic field responses for different applied current ramps are shown in Fig. 4.24. The magnetic field does not drop to zero when no current is applied due to the Earth's magnetic field [97]. However, in Fig. 4.24 the offset is substracted such that the magnetic field B

only describes the magnetic field generated by the coils. For a supplied current ramp with slope 100 A/s, the magnetic field can follow the current trace nicely with no time delay. It reaches its maximal measured B field strength of 2.9 G with a B field slope of ~ 4.0 G/s. The value of the maximal measured magnetic field strength is not relevant for this analysis, as it mostly depends on the sensor distance and orientation to the coil. The magnetic field response is already delayed for a current ramp with a slope of 500 A/s. The magnetic field is ~ 8% slower than expected for an instantaneous response. When further increasing the slope of the current control signal, the magnetic field response gets further delayed, up to a ~ 78% slower response for a current slope of 10 000 A/s. For the falling slope it is also visible that the magnetic field response is not linear anymore due to the induced eddy currents. Thus, one can conclude that up to current slopes around 500 A/s the magnetic field can follow the control current sufficiently with small deviations of ~ 8%. However, the analyzed optimization run featured many parameters with maximal current slopes in the range of 1 000 A/s-8 000 A/s, where the magnetic field response is significantly delayed. Thus, for optimization runs with frequency components in the range of $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$, the magnetic field can not perfectly follow the control signal. For the simple error function y_{err,V_0} a maximal current slope of 838 A/s appears in the current trace of coil 1. The generated magnetic fields are thus already noticeably altered by the eddy currents for the base function in use. The real trajectory of the potential minium during the magnetic transport will therefore differ from the theoretical trajectories y_{V_0}.

This again highlights the benefit of using a machine learning algorithm for optimization. The machine learner optimizes based on the received feedback and learns about the best realizations by indirectly taking into account the difference between the theoretical and real trajectories. For optimization runs with higher frequency components, maximal current slopes above 1 000 A/s occur more often. Although the algorithm can learn about the influence of different frequency ranges, trying out higher modulation frequencies does not seem promising overall and would reduce the efficiency of the optimization process.

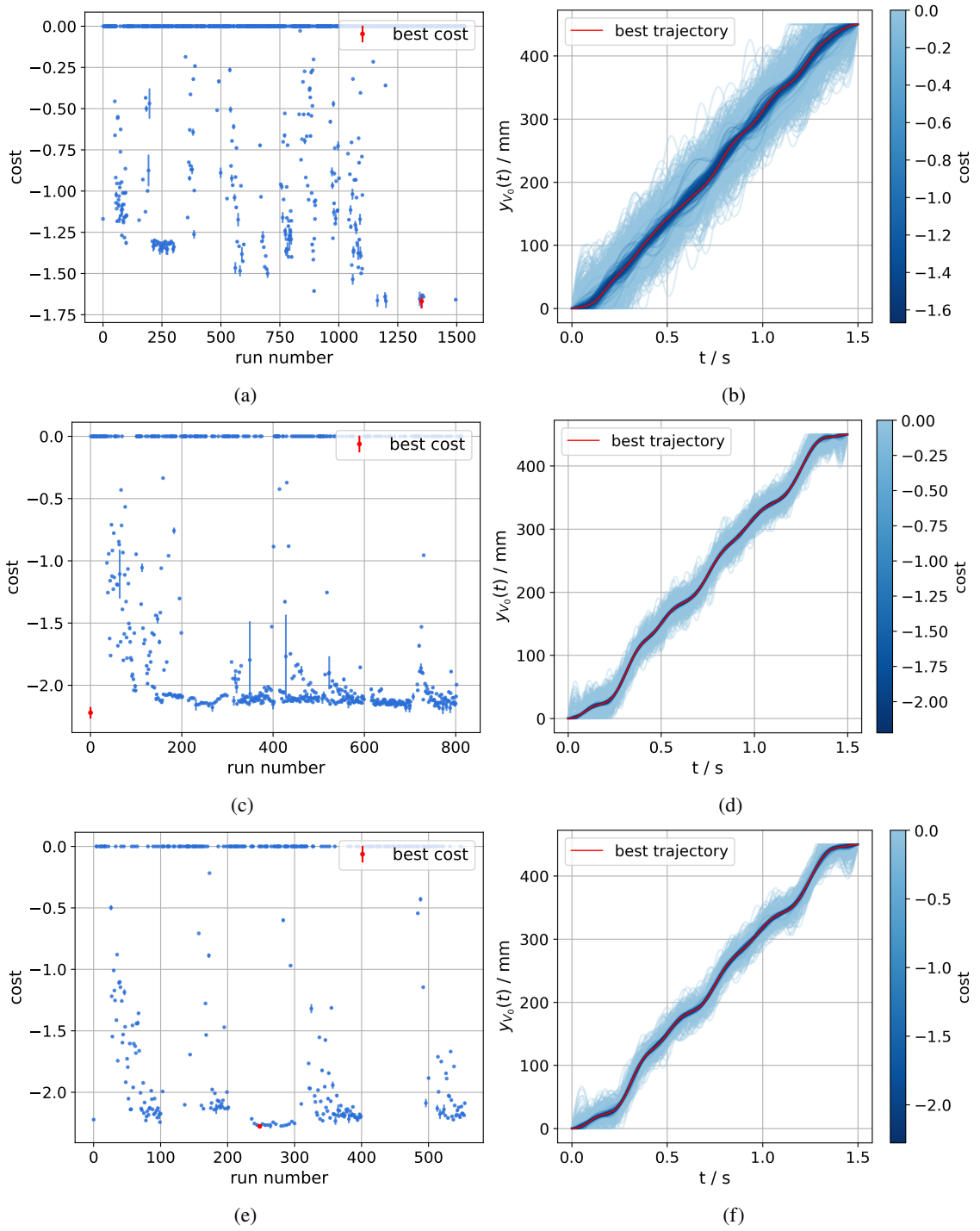


Figure 4.22: Optimization results for the 1.5 s magnetic transport with the parametrization from Eq. (4.3), the cost weight ratio $a : b = 1 : 2$ and different frequency ranges. The left column shows the cost curves. Each run number corresponds to a new parameter set. The right column shows all tested trajectories during the corresponding optimization. The optimizations were conducted with frequency ranges $f = \left\{\frac{k\pi}{1.5s}\right\}$ with **(a)+(b)** $\{k \in \mathbb{N} | 1 \leq k \leq 20\}$ **(c)+(d)** $\{k \in \mathbb{N} | 16 \leq k \leq 25\}$ **(e)+(f)** $\{k \in \mathbb{N} | 11 \leq k \leq 20\}$. The large uncertainties for some of the runs are due to some unknown disturbance in the experiment that sometimes lead to a failed experimental cycle.

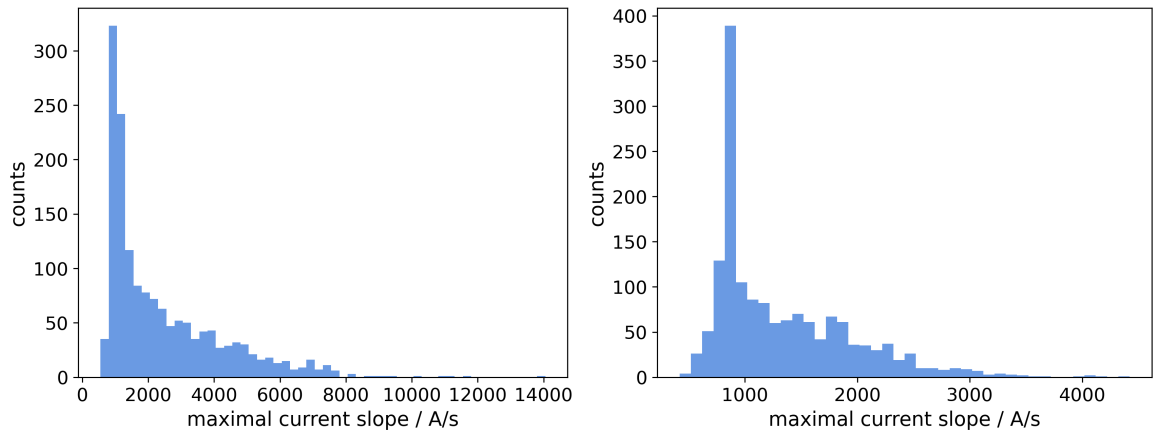


Figure 4.23: Distribution of the maximal current slopes of **(a)** coil pair 1 and **(b)** SC coil for all current traces generated during the 1.5 s magnetic transport optimization run with $a : b = 1 : 2$ and $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$. For each tested parameter set during the optimization, the current traces were generated and the maximal slopes extracted.

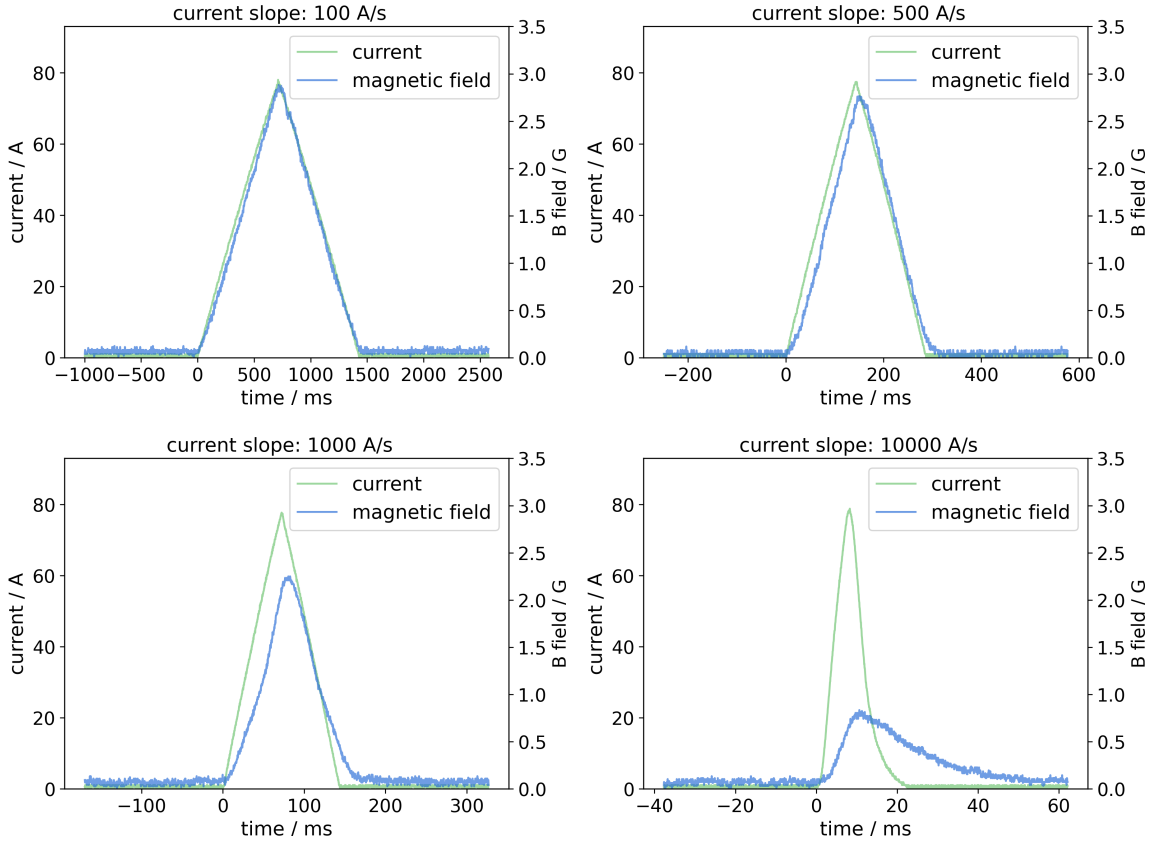


Figure 4.24: Magnetic field response of coil 1 to different applied current slopes. For the measurement a triangular current ramp is generated by the analog ADwin output and applied to coil 1. The current applied to coil 1 was measured with a current clamp (green curve) and the magnetic field generated by coil 1 was measured with a fluxgate magnetic sensor placed above the coil (blue curve). Uncertainties are not displayed for the sake of clarity.

4.3.3.3 Influence of Magnetic Transport Duration

For the previous optimization runs the transport time was fixed at $T_{\mathrm{MT}} = 1.5\,\mathrm{s}$. This section will discuss if a longer magnetic transport time of $T_{\mathrm{MT}} = 2\,\mathrm{s}$ may be more beneficial. The characterization measurements for the 2 s magnetic transport with the error function $y_{\mathrm{err},V_0}(t, 2\,\mathrm{s})$ as parametrization, described by the parameter set $X_{0,2\,\mathrm{s}}$, is shown in Fig. 4.26 and summarized in Table 4.3. The sloshing

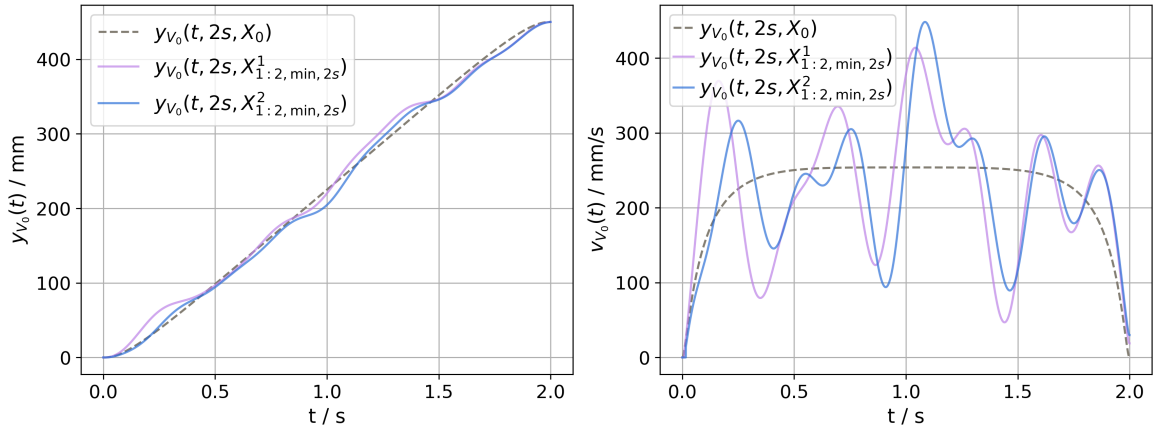
X	η_T	$\frac{\sigma_x(X_{0,T_{\mathrm{MT}}})-\sigma_x(X_{\mathrm{min}})}{\sigma_x(X_{0,T_{\mathrm{MT}}})}$	$\frac{\sigma_y(X_{0,T_{\mathrm{MT}}})-\sigma_y(X_{\mathrm{min}})}{\sigma_y(X_{0,T_{\mathrm{MT}}})}$	$T_x(X_{\mathrm{min}})/\mu\mathrm{K}$	$T_y(X_{\mathrm{min}})/\mu\mathrm{K}$	$S_{y,\mathrm{max}}(X_{\mathrm{min}})/\mathrm{mm}$
$X_{0,2\,\mathrm{s}}$	$(26.9 \pm 2.9)\%$	-	-	136 ± 6	271 ± 19	~ 0.17
$X^1_{1:2,\mathrm{min},2\,\mathrm{s}}$	$(26.5 \pm 2.9)\%$	9.5%	13.3%	134.7 ± 1.4	203 ± 8	~ 0.24
$X_{0,1.5\,\mathrm{s}}$	$(25.0 \pm 2.5)\%$	-	-	117 ± 5	410 ± 70	~ 0.38
$X_{1:2,\mathrm{min},1.5\,\mathrm{s}}$	$(26.5 \pm 2.6)\%$	10.1%	18.1%	120.6 ± 2.0	189 ± 9	~ 0.20

Table 4.3: Comparison of characterization measurements for the 2 s and 1.5 s magnetic transport. $X_{1:2,\mathrm{min},T_{\mathrm{MT}}}$ is the best found parameter set for the optimization run with magnetic transport duration T_{MT}, frequency range $f = \left\{\frac{k\pi}{2\,\mathrm{s}}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and cost weights $a : b = 1 : 2$. $T_{x,y}$ are the temperatures along the x- and y-direction and $S_{y,\mathrm{max}}$ describes the maximal sloshing amplitude along the transport axis, extracted from the characterization measurements. $\frac{\sigma_{x,y}(X_{0,T_{\mathrm{MT}}})-\sigma_{x,y}(X_{\mathrm{min}})}{\sigma_{x,y}(X_{0,T_{\mathrm{MT}}})}$ describes the relative improvement of the cloud widths for the best found trajectories compared to the magnetic transport realization with the initial parameter set $X_{0,T_{\mathrm{MT}}}$, describing the base function y_{err,V_0} for different magnetic transport times T_{MT}.

amplitude is reduced compared to the 1.5 s transport. The temperature in y-direction is also significantly reduced while the temperature for the x-direction is slightly increased. The lower temperature and sloshing in transport direction could be a result of the lower maximal trap potential velocity. For 1.5 s transport the maximal velocity is 339 mm/s, while for the 2 s the trapping potential is accelerated to a maximal velocity of 254 mm/s. The atoms are transported slower along the transport axis, inducing less sloshing when the potential minimum stops in the science chamber. Furthermore, the transport efficiency for the 2 s transport is slightly increased to the transport efficiency for the 1.5 s transport. This shows that the loss of atoms from the trap is not dominated by their finite lifetime, since in this case a longer trapping duration would correspond to fewer atoms remaining in the trap. Therefore, other loss mechanisms which are influenced by the acceleration of the potential along the transport axis must be dominating. In summary, the 2 s long magnetic transport parametrized with the error function performs better than the magnetic transport with the same parametrization but a shorter duration of 1.5 s.

However, the optimized magnetic transport trajectory $y_{V_0}(t, 1.5\,\mathrm{s}, X_{1:2,\mathrm{min}})$ gives comparable results to the 2 s magnetic transport with $y_{\mathrm{err},V_0}(t, 2\,\mathrm{s})$. The transport efficiency is similar, and the temperature is even lower for the optimized 1.5 s transport curve. This shows, that the magnetic transport with a duration of 2 s is not necessarily better.

The results of the machine learning based optimization with $T_{\mathrm{MT}} = 1.5\,\mathrm{s}$ showed that the error function y_{err,V_0} is not the best parametrization for the transport. To see if the performance of the 2 s transport can also be further improved by modifying the error function, the optimization routine was repeated for the longer transport. The same parametrization, as defined in Eq. (4.3), was chosen. The frequency range was set to $f = \left\{\frac{\pi k}{2\,\mathrm{s}}\right\}$, with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$. The cost factors are weighted with $a : b = 1 : 2$. The optimization was repeated two times, to check the reproducibility of the optimization results. The respective cost curves of the optimization are given in Fig. A.3(a) and Fig. A.3(b) in the appendix. The best found transport trajectories with their respective velocity profile can be found in Figs. 4.25(a)



(a) (b)

Figure 4.25: Comparison of the optimization results for the 2 s magnetic transport with frequency range $f = \left\{\frac{k\pi}{2\,\mathrm{s}}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and cost factor weights $a : b = 1 : 2$. **(a)** Best found trajectories during both optimizations, $X^1_{1:2,\mathrm{min},2\,\mathrm{s}}$ and $X^2_{1:2,\mathrm{min},2\,\mathrm{s}}$. **(b)** Velocity profiles for $X^1_{1:2,\mathrm{min},2\,\mathrm{s}}$ and $X^2_{1:2,\mathrm{min},2\,\mathrm{s}}$. The parameter set X_0 describes the error function y_{err,V_0}.

and 4.25(b). Both optimization runs converged to quite similar trajectories with the same cost values within their uncertainties. One curve result gives a cost of $C(X^2_{1:2,\mathrm{min},2\,\mathrm{s}}) = -2.23 \pm 0.06$, while the second curve returns a cost of $C(X^1_{1:2,\mathrm{min},2\,\mathrm{s}}) = -2.129 \pm 0.011$. Thus, both curves are a good realization of the magnetic transport and can further reduce the cost function compared to the error function y_{err,V_0}.

The trajectories overlap almost perfectly for the last 0.5 s. Similar to the results of the 1.5 s transport, the end of the trajectory is therefore the most crucial part. During the beginning and middle of the transport the points of acceleration and deceleration are slightly shifted in time and magnitude but follow a similar shape. When comparing the curves to the error function, it is striking that they show a steeper S-curve shape at the transport end with no plateau. For the 1.5 s optimization run a flat S-curve like shape seemed to be a cause for a lower measured T_y value. Since the 2 s error function already leads to lower T_y values, this may have less of an impact for this optimization.

The characterization measurements for $y_{V_0}(t, 2\,\mathrm{s}, X^1_{1:2,\mathrm{min}})$ (see Figs. 4.26(d) to 4.26(f)) lead to a T_y of (203 ± 8) µK. This is comparable, within the uncertainties, to the temperatures reached by the optimized 1.5 s trajectory with $X_{1:2,\mathrm{min},1.5\,\mathrm{s}}$. However, the sloshing amplitude for $X^1_{1:2,\mathrm{min},2\,\mathrm{s}}$ is increased compared to the 2 s error function. This is presumably due to the fact that the velocity of the potential minimum is not smoothly decreased. This can be seen in Fig. 4.25(b), where the velocity curve does not reach zero at the end of the transport. The potential minimum is thus stopped abruptly at the end, leading to a larger sloshing of the atoms. Furthermore, it was not possible to improve the atom number with the machine learning optimization. Since the algorithm converged twice to quite similar trajectories, it seems that it was not possible to find a better parametrization with the given configuration.

One can thus conclude that the error function for the 2 s magnetic transport, is better suited for the transport than the same parametrization for a shorter transport duration. It also seems to be closer to the achievable optimum. The machine learning optimization could only further improve the temperature value along the transport axis T_y. However, the simple error function $y_{\mathrm{err},V_0}(t, 2\,\mathrm{s})$ and the trajectory $y_{V_0}(t, 2\,\mathrm{s}, X_{1:2,\mathrm{min}})$ found by the machine learner could not further improve the results gained with a

magnetic transport duration of 1.5 s. Since it is beneficial to have a shorter total experimental cycle, the optimized 1.5 s magnetic transport is a better choice for the experiment.

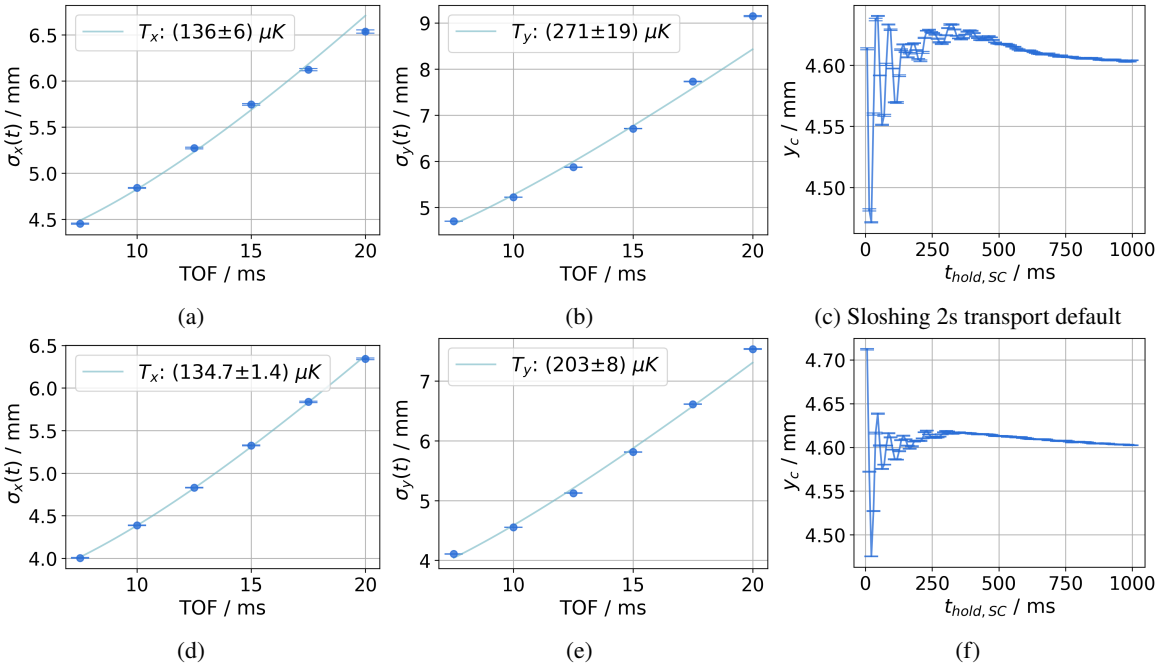


Figure 4.26: Characterization measurements quantifying the performance of the 2 s magnetic transport with **(a)-(c)** $y_{\mathrm{err},V_0}(t, 2\,\mathrm{s})$ and **(d)-(f)** $y_{V_0}(t, 2\,\mathrm{s}, X^1_{1:2,\mathrm{min}})$. The characterization measurements consist of a TOF measurement in the science chamber and the measurement of the displacement of the cloud center y_c along the transport axis for different holding times $t_{\mathrm{hold,\,SC}}$ in the SC trap after the magnetic transport. $\sigma_{x,y}$ describes the cloud width along the x- and y-axis. The temperatures are extracted by fitting Eq. (2.1) to the data. The sloshing measurements are averaged over 10 loops and the TOF measurements are averaged over 20 loops.

CHAPTER 5

Characterization of Rydberg Ion Detection

The successful Rydberg excitation can be verified by field ionizing the Rydberg atoms and detecting the generated ions with a microchannel plate (MCP). The experimental setup and sequence for Rydberg ionization and subsequent detection were introduced in Section 2.3. In Section 5.1 the ion detection setup is described in more detail. After this the characterization of the MCP is discussed in Section 5.2. The ion detection characterization was done with the aim to reliably determine the detection efficiency of the setup.

5.1 Ion Detection Setup

A Rydberg ion generated in the science chamber by field ionization is detected with the Microchannel plate (F4655-11 from Hamamatsu Photonics K.K. [68]). The setup for ionization and subsequent detection via the MCP is shown in Fig. 2.14. The MCP is a two-dimensional, multi-channel electron multiplier which can detect single ions, electrons, UV rays, X-rays and gamma rays in vacuum [68]. The working principle of the MCP is shown in Fig. 5.1. The MCP contains several channels placed next to each other on a circular plate. Each channel acts as an independent electron multiplier. An ion entering a channel causes the emission of secondary electrons. These electrons are accelerated due to the potential

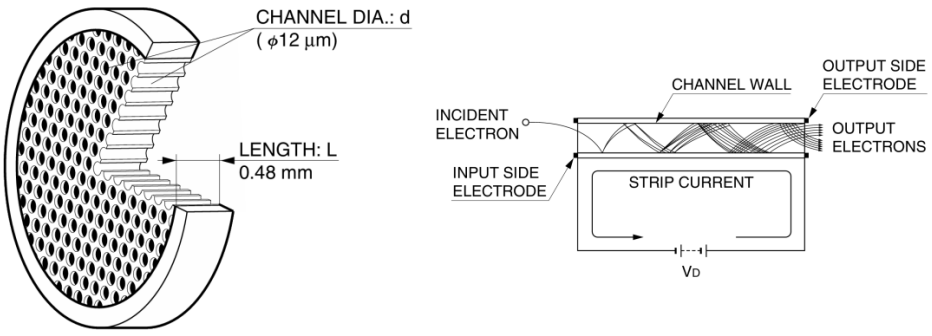


Figure 5.1: Working principle of the microchannel plate (MCP) used for the detection of the ion signals. The left image shows the MCP, consisting of several channels placed next to each other on a circular plane. The right image shows the amplification of an incident electron in a single channel. The images are taken from [68].

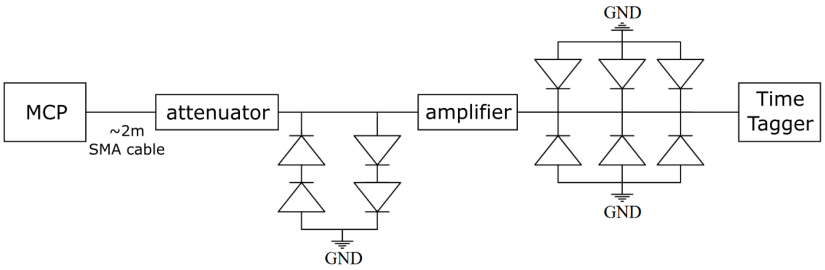


Figure 5.2: Signal path of the MCP signal to the Time Tagger. The MCP signals are attenuated and then pass a first stage of protection diodes, used to clamp positive and negative voltage spikes. After being amplified they pass the second protection diodes, which protect the Time Tagger from negative voltage spikes.

gradient between the channel input and output side generated by the applied voltage V_D. While travelling through the channel, the electrons hit the channel wall repeatedly and produce more secondary electrons leading to an electron cascade. The amplified signal is then read out by an anode placed behind the MCP [98]. Fig. 5.1 shows the working principle of a single channel and indicates that the channels are perpendicular to the plate surface. In reality, the channels are angled with a bias angle that optimizes the detection efficiency. The MCP used in the experiment is a two-stage MCP. It consists of two MCP units placed behind each other. Their channel axes are enclosing an angle corresponding to their bias angle. A two-stage MCP can minimize the signal-to-noise ratio compared to a single-stage MCP. The channels of the MCP have diameters of 12 µm and the maximal supply voltage is 2 kV [68].

In the experimental setup, the readout signal from the anode is connected to the Time Tagger. Ions are detected by measuring voltage spikes in the anode signal resulting from the electron cascade. Therefore, the voltage spikes have negative amplitudes. As the voltage spikes may exceed the damage threshold of the Time Tagger, the MCP signals cannot be sent directly to the Time Tagger. The Time Tagger takes as input signals between −0.3 V and 5 V. It is recommended that the input signal is within the range of 0 V − 3 V. The Time Tagger trigger level can be set to values between 0 V − 2.5 V.

To prevent damaging the Time Tagger a protection circuit (see Fig. 5.2), built by Cedric Wind, is inserted between the MCP and the Time Tagger. The first set of protection diodes (1SS315(TPH3,F) from Toshiba) clamp positive and negative voltage spikes, to protect the RF amplifier (ZX60-6013E-S+ from Mini-Circuit). The amplifier inverts the negative MCP signals and amplifies them. The amplified signals then pass a second set of 6 protection diodes. This protects the Time Tagger from voltage undershoots that exceed the allowed voltage range.

For signals outside the amplifier's bandwidth, impedance matching is not ensured anymore what leads to reflections at the amplifier. The reflected signal travels back to the MCP along the ∼ 2 m long SMA cable, is reflected a second time at the MCP due to impedance mismatch and thus returns to the protection circuit. The signal can thus be detected a second time by the Time Tagger. With a typical relative velocity of propagation of around 69% in such an SMA cable and an additional propagation length around ∼ 4 m, the back reflected signal is expected to be delayed by ∼ 20 ns. An exemplary MCP signal triggered by a Rydberg ion and passing the protection circuit is shown in Fig. 5.3. The signal was measured with an oscilloscope which is connected to the end of the protection circuit instead of the Time Tagger. The signal resulting from a real ion count is the one that occurs at 0 ns. The signal around 25 ns has a significantly lower positive amplitude and causes an undershooting. The delay time is of the same order as the expected delay time of a back reflected signal. It can therefore be assumed

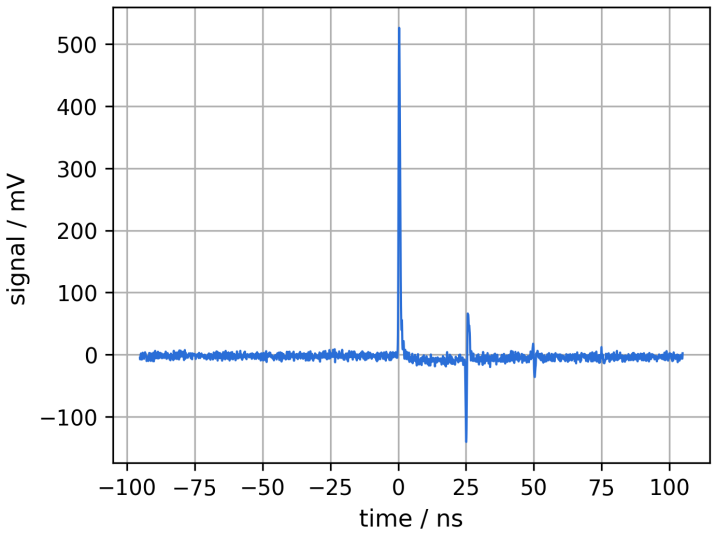


Figure 5.3: Example MCP signal triggered by a Rydberg ion. The signal was measured with the setup shown in Fig. 5.2, but instead of using the Time tagger to count the signals, an oscilloscope was connected to measure the analog signal. The attenuator in the protection circuit was removed for this measurement. The signal at 0 ns is the real ion signal. The back reflection of the ion signal is measured with a delay of 25 ns.

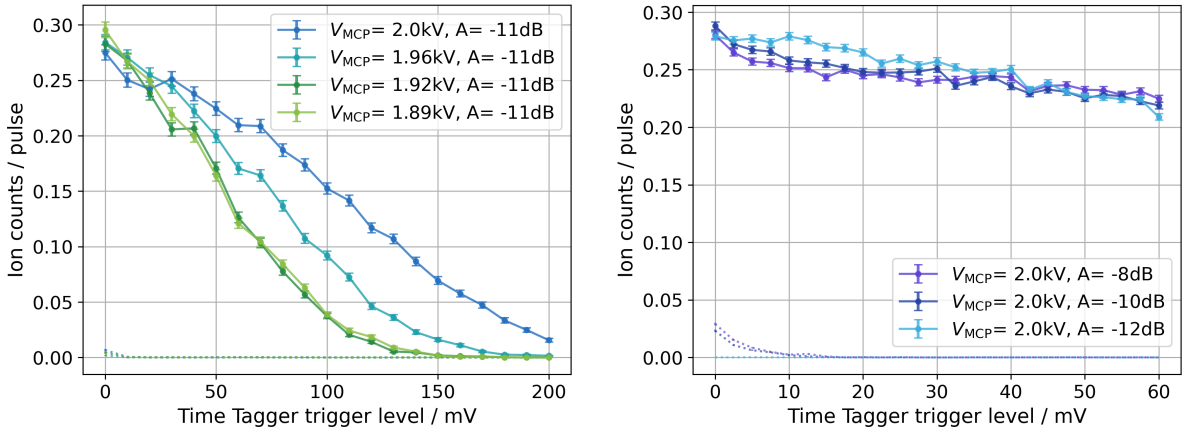
that the second signal is triggered by the back reflection of the real ion signal. A third signal with an even lower amplitude is detected around 50 ns and probably arises from a second back reflection process. These back reflections were observed for the majority of the ion signals during the measurement with the oscilloscope. These signals can contribute to fake counts, which can falsify the Time Tagger's counting statistics. Furthermore, some of the reflected signals featured an undershooting signal with an amplitude smaller than −0.3 V. To suppress the back reflected signal an attenuator is inserted before the protection circuit, as can be seen in Fig. 5.2. Since the back reflected signal passes the attenuator two more times compared to the real ion signal, it gets further suppressed.

The diode protection circuit was tested and characterized by Cedric Wind. The influence of the first amplifier stage on the signal amplitudes and subsequent data acquisition with the Time Tagger was analyzed in this thesis and will be discussed in the following.

5.2 Characterization Measurements

Fig. 5.3 shows the output signal from the MCP anode after the protection circuit. From the figure it is apparent that the measured ion counts stem from real ions and from back reflection spikes. The number of counts depends on different system parameters of the detection setup and the signal processing. The Time Tagger trigger voltage determines how many voltage spikes are counted, and whether noise gets filtered out. The MCP supply voltage affects the internal gain of the ion signal, thereby influencing the amplitude of the MCP signal [68]. Additionally, the choice of MCP signal attenuation determines the overall height of the MCP signal as well as the relative height between the main peak and the back reflection peaks. Therefore, in the next step, the number of detection events is analyzed for different settings of these system parameters.

To do so the Rydberg excitation and detection cycle was started and the Time Tagger trigger level was

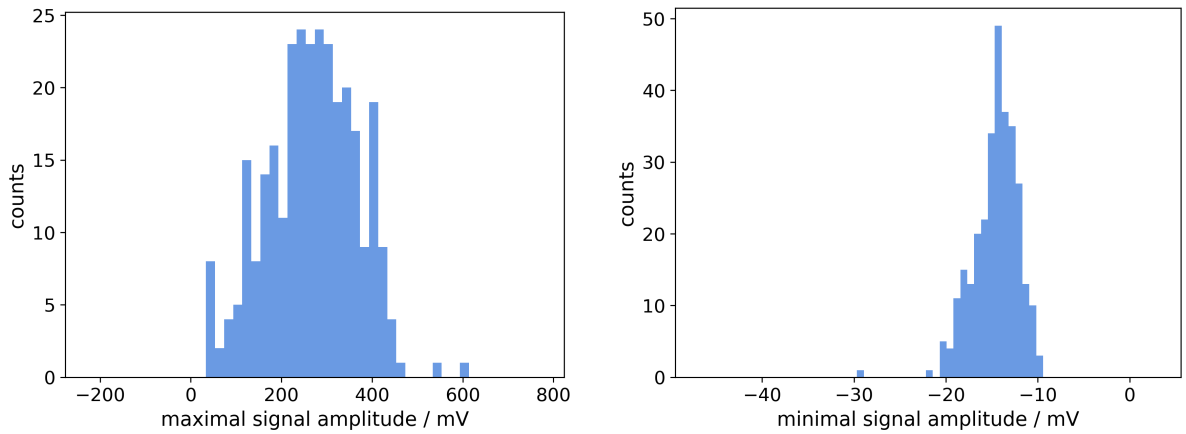


(a) (b)

Figure 5.4: Ion counts per Rydberg excitation pulse, measured with the Time tagger. For the measurement the sequence described in Fig. 2.15 was started, and the Time tagger trigger level was scanned within it. The dashed line shows the background signal, measured during the second 1000 pulses in the Rydberg sequence. **(a)** Measurement results with different MCP voltages V_{MCP}. For all measurements an attenuator with attenuation factor $A = -11\,\mathrm{dB}$ was used. The datapoints are averaged over 6000 pulses. **(b)** Measurement results with the MCP voltage set to $V_{\mathrm{MCP}} = 2\,\mathrm{kV}$ and different attenuators A. The datapoints are averaged over $28 \cdot 10^3$ pulses.

scanned during the sequence. The results are shown in Fig. 5.4(a). The measurement window of the Time Tagger is indicated in Fig. 2.15 by the duration of the Time Tagger trigger. The control laser was detuned by $\Delta_C = -45\,\mathrm{MHz}$ and the probe laser was detuned by $\Delta_P = 44.7\,\mathrm{MHz}$ for the measurement, to drive the off resonant two-photon Rydberg excitation. The ionization voltage was set to 490 V and the deflection voltage to 184 V during the ionization sequence. The measurement was repeated with different MCP supply voltage and a $-11\,\mathrm{dB}$ attenuator. The solid datapoints describe the averaged number of ion counts measured per pulse. One pulse corresponds to one Rydberg excitation pulse and subsequent detection of the Rydberg ion with the MCP. The dashed lines correspond to the averaged background counts measured per pulse when no atoms are present anymore. The datapoints were averaged over 6000 measurements. As expected, the counting rate decreases with increasing trigger level. For trigger levels above $\sim 30\,\mathrm{mV}$ a higher supply voltage results in a larger number of counts. As mentioned before, a higher MCP supply voltage causes an increase of the signal gain in the MCP [68]. Thus, small signals can still be amplified enough such that they lie above the trigger level and are detected by the Time Tagger. However, the smaller the chosen threshold the less difference an increased signal gain makes and thus all curves in Fig. 5.4(a) approach a value between 0.28 and 0.29 counts per pulse. For trigger thresholds below 10 mV the noise is not filtered out anymore, as indicated by the small increase in background counts. The MCP supply voltage does not significantly influence the noise contribution. As the ion counts do not saturate at higher MCP voltages and the noise does not increase significantly with the supply voltage, it was decided to operate at the maximum MCP supply voltage of 2.0 kV to maximize the detection efficiency.

In the next step, the influence of different attenuators is analyzed to further improve the signal-to-noise ratio. The trigger level scan was repeated for a smaller scan range with a MCP supply voltage of 2.0 kV. Three measurements with the different attenuators $A = \{-8\,\mathrm{dB}, -10\,\mathrm{dB}, -12\,\mathrm{dB}\}$ are shown in Fig. 5.4(b). The count rates for all measurement decrease similarly as the trigger threshold is lowered.



(a) (b)

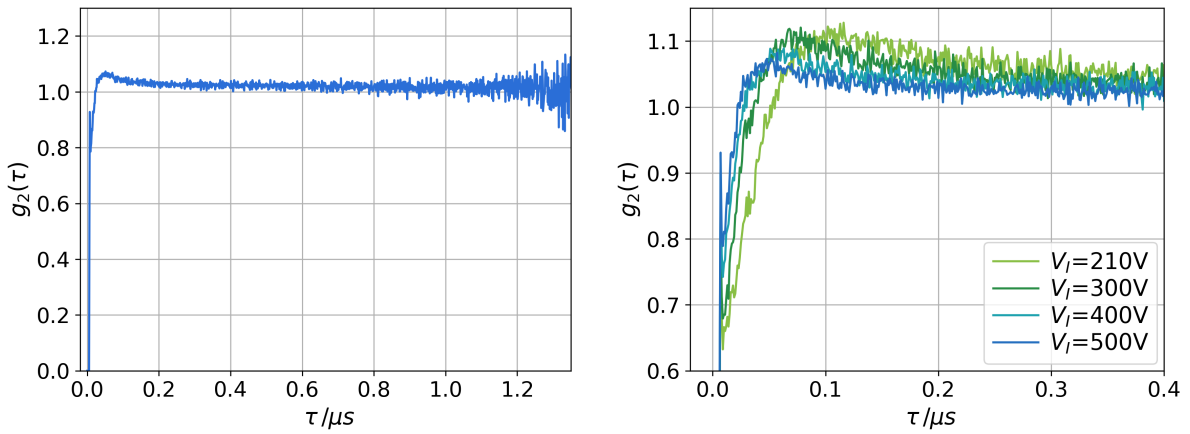
Figure 5.5: The maximal **(a)** and minimal **(b)** amplitudes of the MCP signals measured after the protection circuit with an 12 dB attenuator and an MCP voltage of 2 kV. The trigger level of the oscilloscope was set to 25 mV. The maximal and minimal amplitudes correspond to the amplitudes measured in a time window of 100 ns after the trigger signal. 200 events were triggered during the measurement time.

The three cases mainly differ for trigger levels below 10 mV. In this regime the noise contributes with up to 0.025 counts per pulse, for attenuation factors of $A = -8$ dB and $A = -10$ dB. However, for higher attenuation the noise is sufficiently suppressed, without decreasing the real ion counts per pulse. Thus, the best signal-to-noise ratio is achieved with an MCP supply voltage of 2 kV and a -12 dB attenuator before the protection circuit. A Time Tagger trigger level of 10 mV seems suitable for these settings, since the number of counts does not increase significantly at lower trigger thresholds.

With the discussed parameters for gain, threshold and attenuation a histogram of maximal and minimal signal amplitude behind the protection circuit was measured with an oscilloscope (WavePro 735Zi Oscilloscope from TELEDYNE LECROY [99]). The results are shown in Fig. 5.5. The trigger threshold of the oscilloscope for starting the measurement was set to 25 mV, since lower thresholds were not possible. The maximum and minimum amplitudes correspond to the maximum and minimum amplitudes measured during a time interval of 100 ns after the trigger, including the trigger signal itself. The measurements show that, the protection circuit combined with a -12 dB attenuator and an MCP voltage of 2 kV does not cause the signal amplitudes to exceed the Time Tagger's damage threshold, which ranges from -0.3 V to 5 V. This configuration thus fulfills the requirements for protecting the Time Tagger and at the same time maximize the detection efficiency.

However, the characterization does not yet rule out the occurrence of fake counts due to back reflections. To further analyze possible correlations in the ion signals, the $g_2(\tau)$ function for the time dependent Time Tagger signal $e(t)$ is measured. The Time Tagger counts the number of pulses received during a set time window after being triggered, with a time resolution of 60 ps. It stores the counts measured during the time window in a histogram, enabling time-resolved measurement. $e(t)$ describes the number of counts measured in time bin t. The starting point of the measurement with the Time Tagger is set to $t = 0$. The $g_2(\tau)$ function, inspired by the photon second-order correlation function [100], is defined as

$$g_2(\tau) = \frac{\sum_{t,t'|\tau=t-t'} \langle e(t) \cdot e(t') \rangle_\mathrm{P}}{\sum_{t,t'|\tau=t-t'} \langle e(t) \rangle_\mathrm{P} \cdot \langle e(t') \rangle_\mathrm{P}}, \tag{5.1}$$



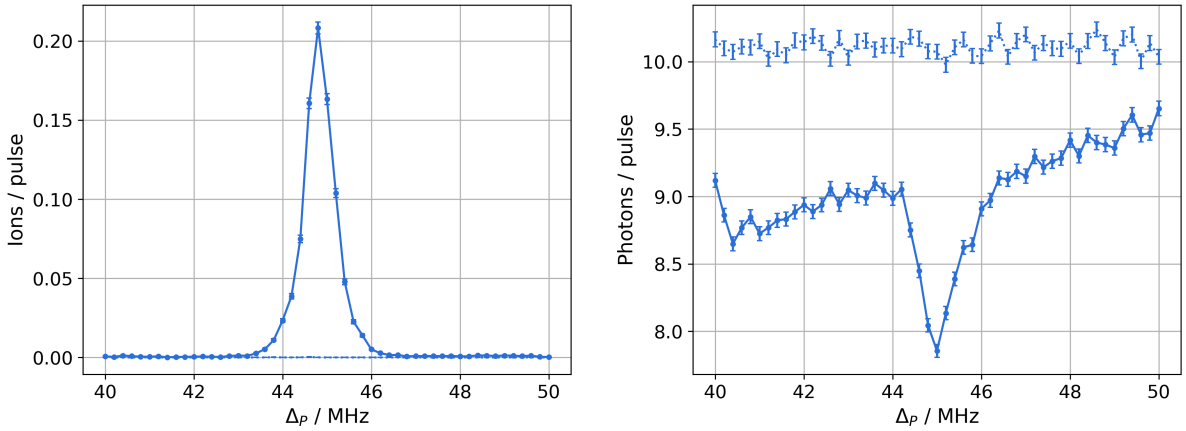
(a) (b)

Figure 5.6: Measurements of the $g_2(\tau)$-function (see Eq. (5.1)) for the time dependent Time Tagger signal. **(a)** Result for a measurement with an ionization voltage of $V_\mathrm{I} = 500\,\mathrm{V}$. The datapoints are averaged over $360 \cdot 10^3$ Rydberg pulses. **(b)** Close up of the g_2 function for different ionization voltages V_I. The datapoints are averaged over $240 \cdot 10^3$ Rydberg pulses.

where $\langle\cdot\rangle_\mathrm{P}$ describes the average over all pulses during the measurement sequence. The $g_2(\tau)$ function can be interpreted as the conditional probability of detecting two ions with a time difference of τ, normalized by the probability of detecting them independently at their respective times. If ions arrive independently of each other with a time delay of τ, then $g_2(\tau) = 1$. When $g_2(\tau) > 1$ the probability of detecting two ions with a time difference of τ is increased, whereas when $g_2(\tau) < 1$ it is decreased. By definition $g_2(\tau = 0) = \frac{\sum_t \mathrm{Var}[e(t)] + \langle e(t)\rangle_\mathrm{P}^2}{\sum_t \langle e(t)^2\rangle_\mathrm{P}} \gg 1$, since the variance $\mathrm{Var}[e(t)] > 0$. The $g_2(\tau)$ function for a measurement with $V_\mathrm{MCP} = 2\,\mathrm{kV}$, a $-12\,\mathrm{dB}$ attenuator and a Time Tagger threshold of 10 mV is shown in Fig. 5.6(a). The measurement was repeated for $360 \cdot 10^3$ Rydberg excitation/detection pulses. A potential of 500 V was applied to the ionization electrode, and 185 V to the deflection electrode.

The g_2-function is 0 for times smaller than 6 ns. This is expected, since the Time Tagger has a dead time of 6 ns. Therefore, ions arriving with a time delay shorter than this can not be detected. At $\tau = 7\,\mathrm{ns}$ the g_2-function features a peak. Fig. 5.6(b) shows a close up of the g_2 function, for different applied ionization voltages, where the peak is more clearly visible. When analyzing the analog MCP signals on an oscilloscope, it was never observed that two signals arrived with a time delay of $\sim 7\,\mathrm{ns}$. Therefore, it is not expected that this peak is due to back reflections. The measurement was repeated with different trigger levels, but the peak did not disappear. Furthermore, a similar peak was observed in the MCP signal of the 'Ytterbium Quantum Optics' (YQO) laboratory of the NQO research group [101]. This experiment uses a similar MCP and the same Time Tagger. It is therefore assumed that the peak may be caused by the Time Tagger itself and is possibly a consequence of dead time induced artifacts.

In the time interval $7\,\mathrm{ns} < \tau < 50\,\mathrm{ns}$ the g_2 value is slowly increasing, crosses 1 around 25 ns and reaches a maximum around 50 ns with $g_2(50\,\mathrm{ns}) > 1$. For longer delay times the g_2 function decreases again and slowly approaches 1. The reason for this could be that ions created close together repel each other, causing the ions to arrive with a time delay τ'. Thus, when detecting one ion at the MCP it is less probable to detect a second ion shortly afterward, leading to $g_2(\tau < \tau') < 1$. The second ion will arrive with a higher probability after τ', leading to $g_2(\tau \approx \tau') > 1$. The delay time τ' depends on the ions



(a) (b)

Figure 5.7: Results for a Rydberg sequence with control detuning set to $\Delta_C = -45\,\text{MHz}$. **(a)** Number of ions per pulse detected with the MCP. **(b)** Calibrated number of photons measured with the SPCM. These measurements are used for extracting the detection efficiency of the ion detection setup. The datapoints are averaged over $16 \cdot 10^3$ pulses.

initial positions and the interaction time. The latter can be influenced by changing the ionization voltage V_I. A lower ionization voltage corresponds to a lower acceleration of the ions and thus leads to a longer time of flight to the MCP. This results into a longer interaction time between ions and presumably a larger delay τ' between ions created close to each other. Fig. 5.6(b) shows the g_2 function measured for four different ionization voltages between 500 V and 210 V. The deflection voltages had to be adapted accordingly for each ionization voltage to ensure that the ions are still deflected towards the MCP. The point where the g_2 function crosses the 1 is shifted to larger times for lower ionization voltages. This is expected, as it corresponds to a larger delay time τ'. The same holds for the point where g_2 reaches its maximum. This measurement is an indication that the assumption about the increase in g_2 up to a value above 1 could result from the Coulomb repulsion between the ions. However, for a more quantitative comparison the experimentally measured g_2-function should be compared to simulations of the ion trajectories while taking into account interactions between the different ions.

Furthermore, it must be noted that no additional peak occurs around $\tau \approx 25\,\text{ns}$. This would be expected if the back reflected signals would still contribute to the counting statistic of the Time Tagger. To ensure that a possible peak around $\tau \approx 25\,\text{ns}$ is not hidden under the correlation bump due to the Coulomb repulsion, the measurement was repeated with a 4 m SMA cable instead of 2 m long cable. Back reflected signals would now occur after $\sim 50\,\text{ns}$. However, the measurement did not show an additional increase in the g_2 function around 50 ns. Thus, one can conclude that the chosen configuration for the ion detection successfully suppresses fake counts caused by back reflections.

In the next step, the detection efficiency of the ion detection setup can be determined. This is done by comparing the probe transmission signal, measured with the SPCM, with the ion signal for off resonant two-photon Rydberg excitation. Fig. 5.7(b) shows the probe transmission signal with the control detuning set to $\Delta_C = -45\,\text{MHz}$, averaged over $16 \cdot 10^3$ pulses. The photon number is calibrated by taking into account the fiber coupling efficiencies to the SPCM and the quantum efficiency of the counter. The probe transmission dip at $\Delta_P = 45\,\text{MHz}$ results from the off resonant Rydberg excitation. The dip lies inside the broad probe transmission valley. A full transmission valley measured with the given setup can

be found in Julia Gamper's Master's thesis [40]. Because of this, the photon number on both sides of the dip is smaller than for the reference measurement, shown by the dotted lines. At $\Delta_P = 45\,\text{MHz}$ the photon number per pulse is reduced by additional $N_{\text{Ph},\Delta_P=45\,\text{MHz}} = 1.24 \pm 0.09$ photons, compared to measurements with control light turned off. Thus, approximately 1.24 Rydberg atoms are created per pulse. However, only $N_{\text{I},\Delta_P=45\,\text{MHz}} = 0.208 \pm 0.004$ signals are measured by the Time Tagger per pulse at $\Delta_P = 45\,\text{MHz}$. This results in a detection efficiency of $D = \frac{N_{\text{I},\Delta_P=45\,\text{MHz}}}{N_{\text{Ph},\Delta_P=45\,\text{MHz}}} = 0.168 \pm 0.012$ of the ion detection setup composed of MCP, protection circuit and Time Tagger. 16.8% of the Rydberg atoms created inside the science chamber can thus be detected with the current ion detection setup.

The measured detection efficiency D includes the ionization probability, the efficiency for deflecting the ions towards the MCP, the MCP detection efficiency and the efficiency to convert MCP signals into Time Tagger counts. The detection efficiency of the MCP itself is only given for other ion species than Rubidium and is ranging between 4% and 85% for these species [98]. Therefore, it is difficult to ascertain whether the determined detection efficiency falls within an appropriate range. In the YQO experiment, with a similar MCP and Time Tagger setup, a detection efficiency of around 57% was achieved for the detection of Ytterbium ions. They also observed that the MCP detection efficiency differs across the active area of the MCP [101]. Since no additional steering electrodes are implemented in the HQO room temperature ion detection setup, it is not possible to optimize the detection with respect to that. However, the design for the ionization inside the cryogenic environment will feature additional electrodes. Thus, further characterization and optimization of the ion detection must be conducted with the future setup.

CHAPTER 6

Conclusion

In this thesis, a machine learning based optimization method was implemented into the HQO experiment and was applied for the optimization of the MOT and the magnetic transport. Additionally, the detection of the Rydberg atoms through ionization and subsequent detection via a microchannel plate was characterized.

The implemented machine learning online optimization routine is based on the M-LOOP Python package [21]. In order to apply the optimization algorithms, implemented in the package, the internal machine learning cycle and the experiment cycle of the HQO experiment had to be coordinated. Communication between the machine learner controller and the experiment control was achieved by introducing an additional interfacing layer, to which both controllers have access. The aim of the information exchange between both controllers is to continuously generate new data points during the optimization process. Each new datapoint consists of a parameter set X, which describes the new optimization parameter values generated by the machine learner, and a respective cost value $C(X)$, which is produced by the experiment. The input parameters X are transferred to the experiment by allowing the machine learner controller to overwrite parameter files, which are used by the experiment control to build a new model of the experimental sequence. The corresponding cost value is extracted from the absorption images taken at the end of the respective experimental run, to characterize its performance. The cost value is conveyed to the machine learner by saving the absorption images in a database accessible to the machine learner controller. This made it possible to successfully combine the machine learning cycle and the experimental cycle, achieving online optimization.

The optimization routine was first tested for the MOT optimization, describing a simple optimization problem with the aim to maximize the number of trapped atoms. The parameters that were optimized included the frequency and power of the Cooler and Repumper lasers, and the MOT magnetic field gradient. The optimization process was conducted for the two machine learning algorithms implemented in the M-LOOP package: the neural network and the Gaussian process. Both algorithms returned results that are similar to the one found during manual optimization. However, the Gaussian process converged more quickly than the neural network.

In the next step the machine learning online optimization was applied to the optimization of the magnetic transport in the HQO experiment. The transport connects the MOT chamber, where the atoms are loaded from Rubidium background gas and initially cooled down, with the Science chamber, which will host the atom chip. The magnetic transport is realized by trapping the atoms in a magnetic field and displacing the trapping potential along the transport axis. The implementation of the magnetic

transport in the experiment allows controlling the transport by defining a time dependent trajectory for the magnetic trap center along the transport axis. Therefore, the optimization of the magnetic transport can be formulated as the task of finding an optimal time profile for the potential minimum trajectory. The parametrization of the potential minimum trajectory for the optimization process was built by defining a base function, adapted from [55], and enabling the machine learner to add modulations in the form of a finite sine series. Furthermore, a cost function had to be defined to characterize the performance of the experiment with respect to the optimization task. The magnetic transport optimization had two goals, namely to increase the transfer efficiency η_{T} from MOT chamber to science chamber, and secondly to decrease heating effects and sloshing at the end of the transport. The transfer efficiency can be quantified by extracting the atom number from the absorption images taken shortly after the transport. The atom cloud widths, extracted from the same absorption images, can be used as a measure for the sloshing and temperature of the atom cloud after the transport. The cost function was then defined as the product of the scaled atom number and the inverse of the scaled cloud widths. Furthermore, the weighting factors a and b, for atom number cost factor and cloud width cost factor respectively, were introduced. This allows to control the contribution of each of the two cost factors to the total cost independently.

The influence of different weighting factors on the optimization process was analyzed. For balanced weight factors $a = b = 1$, the optimizer found a new potential minimum trajectory with the highest transfer efficiency of $\eta_{\mathrm{T}} = 30\%$, what corresponds to $3 \cdot 10^8$ atoms arriving in the science chamber. Giving more weight to the cost factor of the cloud width, by choosing $a : b = 1 : 4$, could reduce the temperature fit value T_{y} for the y-axis by a factor of ~ 3.2, compared to the base function. However, the best trajectory found in this optimization run led to 80% more atom loss. An additional atom loss mechanism may be introduced by larger acceleration phases during this realization, what was analyzed with a simplified 1D Monte Carlo simulation. However, due to the complexity of the transport system, the cause for the observed atom loss could not be determined with certainty. The results gained during an optimization run with weight factor ratio $1 : 2$ are a good trade-off, resulting in a transfer efficiency of $\eta_T = (26.5 \pm 2.6)\%$, 50% reduced maximal sloshing amplitude at the transport end and less heating effects leading to $T_x = (120.6 \pm 2.0)\,\mu\mathrm{K}$ and $T_y = (189 \pm 9)\,\mu\mathrm{K}$. Furthermore, different frequency ranges for the modulation terms of the finite sine series and a longer transport time were tested. However, this did not lead to further significant improvement.

Furthermore, the Rydberg ion detection setup was characterized. The Rydberg ions are measured via an MCP, and the signals are processed and counted with a streaming time-to-digital converter, called Time Tagger. The setup was analyzed with respect to the MCP supply voltage, the Time Tagger trigger threshold and different attenuators inserted in the signal path between MCP and Time Tagger. After the characterization, a detection efficiency of $(16.8 \pm 1.2)\%$ was determined.

In the next steps, the successfully integrated machine learning online optimization routine can be easily adapted for the optimization of different parts in the experimental sequence. This will significantly reduce the time needed for optimizations in the HQO experiment and make them more efficient. One application that could benefit from machine learning based optimization is the loading sequence for transferring atoms from the last quadrupole trap to the Z-wire trap on the atom chip. The transfer sequence was already simulated in Leon Sadowski's Master's thesis [60]. However, using the direct feedback from the experiment to modify the results gained from the simulation could lead to further improvements in the transfer process.

APPENDIX A

Magnetic Transport Optimization Measurements

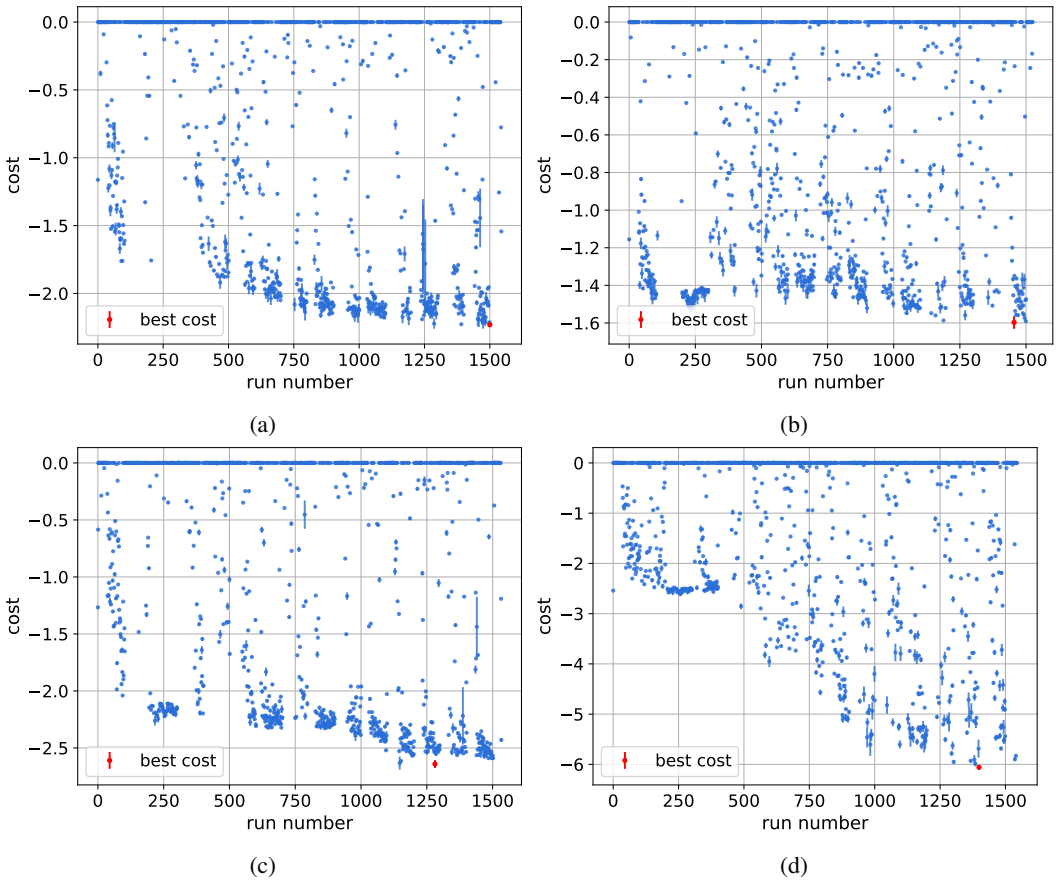
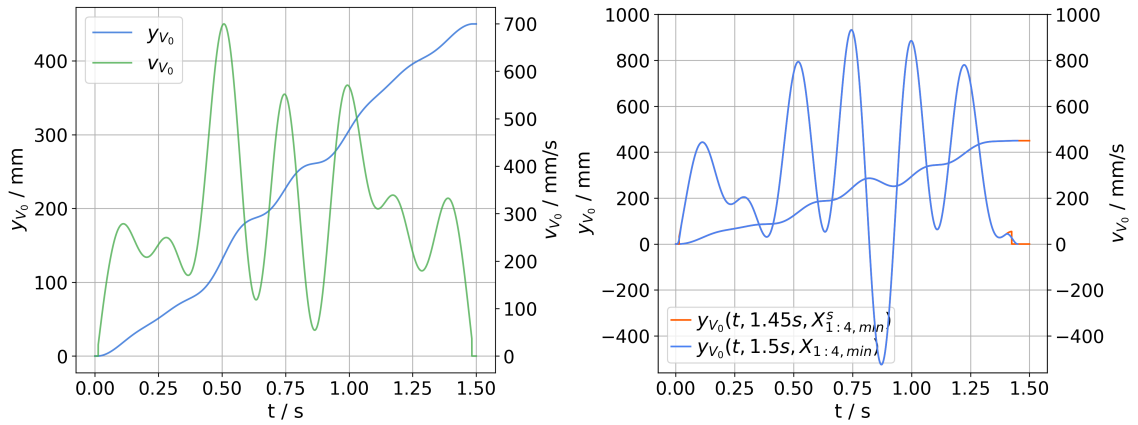
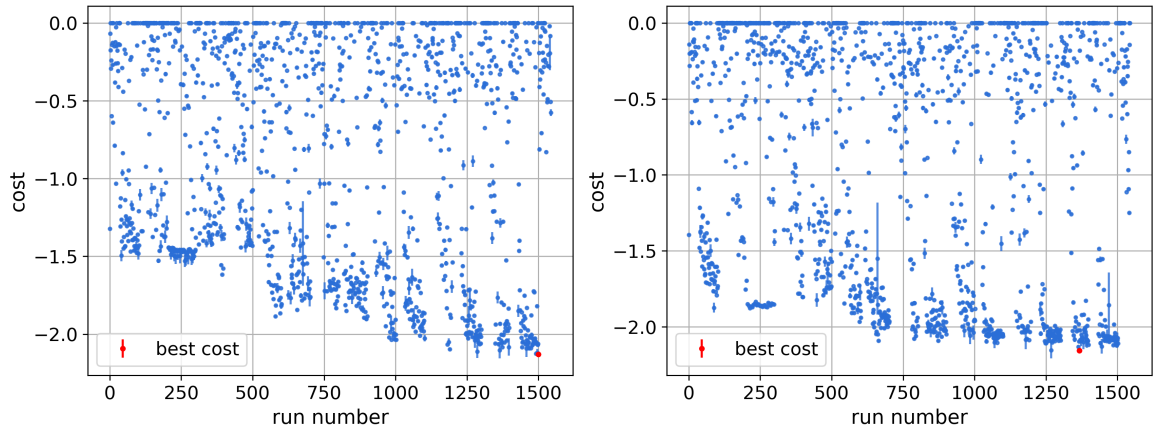


Figure A.1: Results for different 1.5 s magnetic transport optimization processes with $f = \left\{\frac{k\pi}{2s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and different cost weight ratios $a : b$. Optimization with y_{err,V_0} as base function and **(a)** $a : b = 1 : 2$, **(b)** $a : b = 1 : 1$, **(c)** $a : b = 1 : 4$. **(c)** Optimization with the best found parametrization from optimization run (c) ($X_{1:4,\mathrm{min}}$) as base function and $a : b = 1 : 4$. The large uncertainties for some of the runs are due to some unknown disturbance in the experiment that sometimes lead to a failed experimental cycle.



(a) (b)

Figure A.2: **(a)** Optimization results for the 1.5 s magnetic transport with y_{err,V_0} as base function, $f = \left\{\frac{k\pi}{1.5\,\mathrm{s}}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and $a : b = 1 : 4$. This is the result of the first optimization with $a : b = 1 : 4$. **(b)** Results for the second optimization of the 1.5 s magnetic transport with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and $a : b = 1 : 4$. For this optimization parametrization Eq. (4.5) was used. The trajectory shown in (a) was used as base function. The blue curve shows the result of the optimization run. The orange curve shows the smoothed trajectory with a shorter magnetic transport time of 1.45 s.



(a) (b)

Figure A.3: Results for the 2 s magnetic transport optimization processes with frequency range $f = \left\{\frac{k\pi}{2s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and cost weight ratio $a : b = 1 : 2$. Each run number corresponds to a new parameter set that was tested. The best found parametrizations correspond to the **(a)** blue curve and **(b)** purple curve in Fig. 4.25. The large uncertainties for some of the runs are due to some unknown disturbance in the experiment that sometimes lead to a failed experimental cycle.

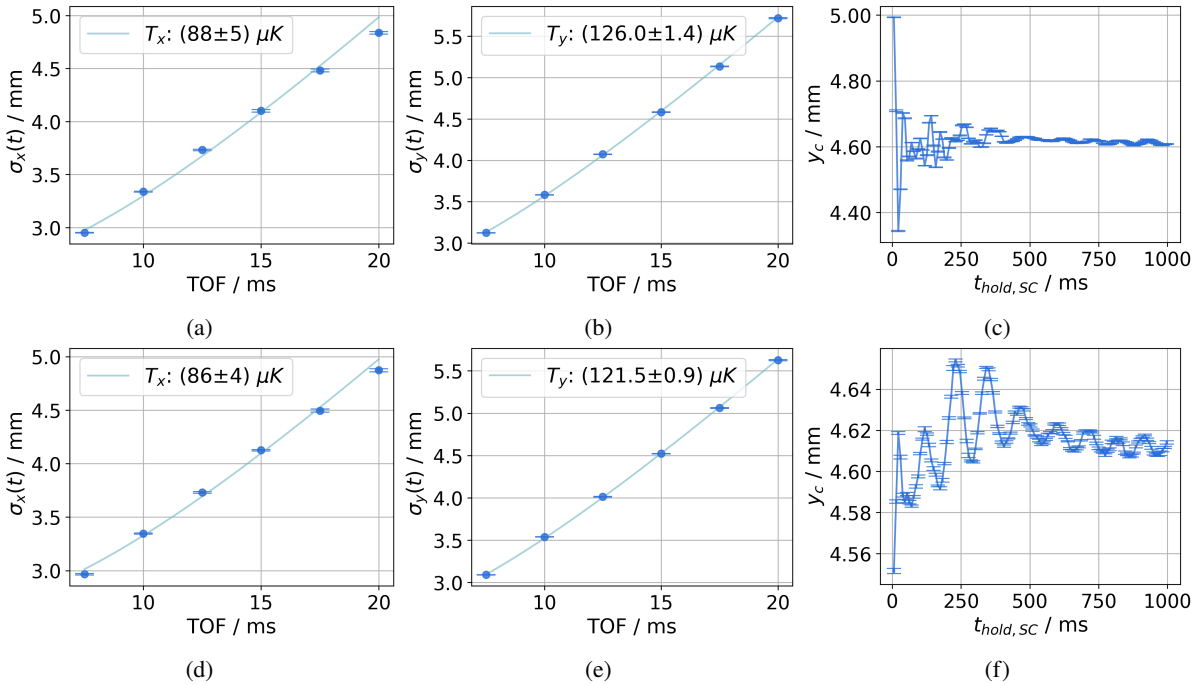


Figure A.4: Characterization measurements quantifying the performance of the magnetic transport with the best found trajectories from the optimization run with the frequency range set to $f = \left\{\frac{k\pi}{1.5s}\right\}$ with $\{k \in \mathbb{N} | 1 \leq k \leq 15\}$ and cost factor weights $a : b = 1 : 4$. **(a)-(c)** Results for the magnetic transport with $X_{1:4,\text{min}}$ parametrization with a reduced transport time of $T_{\text{MT}} = 1.43\,\text{s}$. **(d)-(f)** Results for the magnetic transport with $X^s_{1:4,\text{min}}$ parametrization with a reduced transport time of $T_{\text{MT}} = 1.45\,\text{s}$ and smoothed transport end. The characterization measurements consist of a time-of flight (TOF) measurement in the Science chamber after 500 ms holding time in the last transport trap. $\sigma_{x,y}$ describes the cloud width along the x- and y-axis. The temperatures are extracted by fitting Eq. (2.1) to the data. The second characterization measurement measures the displacement of the cloud center y_c along the transport axis for different holding times $t_{\text{hold, SC}}$ in the SC trap after the magnetic transport. The sloshing measurements are averaged over 10 loops and the TOF measurements are averaged over 20 loops.

Bibliography

[1] R. P. Feynman, *Quantum mechanical computers.*, Found. Phys. **16** (1986) 507.

[2] F. Arute et al., *Quantum supremacy using a programmable superconducting processor*, Nature **574** (2019) 505.

[3] S. Pirandola et al., *Advances in quantum cryptography*, Advances in optics and photonics **12** (2020) 1012.

[4] N. Aslam et al., *Quantum sensors for biomedical applications*, Nature Reviews Physics **5** (2023) 157.

[5] J. Biamonte et al., *Quantum machine learning*, Nature **549** (2017) 195.

[6] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin and X. Yuan, *Quantum computational chemistry*, Reviews of Modern Physics **92** (2020) 015003.

[7] R. D. Somma, S. Boixo, H. Barnum and E. Knill, *Quantum simulations of classical annealing processes*, Physical review letters **101** (2008) 130504.

[8] Y. Cao, J. Romero and A. Aspuru-Guzik, *Potential of quantum computing for drug discovery*, IBM Journal of Research and Development **62** (2018) 6.

[9] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol and D. M. Lucas, *High-fidelity quantum logic gates using trapped-ion hyperfine qubits*, Physical review letters **117** (2016) 060504.

[10] L. Serino et al., *Realization of a multi-output quantum pulse gate for decoding high-dimensional temporal modes of single-photon states*, PRX quantum **4** (2023) 020306.

[11] C. S. Adams, J. D. Pritchard and J. P. Shaffer, *Rydberg atom quantum technologies*, Journal of Physics B: Atomic, Molecular and Optical Physics **53** (2019) 012002.

[12] M. Saffman, T. G. Walker and K. Mølmer, *Quantum information with Rydberg atoms*, Reviews of modern physics **82** (2010) 2313.

[13] Z.-L. Xiang, S. Ashhab, J. You and F. Nori, *Hybrid quantum circuits: Superconducting circuits interacting with other quantum systems*, Reviews of Modern Physics **85** (2013) 623.

[14] G. Kurizki et al., *Quantum technologies with hybrid systems*, Proceedings of the National Academy of Sciences **112** (2015) 3866.

[15] T. F. Gallagher, *Rydberg atoms*, Springer, 1994 231.

[16] Y. Chu et al., *Quantum acoustics with superconducting qubits*, Science **358** (2017) 199.

[17] P. Kharel et al., *Ultra-high-Q phononic resonators on-chip at cryogenic temperatures*, Apl Photonics **3** (2018).

[18] A. N. Cleland and M. R. Geller, *Superconducting qubit storage and entanglement with nanomechanical resonators*, Physical review letters **93** (2004) 070501.

[19] M. Forsch et al., *Microwave-to-optics conversion using a mechanical oscillator in its quantum ground state*, Nature Physics **16** (2020) 69.

[20] M. Bild et al., *Schrödinger cat states of a 16-microgram mechanical oscillator*, Science **380** (2023) 274.

[21] P. B. Wigley et al., *Fast machine-learning online optimization of ultra-cold-atom experiments*, Scientific reports **6** (2016) 25890.

[22] P. Mehta et al., *A high-bias, low-variance introduction to machine learning for physicists*, Physics reports **810** (2019) 1.

[23] Z. Yao et al., *Machine learning for a sustainable energy future*, Nature Reviews Materials **8** (2023) 202.

[24] Y. Almalioglu, M. Turan, N. Trigoni and A. Markham, *Deep learning-based robust positioning for all-weather autonomous driving*, Nature machine intelligence **4** (2022) 749.

[25] X. Qian et al., *A multimodal machine learning model for the stratification of breast cancer risk*, Nature Biomedical Engineering **9** (2025) 356.

[26] S. Azizi et al., *Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging*, Nature Biomedical Engineering **7** (2023) 756.

[27] N. Sapoval et al., *Current progress and open challenges for applying deep learning across the biosciences*, Nature Communications **13** (2022) 1728.

[28] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev and A. Walsh, *Machine learning for molecular and materials science*, Nature **559** (2018) 547.

[29] G. Carleo et al., *Machine learning and the physical sciences*, Reviews of Modern Physics **91** (2019) 045002.

[30] D. Guest, K. Cranmer and D. Whiteson, *Deep learning and its application to LHC physics*, Annual Review of Nuclear and Particle Science **68** (2018) 161.

[31] F. Lanusse et al., *Cmu deeplens: deep learning for automatic image-based galaxy–galaxy strong lens finding*, Monthly Notices of the Royal Astronomical Society **473** (2018) 3895.

[32] A. J. Barker et al., *Applying machine learning optimization methods to the production of a quantum gas*, Machine Learning: science and technology **1** (2020) 015007.

[33] E. T. Davletov et al.,
Machine learning for achieving Bose-Einstein condensation of thulium atoms,
Physical Review A **102** (2020) 011302.

[34] A. D. Tranter et al., *Multiparameter optimisation of a magneto-optical trap using deep learning*,
Nature communications **9** (2018) 4360.

[35] J. Popp, *Construction and Preparation of an Experimental Setup for Excitation and Detection of Rydberg Atoms in a Cryogenic Environment*,
Masters's Thesis: Rheinische Friedich-Wilhelms Universität, 2024.

[36] A. J. Matthies et al.,
Long-distance optical-conveyor-belt transport of ultracold Cs 133 and Rb 87 atoms,
Physical Review A **109** (2024) 023321.

[37] G. Reinaudi, T. Lahaye, Z. Wang and D. Guéry-Odelin,
Strong saturation absorption imaging of dense clouds of ultracold atoms,
Optics letters **32** (2007) 3143.

[38] T. M. Brzozowski, M. Maczynska, M. Zawada, J. Zachorowski and W. Gawlik,
Time-of-flight measurementof the temperature of cold atoms for short trap-probe beamdistances,
Journal of Optics B: Quantum and Semiclassical Optics **4** (2002) 62.

[39] H. Busche,
Efficient loading of a magneto-optical trap for experiments with dense ultracold Rydberg gases,
University Heidelberg, 2011.

[40] J. Gamper, *Atom Preparation and Rydberg Excitation of Rubidium Atoms*,
Masters's Thesis: Rheinische Friedich-Wilhelms Universität, 2024.

[41] J. Gamper, *Frequenzstabilisierung zur Laserkühlung von Rubidium*,
Bachelor's Thesis: Rheinische Friedich-Wilhelms Universität, 2022.

[42] S. Germer, *Frequency stabilization of a laser and a high resolution optical setup for excitation of ultracold Rydberg atoms*, Bachelor's Thesis: Rheinische Friedich-Wilhelms Universität, 2023.

[43] V. Mauth, *Realization of a ^{87}Rb magneto-optical trap*,
Bachelor's Thesis: Rheinische Friedich-Wilhelms Universität, 2023.

[44] VACOM, `https://www.vacom.net/de/home.html`, visited on 13.05.2025.

[45] C. J. Foot, *Atomic physics*, vol. 7, Oxford university press, 2005.

[46] E. L. Raab, M. Prentiss, A. Cable, S. Chu and D. E. Pritchard,
Trapping of neutral sodium atoms with radiation pressure,
Physical review letters **59** (1987) 2631.

[47] K. Krzysztof, K. D. Xuan, G. Małgorzata, B. N. Huy, S. Jerzy et al.,
Magneto-optical trap: fundamentals and realization, CMST (2010) 115.

[48] W. Demtröder, *Experimentalphysik 3: Atome, Moleküle und Festkörper*, Springer-Verlag, 2016.

[49] D. A. Steck, *Rubidium 87 D line data*, (2001).

[50] E. D. Black, *An introduction to Pound–Drever–Hall laser frequency stabilization*,
American journal of physics **69** (2001) 79.

[51] ANALOG DEVICES, *ADF4007 high frequency divider/PLL synthesizer*, https://www.analog.com/en/resources/evaluation-hardware-and-software/evaluation-boards-kits/eval-adf4007.html, visited on 20.05.2025.

[52] J. Dalibard and C. Cohen-Tannoudji, *Laser cooling below the Doppler limit by polarization gradients: simple theoretical models*, Journal of the Optical Society of America B **6** (1989) 2023.

[53] W. Ketterle, D. S. Durfee and D. Stamper-Kurn, *Making, probing and understanding Bose-Einstein condensates*, (1999) 67.

[54] C. Sukumar and D. M. Brink, *Spin-flip transitions in a magnetic trap*, Physical review A **56** (1997) 2451.

[55] T. Badr et al., *Comparison of time profiles for the magnetic transport of cold atoms*, Applied Physics B **125** (2019) 1.

[56] C. Adams and E. Riis, *Laser cooling and trapping of neutral atoms*, Progress in quantum electronics **21** (1997) 1.

[57] R. W. Moore et al., *Measurement of vacuum pressure with a magneto-optical trap: A pressure-rise method*, Review of Scientific Instruments **86** (2015).

[58] V. Gomer et al., *Magnetostatic traps for charged and neutral particles*, Hyperfine Interactions **109** (1997) 281.

[59] M. Greiner, I. Bloch, T. W. Hänsch and T. Esslinger, *Magnetic transport of trapped cold atoms over a large distance*, Physical Review A **63** (2001) 031401.

[60] L. Sadowski, *Design of a superconducting atom chip for interfacing Rydberg atoms with microwave resonators*, Masters's Thesis: Rheinische Friedich-Wilhelms Universität, 2024.

[61] R. Löw et al., *An experimental and theoretical guide to strongly interacting Rydberg gases*, Journal of Physics B: Atomic, Molecular and Optical Physics **45** (2012) 113001.

[62] M. Fleischhauer, A. Imamoglu and J. P. Marangos, *Electromagnetically induced transparency: Optics in coherent media*, Reviews of modern physics **77** (2005) 633.

[63] Laser Components, *COUNT-250C-FC*, https://www.lasercomponents.com/de/produkt/countr/, visited on 06.06.2025.

[64] *Time Tagger 20 from Swabian instruments*, https://www.swabianinstruments.com/time-tagger/, visited on 06.06.2025.

[65] V. C. Gregoric et al., *Quantum control via a genetic algorithm of the field ionization pathway of a Rydberg electron*, Physical Review A **96** (2017) 023403.

[66] I. Beterov, D. Tretyakov, I. Ryabtsev, A. Ekers and N. Bezuglov, *Ionization of sodium and rubidium n S, n P, and n D Rydberg atoms by blackbody radiation*, Physical Review A—Atomic, Molecular, and Optical Physics **75** (2007) 052720.

[67] A. Gürtler and W. Van Der Zande, *l-state selective field ionization of rubidium Rydberg states*, Physics Letters A **324** (2004) 315.

[68] Hamamatsu Photonics K.K., *Microchannel Plate F4655-11*, https://www.hamamatsu.com/eu/en/product/optical-sensors/electron_ion-sensor/mcp/circular_fast-time-response_compact_demountable/F4655-11.html, visited on 06.06.2025.

[69] F. Engel, *Entwicklung eines FPGA basierten Pulsgenerators mit Nanosekunden-Auflösung für schnelle Rydberg-Experimente*, Bachelor's Thesis: Universität Stuttgart, 2013.

[70] CGC INSTRUMENTS, *AMXT-500-EF High Voltage Switch*, https://www.cgc-instruments.com/en/Products/Switches/AMXT/Preconfigured/AMXTEF500-4, visited on 07.06.2025.

[71] *GitHub repository for the experiment control system*, https://github.com/coldphysics, visited on 25.06.2025.

[72] ADwin, https://www.adwin.de/de/produkte/proII.html, visited on 17.05.2025.

[73] M. R. Hush, *M-LOOP*, visited on 2025-05-07, 2016, URL: https://m-loop.readthedocs.io.

[74] J. A. Snyman, D. N. Wilke et al., *Practical mathematical optimization*, vol. 97, Springer, 2005.

[75] R. M. Karp, *On-line algorithms versus off-line algorithms: How much is it Worth to Know the Future?*, Algorithms, Software, Architecture: Information Processing **92** (1992) 416.

[76] J. Roslund and H. Rabitz, *Gradient algorithm applied to laboratory quantum control*, Physical Review A—Atomic, Molecular, and Optical Physics **79** (2009) 053417.

[77] B. Amstrup, G. J. Toth, G. Szabo, H. Rabitz and A. Loerincz, *Genetic algorithm with migration on topology conserving maps for optimal control of quantum systems*, The Journal of Physical Chemistry **99** (1995) 5206.

[78] W. Rohringer et al., *Stochastic optimization of a cold atom experiment using a genetic algorithm*, Applied Physics Letters **93** (2008).

[79] M. Tsubouchi and T. Momose, *Rovibrational wave-packet manipulation using shaped midinfrared femtosecond pulses toward quantum computation: Optimization of pulse shape by a genetic algorithm*, Physical Review A—Atomic, Molecular, and Optical Physics **77** (2008) 052326.

[80] A. Dawid et al., *Modern applications of machine learning in quantum sciences*, arXiv preprint arXiv:2204.04198 (2022).

[81] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, http://www.deeplearningbook.org, MIT Press, 2016.

[82] E. Bartz and T. Bartz-Beielstein, *Online Machine Learning*, Springer, 2024.

[83] V. Losing, B. Hammer and H. Wersing, *Incremental on-line learning: A review and comparison of state of the art algorithms*, Neurocomputing **275** (2018) 1261.

[84] R. Storn and K. Price, *Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces*, Journal of global optimization **11** (1997) 341.

[85] J. A. Nelder and R. Mead, *A simplex method for function minimization*, The computer journal **7** (1965) 308.

[86] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, MIT press Cambridge, MA, 2006.

[87] R. Garnett, *Bayesian optimization*, Cambridge University Press, 2023.

[88] A. J. Barker et al., *Applying machine learning optimization methods to the production of a quantum gas*, Machine Learning: science and technology **1** (2020) 015007.

[89] C. M. Bishop and H. Bishop, *Deep learning: Foundations and concepts*, Springer Nature, 2023.

[90] DELTA ELKETRONIKA, https://www.delta-elektronika.nl, visited on 16.05.2025.

[91] Mini-Circuits, *Voltage Controlled oscillator ZX95-100-S+*, https://www.minicircuits.com/pdfs/ZX95-100-S+.pdf, visited on 20.05.2025.

[92] ANALOG DEVICES, *AD9959 4 channel 500 MSPS DDS with 10-Bit DACs*, https://www.analog.com/en/products/ad9959.html, visited on 20.05.2025.

[93] H. Hattermann et al., *Detrimental adsorbate fields in experiments with cold Rydberg gases near surfaces*, Physical Review A—Atomic, Molecular, and Optical Physics **86** (2012) 022511.

[94] M. R. Lam et al., *Demonstration of quantum brachistochrones between distant states of an atom*, Physical Review X **11** (2021) 011035.

[95] Hantek, *AC/DC Current Clamp CC-65*, https://hantek.com/products/detail/77, visited on 04.06.2025.

[96] TEXAS INSTRUMENTS, *Automative, Fully Integrated Fluxgate Magnetic Sensor DRV425-Q1*, https://www.ti.com/product/DRV425-Q1, visited on 04.06.2025.

[97] D. Meschede, *Gerthsen Physik*, Springer Berlin, Heidelberg, 2006, ISBN: 978-3-662-07460-2.

[98] Hamamatsu Photonics K.K., *Technical Information for MCP ASSEMBLY*.

[99] TELEDYNE LECROY, *WavePro 735Zi Oscilloscope*, https://www.teledynelecroy.com/oscilloscope/, visited on 09.06.2025.

[100] C. C. Gerry and P. L. Knight, *Introductory quantum optics*, Cambridge university press, 2023.

[101] F. Pausewang, *Preparation and Detection of Ytterbium Rydberg Atoms and Molecules*, Masters's Thesis: Rheinische Friedich-Wilhelms Universität, 2025.

Acknowledgements

I would like to conclude this thesis by saying a big thanks to everyone, who helped me in the past year during my Master's thesis. First, I would like to thank Sebastian Hofferberth for giving me the opportunity to be part of the NQO group. I truly appreciate your professional guidance during the last years, starting with my Bachelor's thesis and now finishing my Master's thesis. I learned a lot from you about physics. I would also like to thank Daqing Wang for being my second supervisor. Furthermore, I want to say a special thank you to my HQO family, for letting me be part of this amazing team. Special thanks to Cedric, for teaching me so much during the last year and all the support when writing this thesis. We always had such a fun time working together in the lab. And also thanks for all our gym sessions in the morning and of course for all our wild discussions about whatever topic. Julia, I am so grateful to have you in the team. Working with you in the lab makes so much fun, and I love how we complement each other in the lab by paying attention to different details. And of course also thanks for the emotional support, all the walks through Bonn and shared meals at 'Pie me' café. Sam, thank you for always offering support and staying calm when something did not work out and of course for always updating us about the best Mensa food for the day. And of course HQO thanks for bringing me the beach to our office, to have a beautiful, motivating view when writing this thesis. I really enjoyed working with all of you. I also want to thank Nina, for all the help with different questions about physics, for proofreading, the emotional support at the end and all the gym sessions. Another thank you to the whole NQO group for the nice working atmosphere, the daily lunches at 11:30 and all the coffee and cake breaks. Thanks for this last year!