Georg Jäger

DOCTORAL THESIS

For obtaining the academic degree of
Doktor der Naturwissenschaften

Optimal Quantum Control of Bose-Einstein Condensates

Supervisor: Ao.Univ.-Prof. Mag. Dr. Ulrich Hohenester

KARL-FRANZENS-UNIVERSITÄT GRAZ
UNIVERSITY OF GRAZ

Institute of Physics
Department of Theoretical Physics

Graz, November 2015

Abstract

Bose-Einstein condensates (BECs) are macroscopic quantum objects, which can be used for many different experiments and applications, the most prominent being matter wave interferometry. Usually the BEC is trapped in a magnetic microtrap on an atom chip and can be controlled externally by magnetic fields. Finding a suitable control field that brings the BEC into a desired state is a challenging task. For most experiments it is not possible to calculate an optimal control analytically. This is where simulation and optimization become necessary.

The main goal of this thesis is to find and compare control strategies and algorithms for BECs and applying them to relevant problems. In particular, we investigate different variants of the Gradient Ascend Pulse Engineering (GRAPE) algorithm and Krotov's method and compare them by applying them to relevant problems like wave function splitting or state inversion. Additionally, we use a combination of optimal control theory and analytic calculation to find ways to produce and trap squeezed states of BECs, which can improve the precision of matter wave interferometers significantly.

We find that although both GRAPE and Krotov's method are well suited to solve most BEC problems, GRAPE has certain advantages that make it more versatile and better suited for experimental realization. Furthermore we report on a promising technique for creating squeezed states called "parametric squeezing amplification", and find controls that produce very high squeezing values on short timescales.

Contents

Introduction

This cumulative thesis is concerned with the optimal quantum control of Bose-Einstein condensates (BECs). BECs are interesting macroscopic objects whose properties are governed by quantum effects. They make it possible to utilize quantum physics for many different applications or experiments. Their complex nature makes it, however, very difficult to control them and theoretical simulation and optimization are often necessary in order to work with them. This thesis will show several ways how to use optimal control theory (OCT) in order to find means to manipulate BECs and showcase some interesting experiments and applications that are possible with the use of OCT, like twin atom beams or matter wave interferometry with precision beyond the classical shot noise limit.

The structure of this thesis is as follows: Chapter 1 introduces in the field of BECs and shows several theoretical models that can be used to describe them. Chapter 2 is about one of the most prominent applications of BECs, namely matter wave interferometry. After a short introduction to conventional interferometry, the advantages and challenges of matter wave interferometers are highlighted. Chapter 3 introduces different techniques for implementing OCT, shows advantages and disadvantages and applies those techniques to BECs. The main part of the thesis are chapters 4-7. Each of these chapters is composed of a paper published or accepted in a refereed journal. Chapter 8 consists of a synopsis of the key aspects of these publications, puts them in scientific context and provides a short outlook on possible future developments.

1. Bose-Einstein condensates

1.1. Introduction

A Bose-Einstein condensate (BEC) is a state of matter that occurs in a gas of bosons at very low temperatures, where nearly all bosons are in the ground state of the system. BECs are fascinating objects: In a way they are a link between the quantum and the macroscopic world. While they show behavior typically associated with quantum objects, like uncertainty, entanglement and the ability to interfere, they are macroscopic in size, typically in the range of μm. This makes them very accessible for experimental and technological purposes.

1.1.1. Basic principles

A basic property of all fermions, like electrons or quarks, is the Pauli exclusion principle. It states that only one fermion can occupy a specific quantum state. If two fermions have the same spatial probability distribution, for example, they must differ in some other property, i.e. spin. Bosons are fundamentally different in that regard: An arbitrary number of bosons can occupy the same quantum state. The most prominent consequence of this is the appearance of a completely new state of matter: a Bose-Einstein condensate.

We are interested in (composite) bosons, i.e. atoms with integer spin, such as ^{87}Rb, as long as the average distance between two atoms is large compared to the size of the atoms. In general, a gas of bosons at room temperature has properties very similar to those of a fermion gas. The atoms are thermally distributed, some have more, some have less energy. The fundamental difference between those two types of particles only becomes apparent at very low temperatures, when the system approaches its ground state. Since bosons are not limited by the Pauli exclusion principle, all bosons, even a macroscopic number of them, can occupy the ground state of the system: A BEC is formed. All the atoms are then in the exact same quantum state, they can essentially be described by a single wave function. They all have the same spatial probability distribution, and quantum effects, like interference and the uncertainty principle, become important. Nevertheless the BEC's extension is of macroscopic dimension, typically in the order of μm. This makes BECs intriguing objects to study.

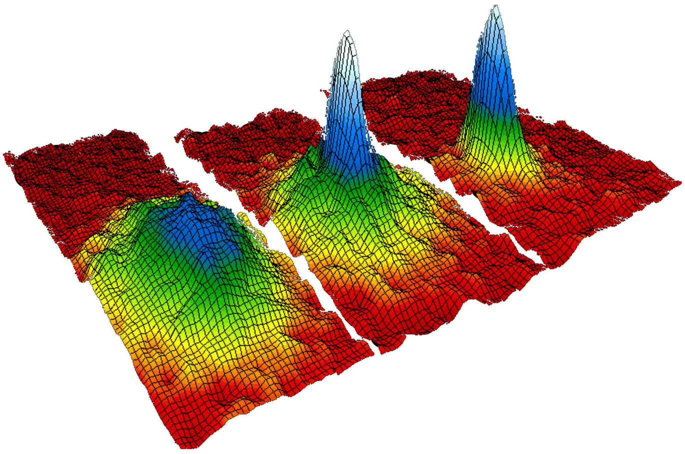

Figure 1.1.: Momentum distribution of a boson gas. For higher temperatures (left) many states are occupied. Lowering the temperature (middle) leads to a dominant occupation of low energy states. For temperatures below the critical temperature (right), only the ground state of the system has a macroscopic occupation. (Source: Based on a graphic from *www.bec.nist.gov*)

Although the most common picture of a BEC is a bosonic gas in the ground state of the system, this is not the definition of a BEC. We call a gas consisting of bosons a BEC, if one or more quantum states are macroscopically occupied, i.e. in the order of the total number of atoms. BECs where only one state has a high occupation are called *simple*. This state can be the ground state, but in principle it can also be any other state. BECs where more than one state is macroscopically occupied are called *fragmented*.

The typical number of atoms of a BEC in a magnetic microtrap is between 1000 and 10000, and additional confinement leads to a “quasi condensate” (see section 1.1.4 for details).

1.1.2. Historical overview

The first notion relevant to BECs was in 1924, when Bose realized that quantum particles are indistinguishable [1]. Shortly after, Einstein [2, 3] concluded that all bosons of a non-interacting gas in its ground state must occupy the same state and that a Bose-Einstein condensation must occur below a critical temperature.

At that point this research was purely theoretical. The critical temperature was far beyond the reach of the experiments at that time. In the next decades new cooling methods became available. The most important ones were laser cooling and evaporative cooling. The principle behind laser cooling is to apply precisely tuned lasers from every spatial direction. The lasers are set up in a way that utilizes the Doppler effect and only interact with atoms that are moving towards them. On average, this cools the gas. Evaporative cooling is achieved by trapping the gas in some form of potential. Then the potential gets lowered, so that hot (= fast) atoms can escape and only the coldest stay in the trap. These two techniques, combined with various other advancements, made the creation of BECs possible, 70 years after their theoretical prediction. The first BEC was produced in 1995 by Cornell and Wieman [4] by cooling rubidium atoms down to 170 nK. In the same year Ketterle demonstrated important BEC properties in a gas of sodium atoms [5]. For their work on BECs the Nobel price was awarded to E. A. Cornell, W. Ketterle and C. E. Wieman in 2001.

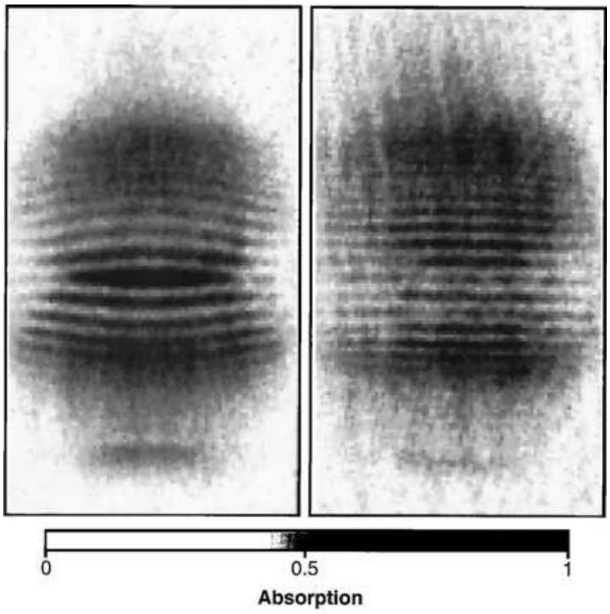

Figure 1.2.: Interference between two macroscopic BECs observed by Ketterle *et at.* (Source: [6])

Between 1995 and today much progress has been made in the field of BECs. It became possible to create BECs with many different types of atoms including helium, caesium and lithium. Another important achievement, especially for matter wave interferometry, was the coherent splitting of a BEC using magnetic fields, first performed by Schmiedmayer *et al.* in 2006 [7]. Today the field is very active and many areas are investigated by various groups.

1.1.3. Current research and applications

Nowadays BECs can be produced in a convenient way on so-called atom chips [8, 9, 10]. These are micro-fabricated chips with wires that allow the generation and precise control of magnetic fields. With these fields it is possible to trap and cool atoms; accurate control allows many interesting experiments and applications.

One of those applications is quantum simulation. Thereby one uses a BEC to simulate other systems. This is not only relevant in the field of solid state physics; even problems in astrophysics or other areas can be mapped on a BEC, for example in the area of gravitational waves [11].

Additionally BECs are often used for metrology and sensing. The wave-like nature of BECs makes it possible to build a matter wave interferometer. Such interferometers are widely investigated [12, 13, 14] and there are many different approaches like Mach-Zehnder interferometers [15, 16], Michelson interferometers [17] or multimode interferometers [18]. Using matter waves instead of light waves has various advantages, the biggest one being that matter has a finite rest mass and is therefore sensitive to gravity. Hence BECs can be used for many experiments that are not possible with conventional interferometry. One prominent example is the detection of gravity waves [19], which could be possible with a sensitive matter wave interferometer. Currently BEC experiments also help in the search for dark energy: In [20] a matter wave interferometer was used to probe the gravitational field of a spherical mass by individual atoms. This helped constrain certain parameters for a large class of dark energy theories. On the other hand BECs can provide alternative approaches to known problems and experiments. In [21] a BEC was used to measure the value of the gravitational constant in an entirely new way. This measurement was not only very precise, but can also help in identifying the systematic errors of conventional techniques, which use macroscopic objects. The second advantage of matter wave interferometry is that it is possible to use non-classical states to improve the precision of an interferometer below the classical shot noise limit, as demonstrated in [22]. Further advancements in this area might enable future matter wave interferometers to outperform conventional ones in terms of accuracy.

1.1.4. BECs in lower dimensions

In a homogeneous system Bose-Einstein condensation is impossible in 1D and 2D [23], since the density of states does not approach 0 for low energies. Nevertheless a 1 dimensional (quasi-)BEC is possible inside magnetic micro traps. The trap fields constrict the gas and change the density of states so that Bose-Einstein condensation can occur. For most experiments reported on in this thesis a quasi-1D BEC is used. In such a BEC the atoms are constricted to a cigar-like shape. Excitations are possible in the longitudinal direction, the Bose-Einstein condensation occurs only in the radial direction, making the BEC quasi 1-dimensional [23, 24, 25].

1.2. Theory

In the following, different models to describe BECs are presented [26, 27, 28]. We will start with the full Hamiltonian. However, since such a many-body Schrödinger equation, with typically thousands of particles is nearly impossible to solve exactly, we will have to rely on approximations. In a mean field approach, the Gross-Pitaevskii equation, we can describe the orbitals very accurately, but cannot calculate many-body properties of the BEC. In order to gain insight in such properties we can use the two-mode model, however this model only works for BECs in double wells and it completely neglects the actual shape of the orbitals. Methods beyond those two are the Bogoliubov approximation [29] or a more sophisticated model called MCTDHB, which will be introduced in order to tackle problems where both number and orbital dynamics are important.

1.2.1. Full Hamiltonian

A BEC consisting of N atoms, trapped in a magnetic microtrap at $T = 0$, can be described by the many-body Schrödinger equation for N bosons:

$$i\frac{\partial\Psi(\vec{r}_1,\vec{r}_2,\dots\vec{r}_N;t)}{\partial t} = \hat{H}(\vec{r}_1,\vec{r}_2,\dots\vec{r}_N;t)\Psi(\vec{r}_1,\vec{r}_2,\dots\vec{r}_N;t) \tag{1.1}$$

$$\hat{H}(\vec{r}_1,\vec{r}_2,\dots\vec{r}_N;t) = \sum_{i=1}^{N}\hat{h}(\vec{r}_i,t) + U_0^{3D}\sum_{k<q}\delta(\vec{r}_k-\vec{r}_q) \tag{1.2}$$

U_0^{3D} is the contact potential between the atoms, dominated by s-wave scattering [26, 27]:

$$U_0^{3D} = \frac{4\pi\hbar^2 a_s}{m}\,, \tag{1.3}$$

with the atom mass m and the s-wave scattering length a_s. For ^{87}Rb, for example, $a_s = 5.77$ nm [30]. We also introduced the single-particle Hamiltonian $\hat{h}$

$$\hat{h}(\vec{r},t) = -\frac{\hbar}{2m}\vec{\nabla}^2 + V_{ext}(\vec{r})\,, \tag{1.4}$$

with the confinement potential V_{ext}. Since we are dealing with bosons, the wave function Ψ is symmetric.

An important concept we need to define are the so-called p-particle reduced density matrices [31]:

$$\rho^{(p)}(x_1,\dots,x_p|x'_1,\dots,x'_p;t) = \\ \frac{N!}{(N-p)!}\int\Psi(x_1,\dots,x_p,x_{p+1},\dots,x_N;t)\Psi^*(x'_1,\dots,x'_p,x_{p+1},\dots,x_N;t)dx_{p+1}\dots dx_N \tag{1.5}$$

The interpretation of these matrices is straight forward:
The diagonals $\rho^{(p)}(x_1,\dots,x_p|x_1,\dots,x_p;t)$ are proportional to the p-particle probability distributions. The one-body density matrix is given by

$$\rho(\vec{r},\vec{r}\,';t) = \langle\Psi(t)|\hat{\Psi}^\dagger(\vec{r}\,')\hat{\Psi}(\vec{r})|\Psi(t)\rangle\,. \tag{1.6}$$

This can also be written in terms of eigenvalues and eigenfunctions.

$$\rho(\vec{r}, \vec{r}'; t) = \sum_i n_i(t) \phi_i^*(\vec{r}'; t) \phi_i(\vec{r}; t) \tag{1.7}$$

Here, the ϕ_i are the natural orbitals and the n_i the natural occupations. The occupations also tell us if the BEC is *simple* or *fragmented*. If only one eigenvalue is macroscopic (i.e. of the order N), the BEC is simple, if there are more macroscopic eigenvalues, the BEC is fragmented [32].

Higher reduced densities can be calculated according to (1.5), for example the often used two-particle density matrix $\rho^{(2)}$ is given as

$$\begin{aligned} &\rho^{(2)}(x_1, x_2 | x_1', x_2'; t) = \\ &N\,(N-1) \int \Psi(x_1, \ldots, x_N; t) \Psi^*(x_1', x_2', x_3, \ldots, x_N; t) dx_3 \ldots dx_N \,. \end{aligned} \tag{1.8}$$

We always assume zero temperature, which cannot be realized in experiments, but the achieved temperatures are low enough so that finite temperature effects can often be neglected. Another feature of most experiments is that we do not work with a 3-dimensional BEC, but rather one of lower dimensionality, where additional longitudinal excitations are possible (see section 1.1.4).

1.2.2. Gross-Pitaevskii equation

Another way to describe BECs is the Gross-Pitaevskii equation [33, 34, 35]. It is a mean field approximation that is based on the assumption that all atoms are in the same state and can therefore be described with the same wave function. This assumption is valid for weakly interacting gases, if the gas is diluted enough that there is on average less than one atom in the s-wave scattering length [33]. The Gross-Pitaevskii equation looks very similar to the Schrödinger equation, but has an additional, nonlinear term that describes the interaction between the atoms. It reads

$$i\hbar\dot{\Psi}(r,t) = \left(-\frac{\hbar^2}{2m} \nabla^2 + V_{ext}(r) + \kappa \left| \Psi(r,t) \right|^2 \right) \Psi(r,t) \,, \tag{1.9}$$

with some external potential $V_{ext}(r)$ (e.g. a magnetic microtrap) and the nonlinearity parameter κ that describes the interaction between the atoms. This parameter κ can be calculated within the Gross-Pietaevskii model and is also experimentally accessible.

The Gross-Pitaevskii equation can be derived from the Hartree ansatz of the full system or by a variational principle.

Calculation of tunneling and charging energy

For a condensate in a double-well potential it is possible within the Gross-Pitaevskii framework to calculate the tunneling energy Ω and the charging energy κ, both introduced in section 1.2.3. In order to do this we start by the variational ansatz

$$\Psi(r,t) = \psi_1(t) \phi_l(r) + \psi_2(t) \phi_r(r) \tag{1.10}$$

This means that the whole wave function $\Psi(r,t)$ can be written in terms of the time dependent parameters ψ_1 and ψ_2 and the two wave functions ϕ_l and ϕ_r, roughly describing the condensate in the left and right well, respectively. ϕ_l and ϕ_r can be calculated from the gerade and ungerade states ϕ_g and ϕ_u, which are the exact symmetric and antisymmetric stationary eigenstates of the Gross-Pitaevskii equation [36]:

$$\phi_l(r) = \frac{\phi_g + \phi_u}{\sqrt{2}}\,, \tag{1.11}$$

$$\phi_r(r) = \frac{\phi_g - \phi_u}{\sqrt{2}}\,. \tag{1.12}$$

This ensures

$$\int \phi_l(r)\,\phi_r(r)dr = 0\,. \tag{1.13}$$

Using equations (1.10-1.13) in the Gross-Pitaevskii equation leads to the Boson Josephson junction (BJJ) equations [36], and we also find expressions for the tunneling and the charging energy:

$$\kappa = g_0 \int |\phi_{l/r}|^4 dr, \tag{1.14}$$

$$\Omega \simeq -\int \left(\frac{\hbar^2}{2m} (\nabla\phi_l \nabla\phi_r) + \phi_l V_{ext} \phi_r \right) dr\,, \tag{1.15}$$

where $g_0 = 4\pi\hbar^2 a/m$, with the atomic mass m and the s-wave scattering length a [36].

1.2.3. Two-mode model

For BECs in double wells another simple approximation is possible: the two-mode model. Here we view the left and the right well as two possible states of the system and neglect the actual shape of the potential. Important features of the trap geometry are only included via the parameters κ and Ω. This description has many similarities to Josephson junctions [37] and is therefore often referred to as a Josephson model [38]. Each atom of the BEC can either be in the left or in the right well of the trapping potential. If the BEC consists of N atoms this would correspond to a spin N/2 system. By using the annihilation and creation operators of states in the left and right well $a_{l/r}$ and $a^\dagger_{l/r}$ we can define pseudospin operators [39]:

$$J_x = \frac{1}{2}\left(a_l^\dagger a_r + a_r^\dagger a_l\right) \tag{1.16}$$

$$J_y = \frac{i}{2}\left(a_l^\dagger a_r - a_r^\dagger a_l\right) \tag{1.17}$$

$$J_z = \frac{1}{2}\left(a_l^\dagger a_l - a_r^\dagger a_r\right)\,, \tag{1.18}$$

These pseudospin operators are very similar to conventional angular momentum operators. However, they have a different interpretation in the context of double well BECs.

J_x exchanges an atom from the left and the right well, J_y measures the phase difference between the left and right well and J_z the atom number imbalance. These operators can also be used to write the Hamiltonian of the system as [40, 41]

$$H = -\Omega \, J_x + 2\kappa \, J_z^2 \, , \tag{1.19}$$

where Ω is the tunneling energy that describes the tunneling between the left and right well and κ is the charging energy, accounting for the nonlinear interaction between the atoms. Both these parameters can be approximated within the Gross-Pietaevskii model [36] as shown in section 1.2.2. Within the two-mode model it is also possible to represent each state of a double well BEC on the Bloch sphere.

The Bloch sphere

The Bloch sphere [42, 43] is a very intriguing representation of two level quantum mechanical systems. It is well known in the field of qubits, but can also be used to visualize double well BECs. Each state of the system is represented by a distribution on the Bloch sphere. The interpretation is quite intuitive in the case of double well BECs. A state where all the atoms are in the left well is represented on the north pole of the sphere, a state where all atoms are in the right well is on the south pole. States where the atom number imbalance is zero are on the equator. The y-axis tells us about the relative phase between the two wells. Additionally to the mean values of the number imbalance n and the relative phase ϕ, also the variance of those values, Δn and $\Delta \phi$ can be seen, as shown in Fig. 1.3. These values are especially important for squeezed states and quantum interferometry, as will be shown in section 2.3.

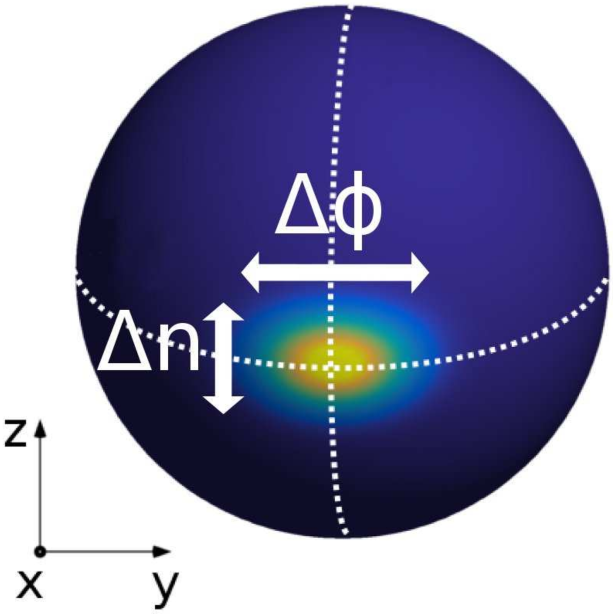

Figure 1.3.: Bloch sphere representation of a double well BEC. States where all atoms are in the left/right well are represented at the poles. States with zero number imbalance, like the one depicted, are centered around the equator. Also the variance of the number imbalance and the relative phase, Δn and $\Delta\phi$, are visible.

Gerade-ungerade basis

In the previous sections we always described the BEC in the left-right basis, meaning that the two orbitals that we built our two-mode model upon are the right orbital, located in the right well, and the left orbital, located in the left well. While this basis is very easy to understand, it has certain disadvantages. Often it is more convenient to use the ground state of the system as one of our orbitals as well as the first excited state. This has the great advantage that most of the time one of these states will have a much higher occupation than the other one. Additionally, the ground state is always gerade, while the first excited state is always ungerade. This is an important feature and can simplify many calculations. We call this basis the gerade-ungerade basis. There is a simple relation between the orbitals of the left-right and gerade-ungerade basis [44]. The ground state of a system without nonlinear interaction is always gerade, the BEC density is distributed symmetrically in the left and right well. This state corresponds to a superposition of the right and the left orbital. The first excited state, which is ungerade, is a superposition with opposite sign, i.e. $\phi_g - \phi_u$. Additionally there could also be a relative phase between the left and the right orbital, which needs to be taken into account. We arrive at

$$\phi_l = \frac{1}{\sqrt{2}}(\phi_g + \tilde{f}\phi_u) \tag{1.20}$$

$$\phi_r = \frac{1}{\sqrt{2}}(\phi_g - \tilde{f}\phi_u)\,, \tag{1.21}$$

where $\tilde{f} = f/|f|$ is the relative phase between the orbitals, with

$$f = \int \theta(x)\ \phi_g^*(x)\phi_u(x)\,dx\,. \tag{1.22}$$

$\theta(x)$ is the Heaviside step function, meaning that we only integrate over half of the position space. If we want to use the pseudospin operators that were defined in 1.2.3, we have to rewrite them with the creation and annihilation operators of the gerade-ungerade basis.

This leads to

$$J_x = \frac{1}{2}\left(\quad a_g^\dagger a_g - \quad a_u^\dagger a_u\right) \tag{1.23}$$

$$J_y = \frac{i}{2}\left(\tilde{f}^* a_u^\dagger a_g - \tilde{f} a_g^\dagger a_u\right) \tag{1.24}$$

$$J_z = \frac{1}{2}\left(\tilde{f}^* a_u^\dagger a_g + \tilde{f} a_g^\dagger a_u\right) \tag{1.25}$$

The gerade-ungerade basis is of special importance for the MCTDHB model, explained in more detail in section 1.2.4.

1.2.4. MCTDHB

The Gross-Pitaevskii model is very accurate in describing the wave function of a BEC. However, since it is only a mean field approach, certain important properties like number squeezing can not be described at all. The two-mode model on the other hand can calculate those properties, yet the shape of the orbitals is only implicitly included in the parameters κ and Ω. In some cases neither of those models is enough to accurately describe a BEC. Especially for parametric squeezing amplification, a method for producing squeezed states introduced in section 2.3, the shape of the orbitals changes so much that it influences the tunneling and charging energy in a way that cannot be ignored. As will be shown in chapter 7, a more sophisticated approach is necessary, one that accounts for both number and wave function dynamics. A suitable model is the MCTDHB model.

MCTDHB [45, 46] stands for multi configurational time dependent Hartree method for bosons and was introduced by Alon *et al.* [47] in 2008. It is based on the multi configurational time dependent Hartree method (MCDTH) developed by H.-D. Meyer *et al.* [48] in 1990.

The main ansatz of MCTDHB is to expand the states into modes according to

$$\hat{\Psi}(r) = \sum_{i=1}^{m} \hat{a}_i(t)\phi_i(x,t)\,, \tag{1.26}$$

where both the modes or orbitals ϕ_i and the occupation a_i are time dependent. For $m \to \infty$ this is not an approximation, but only a general expansion. However, a very effective approximation is possible here: The computational effort of calculations rises exponentially with the number of modes m. This means that truncating the sum at a finite m greatly improves performance. Especially for BECs this works very well. In an ideal BEC only one state is macroscopically occupied and it would even be possible to truncate the sum at $m = 1$. We would arrive at the Gross-Pitaevskii equation. However, in a realistic description of a BEC more than one state can be occupied and it is necessary to include more modes. In principle it is possible to include an arbitrary number of modes, but since the computational effort rises exponentially it is often best to use $m = 2$ for the description of BECs.

A general state can be written as a superposition of symmetrized states. For m orbitals this can be written as

$$|\vec{n}; t\rangle = \frac{1}{\sqrt{n_1!\, n2! \dots n_m!}} \left(a_1^\dagger(t)\right)^{n_1} \left(a_2^\dagger(t)\right)^{n_2} \dots \left(a_m^\dagger(t)\right)^{n_m} |0\rangle \,. \tag{1.27}$$

The whole configuration can now be described by the vector $\vec{n} = (n_1, n_2, \dots, n_m)$. Every state can be written as

$$|\Psi(t)\rangle = \sum_{\vec{n}} C_{\vec{n}}(t)\, |\vec{n}; t\rangle \tag{1.28}$$

We now have two sets of time dependent parameters. On the one hand we have the modes $\{\phi_i(x,t)\}$, which describe the shape of the orbitals, on the other hand the occupation of those orbitals, described by the coefficients $\{C_{\vec{n}}\}$. The time dependency of those parameters can be found via variational calculus. One uses a Lagrange formalism and formulates an action ($\hbar = 1$)

$$S[\{C_{\vec{n}(t)}\}, \{\phi_i(x,t)\}] = \int dt \left\{ \langle\Psi|(\hat{H} - i\frac{\partial}{\partial t})|\Psi\rangle - \sum_{j,k=1}^{m} \mu_{ik}(t)[\langle\phi_i|\phi_k\rangle - \delta_{ik}] \right\} . \tag{1.29}$$

The first term guarantees that the Schrödinger equation is satisfied, the second term ensures orthonormality of the orbitals. For $m = 1$ this variation reproduces the Gross-Pitaevskii equation. The more interesting case, however, is $m = 2$. The two orbitals we use are the gerade orbital ϕ_g and the ungerade orbital ϕ_u. Their time dependency can be calculated by variational calculus [45]. Starting with (1.29) we perform a variation with respect to the orbitals. For this it is convenient to write the Hamiltonian in second quantized form. Performing the variation and eliminating the Lagrange multipliers by using the orthogonality relation of the orbitals yields

$$\begin{aligned} i\dot{\phi}_g &= \hat{P}\left[\hat{h}\phi_g + (\mathfrak{f}_g^g|\phi_g|^2 + \mathfrak{f}_g^u|\phi_u|^2)\phi_g + \tilde{\mathfrak{f}}_g\phi_g^*\phi_u^2\right] \\ i\dot{\phi}_u &= \hat{P}\left[\hat{h}\phi_u + (\mathfrak{f}_u^u|\phi_g|^2 + \mathfrak{f}_u^g|\phi_g|^2)\phi_u + \tilde{\mathfrak{f}}_u\phi_u^*\phi_g^2\right] \,. \end{aligned} \tag{1.30}$$

The coefficients $\mathfrak{f}$, not to be confused with the previously introduced phase f, are given by the elements of the one-body and two-body reduced densities:

$$\mathfrak{f}_g^g = U_0 \{\rho\}_{gg}^{-1} \rho_{gggg} \tag{1.31}$$

$$\mathfrak{f}_u^u = U_0 \{\rho\}_{uu}^{-1} \rho_{uuuu} \tag{1.32}$$

$$\mathfrak{f}_g^u = U_0 \{\rho\}_{gg}^{-1} \rho_{gugu} \tag{1.33}$$

$$\mathfrak{f}_u^g = U_0 \{\rho\}_{gg}^{-1} \rho_{ugug} \tag{1.34}$$

$$\tilde{\mathfrak{f}}_g = U_0 \{\rho\}_{gg}^{-1} \rho_{gguu} \tag{1.35}$$

$$\tilde{\mathfrak{f}}_u = U_0 \{\rho\}_{uu}^{-1} \rho_{uugg} \tag{1.36}$$

The projector $\hat{P} = 1 - |\phi_g\rangle\langle\phi_g| - |\phi_u\rangle\langle\phi_u|$ in (1.30) guarantees orthonormality of the orbitals.

The number distribution C has the following time dependency, found by performing a variation of (1.29) with respect to C:

$$i\frac{\partial C(t)}{\partial t} = \mathcal{H}C\,, \tag{1.37}$$

with

$$\mathcal{H} = -\Omega J_x + \frac{1}{2}\sum_{k,q,l,m}{}' \hat{a}_k^\dagger \hat{a}_q^\dagger \hat{a}_l \hat{a}_m \, W_{kqlm}\,. \tag{1.38}$$

In section 1.2.3 we already rewrote J_x in the gerade-ungerade basis and arrived at $J_x = \frac{1}{2}\left(a_g^\dagger a_g - a_u^\dagger a_u\right)$. Ω is given as $\Omega = \langle\phi_u|\hat{h}|\phi_u\rangle - \langle\phi_g|\hat{h}|\phi_g\rangle$ in the gerade-ungerade basis. The two-body matrix elements are

$$W_{kqlm} = U_0 \int dx\, \phi_k^*(x)\, \phi_q^*(x)\, \phi_l(x)\, \phi_m(x)\,. \tag{1.39}$$

We now have two non-linear, coupled equations: (1.30) and (1.37). These can be solved in parallel.

2. Interferometry

2.1. Introduction

In general, an interferometer is a device that extracts precise information out of waves. Different types of interferometers are used in many fields of science, like astronomy, metrology, spectroscopy and particle physics. Although most interferometers use electromagnetic waves, i.e. light, it is also possible to use matter waves instead [12]. The basic principle for a conventional interferometer is the following: A wave is split by a beam splitter and the two waves then move on different paths. The two paths differ in some way or the other. One possibility is that one path is simply longer or oriented in a different direction. In most cases, however, a sample, which needs to be analyzed, is placed in one of the paths. In that case, one beam acquires some new property (e.g. a phase) through interaction with the sample, while the other beam does not and serves as a reference. In the end the waves are superimposed and interference occurs. The interference gives detailed information about the properties of the sample beam and thereby about the sample and the interaction that took place.

There are many types of interferometers, but we will focus on the the Mach-Zehnder type [49, 50], since its principle can be directly translated to matter wave interferometry.

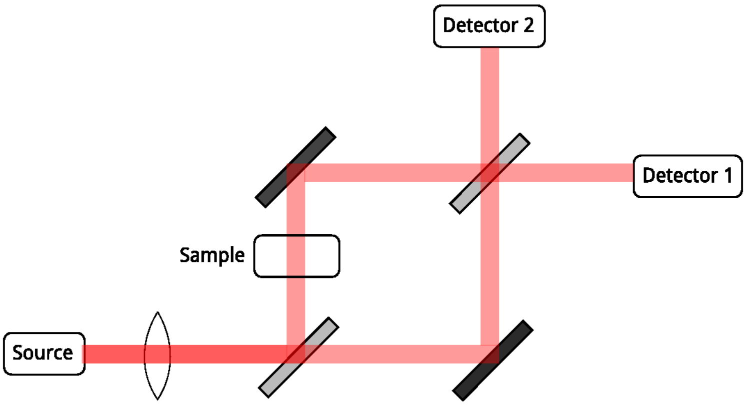

Figure 2.1.: Mach-Zehnder interferometer: A beam is split by a beam splitter into a sample beam (SB) and a reference beam (RF). The sample beam interacts with a sample, afterwards both beams hit a second beam splitter and are measured by detectors.

The set-up of such an interferometer is depicted in Fig. 2.1: First a beam splitter (e.g. a half silvered mirror) splits the beam into a sample beam (SB) and a reference beam (RB). The reference beam then hits a mirror right away, while the sample beam first passes a sample and interacts with it. In the end the two beams hit a second half silvered mirror before reaching the two detectors.

2.2. Interferometry with BECs

Matter wave interferometry is based on the same principles as light wave interferometry. In the following we will introduce the matter wave equivalent of a Mach-Zehnder interferometer as well as the so-called Time-of-flight (TOF) interferometer.

In order to perform any kind of interferometric experiment, the BEC must first be created and prepared on an atom chip [8, 9, 10]. The wires on the atom chip are designed in such a way that they produce a magnetic field that traps the atoms in an elongated cigar shape, as depicted in Fig. 2.2.

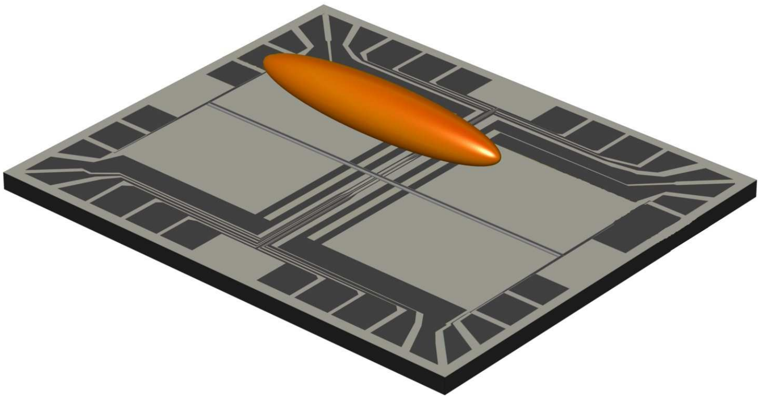

Figure 2.2.: Cigar shaped BEC on an atom chip. The magnetic fields produced by the atom chip trap the atoms in this specific shape. (Source: based on a graphic from *http://atomchip.org*)

The direction that is of interest to us is the radial direction: We ignore the longitudinal direction of the BEC and use the high symmetry of the z-y-plane to reduce the problem to one dimension, in which the 1-d quasi-BEC is located [23]. Fig. 2.3 depicts a possible wire setup of an atom chip: The combination of B-fields in x and in y direction trap the BEC and allow to control it. An radio frequency field is necessary in order to produce a double well potential in the otherwise radially symmetric trap. This RF field can also be used to introduce anharmonicity and anisotropy into a single well trap.

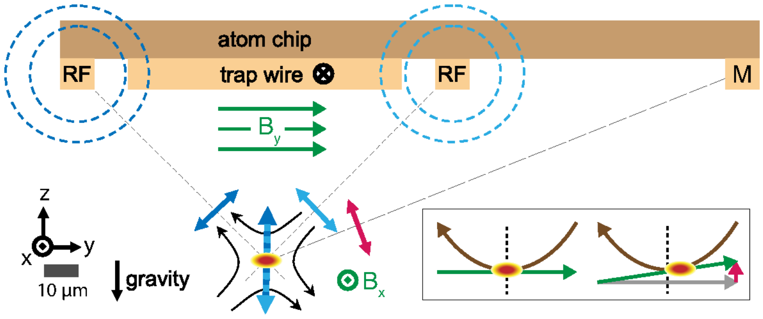

Figure 2.3.: Atom chip: The B-fields trap the BEC and allow for precise control (Source: Paper 2 [51])

In order to have the matter wave equivalent of a beam splitter, the atom chip potential must be able to change from a single well to a double well. This is realized by an additional RF field with variable strength, leading to a Lesanovsky-type potential [52]. Fig. 2.4 shows the potential and the BEC ground state of such a potential for the RF-parameter $\lambda = 0$ (single well) and $\lambda = 1$ (double well). Note that the x-axis denoted with "position" is again the radial direction.

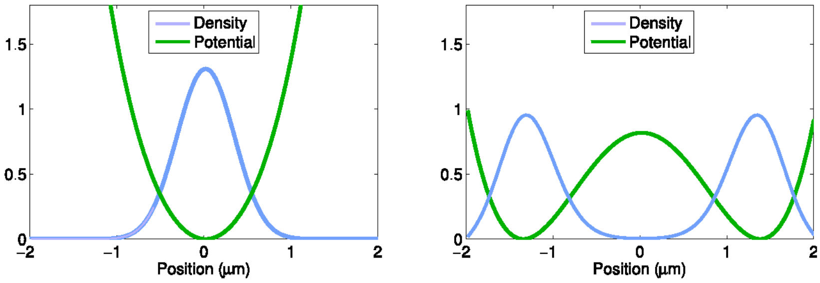

Figure 2.4.: Lesanovsky potential: The potential changes from a single well to a double well, by increasing the parameter λ which describes the strength of a radio frequency field.

On an atom chip of this type, matter wave interferometry is possible and was demonstrated in [15]. The principle is very similar to a Mach-Zehnder type interferometer, however, there are some important differences.

The main stages of the interferometer are the same:

1. Preparation
2. Beam splitting
3. Phase acquisition
4. Recombination
5. Readout

First of all the BEC needs to be trapped and prepared on an atom chip. It needs to be split into two coherent waves which are spatially separated. In the context of matter waves this is done by changing the trapping potential from a single to a double well. This can be done adiabatically, so that the BEC remains in the ground state at all time, or even faster with the use of optimal control theory (see chapter 3). After that separation the BEC density has two maxima, one in the right and one in the left well of the trapping potential.

Next we need some kind of beam splitter that should play the same role as a half silvered mirror in the original design of a Mach-Zehnder interferometer. This cold atom beam splitter is realized by decreasing the barrier height between the two BECs and allowing tunneling.

In the next stage of the experiment one of the BECs is modified in some way. For example one can tilt the potential, so that it accumulates a phase due to the gravitational potential.

After this accumulation stage, the recombination stage begins. In a conventional Mach-Zehnder interferometer this is again done by a half silvered mirror. For matter waves the goal of this process is to convert the relative phase of the condensates into an atom number difference. Several ways have been proposed to do this. One possibility is to use phase dependent heating caused by merging of two condensates [53], a different approach would be to use a quarter of a Josephson oscillation [54, 55]. Another way, presented in [15] is to abruptly decrease the well distance and the barrier height, so that the condensates are accelerated towards each other. Afterwards the wells are separated again to make it possible to count the atoms.

At the beginning of the readout stage the wells get separated even further on a short timescale. This gives the atoms in different wells opposite momenta. The magnetic fields are then switched off, so that the atoms fall down in the gravitational field. Then they can be counted by a fluorescence detector. The atom number difference is the final result of this experiment and gives information about the relative phase of the BECs and thus about the interaction that occurred.

The main steps of a matter wave Mach-Zehnder interferometer are depicted in Fig. 2.5.

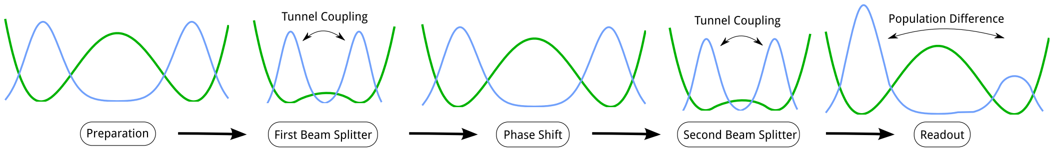

Figure 2.5.: Stages of a matter wave Mach-Zehnder interferometer:
1) A BEC is prepared in a double well.
2) Increasing the tunnel coupling works as a beam splitter.
3) A phase shift is applied.
4) Increasing the tunnel coupling works as a beam splitter
and transforms phase information into a population difference.
5) The population difference can be read out by counting atoms.

When working with matter waves there is also a second important type of interferometer: the time-of-flight (TOF) interferometer. The principle behind it is slightly different from that of a Mach-Zehnder interferometer, since there are no beam splitters. In the preparation stage a BEC is trapped in a double well potential, similar to the first step of the Mach-Zehnder interferometer. Then a phase shift is applied, for example by tilting the trap potential. In the next stage the trap potential is abruptly switched off and the condensates are released. During their free fall they interfere and a fringe pattern can be observed and measured. From this pattern, information about the phase shift can be extracted.

The stages of a TOF interferometer are depicted in Fig. 2.6.

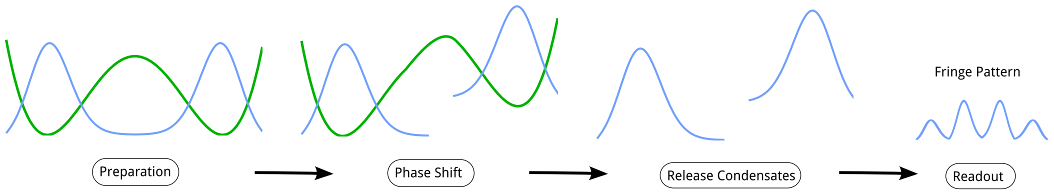

Figure 2.6.: Stages of a TOF interferometer:
1) A BEC is prepared in a double well.
2) A phase shift is applied.
3) The condensates are released.
4) The condensates interfere and fringe patterns can be observed.

2.3. Quantum interferometry

2.3.1. Squeezed states

A big advantage of matter wave interferometry is that non-classical states can be used to increase the precision of measurements [22]. In every interferometer experiment one measures either the atom number in each well or the relative phase between the two wells. The precision of such a measurement is influenced by the quantum uncertainties of the number and the phase Δn and $\Delta\phi$. However, when measuring the atom number, the phase fluctuations $\Delta\phi$ may be arbitrarily large, which leads to the idea to decrease one of those uncertainties at the cost of increasing the other. In principle this is possible due to the Heisenberg uncertainty relation [56]

$$\Delta n \, \Delta\phi \gtrsim \frac{1}{2} \, . \tag{2.1}$$

For a binomial state $\Delta n = \sqrt{N}/2$ and $\Delta\phi = 1/\sqrt{N}$ leading to the standard quantum shot noise limit [57]. States where either Δn or $\Delta\phi$ are smaller than for a binomial state are called squeezed states. States with lower number uncertainty are called number squeezed states, while states with lower phase uncertainty are called phase squeezed states. In order to quantify how much the squeezing can lower the shot noise, several quantities can be introduced. The so-called number squeezing ξ_n is defined as $\xi_n = 2\Delta n/\sqrt{N}$. This means that a binomial state has $\xi_n = 1$ and number squeezed states have $\xi_n < 1$. There is a similar quantity for the phase uncertainty called phases squeezing ξ_ϕ. It is defined as $\xi_\phi = \Delta\phi\sqrt{N}$. The Bloch sphere representation of such squeezed states can be seen in Fig. 2.7.

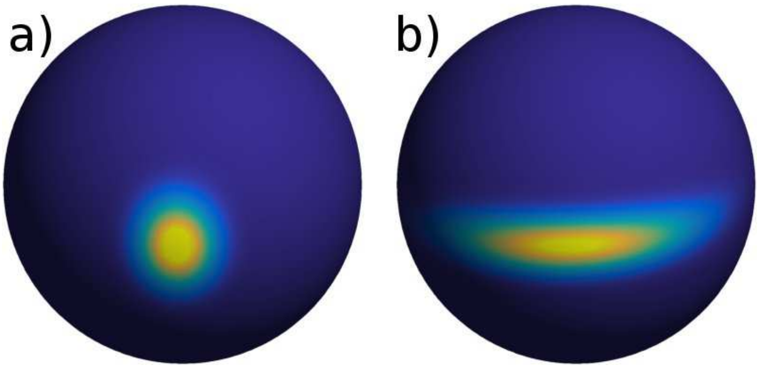

Figure 2.7.: Squeezed states: a phase squeezed (a) and a number squeezed (b) state depicted on the Bloch sphere

However, these two quantities completely neglect coherence, a very important concept for interferometry. Not every number squeezed state improves precision. A simple example is a Fock state with given occupation of the left and right well. In such a system the number squeezing would be very high, since the atoms could not tunnel from one well to the other. Nevertheless interferometry would not be possible with such a system, because of the loss of coherence. This is why it is important to include coherence in squeezing factors. This is often done with the so-called coherence factor [56, 58].

$$\alpha = \frac{2}{N}\sqrt{\langle J_x\rangle^2 + \langle J_y\rangle^2} = \langle \cos\phi\rangle \ , \tag{2.2}$$

It is 1 for a coherent state and 0 for an incoherent state that can not be used for interferometry. Including this factor in the previously introduced squeezing factors leads to the most important squeezing quantity, the so-called useful (or coherent) spin squeezing factor [56]:

$$\xi_S = \frac{\Delta n}{(\sqrt{N}/2)\alpha} = \frac{\xi_n}{\alpha} \, , \tag{2.3}$$

In the following we will always use the coherent spin squeezing factor unless specifically stated otherwise. We refer to a state with a low squeezing factor as “highly squeezed”.

It is very simple to see how much precision improves when using the coherent spin squeezing ξ_S. A binomial state has $\xi_S = 1$ and a minimal phase error of $1/\sqrt{N}$, while using a squeezed state leads to a reduced error of $\xi_S/\sqrt{N}$ [59]. The lower limit of shot noise is the so-called Heisenberg limit $\Delta\phi = 1/N$. Another way of seeing how much precision improves is by giving the squeezing ξ_S^2 in dB. A squeezing value of 0.1 corresponds to $\xi_S^2 = 10\log(0.1^2) = -20$ dB and gives direct insight into the improvement.

Additionally to the improved precision, squeezed states have also other advantages. For example number squeezed states are very robust against dephasing effects [14]. Intuitively this can be understood on the Bloch sphere. Dephasing effects are stronger near the poles of the Bloch sphere and curl the distribution around the x-axis (see Fig. 2.8). A number squeezed state is mainly on or near the equator, where such effects are suppressed.

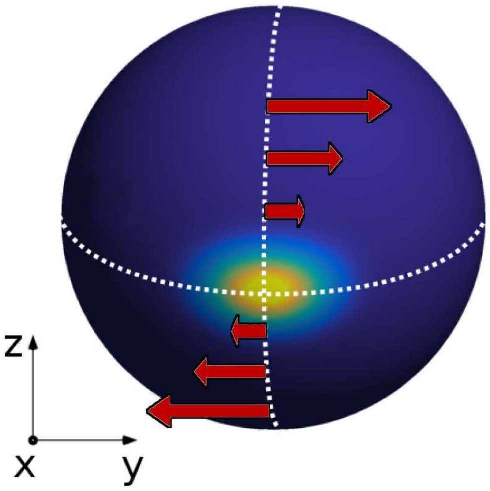

Figure 2.8.: Dephasing on the Bloch sphere: The distribution gets curled around the x-axis. As can be seen from the Hamiltonian (1.19) the nonlinear interaction is proportional to the square of J_z, which measures the atom number imbalance. The absolute atom number imbalance is zero on the equator of the Bloch sphere and increases towards the poles.

2.3.2. Production of squeezed states

There are many different ways of producing squeezed states. The simplest one is to split the trap quasi adiabatically from a single to a double well. This leads to reduced tunneling and the BEC becomes number squeezed. However, this process is relatively slow.

A method that can produce squeezed states much faster is the so-called two parameter optimization method [60, 61, 44]. Similar to adiabatic splitting the BEC is brought from a single to a double well state not in a linear fashion, but according to

$$\Omega(t) = \Omega_0 \left(1 - \frac{\Omega_c}{\Omega_0}\right) \exp\left(\frac{-t}{t_c}\right) + \Omega_c \,, \tag{2.4}$$

This means that there will be a splitting stage, starting with a tunnel coupling of Ω_0 at $t = 0$, followed by a stage where the tunnel coupling approaches the final value Ω_c. The ratio between those two stages is determined by the time t_c, which can be optimized analytically to produce the best squeezing [62].

A more sophisticated way to produce squeezed states is called parametric squeezing amplification. It is based on modulating the tunnel coupling at a certain resonance frequency to increase number or phase squeezing. In order to understand how parametric amplification works we first have to look at the Hamiltonian of the system [40, 41].

$$H = -\Omega\, J_x + 2\kappa\, J_z^2\,, \tag{2.5}$$

with the tunneling energy Ω and the charging energy κ. We can rewrite this Hamiltonian in terms of the particle imbalance n and the relative phase ϕ. J_z measures the particle imbalance $\langle J_z \rangle$, and J_x measures the relative phase, so they correspond to n and $\cos\phi$, respectively. We arrive at a model Hamiltonian

$$H = -\Omega\, \cos\phi + 2\kappa\, n^2 \tag{2.6}$$

At least in the coupled regime, where we expect very small ϕ, it is possible to approximate $\cos\phi$ as $1 - \phi^2/2$ and neglect higher orders of ϕ. This leads to the following, simplified Hamiltonian:

$$H = \frac{\Omega}{2}\, \phi^2 + 2\kappa\, n^2\,, \tag{2.7}$$

which looks very similar to the Hamiltonian of a harmonic oscillator

$$H_{\mathrm{HO}} = \frac{p^2}{2m} + \frac{1}{2}\, m\, \omega^2 x^2 \tag{2.8}$$

The only difference is that the momentum and position operators are replaced by ϕ and n. Indeed, if we look at the commutator relation of the spin operators we find that ϕ and n are canonically conjugate variables and obey $[\phi, n] = i$, much like the momentum and position operator of the harmonic oscillator [63]. Therefore we can describe a BEC in a double well similar to the way we describe a harmonic oscillator. We can even identify properties like the mass m and the frequency ω. If we compare (2.7) and (2.8) we find that $m = 1/\Omega$ and $\omega = 2\sqrt{\kappa\Omega}$. This frequency is identical to the Josephson frequency. Now that we know that our system is similar to a harmonic oscillator with frequency ω_J we can use this to produce squeezed states.

In general, the amplitude of a harmonic oscillator can by increased by two mechanisms. One way would be to simply apply a force and thereby produce a driven oscillator. An example for such a process would be pushing a child on a swing: the amplitude increases because of an applied force. However, this process is not well suited for a BEC system and we have to rely on a different method: The second way to increase an oscillators amplitude is by parametric amplification. It occurs when parameters of the oscillator (like frequency) change at twice the frequency of the oscillator. In our mechanical picture that corresponds to "pumping" on a swing. Also here the parameters of the system (i.e. the center of mass) are changed periodically and the amplitude increases without an external force. There are two main differences between a driven and a parametric oscillator. First of all parametric amplification requires non-zero amplitude to work.

Secondly, in order to get a resonance, the modulation of a parameter must occur at twice the natural frequency of the system.

This way of driving a harmonic oscillator can be used for BECs. It is only necessary to change the parameters of the system periodically, in particular its frequency ω. Since ω is directly related to the tunneling energy Ω, which in turn can be easily controlled by the distance between the two wells, we can indeed modulate it very easily. Also the initial amplitude that is needed for parametric amplification is no problem in the context of BECs. Simply changing the distance between the wells fast enough so that the BEC leaves the ground state induces a breathing mode, the BEC rotates on the Bloch sphere around the x-axis, changing from a number to a phase squeezed state. This small oscillation serve as a starting point for parametric amplification. Note that there is an important difference to the mechanical picture: In the mechanical picture the expectation values of some observables oscillate and become amplified, here we do not amplify the expectation values but the fluctuations. Fig. 2.9 shows parametric squeezing amplification on the Bloch sphere.

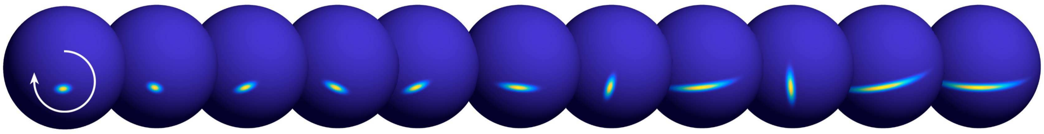

Figure 2.9.: Parametric squeezing amplification: The distribution rotates around the x-axis and the squeezing increases over time.

Details and results of this process are reported in chapter 7 as well as a way to use OCT in order to trap those squeezed states.

3. Optimal control theory

3.1. Introduction

Optimal control theory (OCT) [64, 65, 66] is a powerful mathematical tool that can be used for many different problems, like optimizing production processes or finding solutions to economic problems. Additionally OCT has many applications in physics. Interesting areas in which it was particularly successful are semiconductors [67], controlling atoms [68, 69] or quantum chemistry [70, 71, 72].

In general OCT can solve so-called inverse problems. The difference between an inverse and a conventional problem is what we are interested in. For conventional problems we know the system at the initial time T_0 and know how it develops over time (in form of some differential equations, boundary conditions or other information). We are interested in the state of the system at a terminal time T. For inverse problems the situation is different. Again we know the system at an initial time T_0, but we also have a desired state at time T. The question is now how the system transitions from one state to the other. In general inverse problems are more difficult to solve. They may have no solution at all or even more than one.

The main idea of OCT is to find a time development of the system that brings it as close as possible to the desired state. This time development is often described by one or more parameters that can be controlled externally, called "control" or "control field" λ. These parameters can be time dependent factors in the Hamiltonian of the system or more general concepts. In order to rate the success of each control mathematically, the so-called cost functional maps every control to a real number, called the "cost". A cost of 0 means that the desired and the achieved state are identical, higher costs mean that there is some kind of deviation. The actual form of the cost functional is different for each problem, and there are many different ways to define a cost functional for a given problem, although some are better suited than others.

In the following chapter the general structure of OCT is first shown, and several ways to implement OCT are then introduced. Finally the application to a real BEC problem is discussed for two OCT methods.

3.2. General structure

Suppose we have a dynamic variable $\Psi(t)$ governed by an equation $i\,\dot{\Psi} = F(\Psi)$ and a boundary condition $\Psi(0) = \Psi_0$. The goal of our optimization is now to bring Ψ to the desired state Ψ_{des} at a terminal time T. We can control the time development of Ψ due to a control field $\lambda(t)$. To rate the success of each control, a cost functional $\mathcal{J}$ is introduced. In general, $\mathcal{J}$ will depend on the dynamic variable Ψ and the control field λ. In order to guarantee that the equation of motion is satisfied at all times we use a Lagrange multiplier p.

$$\mathcal{L}(\Psi,\lambda,p) = \mathcal{J}(\Psi,\lambda) + p\left(i\dot{\Psi} - F(\Psi)\right) \tag{3.1}$$

Deriving this functional with respect to each function, and demanding that these derivatives vanish, leads to three equations that need to be satisfied simultaneously in order to find and optimal control. An OCT implementation for BECs will be discussed in 3.4. In general it is not possible to find such solutions analytically and one has to resort to numerical techniques.

3.3. Optimization methods

3.3.1. Gradient Ascend Pulse Engineering

One of the most widely used and versatile OCT algorithms is called "gradient ascend pulse engineering" (GRAPE). As the name suggests, it is based on calculating the gradient of a given control and then "ascend" towards a better (=lower) cost. It should be noted that GRAPE, like most other OCT methods, is only able to find local minima and in general it is hard to prove that such a minimum is a global minimum. However, especially in the field of quantum physics it can be shown that the minima of the control space have usually the same cost, and there are very few traps, i.e. local minima with a suboptimal cost [73].

The first step in using GRAPE is defining a control parameter λ and a cost functional $\mathcal{J}(\lambda)$. In the context of BECs, λ is often the shape or time-dependency of a magnetic field, but in general λ can describe anything that can be controlled and influences the system in some way. The cost functional $\mathcal{J}(\lambda)$ now maps every λ to a real number, i.e. the cost. A cost of 0 means that the goal or target is perfectly reached, higher costs indicate worse success. Naturally, the goal is to find a control that produces a cost that is as low as possible. In many cases a cost of exactly zero can not be achieved, however, one can always find a control that corresponds to a local minimum. This control is called optimal control.

In order to find this optimal control, one has to perform functional derivatives of $\mathcal{J}(\lambda)$ with respect to all the functions it depends on. Demanding that those derivatives equal zero leads to a set of equations that all need to be satisfied simultaneously. In general, finding a control that does that is not trivial. One can, however, start with an initial guess for λ and use the derivatives of $\mathcal{J}(\lambda)$ to find the gradient of the control λ. This gradient has information about the direction along which one has to change λ in order to improve it and find a better control. The new, improved control can in turn be used to obtain a gradient again. This iterative process can be continued until convergence is reached and an optimal control is found.

The crucial step in this process is updating the control. One can only calculate the gradient of the control field, the step size must be chosen by other means. There are many different ways to perform this update, the most common ones will be shown in the next sections.

Choosing a step size

The proper choice of the step size is very important for reaching fast convergence. A small step size makes convergence very slow, while a too big step size may even lead to a failure of the numerical minimization algorithm. In general, the optimal step size can change between two iteration steps by several orders of magnitude, which makes choosing a perfect step size even more difficult. One way to circumvent this problem is by determining the step size in each iteration step by a line search. In its simplest form the update can be written as

$$\lambda_{n+1} = \lambda_n + s\,\Delta\lambda_n\,, \tag{3.2}$$

with the gradient $\Delta\lambda_n$ and the step size s. The line search now consists of using different s, calculating the cost of the resulting control, and finding a minimum of this cost with respect to s. This optimal step size is then used to calculate a new control and in turn a new gradient. For the next iteration step a new line search is performed that most likely results in a different optimal step size. This method can be very time consuming, especially if evaluating the cost of a given control is computationally difficult. On the other hand the optimization of the step size via a line search can be done with arbitrary precision and a well optimized step size can drastically decrease the number of iterations that are needed for convergence.

Steepest descend

The simplest method of updating the control field λ is “steepest descend”. The idea behind it is to always change λ in the direction of the largest gradient. For a simple 2 dimensional control space this method works very well, as shown schematically in Fig. 3.1. However, for a control space of higher dimensionality this method is not well suited. Especially if the control space has complicated features, like flat areas or saddle points, steepest descend works very poorly. Implementation of this method is, however, very simple and also computationally cheap.

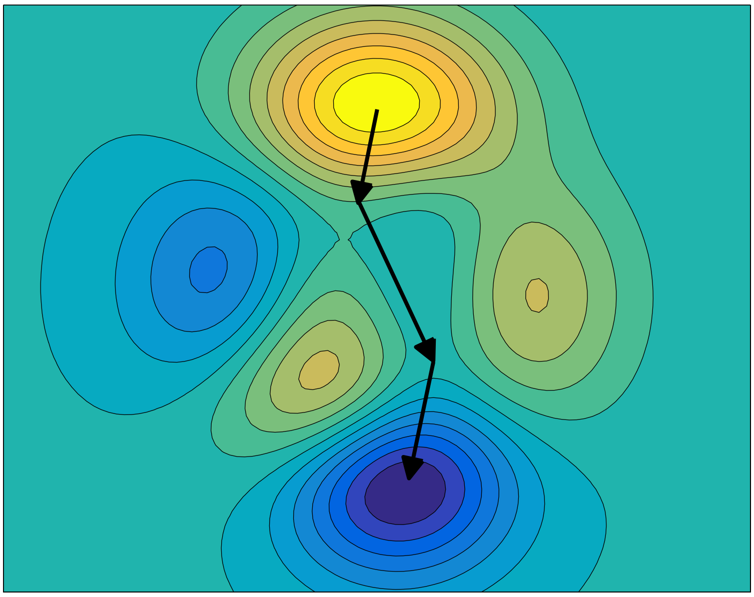

Figure 3.1.: Steepest descend for a simple 2d control space. The minimum is found by always updating the control in the direction of the largest gradient. For simple problems this technique is very successful.

Conjugate gradient

A more sophisticated way to update λ is using the conjugate gradient method [74]. It is similar to steepest descend but tries to avoid going in the same direction twice. This is done by utilizing the previously used update directions. Conjugate gradient converges faster than steepest descend, as illustrated in Fig. 3.2.

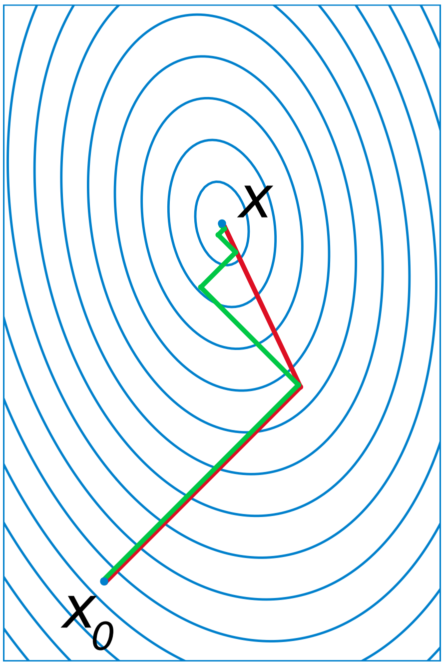

Figure 3.2.: Steepest descend (green) needs 5 steps to find the minimum. By utilizing information from the first search direction, conjugate gradient (red) can find a better update direction than the direction with the largest gradient, and can find the minimum in only 2 steps.

While this method works very well with many OCT problems it has similar problems with high dimensional, complicated control spaces. For some problems the control space has a shape similar to a half pipe, where the direction of the steepest descent is nearly orthogonal to the direction of the minimum. Especially for high-dimensional problems the conjugate gradient method can converge very slowly.

BFGS

The Broydon-Fletcher-Goldfarb-Shanno (BFGS) method [75, 76, 77, 78, 79] solves the aforementioned problems by calculation of additional information from the gradients collected during the iterative process. In particular, BFGS approximates the Hessian. This information about the second derivative of the control speeds up convergence significantly. Especially for complicated control spaces the second derivative can help finding an update direction that is much better than steepest descend. Half pipe shaped areas do not pose a problem when using BFGS and also flat regions of the control space can be traversed quickly. The difference in convergence between BFGS and conjugate gradient can be drastic. Fig. 3.3 shows such an example, taken from the optimization of a BEC process (for a detailed discussion see below). The conjugate gradient method (red) seems to hit a plateau after around 20 iterations and finds a control with a cost of roughly 0.1. BFGS on the other hand finds a control that has a 100 times lower cost in the same amount of iterations.

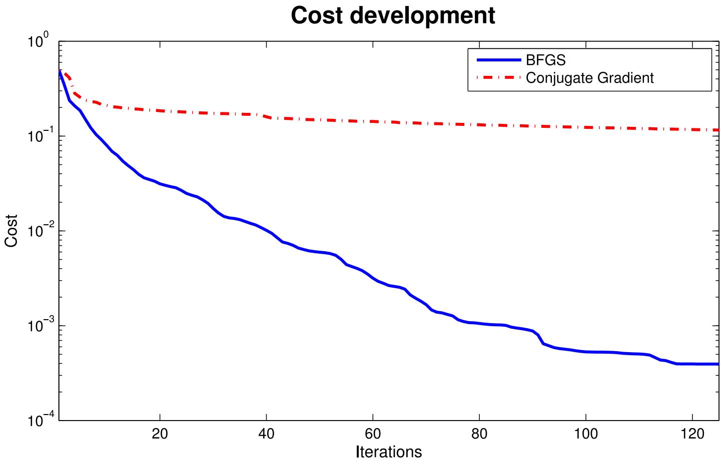

Figure 3.3.: BFGS can perform much better than conjugated gradient in the context of BECs. The optimization problem of this calculation is similar to the one presented in paper 1 [51].

BFGS calculates the approximated Hessian B in the following way [80]: We start by obtaining a search direction d_i from the control equation and an optimal step size α_i from a line search. λ is then updated according to

$$\lambda_{i+1} = \lambda i + \alpha_i \, d_i \tag{3.3}$$

Then we need to define two auxiliary variables: s_i, the last modification of λ, and y_i, the difference between the last two gradients.

$$s_i = \alpha_i \, d_i \tag{3.4}$$

$$y_i = \Delta\lambda_{i+1} - \Delta\lambda_i \tag{3.5}$$

With these auxiliary variables a new approximation for the Hessian can be calculated from the approximation from the last step. At the first step, where no previous Hessian is available, one can take $B = 1$ as a starting point. The update is then performed in the following way:

$$B_{i+1} = B_i + \frac{y_i \, y_i^T}{y_i^T \, s_i} - \frac{B_i \, s_i \, s_i^T \, B_i}{s_i^T \, B_i \, s_i} \tag{3.6}$$

With this method the approximation of the Hessian gets better at each iteration step and even extremely flat regions of the control space can be crossed rather quickly. This is the reason why BFGS is the method of choice for most BEC optimization problems.

3.3.2. Krotov's method

A different approach to optimization is Krotov's method [81, 82, 83]. Similar to GRAPE this method begins with defining a control λ and a cost functional $\mathcal{J}$ that maps every control to a cost. To this cost functional an additional term, the Krotov term, is added. This must be done in such a way, that the position of the minimum does not change. If done correctly, the new term makes it possible to divide the cost functional into two parts that can be optimized separately. Now one can devise a scheme that improves the control, always leading to a lower cost. This control can in turn by improved by the same scheme until convergence is reached. Since it is guaranteed that the control improves each iteration step, Krotov's method is monotonically convergent.

The update with Krotov's method is sequential, which means that every point in time gets updates separately, in contrast to a concurrent algorithm, like GRAPE. On the one hand this can be an advantage, since the updated points are used right away for the calculation of the next update, in a way that more recent information is available. On the other hand, there is less information available about the control as a whole.

There are several parameters on which Krotov's method relies, which can be changed in order to improve performance. For example the step size of the Krotov update can be finetuned in order to speed up convergence. Also the Krotov term has three parameters, α, β and γ [83], which must be chosen in some way and these parameters often influence convergence drastically. Additionally, a shape function can be introduced in order to produce controls that have special features, like a slow switching on or off of the control fields. In general, choosing all these parameters is not trivial, but finetuning them has the potential to lead to very fast convergence.

3.3.3. CRAB

Another interesting optimization method is called "Chopped Random Basis" (CRAB) [84]. While the control field is usually discretized in time for GRAPE and Krotov's method, CRAB expands the control in a function basis characterized by some parameters Ω_j:

$$\lambda_j = \sum_k c_j^k \hat{\lambda}_j^k(\Omega_j^k) \tag{3.7}$$

This expansion is "chopped" in the sense that there is only a finite number of basis functions and "random", meaning that the basis functions get randomized according to

$$\hat{\lambda}_j^k \to \hat{\lambda}_j^k[\Omega_j^k(1 + r_j^k)]\,, \tag{3.8}$$

where r_j^k is a random number. It can be shown that this form of randomization breaks the orthonormality of the functions $\hat{\lambda}_j^k$, but on the other hand leads to faster convergence [84].

The problem now reduces to a multivariable optimization of the cost function, which can be solved by methods like steepest descend, conjugate gradient or even direct search methods [85].

Advantages and disadvantages

The big advantage of CRAB is that the control space has a much lower dimensionality than the control space of GRAPE or Krotov's method. This makes convergence faster and reduces the computational effort. A downside of this simplification is that it is also very restrictive in terms of allowed control fields. The choice of basis functions is crucial for the success and the final cost of this optimization technique, and without prior knowledge about the system it is often very difficult to find well suited basis functions.

3.3.4. Genetic algorithms

Another approach for solving OCT problems are so-called genetic algorithms [86]. Hereby one tries to use the same principles as biological evolution in order to find good controls. First of all a population of n random controls is generated. In principle it is possible to use discretized control fields, but very often it is more convenient to parametrize them in some form, e.g. by expanding them in base functions. The next step is calculating the cost for each of these controls. The controls with the lowest costs are then used to produce controls for the second population. There are many different ways of doing so. One example would be to choose the parameters for a new control randomly from a pair of good controls or combining the parameters of good controls in some other way. This generates a second generation, again with a population of n, which ideally now has a lower average cost than the last population. They are in turn tested for their fitness and the best ones are used to produce a third generation. This can be repeated as often as necessary, as depicted in Fig. 3.4.

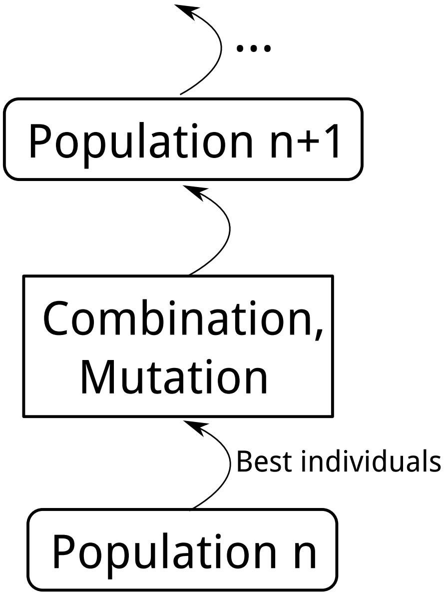

Figure 3.4.: Genetic algorithm: The best individuals of each population are combined to produce a new population.

Advantages and disadvantages

Genetic algorithms have problems similar to CRAB. Parametrizing the control fields excludes many controls from the control space, possibly those with low costs. Discretization on the other hand leads to many parameters for each control, which poses a big problem for genetic algorithms and slows down or even inhibits convergence. The second problem is that this approach relies heavily on randomly generated numbers and controls. For problems where randomly generated controls nearly always lead to high costs it is very difficult to produce controls with low cost by means of a genetic algorithm.

3.4. OCT applied to BECs

3.4.1. Gradient Ascend Pulse Engineering

The main ingredients one needs for GRAPE are a control, a cost functional and the equations of motion that govern the system. For BECs the control parameter λ often describes a magnetic field that can be controlled externally. Additionally to this control we also define a cost functional. Suppose that the problem we are interested in is how to bring a BEC from its initial state Ψ_0 at time 0 to a desired state Ψ_{des} at time T. The cost functional can then simply be defined as

$$\mathcal{J} = \frac{1}{2} \, ||\Psi(T) - \Psi_{\mathrm{des}}||^2 \, . \tag{3.9}$$

This would yield $\mathcal{J} = 0$ for $\Psi_0 = \Psi_{\text{des}}$ and $\mathcal{J} > 0$ for $\Psi_0 \neq \Psi_{\text{des}}$. However, this cost functional has disadvantages. Suppose we find a state $\Psi(T)$ that only differs from Ψ_{des} by a global phase. The cost functional would still not yield 0, although the observables $|\Psi_{\text{des}}|^2$ and $|\Psi(T)|^2$ would match perfectly. To remedy that we can use a different cost functional in the form of

$$\mathcal{J} = \frac{1}{2}\left(1 - |\langle\Psi(T)|\Psi_{\text{des}}\rangle|^2\right) . \tag{3.10}$$

It is apparent that there are many different ways to define a cost functional and some thought is required to find a well suited one for the problem at hand. Especially for problems that have an experimental application it is often a good idea to include a term in the cost functional that favors smooth control parameters, since they are often easier to implement experimentally. One way of including such a term is

$$\mathcal{J} = \frac{1}{2}\left(1 - |\langle\Psi(T)|\Psi_{\text{des}}\rangle|^2\right) + \frac{\nu}{2}\int_0^T \dot{\lambda}(t)^2 dt\,. \tag{3.11}$$

ν serves as a weighting parameter and is chosen very small (e.g. 10^{-9}) in order to ensure that the cost value is primarily determined by Ψ_{des}.

The next step in order to construct an optimality system is to include the equations of motion in the cost functional. For the example of a BEC this could be the Gross-Pitaevskii equation [33, 34, 35] or the more complicated MCTDHB equations [45]. The equations of motion are then included via Lagrange multipliers. For the Gross-Pitaevskii equation

$$i\dot{\Psi} = \left(-\frac{1}{2}\nabla^2 + V(\lambda) + \kappa|\Psi|^2\right)\Psi \tag{3.12}$$

the full Lagrangian is given by

$$\mathcal{L} = \mathcal{J} + \text{Re}\left\langle p, \left(-\frac{1}{2}\nabla^2 + V(\lambda) + \kappa|\Psi|^2\right)\Psi\right\rangle \tag{3.13}$$

Here $\langle a, b\rangle$ is shorthand notation for $\int dx \int_0^T dt\ a^* b$ and p is the Lagrange multiplier. The full cost functional given in (3.13) depends on three functions, namely the wave function Ψ, the control parameter λ, and the Lagrange multiplier p:

$$\mathcal{L} = \mathcal{L}(\Psi, \lambda, p) \tag{3.14}$$

In order to find the minimum of this functional we have to perform functional derivatives with respect to each variable and set them zero, similar to the procedure of finding the minimum of a function. So we are looking for the three derivatives $\frac{\partial\mathcal{L}}{\partial\Psi}$, $\frac{\partial\mathcal{L}}{\partial\lambda}$ and $\frac{\partial\mathcal{L}}{\partial p}$. These three derivatives lead to three equations. $\frac{\partial\mathcal{L}}{\partial p}$ yields the equation of motion, in this case the Gross-Pitaevskii equation. $\frac{\partial\mathcal{L}}{\partial\Psi}$ yields the so-called adjoint equation, an equation for the Lagrange multiplier p that runs backwards in time, i.e. it starts at time T. This has to do with the boundary condition (3.15g) in which the value of p at time T depends only on $\Psi(T)$. The last equation comes from $\frac{\partial\mathcal{L}}{\partial\lambda}$. It is the so-called control equation and gives the gradient of the control λ as a function of $\Psi(t)$ and $p(t)$. The derivation of this optimality system can be found in appendix A, and we obtain:

$$i\dot{\Psi} = \left(-\frac{1}{2}\nabla^2 + V_\lambda + \kappa\left|\Psi\right|^2\right)\Psi \tag{3.15a}$$

$$\dot{p} = \left(-\frac{1}{2}\nabla^2 + V_\lambda + 2\kappa|\Psi|^2\right)p + \kappa\Psi^2 p^* \tag{3.15b}$$

$$\gamma\ddot{\lambda} = -\operatorname{Re}\left\langle p\left|\frac{\delta V_\lambda}{\delta\lambda}\right|\Psi\right\rangle \tag{3.15c}$$

$$\lambda(0) = \lambda_0 \tag{3.15d}$$

$$\lambda(T) = \lambda_T \tag{3.15e}$$

$$\Psi(0) = \Psi_0 \tag{3.15f}$$

$$ip(T) = -\left\langle\Psi_{\text{des}}\left|\Psi(T)\right.\right\rangle\Psi_{\text{des}} \tag{3.15g}$$

If all three equations of the optimality system yield zero we have found a minimum of the cost functional. However, the control equation has another useful property: If the Gross-Pitaevskii equation and the adjoint equation are both satisfied, the control equation gives us the gradient of λ, i.e. the direction along which we need to change the control in order to improve it. This information is crucial for GRAPE. With these equations, namely the equation of motion, the adjoint equation and the control equation, the GRAPE algorithm works like this:

1. Start by guessing an initial control field $\lambda(t)$
2. Solve the equation of motion to obtain $\Psi(t)$
3. Solve adjoint equation to obtain $p(t)$
4. Solve control equation to obtain the gradient of λ
5. update λ
6. go to 2 until convergence is reached

This process is also depicted in Fig. 3.5.

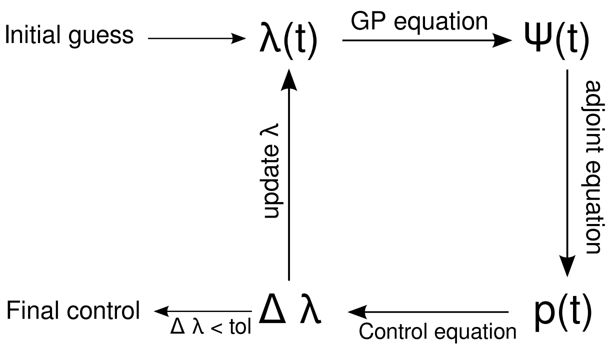

Figure 3.5.: Schematic of the GRAPE algorithm

This algorithm allows to find a suitable control for many different problems. An important part hereby is step 5, the update of λ. In principle many of the methods described in section 3.3.1 can be used to perform this update, however, BFGS usually leads to the best results. Further improvements can be made by adapting the cost functional, as shown in the next section.

H1-norm

Let us review the cost functional (3.11).

$$\mathcal{J} = 1 - |\langle\Psi(T)|\Psi_{\text{des}}\rangle|^2 + \nu \int_0^T \dot{\lambda}(t)^2 dt \tag{3.16}$$

The last term can be interpreted as the norm of the derivative of the control parameter λ. Ordinarily we use the L_2 norm.

$$\langle\lambda|\lambda\rangle_{L_2} = \int_0^T \lambda^* \, \lambda \, dt \tag{3.17}$$

However, this is not the only possibility and other choices can be more physical. For example we can choose λ to be a H_1 function instead.

$$\int_0^T \dot{\lambda}^2 dt = \langle\dot{\lambda}|\dot{\lambda}\rangle_{L2} = \langle\lambda|\lambda\rangle_{H1} \tag{3.18}$$

This has no impact on the cost functional, but changes its interpretation and the way we numerically calculate the update of λ.

$$\begin{aligned} \nabla_\lambda \mathcal{J} &= -\gamma\ddot{\lambda} - \text{Re}\langle p|\frac{\partial V}{\partial \lambda}|\psi\rangle \text{ for } L^2 \text{ norm} && (3.19) \\ -\frac{d^2}{dt^2}[\nabla_\lambda \mathcal{J}] &= -\gamma\ddot{\lambda} - \text{Re}\langle p|\frac{\partial V}{\partial \lambda}|\psi\rangle \text{ for } H^1 \text{ norm} && (3.20) \end{aligned}$$

Using the H_1 norm is equivalent to solving an additional Poisson equation. The interpretation of $\int_0^T \dot{\lambda}(t)^2 dt$ as H_1 norm is still mathematically valid [87] and has several advantages. First of all, convergence can be much faster. Secondly, because of the additional Poisson equation, the resulting control fields are very smooth. This is good for an experimental realization, since smooth controls are easier to implement. For most BEC optimization problems it is beneficial to use the H_1 norm in this case.

3.4.2. Krotov's method

Applying Krotov's method to the previous example of bringing a BEC from a initial state into a desired state again begins with defining a cost functional. In general, the cost functional for Krotov's method includes a vanishing term Φ that depends on the Krotov function σ, defined in (3.24).

$$\mathcal{J} = J_T\left(\Psi(T)\right) + \Phi(\Psi(T), T) - \Phi(\Psi_0, 0) - \int_0^T R(\Psi, \lambda, t) dt \tag{3.21}$$

A possible way to define that cost functional in the context of BECs would be

$$\mathcal{J} = \frac{1}{2}\left(1 - |\langle \Psi(T) | \Psi_{\text{des}} \rangle|^2\right) + \int_0^T \frac{[\lambda(t) - \lambda_{\text{ref}}(t)]^2}{S(t)} dt\,. \tag{3.22}$$

Here λ_{ref} is a reference field, typically chosen to be the control field of the previous iteration. This ensures small changes during one iteration step while the contribution of this term decreases over the optimization process, as the difference of controls becomes smaller and finally nearly vanishes if convergence is reached. $S(t)$ is a so-called shape function that is used to enable a smooth turning on and off of the control.

With this cost function it is possible to derive Krotov's equation for updating the control field:

$$\lambda^{(i+1)} = \lambda^{(i)} + S(t)\text{Re}\Big[\langle p^{(i)} | \frac{\partial V}{\partial \lambda} | \psi^{(i+1)} \rangle + \frac{\sigma(t)}{2i} \langle \Delta\psi | \frac{\partial V}{\partial \lambda} | \psi^{(i+1)} \rangle\Big] \tag{3.23}$$

From this equation it is apparent why Krotov's method is sequential, in contrast to GRAPE. Ψ^{i+1} appears on the right hand side of the equation, so for a given time t, the update depends on all updates at earlier times. An important part of the update equation is the σ term [83, 82]. It is needed to ensure convergence, but is very difficult to choose. In general σ has the form

$$\sigma(t) = \alpha(e^{\gamma(t-T)} - 1) + \beta\,, \tag{3.24}$$

where α, β and γ are the so-called Krotov parameters. Suitable Krotov parameters need to be found for each individual problem. Too high values for the parameters lead to very slow convergence, while too small values lead to an increase in cost from one step to the next, which makes optimization impossible. There are different ways to find well suited Krotov parameters. It is possible to find bounds through analytical calculations [83], but sometimes it is more convenient to simply try to start with small values and increase them until the cost starts to decrease. For some problems σ can be neglected altogether. Once σ is chosen properly, the optimization can begin.

The algorithm for Krotov's method is different from the one of GRAPE (see Fig. 3.6):

1. Start by guessing an initial control field $\lambda(t)$
2. Solve the equation of motion to obtain $\Psi(t)$
3. Solve adjoint equation to obtain $p(t)$
4. Solve equation of motion simultaneously with the equation for the new control (3.23)
5. go to 3 until convergence is reached

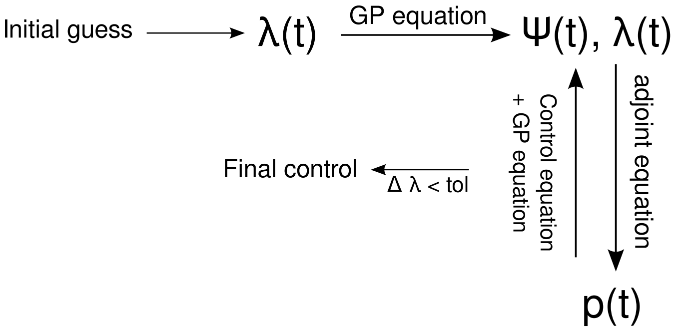

Figure 3.6.: Schematic of Krotov's method

Also Krotov's method is well suited for all kinds of problems in the field of BECs and a detailed comparison between the optimization methods can be found in chapter 6 and section 8.1.

4. Paper 1: Vibrational state inversion of a Bose–Einstein condensate: optimal control and state tomography

Published in: Journal of Physics B: Atomic, Molecular and Optical Physics **46** (2013)

IOP Publishing — Journal of Physics B: Atomic, Molecular and Optical Physics

J. Phys. B: At. Mol. Opt. Phys. **46** (2013) 104012 (16pp) doi:10.1088/0953-4075/46/10/104012

Vibrational state inversion of a Bose–Einstein condensate: optimal control and state tomography

Robert Bücker[1], Tarik Berrada[1], Sandrine van Frank[1], Jean-François Schaff[1], Thorsten Schumm[1], Jörg Schmiedmayer[1], Georg Jäger[2], Julian Grond[2] and Ulrich Hohenester[2]

[1] Vienna Center for Quantum Science and Technology, Atominstitut, TU Wien, Stadionallee 2, A-1020 Vienna, Austria
[2] Institut für Pysik, Karl–Franzens–Universität Graz, Universitätsplatz 5, A-8010 Graz, Austria

E-mail: schmiedmayer@atomchip.org

Received 17 December 2012, in final form 31 January 2013
Published 9 May 2013
Online at stacks.iop.org/JPhysB/46/104012

Abstract

We present theoretical and experimental results on high-fidelity transfer of a trapped Bose–Einstein condensate into its first vibrationally excited eigenstate. The excitation is driven by mechanical motion of the trap, along a trajectory obtained from optimal control theory. Excellent agreement between theory and experiment is found over a large range of parameters. We develop an approximate model to map the dynamics of the many-body condensate wave function to a driven two-level system.

(Some figures may appear in colour only in the online journal)

1. Motivation

The precise control over quantum systems represents a major challenge in modern physics. Successful implementation of quantum technologies may lead to the construction of devices such as quantum simulators, quantum cryptography devices, and quantum computers. For such applications, one needs to produce arbitrary quantum states, e.g. strongly entangled many-body states, or states which are far from thermal equilibrium or the ground state of the system.

In this paper we report on highly efficient preparation of a non-classically excited motional state of a Bose–Einstein condensate (BEC), by a modulation of the trapping potential, as obtained from optimal control theory (OCT) [1]. Fast changes of the potential are routinely used in BEC laboratories, for instance as ways to probe the gas by exciting collective excitations [2, 3], or to displace the samples for further manipulation. Recently, controlled modulations of the trapping potential were achieved in order to quickly displace or deform BECs while keeping them in their ground state. These ‘shortcuts to adiabaticity’ [4–9] take advantage of the self-similar dynamics of interacting BECs trapped in time-dependent harmonic potentials [10, 11].

For more general desired states, like excited stationary states [12], for which no exact solutions are found, one needs to use numerical methods such as OCT, which provides a framework for determining optimized driving sequences to actively manipulate a quantum system. One typical task is the precise control of spin evolution in an NMR spectroscopy experiment [13–15]. Other exciting applications are the shaping of laser pulses for controlling chemical reactions [16–19], or optimizing solid-state quantum gates [20–25]. For rapid quantum state transitions, the *quantum speed limit* [26–28] may be approached, which sets the upper limit to how fast a system can evolve in its Hilbert space. This may be highly relevant for quantum computation, where fast gate times are required to outrun decoherence.

In the context of ultracold atoms, such approaches were investigated theoretically for the splitting of BECs [29], to optimize the transport of atoms in optical lattices for quantum gate operations [30], cooling [31], generation and storage of entanglement [32, 33] or number-squeezed states [34].

Here, we aim for a vibrational state inversion, where the entire population of the condensate is transferred to the first excited state of its motional degree of freedom. Such an inverted state can be used as a source for the amplified

0953-4075/13/104012+16$33.00

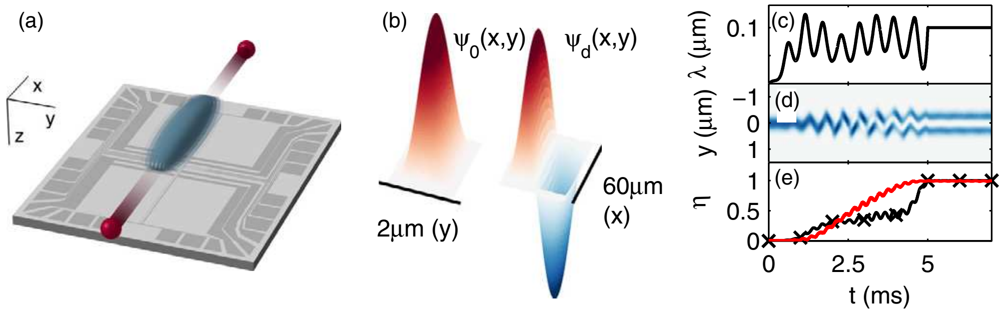

Figure 1. Schematic representation of our excitation scheme. (a) Illustration of the sequence: an elongated Bose–Einstein condensate (blue) is trapped on an atom chip (chip and condensate aspect ratio not drawn to scale). Mechanical motion along the y-axis pumps the gas into a vibrationally excited state, which decays by directed emission of twin-atom beams along x (red). (b) Initial and excited condensate wave functions in the (x, y)-plane. The state $\psi_d(x, y)$ (desired state) is the first excited state of the Gross–Pitaevskii equation (1) along y. (c) Trajectory $\lambda(t)$ of the trap minimum along y. (d) Time-dependent density $n(y, t) = |\psi(y, t)|^2$ of the condensate wave function under the influence of the excitation process. (e) Population (simulated) of the excited state $\psi_d(y)$ derived from wave-function overlap (black, with markers) and using the two-mode model as introduced in section 4.3 (red, solid).

emission of matter-wave twin beams [35, 36], similar to a pumped gain medium in a laser or an optical parametric amplifier (figure 1(a)). We start from a condensate in the ground state along the strongly confined (transverse) directions of an elongated trapping potential. We then use OCT on a controlled displacement of the trap centre (transverse 'shaking' of the cloud, figure 1(c)), in order to transfer the BEC to the first antisymmetric stationary state as given by the Gross–Pitaevskii equation (GPE), which is governing the system's dynamics (figure 1(b)). The efficiency of this process is close to 100%, which corresponds to the desired vibrational inversion of the atomic cloud. Since a harmonic potential (where all energy levels are equidistant) would not allow the transfer to an excited stationary state by means of displacement [37–41], we use here an anharmonic potential [25, 23, 41] generated by a radio-frequency (RF) dressed magnetic trap [42]. In the limit of vanishing atom–atom interactions, the final state would simply correspond to all atoms residing in the first excited Fock state of the trap (quantum numbers $n_x = 0, n_y = 1, n_z = 0$). However, in our many-body wave function inter-atomic interactions are an essential ingredient of the system's dynamics and cannot be neglected in the optimization. In the experiment they manifest themselves in energy shifts due to the atomic mean-field, and a decay of the excited state by means of inelastic two-body scattering [35, 36].

Another aspect we will address is the interpretation of our results beyond a simple comparison of calculated and measured wave function dynamics. While such a comparison benchmarks the accuracy to which experiments and theory are matched, it provides only limited insight into the nature of the excitation mechanism, and the structure of the quantum state during and after the excitation process. To this end, we will deduce an approximate description, that allows to map the many-body wave function in a weakly anharmonic confinement to a driven two-level system, where the excitation process corresponds to a π-pulse that transfers all population to the excited state.

The paper is structured as follows: section 2 presents the theoretical description of the problem and the OCT algorithm used to obtain the excitation trajectory of the trap centre, section 3 details the experimental implementation, and finally, section 4 discusses the results, with an emphasis on how the behaviour of key observables can be captured by a two-level model. To our knowledge, this excitation sequence represents the first successful use of OCT for the preparation of exotic many-body states of BECs.

2. Optimal control theory

OCT is a mathematical tool that allows an optimal control sequence for a given control problem to be determined [43, 44, 1]. In the following we review the basic ingredients of OCT, taking the example of the shaking process which brings the condensate from the ground state to its first vibrationally excited eigenstate. Our analysis closely follows the presentation given in [29, 36].

As will be discussed in section 3, the condensate can be described by a one-dimensional GPE for the transverse coordinate y, along which the condensate is displaced, according to[3]

$$\mathrm{i}\frac{\partial\psi(y,t)}{\partial t} = \left(-\frac{1}{2}\frac{\partial^2}{\partial y^2} + V_\lambda(y,t) + g\,|\psi(y,t)|^2\right)\psi(y,t). \tag{1}$$

Furthermore, g is a nonlinearity parameter accounting for the repulsive atom–atom interactions [29, 34]. The anharmonic confinement potential $V_\lambda(y, t) = V_6(y - \lambda(t), 0)$ (see section 3.1) follows a control parameter $\lambda(t)$, in our case the displacement of the potential minimum, and provides the means for exciting the condensate.

One might wonder how a mere displacement of the symmetric confinement potential $V_6(y)$ can induce transitions between states of opposite parity. This becomes obvious by performing a unitary transformation $\exp[\mathrm{i}\,\lambda(t)\,P]$, where P is

---

[3] In this section, we will use dimensionless coordinates, which are matched with the typical scales of the problem; lengths are given in units of $l_0 \equiv 1\,\mu$m, times in units of $\tau_0 \equiv m l_0^2/\hbar \approx 1.37$ ms, and energies in units of $\hbar/\tau_0 \approx \hbar \cdot 0.73\,\mathrm{ms}^{-1}$, where m is the mass of the ^{87}Rb atoms.

the momentum operator along y, to obtain a frame of reference that moves with the potential minimum at $\lambda(t)$. In the shifted frame, the GPE (1) contains a static potential, but acquires another term $-\dot{\lambda}(t)P$ (in the right-hand side parentheses), which is linear in the momentum operator P and proportional to the velocity of the trap minimum. This additional term breaks the inversion symmetry and thus induces transitions between states of opposite parity.

The objective of the control problem can now be formulated as follows. Let λ_0 be the control parameter at the initial time $t = 0$, and λ_1 the control parameter at the final time $t = T$ of the control process. Likewise, we denote the initial ground state of the GPE with $\psi_0(y)$ and the *desired* final wave function (in our case the first excited state of the GPE in the anharmonic trap) with $\psi_d(y)$. OCT then seeks for the optimal time variation of $\lambda(t)$ that brings the final wave function as close as possible to the desired state ψ_d.

To gauge the success of the excitation process for a given control field $\lambda(t)$, we define a cost function

$$J(\psi(T),\lambda) = \frac{1}{2}[1 - |\langle\psi_d|\psi(T)\rangle|^2] + \frac{\gamma}{2}\int_0^T [\dot{\lambda}(t)]^2\,\mathrm{d}t. \tag{2}$$

The first term of the cost function becomes minimal when the final wave function precisely matches the desired wave function, apart from a global (irrelevant) phase. The second term favours smooth control fields and is needed to make the OCT problem well posed [45]. γ is a parameter that weights the relative importance of the two control objectives of smooth control fields and of wave function matching. As our experimental implementation allows fast and precise control of $\lambda(t)$, the parameter γ can be set such that the control penalization is always much smaller than the first term in equation (2). OCT is now seeking for an 'optimal control' that minimizes the cost function $J(\psi(T),\lambda)$, under the condition that the final wave function $\psi(T)$ has to be obtained from the GPE (1) with the initial wave function $\psi_0(y)$. To turn this constrained minimization problem into an unconstrained one, within the OCT framework one introduces a Lagrange function

$$L(\psi,p,\lambda) = J(\psi,\lambda) + \Re e\int_0^T \left\langle p \left| \mathrm{i}\frac{\partial\psi}{\partial t} - \left(-\frac{1}{2}\frac{\partial^2}{\partial y^2} + V_\lambda + g|\psi|^2\right)\psi \right.\right\rangle \mathrm{d}t,$$

where the adjoint function $p(y,t)$ acts as a generalized Lagrange parameter. Here and in the following we will, for the sake of brevity, often omit parameters y and t. At the minimum of $J(\psi,\lambda)$ the Lagrange function has a saddle point, where all three derivatives $\delta L/\delta\psi$, $\delta L/\delta p$ and $\delta L/\delta\lambda$ must vanish. Performing the usual functional derivatives, we obtain after some variational calculation the following optimality system:

$$\mathrm{i}\frac{\partial\psi}{\partial t} = \left(-\frac{1}{2}\frac{\partial^2}{\partial y^2} + V_\lambda + g|\psi|^2\right)\psi \tag{3a}$$

$$\mathrm{i}\frac{\partial p}{\partial t} = \left(-\frac{1}{2}\frac{\partial^2}{\partial y^2} + V_\lambda + 2g|\psi|^2\right)p + g\,\psi^2\,p^* \tag{3b}$$

$$\gamma\ddot{\lambda} = -\Re e\left\langle\psi\left|\frac{\partial V_\lambda}{\partial\lambda}\right|p\right\rangle, \tag{3c}$$

which has to be solved together with the initial condition $\psi(0) = \psi_0$, as well as with the constraints on the control field $\lambda(0) = \lambda_0$ and $\lambda(T) = \lambda_1$. To obtain the equation for the adjoint function p, we have performed an integration by parts for the term involving the time derivative of ψ prior to working out the functional derivative $\delta L/\delta\psi$. This procedure gives, in addition to equation (3*b*), the terminal condition

$$\mathrm{i}\,p(y,T) = -\langle\psi_d|\psi(T)\rangle\,\psi_d(y). \tag{4}$$

Quite generally, the Lagrange parameter determines the sensitivity of the system with respect to the external control. In our case, the dynamic equation (3*b*) describes the propagation of fluctuations around the Gross–Pitaevskii solution and is closely related to the usual Bogoliubov–de Gennes equations [46].

In most cases of interest it is impossible to guess $\lambda(t)$ such that equations (3*a*)–(3*c*) are fulfilled simultaneously, and one has to employ a numerical solution scheme. Suppose that $\lambda(t)$ is some guess for a viable control field. We can now solve equation (3*a*) forward in time to obtain the final wave function $\psi(T)$, which, in turn, allows us to compute the adjoint function $p(T)$ from equation (4). In the ensuing step, the time evolution of $p(t)$ is solved backwards in time. Since $\lambda(t)$ is not the optimal control, equation (3*c*) is no longer fulfilled. However, the functional derivative

$$\frac{\delta L}{\delta\lambda} = -\gamma\ddot{\lambda} - \Re e\left\langle\psi\left|\frac{\partial V_\lambda}{\partial\lambda}\right|p\right\rangle \tag{5}$$

provides us with a search direction for $\lambda(t)$. Adding a fraction of $\delta L/\delta\lambda$ to $\lambda(t)$ leads to a control that performs better and brings the final wave function $\psi(T)$ closer to the desired one. The improved control field is then used in the next iteration. In our simulations we typically perform a time discretization of the interval $[0,T]$ and use a generic optimization routine, such as the nonlinear conjugate gradient [47] or a quasi-Newton method, together with equation (5) for computing the appropriate search directions. One shortcoming of equation (5) is that in general $\delta L/\delta\lambda$ does not vanish at the boundary points of the time interval, although the control field is fixed to the values of λ_0 and λ_1 there. To overcome this problem, one rewrites the penalization term of the control field $(\gamma/2)\,(\dot{\lambda},\dot{\lambda})_{L^2}$ as $(\gamma/2)\,(\lambda,\lambda)_{H^1}$, where the definition of the H^1 inner product is $(u,v)_{H^1} = (\dot{u},\dot{v})_{L^2}$ [48]. It is important to realize that this different norm does neither affect the value of the cost function nor the Gross–Pitaevskii or adjoint equations. However, it does affect the equation for the control field in the case of a non-optimal $\lambda(t)$, which now satisfies a Poisson equation

$$-\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\delta L}{\delta\lambda} = -\gamma\frac{\mathrm{d}^2\lambda}{\mathrm{d}t^2} - \Re e\left\langle\psi\left|\frac{\partial V_\lambda}{\partial\lambda}\right|p\right\rangle. \tag{6}$$

The advantages of equation (6) are that the boundary conditions for $\lambda(t)$ are automatically fulfilled and that changes due to large values of the second term on the right-hand side are distributed, through the solution of the Poisson equation,

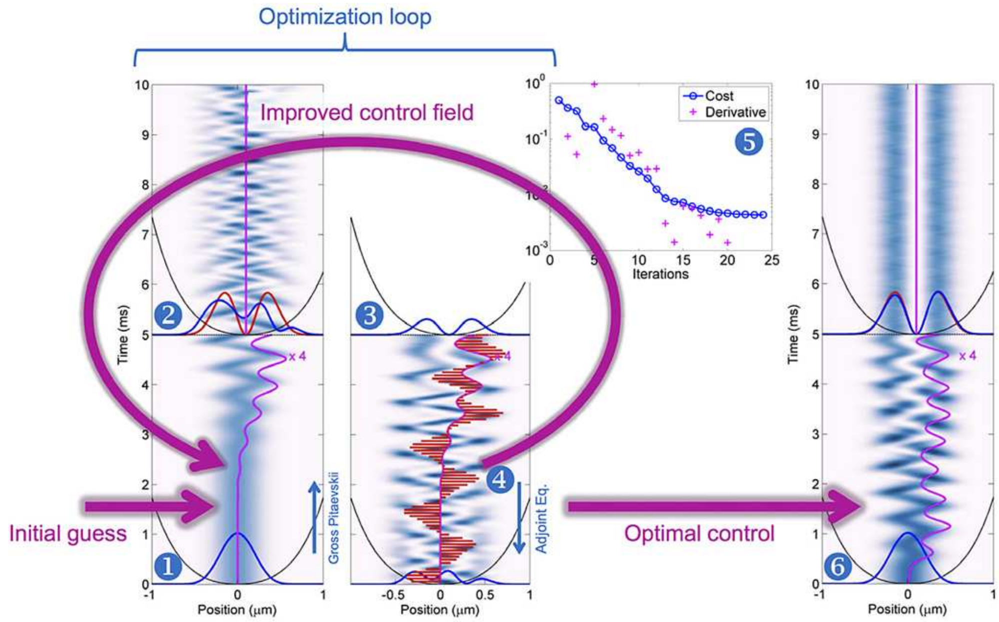

Figure 2. Schematics of the OCT optimization loop, which starts with an initial guess for the control field $\lambda(t)$ associated with the displacement of the minimum of the confinement potential. First, the Gross–Pitaevskii equation with the (panel 1) initial condensate wave function $\psi_0(y)$ is solved forwards in time, to obtain (2) the final wave function $\psi(y, T)$ at the terminal time $T = 5$ ms of the control process, which in general deviates significantly from the desired, first excited wave function $\psi_\mathrm{d}(y)$. The density plots in the different panels report the time evolution of the square moduli of the different functions. From the knowledge of $\psi(y, T)$ and $\psi_\mathrm{d}(y)$ we can compute the (3) terminal value of $p(y, T)$ via equation (4), and solve the adjoint equation (3*b*) backwards in time (4), to finally come up with a new search direction for the optimal control field (equation (6)) that is used in the next iteration of the optimization loop. The solid lines superimposed on $\lambda(t)$ in the panel of the adjoint equation depict the search directions. The inset (5) shows how the cost and derivative for a given control decrease with increasing iterations, until (6) an optimal control is obtained. Here $\lambda(t)$ (magnified by a factor 4) steers the system from $\psi_0(y)$ to the desired wave function at the terminal time of the control process.

over the whole time interval. In all of our OCT calculations we use equation (6) instead of equation (5).

Our OCT implementation relies on a numerical optimization routine and a differential equation solver. As for the optimization routine, one can use any generic code that, starting from some initial guess for the control field, requires a function value (the cost function) together with the derivative of the evaluated function $\delta L/\delta\lambda$ to compute a new, improved $\lambda(t)$. When using the H^1 norm of equation (6) one must ensure that all inner products in the generic code are evaluated as $(u, v)_{\mathrm{H}^1}$ rather than $(u, v)_{\mathrm{L}^2}$. In general we observed the best performance for the quasi-Newton BFGS optimization [49], which outperforms the nonlinear conjugate gradient method for larger number of iterations in the optimization loop. As for the differential equation solver, we usually employ a split operator technique [29] because of its robustness and simplicity.

The OCT optimization is schematically depicted in figure 2. One starts with some initial guess for the control field. In general, the outcome of the OCT loop does not depend critically on the initial $\lambda(t)$ and one can use any reasonable guess, such as in our case some interpolating function between the boundary values of $\lambda_0 = 0$ and $\lambda_1 = 0.1$ μm at the terminal time $T = 5$ ms. Next, (panel 1) the condensate wave function $\psi(y, 0) = \psi_0(y)$ is set to the ground state $\psi_0(y)$ of the anharmonic trap, including the nonlinear term of the GPE [29], and equation (3*a*) is solved forwards in time to obtain (2) the terminal wave function $\psi(y, T)$. For the initial guess of the control, $\psi(y, T)$ differs significantly from the desired, first excited state $\psi_d(y)$ of the anharmonic trap, which has a node in the middle, as can also be inferred from the ensuing time evolution where the trap displacement is held constant. From equation (4) we can compute the (3) terminal condition for $p(t)$, and (4) solve the adjoint equation (3*b*) backwards in time. Finally, the knowledge of the complete history of $\psi(y, t)$ and $p(y, t)$ allows us to compute the new search direction through equation (6), and to pass this direction to the optimization routine which will come up with a new, improved $\lambda(t)$, which can be used in the next iteration of the optimization loop.

In the inset (5) of figure 2 we show how the cost function $J(\psi, \lambda)$ and the derivative measure $|\delta L/\delta\lambda|$ evolve with increasing iterations. Note that the 'optimal control' corresponds to a minimum of the control landscape, associated with a derivative equal to zero, but it is generally not guaranteed that the cost is also small there. However, there are indications that under quite broad conditions the OCT loop will come up with a $\lambda(t)$ that fulfils the control objective of wave function matching almost perfectly [50]. In our simulations we typically stop after a given number of iterations or when the derivative has become sufficiently small. The resulting $\lambda(t)$ sequence is then called the optimal control. As can be seen from the solution of the GPE on in (6), with this control we closely match the desired wave function at the terminal time, with a fidelity of $|\langle\psi_\mathrm{d}|\psi(T)\rangle|^2 \approx 1$–$3 \times 10^{-3}$. Up to a global phase, the wave function remains stationary for $t > T$.

3. Experimental implementation

The vibrational state control scheme is realized using an ultracold Bose gas trapped on an atom chip. The experimental procedure is very close to that described in [35, 36]. In brief, a laser-cooled cloud of rubidium-87 atoms in the $|F = 1, m_F = -1\rangle$ Zeeman level is loaded into a strongly elongated atom chip wire trap [51, 52]. Using forced evaporative cooling, the gas is brought to a temperature close to quantum degeneracy. Then, by means of RF dressing [53, 42], the external confinement along the two tightly trapped axes y, z is deformed from a harmonic to an anharmonic and anisotropic potential (see the next section for details). In the anharmonic trap, further cooling down to a temperature T well below 50 nK is performed. With $N \sim 800$ atoms, the gas is now in a quasi-condensate [54, 55] regime, where phase fluctuations prevent true condensation of the matter wave along the longitudinal (elongated) direction x. However, the energy scales corresponding to both temperature ($k_BT < h \times 1$ kHz, with Planck and Boltzmann constants denoted as h and k_B, respectively) and atom interactions (chemical potential $\mu \sim h \times 600$ Hz) are well below the trap level spacing ($\sim h \times 2$ kHz) along the transverse directions y, z. Thus, thermal excitations in the transverse degrees of freedom are frozen out, and almost all atoms occupy a single transverse mode; along y and z, the system is hence appropriately described by a single condensate wave function. In this paper, we are mostly concerned with the dynamics along the transverse direction y, and hence neglect the longitudinal mode structure of the quasi-condensate.

Having prepared the system in this way, we apply the control sequence, while monitoring the momentum space distribution of the condensate, as will be described in the following sections.

3.1. Trap preparation

In a quantum harmonic oscillator, all states that can be addressed by simple displacement of the potential are quasi-classical coherent states [38]. This statement also holds for a harmonically trapped interacting many-body system, where a quasi-classical collective oscillation at the trap frequency fully decouples from more complex internal dynamics [39, 40], and the shape of the wave function is preserved [41]. Hence, transferring the condensate population into an excited, stationary state necessitates an *anharmonic* potential along the displacement direction y, where the decoupling of collective and internal dynamics breaks down. Furthermore, to be robust against excitations in the perpendicular direction z, anisotropy in the transverse plane of the potential is required, causing a detuning of trap levels between the directions.

Initially, the Ioffe–Pritchard field configuration as created by the chip wires (plus external offset fields, see figure 3) is rotationally symmetric, and provides harmonic trapping along the transverse directions y, z. For the parameters chosen in our experiment, the transverse trap frequency is $\nu_0 = 4.1$ kHz in both directions, whereas the longitudinal frequency is of the order of 30 Hz. To introduce anharmonicity and anisotropy, we apply RF dressing [56, 42, 53]. Using two chip wires running in parallel to the trapping wire as antennae, the atoms are irradiated by a RF near field with linear magnetic polarization, which is red-detuned by tens of kHz with respect to the atomic Larmor frequency due to the static magnetic field at the trap centre. The RF field adiabatically mixes the Zeeman levels of the $F = 1$ hyperfine manifold, coupling them to dressed states. This gives rise to an energy shift that depends on detuning from the Larmor frequency $\Delta(\mathbf{r})$ and coupling strength (Rabi frequency) $\Omega(\mathbf{r})$. Both quantities are position-dependent, the latter because of the changing RF polarization with respect to the local magnetic field that modulates the coupling strength. In the rotating-wave approximation [42], the resulting potential landscape up to a constant is given by

$$V(\mathbf{r})/h = \sqrt{\Omega(\mathbf{r})^2 + \Delta(\mathbf{r})^2}.$$

The dressing is most effective along the direction perpendicular to the RF polarization; in our case, applying a polarization along the vertical axis z leads to a deformation mostly along y. In figure 4(a), the potential along y is shown as a function of dressing strength, expressed as coupling Ω_0 near the trap centre. At sufficiently strong coupling, splitting of the potential into a double well occurs, which is the typical application of RF-dressed potentials [53, 57–60]. However, at lower coupling, this technique also allows for the introduction of anharmonicity and anisotropy to a single trap, as needed for our scheme. In the experiment, we apply a RF field of ~0.84 G peak-to-peak amplitude, leading to a coupling $\Omega_0 = 147$ kHz, at a frequency red-detuned by $\Delta_0 = -54$ kHz with respect to the Larmor frequency near the trap minimum (824 kHz).

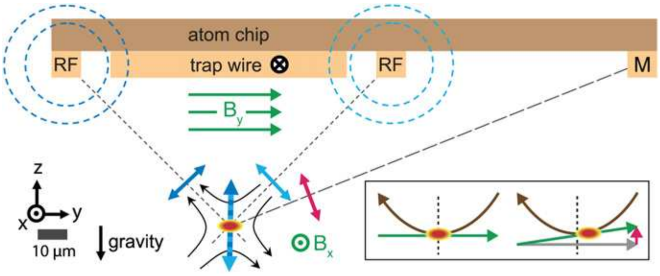

Figure 3. Main figure: schematic of the atom chip layout (see [52] for details). The waveguide potential is formed by the current through the trap wire along $-x$ and a static bias field B_y, adding up to quadrupole field (bent arrows). An external offset field along B_x, perpendicular to the figure plane, defines the Larmor frequency at the trap minimum (Ioffe–Pritchard field configuration). On a separate chip layer, currents in broad wires along y (not shown) provide weak longitudinal confinement. The radio frequency dressing currents are applied to wires (RF) in parallel to the trapping wire, leading to a RF field along z (blue arrows). The resulting anisotropic transverse potential is shown as ellipse in the centre of the quadrupole. Finally, the modulation of the trap position is accomplished by a current in an auxiliary wire (M), leading to a magnetic field, aligned at ~19° with respect to the z-axis (red arrow). Inset: field configuration for trap position modulation. The transverse trap position is defined by cancellation of the chip wire field (brown) and the bias field (green). Adding a weak field along z (red) tilts the bias field slightly, leading to a horizontal shift of the trap minimum.

The resulting potential is shown as a green line in figure 4(a). Even though the rotating-wave approximation holds well for the used dressing strength [61], the high

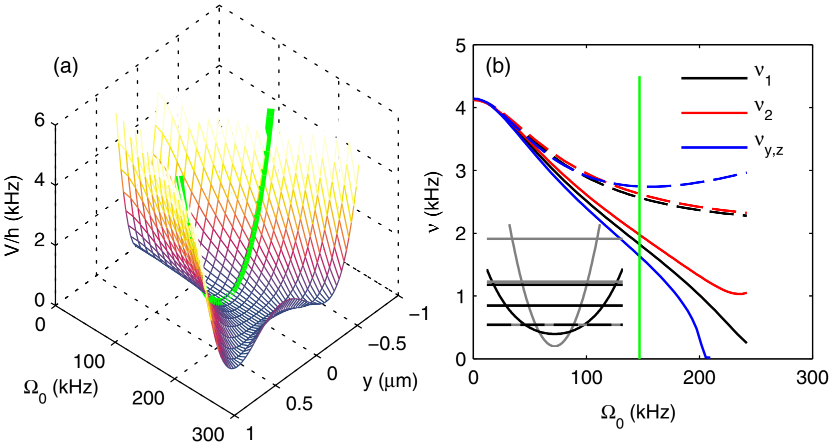

Figure 4. Effects of RF dressing on the transverse trapping potential. (a) Potential along the y (displacement) direction, as a function of RF Rabi frequency. The detuning is $\Delta_0 = -55$ kHz. At dressing strengths above $\Omega_0 \sim 180$ kHz, splitting of the single potential into a double well occurs. (b) Shift of single-particle trap levels versus dressing strength. Solid and dashed lines correspond to perpendicular (y) and parallel (z) directions with respect to the RF polarization, respectively. Blue: frequency of harmonic part, as defined in equation (7). Black, red: first and second level spacing of single-particle eigenstates. Inset: initial (grey) and dressed (black) potential, each with their first three energy levels. The green lines in both panels mark the setting used for the experiments.

sensitivity of the excitation protocol to the exact potential shape calls for an exact calculation by means of a Floquet analysis [62]. Along two transverse directions the result can be approximated by a sixth-order polynomial of the form

$$V_6(y,z)/h = \frac{\nu_y}{2}\left(\frac{y}{l_y}\right)^2 + \sigma_y\left(\frac{y}{l_y}\right)^4 + \xi_y\left(\frac{y}{l_y}\right)^6 \tag{7}$$

$$+\frac{\nu_z}{2}\left(\frac{z}{l_z}\right)^2 + \sigma_z\left(\frac{z}{l_z}\right)^4 + \xi_z\left(\frac{z}{l_z}\right)^6. \tag{8}$$

In this expression, the lengths $l_{y,z} = \sqrt{h/(m\nu_{y,z})/(2\pi)}$ correspond to the characteristic length of the harmonic part. The parameters are given by

$$\begin{aligned} \nu_y &= 1655\ \text{Hz}; & \nu_z &= 2751\ \text{Hz} \\ \sigma_y &= 78.2\ \text{Hz} & \sigma_z &= -69.6\ \text{Hz} \\ \xi_y &= -0.96\ \text{Hz} & \xi_z &= 9.1\ \text{Hz} \\ l_y &= 265\ \text{nm} & l_z &= 206\ \text{nm}. \end{aligned} \tag{9}$$

Along y, the sixth-order term ξ_y is negligibly small, and the description reduces to a Duffing oscillator [63].

By solving the Schrödinger equation, the single-particle trap levels of the dressed potential can be obtained. The first two level spacings $\nu_{1,2}$ along y and z are shown in figure 4(b). For the used parameters (as marked by a green line), the initial degeneracy of the level spacings is lifted, and we obtain the excitation energies (zero-point energy subtracted) $[E_{10}, E_{20}, E_{01}, E_{02}, E_{11}]/h = [1.84, 3.83, 2.58, 5.21, 4.42]$ kHz with E_{ij} denoting the ith and jth state along y and z, respectively. The relevant level spacings along y are given by $\nu_1 = 1.84$ kHz, $\nu_2 = 1.99$ kHz, the first level spacing along z is $\nu_z = 2.58$ kHz. From the corresponding eigenfunction along the z direction, and a Thomas–Fermi approximation [64] of the longitudinal profile for $N = 800$ atoms, we can estimate the coupling constant[4] in equation (1) by averaging as $g = h \times 300$ Hz μm [65].

In the experiment, characterization of the initial harmonic trap is straightforward, using RF spectroscopy and observation of collective oscillations. On the other hand, confirming the (calculated) parameters of the anharmonic dressed trap with sufficient accuracy is difficult. Instead, we optimize the experimental control parameters (RF field strength and detuning) directly, by comparing the response to the control ramp to that determined numerically using those trap parameters. Along the longitudinal x-axis, the harmonic trap frequency $\nu_x = 16.3$ Hz is determined by observation of deliberately excited collective modes of the atom cloud.

3.2. Control of trap motion

The transverse movement of the potential is accomplished by applying a time-dependent current to an auxiliary wire running parallel to the main trapping wire. As shown in the inset of figure 3, the additional magnetic field along z causes a slight tilt of the homogeneous bias field, which is exactly aligned along y initially. The trap minimum position, which is given by the point where the bias field cancels that of the trapping wire, is displaced along y. Additionally, the y-component of the modulation field, which changes the magnitude of the bias field, causes a slight proportional movement along z. However, as confirmed by two-dimensional simulations, the anisotropy of the transversal potential suppresses any significant influence on the excitation along y. From numerical simulations of the field geometry, the movement of the trap minimum caused by the current can be calculated as 26 nm mA^{-1} along y and 9 nm mA^{-1} along z. The geometry of all chip wires and homogeneous offset fields involved in trapping and modulation is shown in figure 3.

3.3. Effect of finite control bandwidth

The current in the modulation wire is driven by a custom-design low-noise current source, which is controlled from

---

[4] Note, that we normalize the wave function to 1, not N, in equation (1). Hence, g incorporates the atom number.

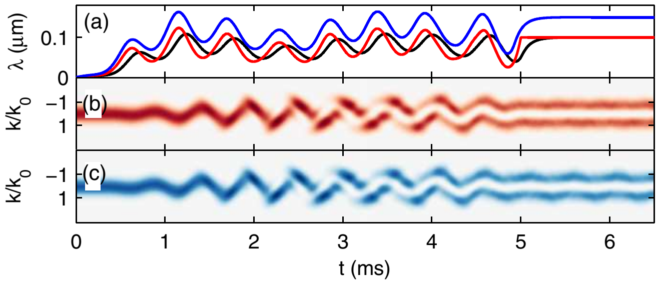

Figure 5. Effect of filtering due to finite electronics bandwidth. k-axes are scaled to $\hbar k_0 = \sqrt{2mh\nu_1}$. (a) Control ramps $\lambda(t)$. Red: original control ramp as derived from OCT. Black: control ramp after applying the electronics filtering. Blue: filtered, rescaled and shifted control ramp. (b) GPE momentum distribution, simulated without accounting for finite bandwidth. (c) GPE momentum distribution, simulation including finite bandwidth, rescaling of the control ramp by a factor of 1.6 and a time shift of 0.08 ms.

an arbitrary waveform generator[5], that outputs the excitation ramp. A slight smoothing of the control sequence is imposed by finite bandwidth of the electronics, which has to be accounted for when comparing experimental and numerical results (see section 4.1). The measured transfer function modulus $|\mathcal{M}(\nu)|$ at a frequency ν can be approximated by an exponential $|\mathcal{M}(\nu)| \approx \mathrm{e}^{\nu/\nu_{\mathrm{co}}}$ with cutoff frequency $\nu_{\mathrm{co}} \approx 4.4$ kHz. Furthermore, a frequency-dependent phase shift is imposed. Effectively, filtering causes a reduction of the driving amplitude near the resonant frequency $\nu_1 \approx 1.8$ kHz by a factor $|\mathcal{M}(\nu_y)|^{-1} \sim 1.6$, and a time delay on the order of 0.1 ms (figure 5(d)). In figure 5 it is shown, that the filtering due to the electronics can be largely cancelled by rescaling and shifting the control sequence by these factors. The difference in the outcome of the simulated momentum distribution is only small and largely given by a slightly enhanced collective oscillation (see section 4.3).

3.4. Measurement of momentum distributions

At a time t after starting the excitation process, we suddenly switch off the trapping potential, implying that for $t < 5$ ms the excitation process is still incomplete. The fast transverse expansion of the cloud due to the tight waveguide confinement causes atom interactions to vanish rapidly, and the ensuing expansion can be considered ballistic. After $t_{\mathrm{exp}} = 46$ ms of expansion, we take a fluorescence image [66] similar to that shown in figure 6(a), fully integrating over the z-direction. In the images, three separate clouds can be observed along the longitudinal x-direction. The two side peaks emerge due to decay of the excited state which has been populated by our excitation protocol. Correlation properties of these (twin beams) and the dynamics of the decay process have been analysed elsewhere [66, 35]. We separately integrate along x over the central and side peaks, respectively (blue dashed lines) to analyse the transverse state of each part of the system. The observed density distributions along y (figure 6(b)) represent the momentary momentum distributions of the trapped cloud at time t, as the initial transverse cloud size (of the order of $l_y \sim 250$ nm) is negligible compared to that after expansion (far field). If we express momenta as wave numbers k_y, a distance δy in the image hence corresponds to $\delta k_y = \alpha\,\delta y$ with $\alpha = m/\hbar t_{\mathrm{tof}} \approx 0.034\,\mu\mathrm{m}^{-2}$. Taking an experiment series where t is scanned, we can thus fully access the momentum distribution dynamics $\tilde{n}_y(k, t)$ along the excitation direction, which we will typically depict as false-colour plot, see e.g. figure 7(a). Our main interest will be the dynamics of the central (source) cloud which is subject to the excitation; however, in section 4.2, also the transverse dynamics of the twin-beam peaks will be of some importance. For each of the experimental series shown in this paper, we averaged over typically 12 realizations to suppress noise and allow for robust comparison to theory results.

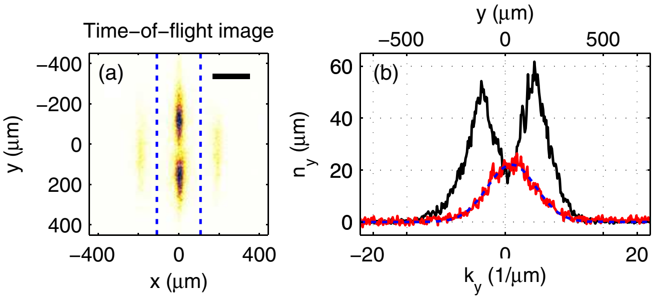

Figure 6. Experimental image analysis. (a) Typical experimental image data for optimal excitation and $t = 5.5$ ms, averaged over 12 shots. As the image is taken after 46 ms of expansion time, it predominantly reflects the initial momentum distribution. The scale bar corresponds to a spatial distance of 187 μm, equivalent to the typical momentum of 5.5 $\mu\mathrm{m}^{-1} \approx k_0$. (b) Transverse momentum distribution inferred from the image in panel (a). Upper (black) line: central peak, inside dashed lines in (a). Lower (red) line: emitted atoms, outside dashed lines in (a). Dashed (blue) line: Gaussian fit to emitted atom momentum.

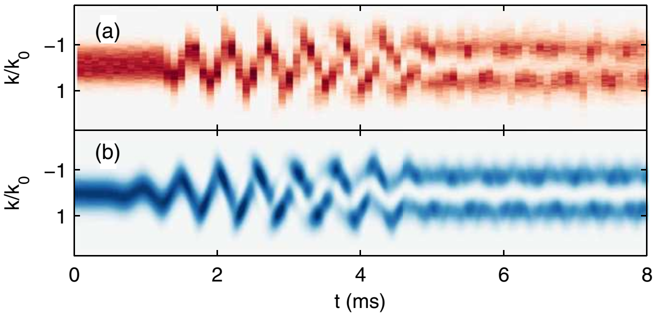

Figure 7. Comparison of momentum distribution dynamics as obtained from experiment and theory for typical parameters. (a) Experiment. Each pixel column in the false-colour plot corresponds to a distribution as shown in figure 6(b). (b) 1d GPE numerics, including finite bandwidth effects (see text).

4. Results

We will analyse the excitation dynamics in various complementary ways, motivated by the goal of developing an effective mapping of the many-body dynamics to a driven two-level system. In section 4.1 we will start by comparing the results obtained in section 2 for the time-dependent momentum distribution of the condensate wave function to experimental observations, varying a range of relevant

---

[5] Tabor Electronics WW5061.

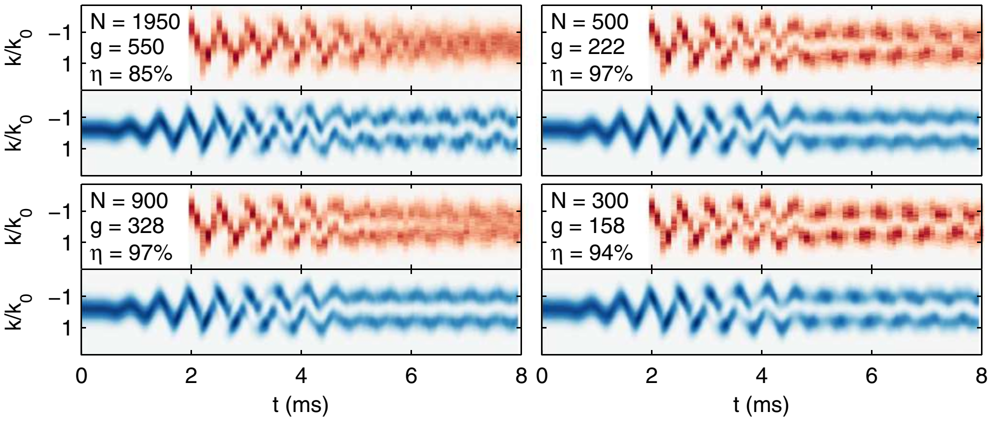

Figure 8. Effects of interactions on excitation dynamics. In each pair of plots (top: experiment, bottom: numerics), the typical experimental atom number is shown alongside the GPE nonlinearity term g (in Hz μm, ordinary frequency). The efficiency η is defined as described in the text.

parameters. While the excellent agreement ensures that the numerics used to obtain the optimized ramp are accurate, this result only gives limited insight into how the excitation process can be understood qualitatively. In section 4.2 a more phenomenological analysis is performed directly on the experimental data, which will give hints about how to develop a two-level description. In section 4.3 the GPE simulations are investigated in more detail, using a description based on Wigner quasi-probability functions, and displaced Fock states. It will become evident that all approaches lead to conceptually similar and quantitatively compatible interpretations, which can finally be unified to obtain a two-level interpretation as sought after initially.

4.1. Comparison of experiment and numerics

Compared to other driven quantum systems, where optimal control techniques may be applicable, a rather unique advantage of cold atoms is the accessibility of the system response, enabled by the relatively large time and length scales and the abundance of powerful imaging techniques. Probing the performance of a control strategy such as that developed in section 2 is not restricted to the final outcome, but the driven system can be monitored even *while* it is being driven, providing direct means to compare experiment and numerical simulations, or apply feedback schemes [67]. As explained in section 3.4, time-of-flight fluorescence images give us direct access to the momentum distribution of the condensate along its transverse axis. In figure 7(a), a typical momentum distribution dynamics plot, as obtained from the experiment, is shown.

Many-body effects. Along the transverse directions, confinement is strong enough ($h\nu_1 \gg \mu$) to make interaction-induced effects comparatively small. Still, to achieve the highest possible fidelity of the excitation, it is crucial to take into account the nonlinear term in equation (1) for optimization. In figure 8, the excitation dynamics is shown for a data set, where the atom number has been varied before starting the excitation sequence. The data is compared to the result of the GPE (1).

It is observed that effective excitation is achieved for a nonlinearity corresponding to an atom number $N \sim 900$, which is close to what has been used in the optimization. For all other atom numbers, stronger residual dynamics after the end of the sequence ($t > 5$ ms) is found, indicating decreased fidelity, as the desired state is stationary. While the GPE simulations reproduce the general tendencies found in the experiment, the agreement is not as good as e.g. for scaled excitations (see below). For the highest atom number, only rather poor qualitative agreement is reached, indicating the insufficiency of a mean-field model such as GPE (necessitating e.g. a MCTDHB ansatz [34, 48]) and strong effects of the rapid decay of the excited state.

Robustness against experiment inaccuracy. In OCT, an aspect of high relevance is the sensitivity of the excitation dynamics to deviations of experimental parameters from the ones used for optimization. In our case, this predominantly applies to parameters affecting the trapping potential. We consider small changes of the potential parameters ν_y, σ_y. In the experiment such deviations arise from variation of the dressing parameters Ω_0, Δ_0, which, in turn, are caused by inaccuracy of the current in the RF antenna wires, and of the external offset field along x (defining the atomic Larmor frequency), respectively.

Numerical results for a range of parameters are shown in figure 9. Panels (c)–(f) correspond to deviations caused by an offset field misalignment of ±2 mG (b), (c) and ±7 mG (e), (f) leading to weaker (positive values) or stronger RF dressing, respectively. It is observed that any deviation leads to a decrease in excitation efficiency, which is defined here as time-averaged overlap with the desired wave function ψ_d, $\eta = \langle |\langle\psi_\mathrm{d}|\psi(t)\rangle|^2\rangle_{t>5\,\mathrm{ms}}$. The similarly defined population of higher excited states ζ becomes strong at trap modifications with weaker dressing $\delta\nu_y > 0$ and $\delta\sigma_y < 0$. This effect can be expected, as the protection against excitation to higher states fades with decreasing anharmonicity, while the excursion of the trap relative to the typical length l_y increases. In panel (d), on top of an offset field mismatch of +3 mG, the current in the RF wire has been adjusted to cancel the effect on ν_y. The weak mismatch in σ_y and ξ_y only leads

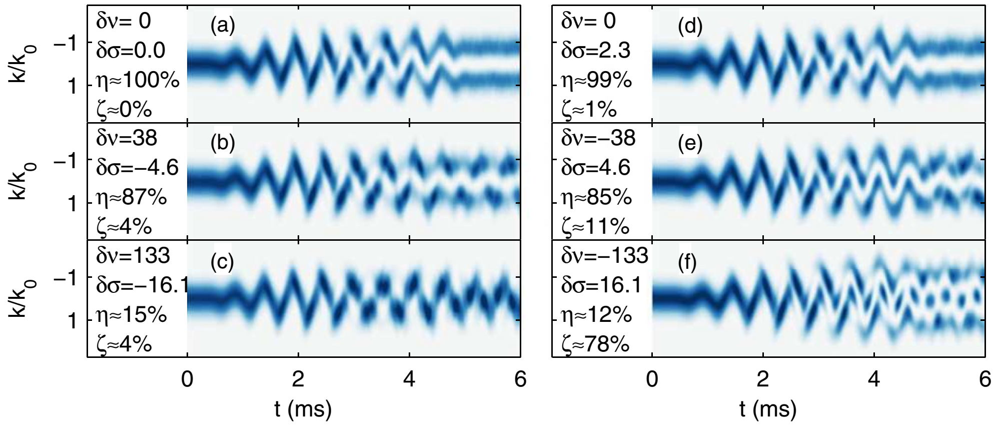

Figure 9. Stability of the excitation sequence against inaccuracy of the trapping potential (numerical result). In each plot, the deviation of the potential terms $\delta\nu_y$, $\delta\sigma_y$ are given (in units of Hz), as well as the efficiency η and the spurious excitation to higher states ζ as defined in the text.

to a slight reduction of efficiency. Consequently, optimizing the experimental parameters for a strong excitation (e.g. by minimizing residual dynamics at $t > T = 5$ ms) may lead to slightly shifted values, which however compensate each other. Using this method, a sensitivity of better than 1 mG (or an equivalent mismatch of the dressing current) can be reached, which is beyond what can be achieved by independent characterization of the trapping potential.

Scaled excitations. In figure 10 the momentum distribution dynamics is shown for a data set with varying excitation efficiency, which will be the main subject of analysis in the remainder of this and the following section, as it covers a very broad range of control sequences. To achieve different efficiencies, the excitation ramp has been scaled in amplitude by factors s with respect to the optimal control result, resulting in strongly varying wave function dynamics. The approach of simple amplitude scaling has been chosen over using separately optimized ramps for different efficiencies, to allow for easier comparison due to the well-defined relation between the used control sequences. Furthermore, our analysis will show that the main spurious effect of this strategy are collective oscillations at reduced scalings. Comparison between GPE and the experimental result (average over ∼12 realizations) shows excellent agreement at early times[6]. At later times, decay of the excited state into twin beams, which is not accounted for in theory, becomes significant (see bottom right panel) and for high values of s, agreement is reduced due to inelastic collisions with the twin beams which reside in a different transverse state. However, for weak excitation, even the shape of single 'beating peaks' after the end of the excitation pulse is precisely captured by numerics. Along the k-axis, the GPE result has been convolved with a Gaussian of $m/(\hbar t_{\mathrm{tof}}) \cdot 40\,\mu\mathrm{m} \approx 1.20\,\mu\mathrm{m}^{-1}$ rms width to account for finite imaging resolution and bulk position fluctuations. Apart from a small shift of the t-axis and a slight re-scaling of the k-axis[7], the scaling factor s is the only free input parameter of the simulation.

Having established the detection method, and verified that the outcome is consistent with the numerics on which the control optimization has been founded, we now proceed to a more qualitative analysis of the experimental result.

4.2. Analysis of experimental momentum dynamics

In this and the following section we will analyse the momentum distribution dynamics beyond a simple comparison to numerical results. The notion underlying the discussion will be that of a few-level system, comprised by the ground, first and occasionally second excited state of the confinement potential along the excitation direction, with the final goal to reduce the anharmonic oscillator to a closed two-level system. This approach may seem inappropriate, as it relies on the superposition principle, which requires a linear equation of motion and is hence not applicable to a mean-field wave-function as described by the GPE. However, in our case the nonlinearity is weak compared to the oscillator energy, and so is the modification of the dynamics due to many-body effects (see figure 8), suggesting that a description in terms of single-particle states may still provide significant insight[8].

Centre-of-mass dynamics. As the simplest possible observable derivable from the momentum dynamics, we start by analysing the transverse centre-of-mass of the experimental images, corresponding to the momentum expectation value $K(t) \equiv \langle k_y(t)\rangle$, see black lines in figure 11 (left panels). In the power spectra of $K(t)$ (centre panels), two strong peaks are observable near the first two transverse level spacings at frequencies $\nu_1 = 1.84$ kHz and $\nu_2 = 1.99$ kHz, and a weak third at $\nu_3 \approx 2.10$ kHz, defined analogously.

---

[6] Note that s has been defined including the necessary re-scaling due to finite electronics bandwidth (see figure 5).

[7] The shift in t is well below the experimental time resolution, and is very probably due to the inaccuracy of the filtering circuit characterization. The necessity for the re-scaling of k (of the order of 10%) might arise from interaction effects causing weak hydrodynamic effects in expansion [68]. The values of both adjustments are consistent among all sets shown.

[8] A more involved, but conceptually similar approach, which is directly based on stationary states of the GPE can be found in [12, 69, 41].

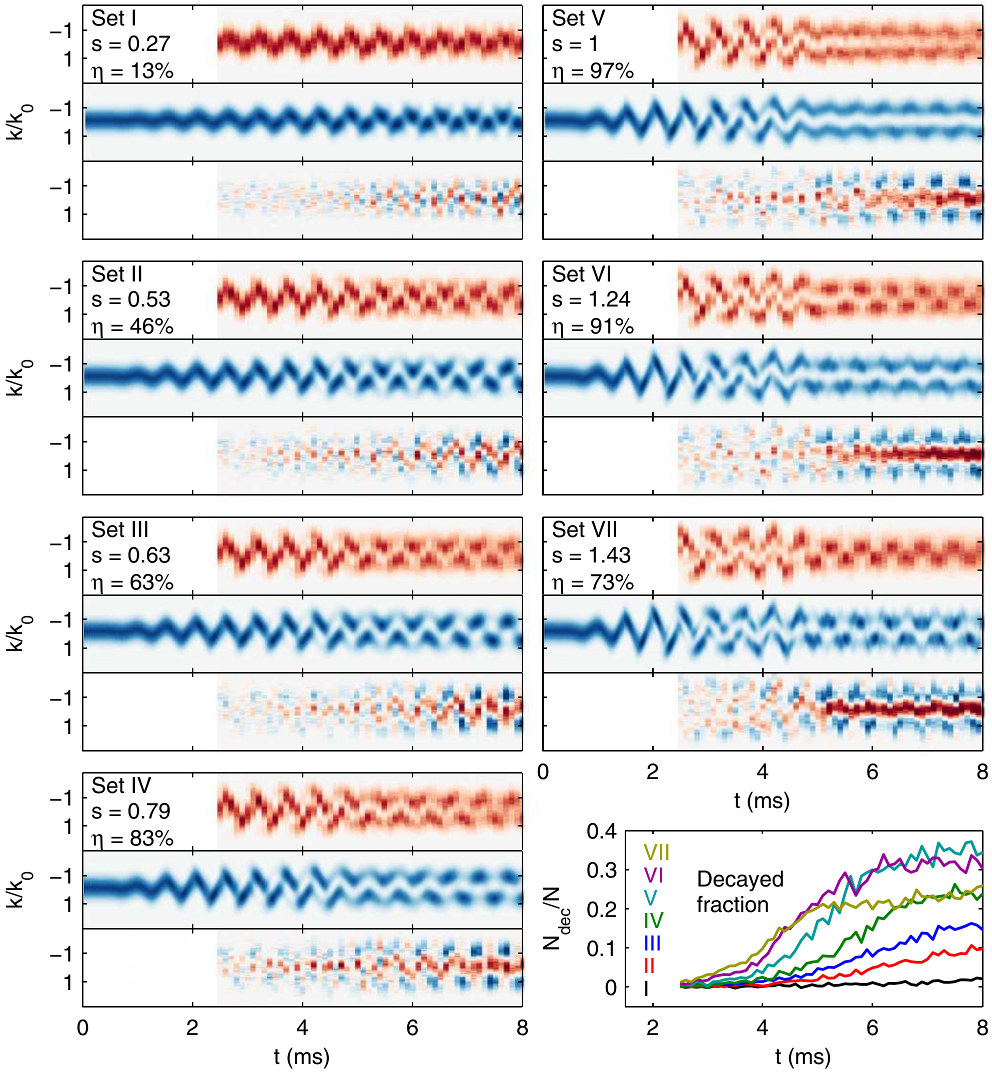

Figure 10. Results for scaled excitation ramps. Mean atom number is 770 for sets I, II, IV–VI, and 856 for sets III and VII. For each of the seven sub-sets, the upper image (red false-colour) is the experimental result, normalized separately for each time step. The middle image (blue false-colour) shows the numerical GPE result, including low-pass filtering and scaling by the factor s as given. The bottom image shows the deviation between experiment and theory, expressed as imbalance $\tilde{n}_{\mathrm{ex}} - \tilde{n}_{\mathrm{th}}$; the colour scale for the imbalance is enhanced by a factor 3. The bottom right inset shows the relative amount N_{dec}/N of atoms that have decayed from the excited state into twin atom pairs.

Assuming a single-particle level picture, these peaks can be interpreted as beating frequencies between populations of the first three levels of the oscillator, where mean-field effects are causing frequency shifts, as described below. Consequently, the magnitude of oscillations is the strongest for intermediate excitation efficiencies (sets II–IV), where the levels are populated most evenly, maximizing the beating contrast (see below).

A crucial observation is that also the transverse profiles of the twin-beam peaks, which are separated in the images longitudinally (see figure 6), exhibit strong oscillations of $K_t(t) \equiv \langle k_y^{(t)}(t)\rangle$. Meanwhile, they fully maintain their Gaussian shape (figure 6(b)). In figure 11, oscillations of the *relative* centre-of-mass $K_r(t) = K(t) - K_t(t)$ (left), and their power spectrum (centre) $f(\nu) = |\mathcal{F}[K_r(t)](\nu)|^2$, are shown as blue lines. It is observed that, while the oscillations are similarly strong as in a fixed frame, all peaks in the power spectrum, except that near ν_1 are suppressed. This suggests that in a reference frame co-oscillating with $K_t(t)$, the dynamics can be understood in terms of two transverse levels, motivating an approach of decomposition into a quasi-classical oscillation, and 'internal' dynamics, which remain unaffected by the bulk oscillation[9]. This interpretation is consistent with our understanding of the decay process [36], where the transverse state of the twin beams, a ground state displaced by $K_t(t)$, defines the appropriate ground state for the internal dynamics. In section 4.3, a more rigorous formalism for the co-oscillating frame will be given, and its position will be independently derived from numerical results.

In the right column of figure 11, spectra are shown which are derived from the oscillations at times $t > 5$ ms only,

---

[9] This decomposition is exactly valid for harmonically confined many-body systems [39, 40]. Obviously, this does not hold for an anharmonic oscillator, which is exactly why our excitation to a non-classical state by displacement can work at all (see above). Being aware of the inconsistency, we still apply the decomposition approach to qualitatively understand the dynamics.

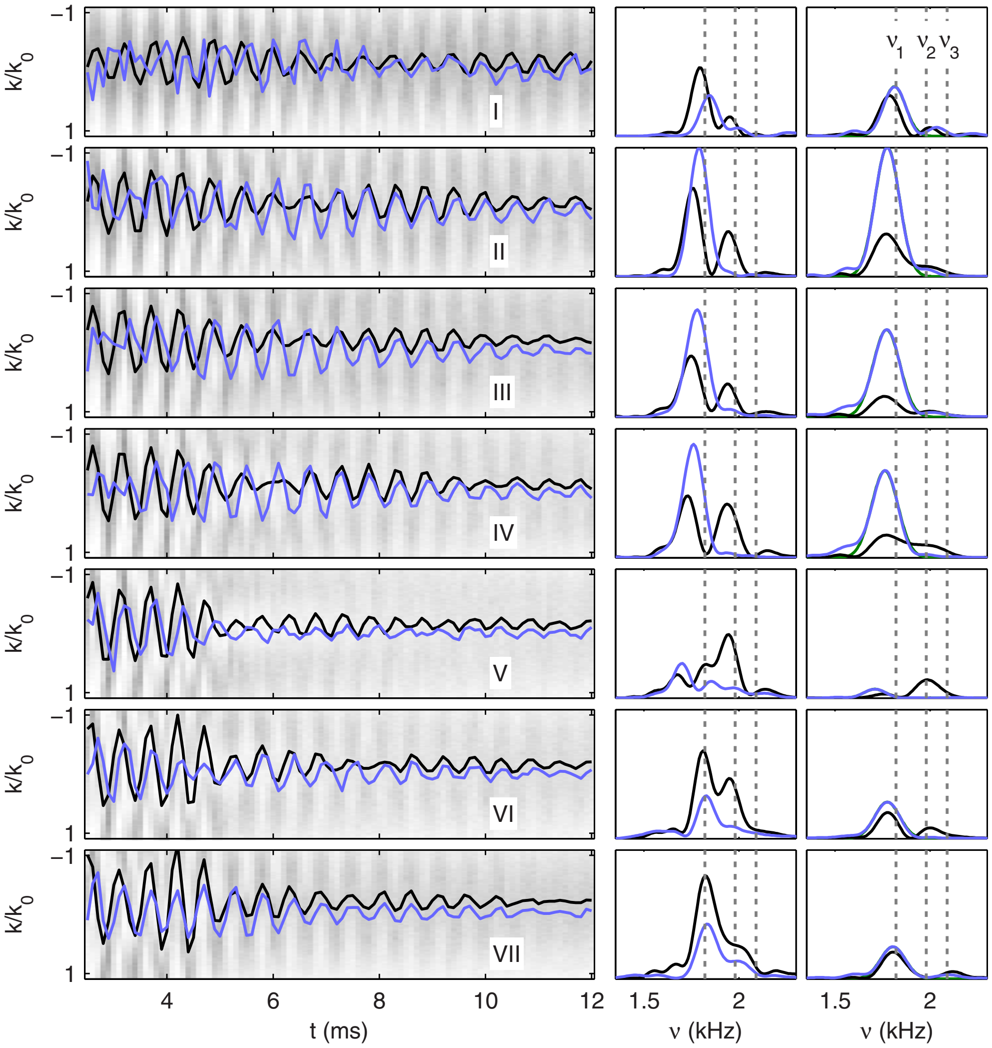

Figure 11. Momentum space centre-of-mass dynamics for data set as shown in figure 10. Left column: centre-of-mass momentum of the source cloud with respect to a fixed frame ($K(t)$, black) and relative to the twin-beam centre-of-mass ($K_r(t)$, blue). (See figure 14 for the twin-beam centre-of-mass.) In the background, the full dynamics is shown (see figure 10). Middle column: corresponding power spectra $f(\nu)$, taken over the entire time span shown. Right column: spectra, taken over a time span starting from $t > T = 5$ ms, i.e. after the end of the excitation. Grey dashed lines in the background indicate the harmonic frequency ν_h, and the first three level spacings, as defined in the previous section. All spectra are in arbitrary units, but normalized identically for each of the columns.

i.e. where no driving occurs anymore. Hence, they provide a characterization of the final state that is reached after the excitation. Qualitatively, the same features are observed as in the full time spectra, however, peaks at ν_2 are smaller, which is consistent with theory, as will be shown below. Also, in the relative centre-of-mass spectrum, the observation of a single-peak structure, with a minimal amplitude for the most efficient excitation is even more evident.

In figure 12(a), the integrated power of the oscillations $P \propto \int f(\nu)\mathrm{d}\nu$, measuring the stationarity of the final state, is shown as a function of the numerically obtained excitation efficiency η (see previous section). Apart from the strongest driving, where higher states may become excited more easily, P shows fair agreement with a curve given by $\eta(1-\eta)$, which is the squared amplitude of the interference term in the momentum-space density of a two-level system with momentum-space wave functions $\tilde{\psi}_0$, $\tilde{\psi}_\mathrm{d}$:

$$\begin{aligned}\tilde{n}(k_y,t;\eta) &= \left|\sqrt{1-\eta}\tilde{\psi}_0(k_y) + \sqrt{\eta}\tilde{\psi}_\mathrm{d}(k_y)\right|^2 \\ &= (1-\eta)|\tilde{\psi}_0(k_y)|^2 + \eta|\tilde{\psi}_\mathrm{d}(k_y)|^2 \\ &\quad + 2\sqrt{\eta(1-\eta)}\Re[\tilde{\psi}_0^*(k_y)\tilde{\psi}_\mathrm{d}(k_y)]\cos(2\pi\nu_1' t).\end{aligned} \tag{10}$$

The positions of the beating peak (obtained from a Gaussian fit) are shown in figure 12(b). For high efficiency, the frequency is shifted downwards from the oscillator level spacing ν_1 (red line). This is explained by the mean field term in the GPE (1). For the boundary case of near-unity efficiency, the shift can be calculated rather easily. As the ground state population is negligible, it does not contribute to the interaction energy, and the chemical potential μ_e for an atom in the excited state is given by the second eigenvalue of the time-independent GPE. The according wave function ψ_d (i.e. the desired state

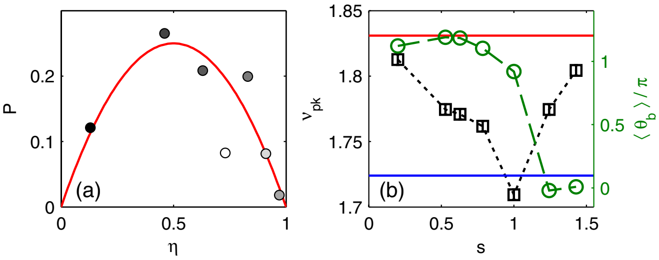

Figure 12. Analysis of post-excitation beating spectra shown in the right column of figure 11. (a) Integrated power of oscillations P. The experimental points have been scaled along the y-axis for best fit to $\eta(1-\eta)$ (red line). η has been derived as described in the previous section. The shading of each point indicates the corresponding scaling s (white is highest). (b) Peak position (black squares, left axes) and cosine of averaged phase (green circles, right axes). Red and blue lines correspond to the single-particle level spacing ν_1, and the mean-field-shifted level spacing ν_1', respectively.

in the optimization process) can now be used to calculate the chemical potential of a single atom in the ground state ψ_0', using a Schrödinger equation with effective potential arising from the mean field of the excited state:

$$\mu_e\psi_0'(y) = \left[-\frac{\hbar}{2m}\frac{\partial^2}{\partial y^2} + V_{\mathrm{ext}}(y) + 2g|\psi_{\mathrm{d}}(y)|^2\right]\psi_0'(y). \qquad (11)$$

The beating frequency is now given by the difference in chemical potential. Instead of the oscillator level spacing $\nu_1 \approx 1.831$ kHz, we obtain $\nu_1' = (\mu_e - \mu_g)/(2\pi\hbar) \approx 1.724$ kHz (blue line). Given the uncertainty in the input parameters of the calculation (such as the assumption of an equilibrium Thomas–Fermi shape longitudinally), this value agrees well with the experimentally obtained one for maximum efficiency (set IV), $\nu_{\mathrm{V}} = 1.709(4)$ kHz.

Finally, we can have a look at the phase of the (relative) centre-of-mass oscillation. When comparing the value of $K_y^{(\mathrm{r})}(t)$ for different scalings at a fixed time in figure 11, it is apparent, that the phase inverts at the point of maximum efficiency. We take the averaged phase from the Fourier transform result, weighted by the Lorentzian fit of the peak, and obtain the curve shown in figure 12(b) (right axes). The inversion is reminiscent of a two-level system subject to a Rabi driving, where the phase of precession inverts after passing the pole of the Bloch sphere at a pulse area larger than π. As will be shown in section 4.3, the excitation process can be understood analogously.

4.3. Interpretation of numerical result: two-level driving model

To understand the physical mechanism governing the optimal excitation protocol, in the following we analyse the time evolution of the condensate wave function in the Wigner representation [38]:

$$W(y,k,t) = \int \mathrm{e}^{-\mathrm{i}ks}\psi\left(y+\frac{s}{2},t\right)\psi^*\left(y-\frac{s}{2},t\right)\,\mathrm{d}s, \qquad (12)$$

which provides a mixed position–momentum distribution. The Wigner function has many appealing features reminiscent of a classical distribution function. Integration over all momenta k gives the spatial probability distribution $|\psi(y,t)|^2$.

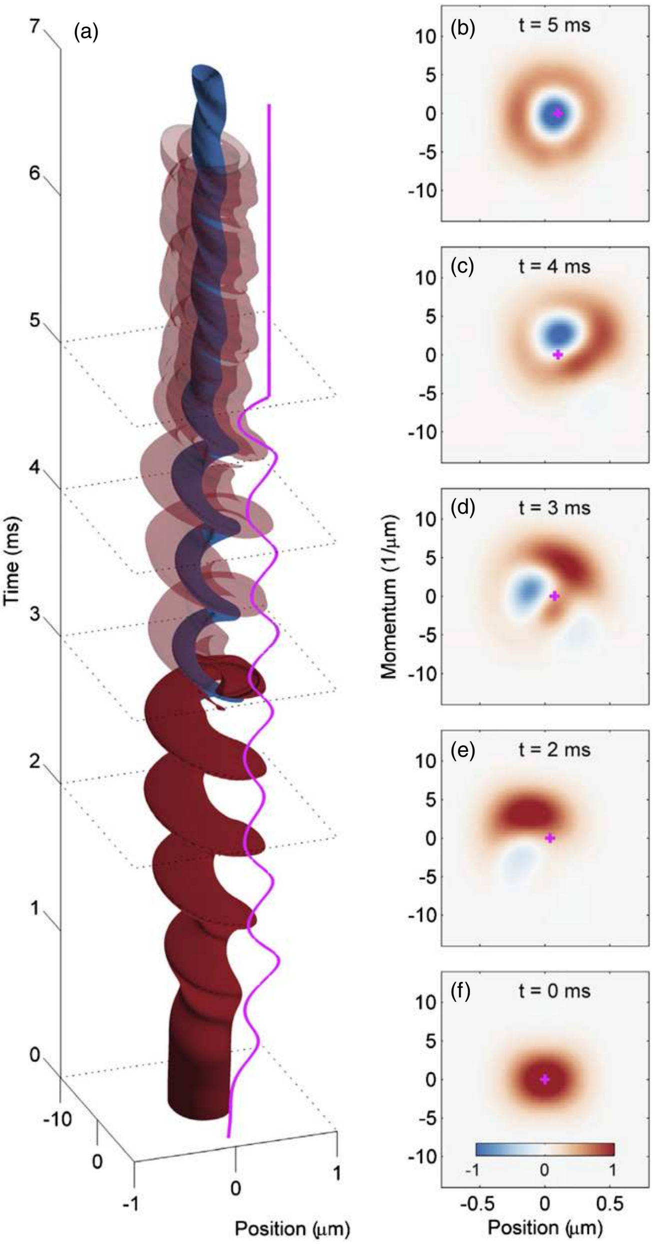

Figure 13. Time evolution of Wigner function. Panel (a) reports the time evolution of the Wigner function, and panels (b)–(f) show snapshots at selected times. In (a) we show the iso-surfaces at ±0.35 times the maximum value of the Wigner function, with transparency added to the iso-surface at time above 3 ms to highlight the appearance of the negative Wigner function part, associated with the first excited state. The magenta line indicates the excitation trajectory $\lambda(t)$. For discussion see text.

Likewise, integration over y gives the momentum probability distribution. The Wigner function of the condensate ground state, figure 13(f), approximately corresponds to the ground state of the harmonic oscillator, with equal uncertainty in position and momentum. In the figure the distribution is slightly elongated along y due to the nonlinear atom–atom interactions. The desired state of the control, figure 13(b), corresponds to the first excited state of the GPE in the anharmonic trap. It has positive and negative values (giving a node at $y = 0$ upon integration over all momenta), and thus differs from a genuine classical distribution function.

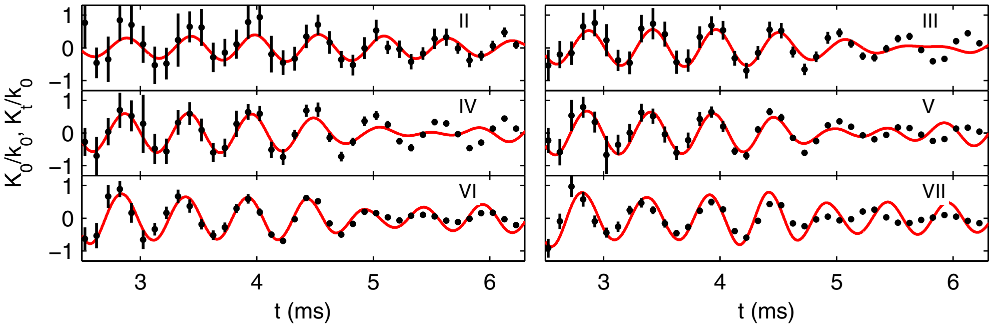

Figure 14. Reference frame for the two-level model. Underlying data are the same as shown in figures 10 and 11. Red lines are the momentum-space displacement $K_0(t)$ of the two-mode basis states, as obtained from applying equation (13) to the GPE result. Black points indicate the experimentally found centre-of-mass position of twin beams that have decayed from the excited state K_t, defining the reference frame for the emission process (see section 4.2). Similar to the momentum space dynamics as shown in figure 10, agreement reduces at later times, where decay into twin beams becomes strong. Data set I has been omitted due to the emission of twin beams being insufficient to determine K_t.

Panel (a) of the figure reports the time evolution of the Wigner function. We plot the iso-surfaces at ±0.35 times the maximum value of the Wigner function. At times later than 3 ms we have added transparency to the iso-surface for positive values to show the appearance of the negative part of the Wigner function, associated with a non-classically excited state. The solid line shows the time variation of the spatial minimum of the confinement potential. Initially, this displacement brings the condensate into collective oscillations, whose frequency is determined by the harmonic part of the confinement potential. For a large enough displacement, the condensate wave function is brought into the region where the anharmonicity of the confinement is sufficiently large to modify the internal structure of the wave function (and not just its displacement). One observes that in addition to the centre-of-mass oscillations in this regime the transfer from the ground to the first excited state occurs. Finally, at the terminal time $T = 5$ ms of the control process the minimum of the confinement potential is shifted to bring the condensate to a complete halt.

We next suggest a procedure to approximately map the excitation dynamics onto a genuine two-level description of ground and excited condensate states. As in section 4.2, the main idea is to separate the wave function dynamics into (i) a collective, quasi-classical oscillation, which is needed to bring the condensate into the anharmonic part of the trap, and (ii) an internal conversion between the ground and first excited state, defined in a co-moving frame. The latter conversion is governed by the anharmonic part of the trap, as explained above.

We define wave functions $\phi_g(y)$ and $\phi_e(y)$ as single-particle eigenfunctions of the *harmonic* part of the trap potential only, i.e. equation (7) with σ_y and ξ_y set to zero. Also, any modifications due to the nonlinear atom–atom interactions are neglected. This simplification allows us to analyse the dynamics in terms of displaced Fock states [70], that capture well the notion of the separation approach. Let $\hat{D}[\alpha(t)] = \exp[\alpha(t)\hat{a}^\dagger - \alpha(t)^*\hat{a}]$ denote the displacement operator of the harmonic oscillator [38], where $\alpha(t) = [l_y^{-1}Y_0(t) + \mathrm{i}l_yK_0(t)]/\sqrt{2}$ determines the position and momentum of the displacement at time t, and $\hat{a}$ denotes the annihilation operator. For a given displacement $\alpha(t)$, we can compute the overlap between the displaced ground and excited states with the condensate wave function according to

$$\chi(t) = \left|\int \left[\hat{D}[\alpha(t)]\phi_g(y)\right]^* \psi(y,t)\,\mathrm{d}y\right|^2 + \left|\int \left[\hat{D}[\alpha(t)]\phi_e(y)\right]^* \psi(y,t)\,\mathrm{d}y\right|^2. \qquad (13)$$

Determining the value $\alpha(t)$ which gives the largest overlap at time t allows us the aforementioned decompositions into (i) centre-of-mass coordinates $Y_0(t)$ and $K_0(t)$, and (ii) probability amplitudes $\langle\hat{D}[\alpha(t)]\phi_g|\psi(t)\rangle$ and $\langle\hat{D}[\alpha(t)]\phi_e|\psi(t)\rangle$ for the ground and excited state within the displaced frame. In all cases we find an overlap $\chi(t)$ well above 90%, which thus justifies the wave function decomposition. In figure 14, the obtained values for $K_0(t)$ are shown as red lines, and compared to experimentally obtained values, as described below. In figure 15 we compare $\chi(t)$ and the obtained excited population $\eta'(t) = |\langle\hat{D}[\alpha(t)]\phi_e|\psi(t)\rangle|^2$ to results from direct projection of $\psi(y,t)$ on the oscillator states $\phi(y)$ (defined in the co-moving frame of the excitation motion). The direct projection leads to strong transient population of higher excited states, and a sudden jump near the end of the excitation (figure 15(a)), where they are depopulated again. This is reminiscent of the fixed-frame centre-of-mass spectra (black lines in figure 11), where a peak near ν_2 is present when regarding the entire sequence (centre column), but mostly vanishes after $t = T$ (right column). In contrast, the two-level approximation in the system displaced by $\alpha(t)$ yields a smooth transition (figure 15(b)), consistent with the continuous appearance of negative values of the Wigner function (figure 13). In the momentum dynamics derived from the two-level model in a similar manner to equation (10) as shown in figure 15(d), a continuous transfer to the excited state is observed, with strong beating at intermediate excited population. Again, this is consistent with the experimental *relative* centre-of-mass spectra (blue lines in figure 11), where only a single peak near ν_1 persists, even during the excitation. Similar to a Rabi pulse with area larger than π, the excited population $\eta(t)$ is decreasing towards $t = T$ for scaling parameters $s > 1$.

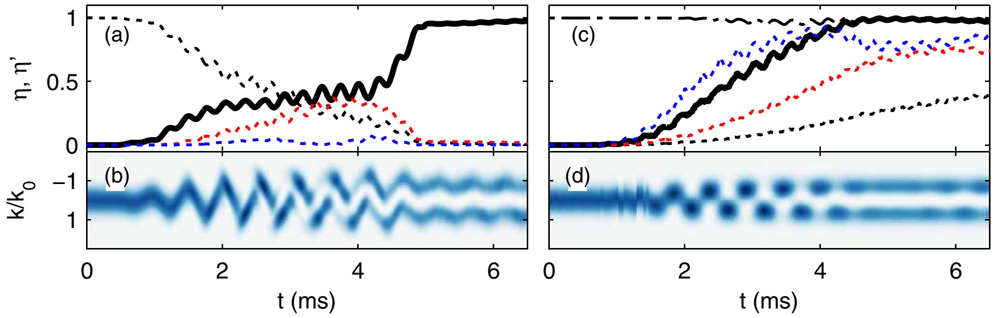

Figure 15. State populations during the excitation process. (a) Populations $\eta(t)$ of excited states of the full anharmonic potential (equation (7)) as arising from direct projection of the GPE result for data set V ($s = 1$). The solid line indicates the population of the first excited state, dotted lines represent the ground (black) and first and second excited (red, blue) states. (b) Corresponding momentum dynamics (identical to figure 10-V). (c) Population of first excited state in the co-oscillating frame within the two-mode model $\eta'(t)$. Solid line: data set V, corresponding to the solid line in (a). Dashed lines: sets I (black, $s = 0.27$), III (red, $s = 0.63$), and VII (blue, $s = 1.43$). The dash-dotted line indicates the total overlap of the two-level model with the GPE result $\chi(t)$ (see equation (13)), which exceeds a value of 0.95 at all times t. (d) Momentum dynamics arising from time-dependent superposition of ϕ_0, ϕ_1 in the co-oscillating frame, data set V. Note the strong beating at intermediate times/excited fractions.

As laid out in section 4.2, the appropriate ground state for the internal conversion dynamics can also be determined in the experiment from the centre-of-mass position $K_t(t)$ of the twin-atom beams which the excited state is decaying into continuously. For times t, where the decayed fraction becomes perceivable, we can compare the experimentally found $K_t(y)$ to $K_0(y)$ as in figure 14, and find good agreement without any free parameter over a large range of settings. Together with the absence of decay products from higher excited states in the experiment, this result confirms the validity of the decomposition approach. In [36] it has been shown, that the obtained populations of the excited state lead to an accurate quantitative description of the ensuing decay process.

5. Conclusion

In conclusion, we have presented successful application of optimal control theory to the problem of preparing a non-classical, strongly out-of-equilibrium motional state of a Bose–Einstein condensate, realizing population inversion with near-unity fidelity. The obtained condensate wave function corresponds to the first excited eigenstate of the Gross–Pitaevskii equation, closely resembling the first odd Fock state of a harmonic oscillator. To manipulate the external state of the Bose–Einstein condensate, we used precisely controlled motion of an anharmonic trap potential along the optimized trajectory. Experimental and numerical results on the momentum distribution dynamics during and after the excitation sequence show excellent agreement over a large range of parameters, including tuning of many-body effects. Moreover, a model of the excitation dynamics based on decomposition into a quasi-classical oscillation and the actual state transfer has been developed, and shown to be consistent with various observations made in both experiment and theory. Using this approach, we were able to deduce an approximate two-level description of the excitation process.

A first application of the vibrational state inversion, using the condensate as a gain medium for matter wave amplification, has been demonstrated in [35, 36]. However, optimal control in condensates is not restricted to high-fidelity preparation of a desired wave function, and more general pulses, that e.g. act on non-stationary initial states in a phase-sensitive manner can be implemented [71]. Also, given the excellent control of the trapping potential, other parameters than a simple one-dimensional displacement are readily accessible to study properties of more complex excited states [41]. State preparation beyond a mean-field description has been proposed, including entanglement generation [32, 33], number-squeezed states [34], or cooling [31], which should be realizable in a similar fashion. More generally, our results highlight the potential of experiments with Bose–Einstein condensates as a test-bed for a large range of quantum control problems, as known from NMR spectroscopy [13–15], solid-state [20–25], atomic [19], or molecular physics [16–18].

Acknowledgments

We acknowledge financial support from the Austrian Science Fund projects P24248, CAP, SFB FoQuS, the FWF doctoral program CoQuS (W 1210), and the EU project AQUTE.

References

[1] Brif C, Chakrabarti R and Rabitz H 2010 Control of quantum phenomena: past, present and future *New J. Phys.* **12** 075008

[2] Jin D, Ensher J, Matthews M, Wieman C and Cornell E 1996 Collective excitations of a Bose–Einstein condensate in a dilute gas *Phys. Rev. Lett.* **77** 420–3

[3] Mewes M-O, Andrews M, van Druten N, Kurn D, Durfee D, Townsend C and Ketterle W 1996 Collective excitations of a Bose–Einstein condensate in a magnetic trap *Phys. Rev. Lett.* **77** 988–91

[4] Couvert A, Kawalec T, Reinaudi G and Guéry-Odelin D 2008 Optimal transport of ultracold atoms in the non-adiabatic regime *Europhys. Lett.* **83** 13001

[5] Muga J G, Chen X, Ruschhaupt A and Guéry-Odelin D 2009 Frictionless dynamics of Bose–Einstein condensates under fast trap variations *J. Phys. B: At. Mol. Opt. Phys.* **42** 241001

[6] Schaff J-F, Song X-L, Capuzzi P, Vignolo P and Labeyrie G 2011 Shortcut to adiabaticity for an interacting Bose–Einstein condensate *Europhys. Lett.* **93** 23001

[7] Schaff J-F, Capuzzi P, Labeyrie G and Vignolo P 2011 Shortcuts to adiabaticity for trapped ultracold gases *New J. Phys.* **13** 113017

[8] Bason M G, Viteau M, Malossi N, Huillery P, Arimondo E, Ciampini D, Fazio R, Giovannetti V, Mannella R and Morsch O 2011 High-fidelity quantum driving *Nature Phys.* **8** 147–52

[9] del Campo A and Boshier M G 2012 Shortcuts to adiabaticity in a time-dependent box *Sci. Rep.* **2** 648

[10] Kagan Y, Surkov E and Shlyapnikov G 1996 Evolution of a Bose-condensed gas under variations of the confining potential *Phys. Rev.* A **54** R1753–6

[11] Castin Y and Dum R 1996 Bose–Einstein condensates in time dependent traps *Phys. Rev. Lett.* **77** 5315–9

[12] Yukalov V, Yukalova E and Bagnato V 1997 Non-ground-state Bose–Einstein condensates of trapped atoms *Phys. Rev.* A **56** 4845–54

[13] Skinner T E, Reiss T O, Luy B, Khaneja N and Glaser S J 2003 Application of optimal control theory to the design of broadband excitation pulses for high-resolution NMR *J. Magn. Reson.* **163** 8–15

[14] Skinner T E, Reiss T O, Luy B, Khaneja N and Glaser S J 2004 Reducing the duration of broadband excitation pulses using optimal control with limited RF amplitude *J. Magn. Reson.* **167** 68–74

[15] Khaneja N, Reiss T, Kehlet C, Schulte-Herbrüggen T and Glaser S J 2005 Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms *J. Magn. Reson.* **172** 296–305

[16] Assion A, Baumert T, Bergt M, Brixner T, Kiefer B, Seyfried V, Strehle M and Gerber G 1998 Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses *Science* **282** 919–22

[17] Schirmer S, Fu H and Solomon A 2001 Complete controllability of quantum systems *Phys. Rev.* A **63** 063410

[18] Tesch C and de Vivie-Riedle R 2002 Quantum computation with vibrationally excited molecules *Phys. Rev. Lett.* **89** 157901

[19] Vasilev G, Kuhn A and Vitanov N 2009 Optimum pulse shapes for stimulated Raman adiabatic passage *Phys. Rev.* A **80** 013417

[20] Hofheinz M, Weig E M, Ansmann M, Bialczak R C, Lucero E, Neeley M, O'Connell A D, Wang H, Martinis J M and Cleland A N 2008 Generation of Fock states in a superconducting quantum circuit *Nature* **454** 310–4

[21] Hohenester U 2006 Optimal quantum gates for semiconductor qubits *Phys. Rev.* B **74** 161307

[22] Safaei S, Montangero S, Taddei F and Fazio R 2009 Optimized single-qubit gates for Josephson phase qubits *Phys. Rev.* B **79** 064524

[23] Jirari H, Hekking F W J and Buisson O 2009 Optimal control of superconducting N-level quantum systems *Europhys. Lett.* **87** 28004

[24] Rebentrost P and Wilhelm F 2009 Optimal control of a leaking qubit *Phys. Rev.* B **79** 060507

[25] Khani B, Gambetta J M, Motzoi F and Wilhelm F K 2009 Optimal generation of Fock states in a weakly nonlinear oscillator *Phys. Scr.* T **137** 014021

[26] Giovannetti V 2003 The quantum speed limit *Proc. SPIE* **5111** 1–6

[27] Caneva T, Murphy M, Calarco T, Fazio R, Montangero S, Giovannetti V and Santoro G E 2009 Optimal control at the quantum speed limit *Phys. Rev. Lett.* **103** 240501

[28] del Campo A, Egusquiza I L, Plenio M B and Huelga S F 2013 Quantum speed limits in open system dynamics *Phys. Rev. Lett.* **110** 050403

[29] Hohenester U, Rekdal P K, Borzì A and Schmiedmayer J 2007 Optimal quantum control of Bose–Einstein condensates in magnetic microtraps *Phys. Rev.* A **75** 023602

[30] De Chiara G, Calarco T, Anderlini M, Montangero S, Lee P, Brown B, Phillips W and Porto J 2008 Optimal control of atom transport for quantum gates in optical lattices *Phys. Rev.* A **77** 052333

[31] Rahmani A, Kitagawa T, Demler E and Chamon C 2012 Cooling through optimal control of quantum evolution (arXiv:1210.5244)

[32] Platzer F, Mintert F and Buchleitner A 2010 Optimal dynamical control of many-body entanglement *Phys. Rev. Lett.* **105** 020501

[33] Caneva T, Calarco T and Montangero S 2012 Entanglement-storage units *New J. Phys.* **14** 093041

[34] Grond J, Schmiedmayer J and Hohenester U 2009 Optimizing number squeezing when splitting a mesoscopic condensate *Phys. Rev.* A **79** 21603

[35] Bücker R, Grond J, Manz S, Berrada T, Betz T, Koller C, Hohenester U, Schumm T, Perrin A and Schmiedmayer J 2011 Twin-atom beams *Nature Phys.* **7** 608–11

[36] Bücker R, Hohenester U, Berrada T, van Frank S, Perrin A, Manz S, Betz T, Grond J, Schumm T and Schmiedmayer J 2012 Dynamics of parametric matter-wave amplification *Phys. Rev.* A **86** 013638

[37] Jacak L, Hawrylak P and Wojs A 1998 *Quantum Dots* (Berlin: Springer)

[38] Walls D F and Milburn G J 2007 *Quantum Optics* 2nd edn (Berlin: Springer)

[39] García-Ripoll J, Pérez-García V and Vekslerchik V 2001 Construction of exact solutions by spatial translations in inhomogeneous nonlinear Schrödinger equations *Phys. Rev.* E **64** 056602

[40] Bialynicki-Birula I and Bialynicka-Birula Z 2002 Center-of-mass motion in the many-body theory of Bose–Einstein condensates *Phys. Rev.* A **65** 063606

[41] Yukalov V, Marzlin K-P and Yukalova E 2004 Resonant generation of topological modes in trapped Bose–Einstein gases *Phys. Rev.* A **69** 023620

[42] Lesanovsky I, Schumm T, Hofferberth S, Andersson L M, Krüger P and Schmiedmayer J 2006 Adiabatic radio-frequency potentials for the coherent manipulation of matter waves *Phys. Rev.* A **73** 33619

[43] Peirce A P, Dahleh M A and Rabitz H 1988 Optimal control of quantum-mechanical systems: existence, numerical application, and applications *Phys. Rev.* A **37** 4950

[44] Rabitz H A, de Vivie-Riedle R, Motzkus M and Kompka K 2000 Wither the future of controlling quantum phenomena *Science* **288** 824

[45] Borzì A, Stadler G and Hohenester U 2002 Optimal quantum control in nanostructures: theory and application to generic three-level system *Phys. Rev.* A **66** 053811

[46] Leggett A 2001 Bose–Einstein condensation in alkali gases: some fundamental concepts *Rev. Mod. Phys.* **73** 307

[47] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 2002 *Numerical Recipes in C++: The Art of Scientific Computing* 2nd edn (Cambridge: Cambridge University Press)

[48] Grond J, von Winckel G, Schmiedmayer J and Hohenester U 2009 Optimal control of number squeezing in trapped Bose–Einstein condensates *Phys. Rev.* A **80** 053625

[49] Bertsekas D P 1999 *Nonlinear Programming* (Cambridge: Athena Scientific)

[50] Rabitz H A, Hsieh M M and Rosenthal C M 2004 Quantum optimally controlled transition landscapes *Science* **303** 1998

[51] Reichel J and Vuletic V (ed) 2011 *Atom Chips* (Weinheim: Wiley-VCH)

[52] Trinker M, Groth S, Haslinger S, Manz S, Betz T, Schneider S, Bar-Joseph I, Schumm T and Schmiedmayer J 2008

Multilayer atom chips for versatile atom micromanipulation *Appl. Phys. Lett.* **92** 254102

[53] Schumm T, Hofferberth S, Andersson L M, Wildermuth S, Groth S, Bar-Joseph I, Schmiedmayer J and Krüger P 2005 Matter-wave interferometry in a double well on an atom chip *Nature Phys.* **1** 57–62

[54] Kheruntsyan K V, Gangardt D M, Drummond P D and Shlyapnikov G V 2003 Pair correlations in a finite-temperature 1D Bose gas *Phys. Rev. Lett.* **91** 40403

[55] Petrov D S, Shlyapnikov G V and Walraven J T M 2000 Regimes of quantum degeneracy in trapped 1D gases *Phys. Rev. Lett.* **85** 3745–9

[56] Zobay O and Garraway B M 2004 Atom trapping and two-dimensional Bose–Einstein condensates in field-induced adiabatic potentials *Phys. Rev.* A **69** 23605

[57] Hofferberth S, Lesanovsky I, Fischer B, Schumm T and Schmiedmayer J 2007 Non-equilibrium coherence dynamics in one-dimensional Bose gases *Nature* **449** 324–7

[58] Hofferberth S, Lesanovsky I, Schumm T, Imambekov A, Gritsev V, Demler E and Schmiedmayer J 2008 Probing quantum and thermal noise in an interacting many-body system *Nature Phys.* **4** 489–95

[59] Betz T *et al* 2011 Two-point phase correlations of a one-dimensional bosonic Josephson junction *Phys. Rev. Lett.* **106** 020407

[60] Gring M, Kuhnert M, Langen T, Kitagawa T, Rauer B, Schreitl M, Mazets I, Adu Smith D, Demler E and Schmiedmayer J 2012 Relaxation and prethermalization in an isolated quantum system *Science* **337** 1318–22

[61] Hofferberth S, Fischer B, Schumm T, Schmiedmayer J and Lesanovsky I 2007 Ultracold atoms in radio-frequency dressed potentials beyond the rotating-wave approximation *Phys. Rev.* A **76** 13401

[62] Shirley J H 1965 Solution of the Schrödinger equation with a Hamiltonian periodic in time *Phys. Rev.* **138** B979–87

[63] Amore P and Sáenz R A 2005 The period of a classical oscillator *Europhys. Lett.* **70** 425–31

[64] Pethick C J and Smith H 2002 *Bose–Einstein Condensation in Dilute Atomic Gases* (Cambridge: Cambridge University Press)

[65] Salasnich L, Parola A and Reatto L 2002 Effective wave equations for the dynamics of cigar-shaped and disk-shaped Bose condensates *Phys. Rev.* A **65** 043614

[66] Bücker R, Perrin A, Manz S, Betz T, Koller C, Plisson T, Rottmann J, Schumm T and Schmiedmayer J 2009 Single-particle-sensitive imaging of freely propagating ultracold atoms *New J. Phys.* **11** 103039

[67] Rohringer W, Bücker R, Manz S, Betz T, Koller Ch, Göbel M, Perrin A, Schmiedmayer J and Schumm T 2008 Stochastic optimization of a cold atom experiment using a genetic algorithm *Appl. Phys. Lett.* **93** 264101

[68] Krüger P, Hofferberth S, Mazets I, Lesanovsky I and Schmiedmayer J 2010 Weakly interacting Bose gas in the one-dimensional limit *Phys. Rev. Lett.* **105** 265302

[69] Yukalov V, Yukalova E and Bagnato V 2002 Nonlinear coherent modes of trapped Bose–Einstein condensates *Phys. Rev.* A **66** 043602

[70] de Oliveira F A M, Kim M S, Knight P L and Buek V 1990 Properties of displaced number states *Phys. Rev.* A **41** 2645–52

[71] van Frank S *et al* 2013 in preparation

5. Paper 2: Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Consideration of filter effects

Published in: Physical Review A **88** (2013)

PHYSICAL REVIEW A **88**, 035601 (2013)

Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Consideration of filter effects

Georg Jäger and Ulrich Hohenester
Institut für Physik, Karl-Franzens-Universität Graz, Universitätsplatz 5, 8010 Graz, Austria

(Received 19 July 2013; published 9 September 2013)

We theoretically investigate protocols based on optimal control theory (OCT) for manipulating Bose-Einstein condensates in magnetic microtraps, using the framework of the Gross-Pitaevskii equation. In our approach we explicitly account for filter functions that distort the computed optimal control, a situation inherent to many experimental OCT implementations. We apply our scheme to the shakeup process of a condensate from the ground to the first excited state, following a recent experimental and theoretical study, and demonstrate that the fidelity of OCT protocols is not significantly deteriorated by typical filters.

DOI: 10.1103/PhysRevA.88.035601 PACS number(s): 03.75.−b, 37.10.Gh, 02.60.Pn

Introduction. Optimal quantum control aims at the manipulation of a quantum-mechanical wave function in a controlled fashion [1–3]. External parameters, such as laser fields, can be controlled at will and allow us to steer the wave function from a given initial to a desired terminal state. Recent years have seen tremendous research efforts in the realm of quantum control [3]. Quantum chemistry implementations often rely on stochastic optimization techniques, which are particulary appealing for experimental implementations [2]. An alternative approach is provided by optimal control theory (OCT) [1,4,5], which performs a numerical optimization of the control fields through an iterative procedure by solving the dynamic system equations.

In Refs. [6,7] we presented an experimental implementation of optimal quantum control for a Bose-Einstein condensate. Ultracold atoms become trapped in the vicinity of an atom chip [8] by the magnetic fields produced by currents running through the wires of the chip, and the magnetic confinement potential can be controlled by changing the currents. We have demonstrated the excitation of the condensate wave function from the ground to the first excited state of an anharmonic potential, where the population transfer has been achieved with an efficiency close to 100% by displacing the potential minimum according to a protocol computed with OCT.

In this Brief Report, we investigate the effects of filter functions that distort the control parameters computed from OCT. Such filters might be due to electronics and are inherent to many experiments. For sufficiently simple control protocols filter effects can be corrected through a simple deconvolution scheme, but in general it is advantageous to incorporate filtering directly in the OCT approach. In this paper we first develop the methodology for OCT with filtered control parameters and then apply our scheme to the condensate shakeup investigated in Refs. [6,7]. We find that for realistic filter functions the fidelity of the control process does not become deteriorated significantly. Although in this paper we focus only on Bose-Einstein condensates, the developed methodology is general and might be useful in a much wider context.

OCT without a filter. We first briefly review the optimal control implementation of Bose-Einstein condensates formulated in Ref. [9] (see also Ref. [10] for related work). Within the framework of the Gross-Pitaevskii equation [11,12] the dynamics of the condensate wave function $\psi(\boldsymbol{r},t)$ is described by ($\hbar = 1$)

$$i\frac{\partial\psi(\boldsymbol{r},t)}{\partial t} = \left(-\frac{\nabla^2}{2M} + V(\boldsymbol{r},\lambda(t)) + \kappa|\psi(\boldsymbol{r},t)|^2\right)\psi(\boldsymbol{r},t). \tag{1}$$

The first term on the right-hand side is the operator for the kinetic energy; the second one is the confinement potential, which can be controlled by some external parameter $\lambda(t)$; and the last term is the nonlinear atom-atom interaction in the mean-field approximation of the Gross-Pitaevskii framework. M is the atom mass and κ is the strength of the atom-atom interactions.

OCT is seeking an "optimal" time variation of the control parameter $\lambda(t)$ in order to fulfill certain control objectives. For instance, the cost function

$$J(\psi,\lambda) = \frac{1}{2}[1 - |\langle\psi_d|\psi(T)\rangle|^2] + \frac{\gamma}{2}\int_0^T [\dot{\lambda}(t)]^2 dt \tag{2}$$

becomes minimal when state $\psi(T)$ at the terminal time T of the control process comes as close as possible to the *desired* state ψ_d, apart from an irrelevant global phase [9]. The second term penalizes strong variations of the control parameter and is needed to make the OCT problem well posed [9,13,14]. Through γ it is possible to weight the importance of wave-function matching and control smoothness, and below we set $\gamma \ll 1$.

In order to bring the system from the initial state ψ_0 to the terminal state $\psi(T)$ we have to fulfill the Gross-Pitaevskii equation, which enters as a *constraint* in our optimization problem. The constrained optimization problem can be turned into an unconstrained one by means of Lagrange multipliers $p(t)$. To this end, we introduce a Lagrange function,

$$L(\psi,p,\lambda) = J(\psi,\lambda) + \mathrm{Re}\left[\int_0^T \langle p(t)|i\dot{\psi}(t) - [H_\lambda + \kappa|\psi(t)|^2]\psi(t)\rangle dt\right], \tag{3}$$

where H_λ is the single-particle Hamiltonian defined through Eq. (1). The Lagrange function has a saddle point at the minimum of $J(\psi,\lambda)$ where all derivatives $\delta L/\delta\psi^*$, $\delta L/\delta p^*$, and $\delta L/\delta\lambda$ become 0. Performing functional derivatives in

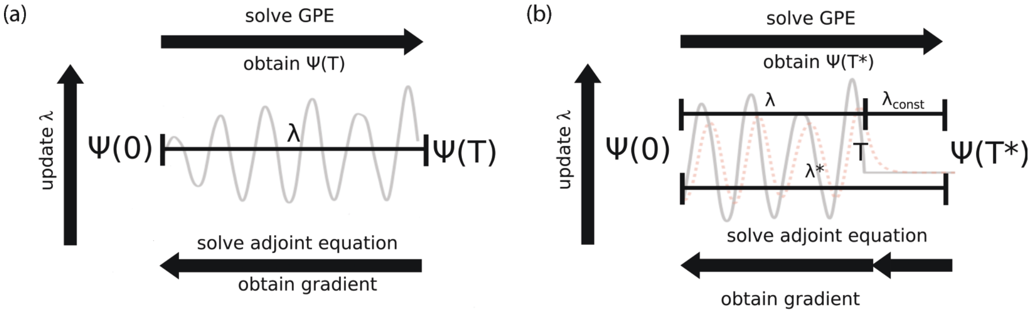

FIG. 1. (Color online) Optimization strategy (a) without and (b) with a filter function. When a filter is considered, the optimization interval $[0,T]$ should differ from the time interval $[0,T^\star]$ of the condensate dynamics to account for the finite filter response time.

the Lagrange function, we then arrive at the following set of equations [9,13]:

$$i\dot\psi = \left(-\frac{\nabla^2}{2M} + V(\boldsymbol{r},\lambda(t)) + \kappa|\psi|^2\right)\psi, \tag{4a}$$

$$i\dot p = \left(-\frac{\nabla^2}{2M} + V(\boldsymbol{r},\lambda(t)) + 2\kappa|\psi|^2\right)p + \kappa\psi^2 p^*, \tag{4b}$$

$$\gamma\ddot\lambda = -\mathrm{Re}\left\langle p\left|\frac{\partial H_\lambda}{\partial\lambda}\right|\psi\right\rangle. \tag{4c}$$

Equation (4a) is the initial value problem $\psi(0) = \psi_0$ of the Gross-Pitaevskii equation, whereas Eq. (4b) is a terminal value problem for the adjoint variable $p(T) = i\langle\psi_d|\psi(T)\rangle\psi_d$. Finally, Eq. (4c) determines the optimal control and is a boundary value problem where both the initial and the terminal values are fixed, $\lambda(0) = \lambda_0$ and $\lambda(T) = \lambda_T$. It has been discussed in Refs. [9,14,15] that Eqs. (4a)–(4c) can be also used for arbitrary control parameters in order to formulate an iterative procedure that successively improves $\lambda(t)$.

OCT with a filter. In this paper we discuss the case where the external control parameter $\lambda(t)$ does not directly influence the confinement potential but becomes distorted by some filter function $h(\tau)$. Such filtering is inherent to many quantum control experiments and has previously been discussed also in the context of laser pulse shaping [16,17]. Figure 2(b) shows a typical filter function due to the finite bandwidth of the electronics [7]. As a consequence, when the optimized control $\lambda(t)$ is sent to the electronics, a distorted (nonoptimal) signal,

$$\lambda^\star(t) = \int_0^t h(\tau)\lambda(t-\tau)\,d\tau, \tag{5}$$

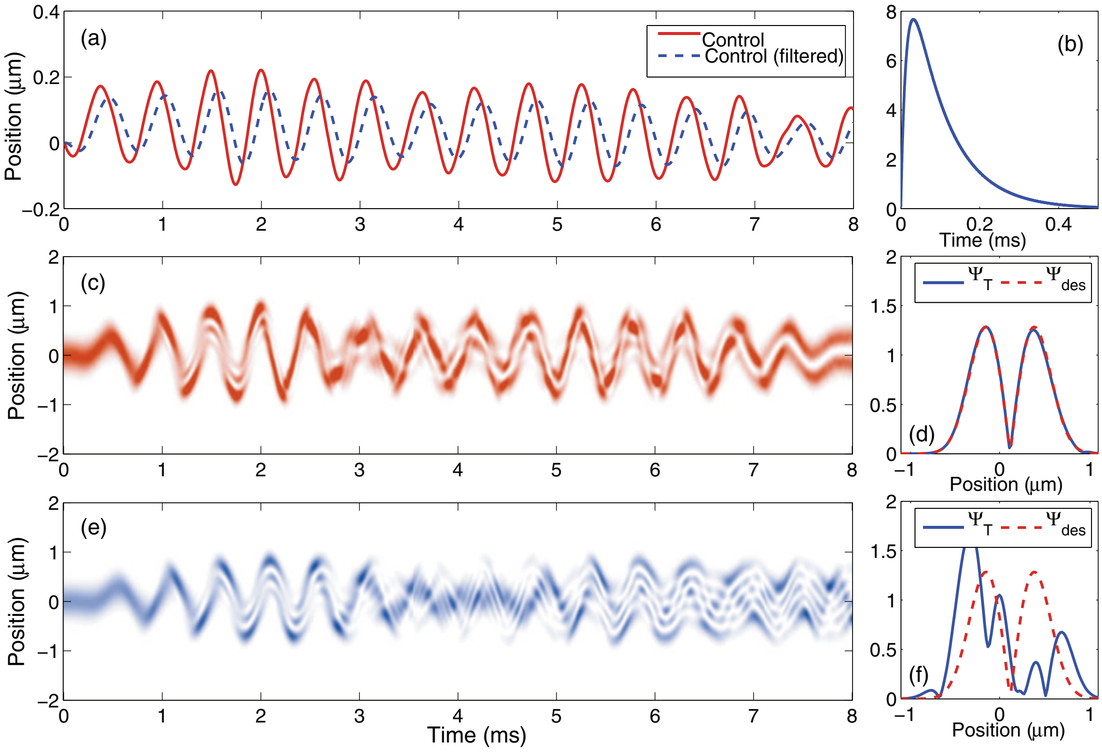

FIG. 2. (Color online) (a) Filtered and unfiltered control for λ obtained *without* the consideration of a filter. (b) Typical filter function $h(t)$ taken from experiment [7]; the impulse response of the electronics is modelled as a second-order linear filter with a 4.4 kHz cutoff frequency and a quality factor of 0.34. (c) Time evolution of the condensate density under the effect of the unfiltered control. (d) Absolute value of desired and terminal wave function for unfiltered control at terminal time T of control process. (e, f) Same as (c, d) but for the filtered control $\lambda^\star(t)$.

determines the time evolution of $V(\boldsymbol{r},\lambda^\star(t))$. It is obvious that the condensate dynamics with the nonoptimized $\lambda^\star(t)$ will no longer bring the condensate wave function to the desired state.

A possible solution is to try to deconvolute Eq. (5) and to find an input signal that produces the proper output signal. However, since h is typically a filter for high frequencies, and the optimal control might include fast time variations, the deconvolution is often doomed to failure. In what follows we formulate a different strategy. We stay with the OCT framework of Eqs. (1)–(3) but replace $\lambda(t)$ with the filtered $\lambda^\star(t)$. Our task now is to determine the optimal control $\lambda(t)$ such that the filtered control of Eq. (5) brings the condensate from ψ_0 to ψ_d. In fact, we can carry over most of the results of the previous discussion. However, in Eqs. (4a) and (4b) we have to replace $\lambda(t)$ with $\lambda^\star(t)$, and the control, Eq. (4c), is changed to the form

$$\gamma\ddot{\lambda} = -\mathrm{Re}\int_t^T h(s-t)\left\langle p(s)\left|\frac{\partial V}{\partial\lambda}\right|\psi(s)\right\rangle ds, \qquad (6)$$

which is now nonlocal in time (note that $\partial V/\partial\lambda$ is evaluated at time s). The optimality system is then formed by Eqs. (4a) and (4b), with λ replaced by $\lambda^\star$, together with Eq. (6).

A slight complication appears at this point. Let us consider the schematic OCT loop depicted in Fig. 1. In the unfiltered case in Fig. 1(a) the optimal control parameter $\lambda(t)$ determines how the condensate wave function is brought to ψ_d. If ψ_d is a stationary state of the Gross-Pitaevskii equation and λ is kept fixed at $t \geqslant T$, the system will remain in this stationary state. On the other hand, in the presence of a filter things behave differently. For $\lambda(t \geqslant T) = \lambda_T$ the filtered response $\lambda^\star(t)$ can still vary at times later than T, because of the finite response time of the filter. Thus, even if the system ends up in the desired state $\psi(T) = \psi_d$ at the terminal time T, the ensuing temporal evolution of $\lambda^\star$ will push the system away from ψ_d.

To account for this, we propose a slight variation of our OCT implementation. We use different time intervals $[0,T]$ and $[0,T^\star]$ for the control optimization and the condensate simulation, respectively. The two end points differ by $T^\star - T = \tau^\star$, as shown in Fig. 1(b). $\tau^\star$ is a time at which the filter function has dropped to 0 (approximately the inverse of the cutoff frequency). Thus, if λ is kept constant at times $t \geqslant T$, the filtered response $\lambda^\star$ will become constant at $t \geqslant T^\star$.

In the presence of filtering, the OCT loop formulated in Refs. [9,13,14] consists of the following steps.

(1) Start with some initial guess for $\lambda(t)$, where λ is kept constant for $t \geqslant T$.

(2) Solve the Gross-Pitaevskii equation, (4a), forward in the time interval $[0,T^\star]$ using the filtered control $\lambda^\star$.

(3) Determine the terminal condition $p(T^\star)$ for the adjoint variable, and integrate the sensitivity, Eq. (4b), backward in time.

(4) Compute a search direction for $\lambda(t)$ via Eq. (6).

(5) Compute an improved control parameter $\lambda(t)$ within the time interval $[0,T]$, and iterate the loop 2–5 until convergence is obtained or a given number of iterations is reached.

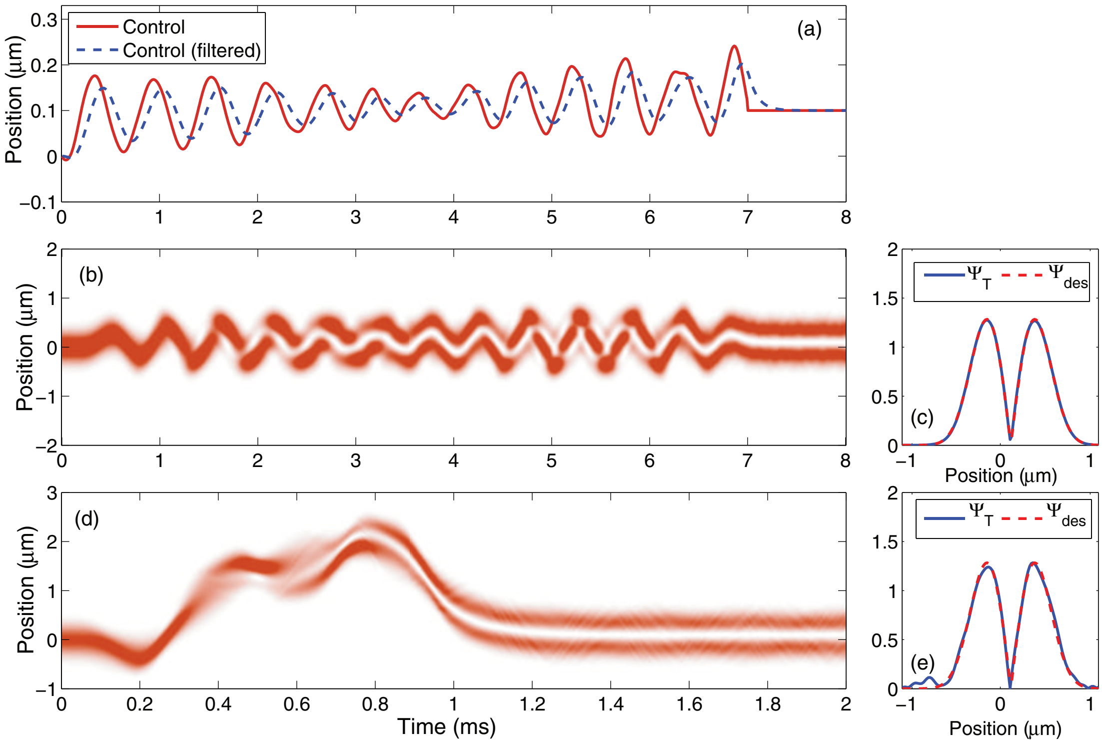

FIG. 3. (Color online) (a) Optimal control *with* consideration of the filter function. The last part of the time evolution of λ is kept constant in order to guarantee $\lambda^\star(T^\star) = \lambda_T$. (b) Time development of the condensate density driven by the optimal control for a transition time of 8 ms. (c) Absolute value of desired and terminal wave function at the final time. (d, e) Same as (b, c), but for the shorter transition time of 2 ms.

Results. We next apply the OCT implementation including filter functions to the case of condensate shakeup described in Refs. [6,7]. Here one starts in the ground state of an anharmonic trap, and the trap center is displaced in an optimized fashion such that the condensate is brought to the first excited state. We only consider the spatial dynamics along the displacement direction, and use the same trap and condensate parameters as in Refs. [6,7]. The solid line in Fig. 2(a) shows the optimized control for a shakeup process, using a time interval of 8 ms and discarding filter effects. Here $\lambda(t)$ directly corresponds to the displacement of the trap center. The density plot in Fig. 2(c) provides details about how the condensate is brought from the initial to the desired state. Figure 2(d) demonstrates that $\psi(T)$ and ψ_d almost perfectly match at the terminal time T. In the presence of a filter function, depicted in Fig. 2(b), the control parameter becomes distorted [dashed line in Fig. 2(a)] and the control process leaves the system in a highly excited and nonstationary state, as shown in Figs. 2(e) and 2(f).

Things change considerably for the OCT implementation with filtering, introduced above. Figure 3(a) shows the optimized control $\lambda(t)$ together with the filtered signal $\lambda^\star(t)$. In our OCT simulations we set the filter response time $\tau^\star$ to 0.5 ms. Figures 3(b) and 3(c) show that with the optimized control parameter the desired state is reached perfectly, even in the presence of filtering. In Fig. 3(d) we finally demonstrate that the shakeup protocol also works for significantly shorter transition times, here 2 ms. In Fig. 4 we investigate the success for the shakeup process for different times $T^\star$. For small transition times, say around 2 ms, the transfer has a high efficiency but the terminal state still differs somewhat from the desired one. With increasing transition time the efficiency increases, and the cost function saturates at later times.

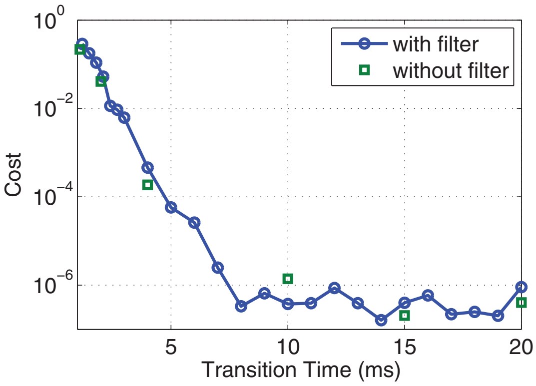

FIG. 4. (Color online) Final cost for different transition times (we set $\gamma = 10^{-9}$). For transition times under 8 ms the final cost begins to rise; for transition times under 2 ms no satisfying control could be obtained.

In conclusion, we have developed a methodology that allows us to incorporate filter effects in optimal quantum control simulations. We have applied our approach to a condensate shakeup process, where the wave function is transferred from the ground to the first excited state of an anharmonic, magnetic microtrap, and have demonstrated that high transfer efficiencies can be achieved even in the presence of filtering. Although in this paper we have only focused on Bose-Einstein condensates, the developed methodology is rather general and might also be useful for other systems.

Acknowledgments. We are grateful to Jörg Schmiedmayer, Robert Bücker, Tarik Berrada, and Thorsten Schumm for most helpful discussions and for providing us with the response function of the control electronics. This work was supported in part by the Austrian science fund FWF under Project No. P24248.

---

[1] A. P. Peirce, M. A. Dahleh, and H. Rabitz, Phys. Rev. A **37**, 4950 (1988).
[2] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompka, Science **288**, 824 (2000).
[3] C. Brif, R. Chakrabarti, and H. Rabitz, New J. Phys. **12**, 075008 (2010).
[4] D. J. Tannor and S. A. Rice, J. Chem. Phys. **83**, 5013 (1985).
[5] W. Zhu and H. Rabitz, Phys. Rev. A **58**, 4741 (1998).
[6] R. Bücker, J. Grond, S. Manz, T. Berrada, T. Betz, C. Koller, U. Hohenester, T. Schumm, A. Perrin, and J. Schmiedmayer, Nature Phys. **7**, 508 (2011).
[7] R. Bücker, T. Berrada, S. van Frank, T. Schumm, J. F. Schaff, J. Schmiedmayer, G. Jäger, J. Grond, and U. Hohenester, J. Phys. B **46**, 104012 (2013).
[8] J. Reichel and V. Vuletic (eds.), *Atom Chips* (Wiley, Weinheim, 2011).
[9] U. Hohenester, P. K. Rekdal, A. Borzi, and J. Schmiedmayer, Phys. Rev. A **75**, 023602 (2007).
[10] S. E. Sklarz and D. J. Tannor, Phys. Rev. A **66**, 053619 (2002).
[11] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, Rev. Mod. Phys. **71**, 463 (1999).
[12] A. Leggett, Rev. Mod. Phys. **73**, 307 (2001).
[13] A. Borzi and U. Hohenester, SIAM J. Sci. Comput. **30**, 441 (2008).
[14] G. von Winckel and A. Borzi, Inverse Probl. **24**, 034007 (2008).
[15] J. Grond, G. von Winckel, J. Schmiedmayer, and U. Hohenester, Phys. Rev. A **80**, 053625 (2009).
[16] K. Sundermann, H. Rabitz, and R. de Vivie-Riedle, Phys. Rev. A **62**, 013409 (2000).
[17] C. Gollub, M. Kowalewski, and R. de Vivie-Riedle, Phys. Rev. Lett. **101**, 073002 (2008).

6. Paper 3: Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Comparison of gradient-ascent-pulse-engineering and Krotov optimization schemes

Published in: Physical Review A **90** (2014)

PHYSICAL REVIEW A **90**, 033628 (2014)

Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Comparison of gradient-ascent-pulse-engineering and Krotov optimization schemes

Georg Jäger,[1] Daniel M. Reich,[2] Michael H. Goerz,[2] Christiane P. Koch,[2] and Ulrich Hohenester[1]
[1]*Institut für Physik, Karl-Franzens-Universität Graz, Universitätsplatz 5, 8010 Graz, Austria*
[2]*Theoretische Physik, Universität Kassel, Heinrich-Plett-Str. 40, 34132 Kassel, Germany*

(Received 11 August 2014; published 25 September 2014)

We study optimal quantum control of the dynamics of trapped Bose-Einstein condensates: The targets are to split a condensate, residing initially in a single well, into a double well, without inducing excitation, and to excite a condensate from the ground state to the first-excited state of a single well. The condensate is described in the mean-field approximation of the Gross-Pitaevskii equation. We compare two optimization approaches in terms of their performance and ease of use; namely, gradient-ascent pulse engineering (GRAPE) and Krotov's method. Both approaches are derived from the variational principle but differ in the way the control is updated, additional costs are accounted for, and second-order-derivative information can be included. We find that GRAPE produces smoother control fields and works in a black-box manner, whereas Krotov with a suitably chosen step-size parameter converges faster but can produce sharp features in the control fields.

DOI: 10.1103/PhysRevA.90.033628 PACS number(s): 03.75.−b, 37.25.+k, 02.30.Yy

I. INTRODUCTION

Controlling complex quantum dynamics is a recurring theme in many different areas of atomic, molecular, and optical (AMO) physics and physical chemistry. Recent examples include quantum state preparation [1,2], interferometry [3] and imaging [4,5], or reaction control [6,7]. The central idea of quantum control is to employ external fields to steer the dynamics in a desired way [8,9]. The fields that realize the desired dynamics can be determined by optimal control theory (OCT) [10,11]. An expectation value that encodes the target is then taken to be a functional of the external field which is minimized or maximized. The target can be simply a desired final state [10] or a unitary operator [12], a prescribed value of energy or position [13], or an experimental signal such as a pump-probe trace [14].

The algorithms that can be employed for optimizing the target functional broadly fall into two categories—those where changes in the field are determined solely by evaluating the functional, such as simplex algorithms [13,15], and those that utilize derivative information, such as Krotov's method [16,17] or gradient-ascent pulse engineering (GRAPE) [18], possibly combined with quasi-Newton methods [19,20]. The solutions that one obtains typically depend not only on the target functional but also on the specific algorithm that is employed and the initial-guess field. This is due to the fact that numerical optimization is always a local search which may find one of possibly many optimal solutions or get stuck in a local extremum. It is thus important to understand which features of an optimal control solution are due to the optimization procedure and which reflect truly physical properties of the quantum system.

For example, when seeking to identify, by use of optimal control theory, the quantum speed limit, i.e., the shortest possible time in which a quantum operation can be carried out [21], the answer should be independent of the algorithm. Moreover, in view of employing calculated solutions in an experiment, conditions such as limited power, limited time resolution, or limited bandwidth need to be met. The way in which the various optimization approaches can accommodate such requirements differ greatly.

Here, we study control of a Bose-Einstein condensate in a magnetic microtrap, comparing several variants of a GRAPE-type algorithm [22,23] with Krotov's method [16,17,24]. We consider two targets—splitting the condensate, which resides initially in the ground state of a single well, into a double well, without inducing excitation, and exciting the condensate from the ground to the first-excited state of a single well. The latter is important for stimulated processes in matter waves, whereas the former presents a crucial step in interferometry [25–27]. A challenging aspect of controlling a condensate is the nonlinearity of the equation of motion which can compromise or even prevent convergence of the optimization [17]. The two methods tackle this problem in different ways: GRAPE by computing the search direction for new control fields within the framework of Lagrange parameters and submitting the optimal control search to generic minimization routines [22,28], Krotov's method by accounting for the nonlinearity of the equations of motion in the monotonicity conditions when constructing the algorithm [16,17,24]. Furthermore, the methods differ in the way in which additional requirements such as smoothness of the control can be accounted for. We compare the two optimization approaches with respect to the solutions they yield as well as their performance and ease of use. Our study extends an earlier comparison of GRAPE-type algorithms with Krotov's method [19] that was concerned with the linear Schrödinger equation and with finite-size (spin-type) quantum systems.

Our paper is organized as follows: After introducing the equation of motion for the condensate dynamics together with the control targets in Sec. II, we briefly review the two optimization schemes in Sec. III. Section IV presents our results for wave-function splitting and shaking. Moreover, we investigate the influence of the nonlinearity, the performance of the two algorithms, and the smoothness of the optimized control in Secs. IV B to IV D. Our conclusions are presented in Sec. V.

II. MODEL AND OPTIMIZATION PROBLEM

In this paper we consider a quasi-one-dimensional (quasi-1D) condensate residing in a magnetic confinement potential $V(x,\lambda(t))$ that can be controlled by some external *control parameter* $\lambda(t)$ [2,22,23,29]. We describe the condensate dynamics within the mean-field framework of the Gross-Pitaevskii equation, where $\psi(x,t)$ is the condensate wave function, normalized to one, whose time evolution is governed by [30] ($\hbar = 1$)

$$i\frac{\partial\psi(x,t)}{\partial t} = \left[-\frac{1}{2M}\frac{\partial^2}{\partial x^2} + V(x,\lambda(t)) + \kappa|\psi(x,t)|^2\right] \times \psi(x,t). \qquad (1)$$

The first term on the right-hand side is the operator for the kinetic energy, the second one is the confinement potential, and the last term is the nonlinear atom-atom interaction in the mean-field approximation. M is the atom mass and κ is the strength of the nonlinear atom-atom interactions, which is related to the effective one-dimensional interaction strength U_0 and the number of atoms N through $\kappa = U_0(N-1)$ [31].

We can now formulate our optimal control problem. Suppose that the condensate is initially described by the wave function $\psi(x,0) = \psi_0(x)$ and the potential is varied in the time interval $[0,T]$. We are now seeking for an optimal time variation of $\lambda(t)$ that brings the terminal wave function $\psi(x,T)$ as close as possible to a *desired* wave function $\psi_d(x)$. To rate the success for a given control, we introduce the cost function

$$J_T(\psi(T)) = \tfrac{1}{2}[1 - |\langle\psi_d|\psi(T)\rangle|^2], \qquad (2)$$

which becomes zero when the terminal wave function matches the desired one up to an arbitrary phase. Optimal control theory aims at a $\lambda_{\mathrm{OCT}}(t)$ that minimizes Eq. (2).

III. OPTIMIZATION METHODS

In this paper, we apply two different optimal-control approaches; namely, a gradient-ascent-pulse-engineering (GRAPE) scheme [18] and Krotov's method [16,17,24], which is discussed separately below. An overview of the control approaches is given in Table I.

A. GRAPE: Functional and optimization scheme

The GRAPE scheme for Bose-Einstein condensates has been presented in detail elsewhere [22,23,29,32], for this reason we only briefly introduce the working equations. Experimentally, strong variations of the control parameter are difficult to achieve. Therefore, we add to the cost function an additional term [22,33,34],

$$J(\psi(T),\lambda) = J_T(\psi(T)) + \frac{\gamma}{2}\int_0^T [\dot\lambda(t)]^2 dt. \qquad (3)$$

Mathematically, the additional term penalizes strong variations of the control parameter and is needed to make the OCT problem well posed [22,33,34]. Through γ it is possible to weight the relative importance of wave-function matching and control smoothness. Below, we set $\gamma \ll 1$ such that J is dominated by the terminal cost J_T.

In order to bring the system from the initial state ψ_0 to the terminal state $\psi(T)$ we have to fulfill the Gross-Pitaevskii equation, which enters as a *constraint* in our optimization problem. The constrained optimization problem can be turned into an unconstrained one by means of Lagrange multipliers $p(x,t)$, whose time evolution is governed by [22]

$$i\dot p = \left(-\frac{1}{2M}\frac{\partial^2}{\partial x^2} + V(x,\lambda(t)) + 2\kappa|\psi|^2\right)p + \kappa\psi^2 p^*, \qquad (4)$$

subject to the terminal condition $p(T) = i\langle\psi_d|\psi(T)\rangle\psi_d$. The optimal control problem is then composed of the Gross-Pitaevskii equation (1) and Eq. (4), which must be fulfilled simultaneously together with [22]

$$\gamma\ddot\lambda = -\mathrm{Re}\langle p|\frac{\partial V}{\partial\lambda}|\psi\rangle \qquad (5)$$

for the optimal control. This expression differs from standard GRAPE [18] and results from minimizing changes in the control; cf. Eq. (3).

This set of equations can be also employed for a nonoptimal control where Eq. (5) is not fulfilled. In this case Eq. (1) is solved forwards in time and Eq. (4) backwards in time, and the search direction $\nabla_\lambda J$ for an improved control is calculated from one of the Eqs. [22,23,34]:

$$\nabla_\lambda J = -\gamma\ddot\lambda - \mathrm{Re}\langle p|\frac{\partial V}{\partial\lambda}|\psi\rangle \quad \text{for} \quad L^2 \text{ norm}, \qquad (6)$$

$$-\frac{d^2}{dt^2}[\nabla_\lambda J] = -\gamma\ddot\lambda - \mathrm{Re}\langle p|\frac{\partial V}{\partial\lambda}|\psi\rangle \quad \text{for} \quad H^1 \text{ norm}. \qquad (7)$$

These two expressions are obtained by interpreting, on the right-hand side of Eq. (3), the integral $\int_0^T[\dot\lambda]^2 dt = \langle\dot\lambda,\dot\lambda\rangle_{L^2} = \langle\lambda,\lambda\rangle_{H^1}$ in terms of an L^2 or H^1 norm [31,34]. The H^1 norm

TABLE I. Optimization approaches used in this paper. For each algorithm, we specify whether a line search is used, which free parameter is available to influence the optimization, the order of the derivative for the determination of the new control parameter, the penalty term that is added to Eq. (2), with $\Delta\lambda = \lambda - \lambda_{\mathrm{ref}}$, the equation for the cost function, the type of the update of the control, and the update equation used in our simulations.

Algorithm	Line search	Free parameter	Deriv.	Penalty	Penalty equation	Update	Update equation
GRAPE grad L2	Yes	γ	1	$\dot\lambda^2$	Eq. (3)	Concurrent	Eq. (6)
GRAPE grad H1	Yes	γ	1	$\dot\lambda^2$	Eq. (3)	Concurrent	Eq. (7)
GRAPE BFGS L2	Yes	γ	2	$\dot\lambda^2$	Eq. (3)	Concurrent	Eq. (6)
GRAPE BFGS H1	Yes	γ	2	$\dot\lambda^2$	Eq. (3)	Concurrent	Eq. (7)
Krotov	No	k	1	$(\Delta\lambda)^2$	Eq. (9)	Sequential	Eq. (10)
KBFGS	No	k	2	$(\Delta\lambda)^2$	Eq. (9)	Sequential	Eq. (38) of Ref. [20]

implies that one additionally has to solve a Poisson equation; see the derivative operator on the left-hand side of Eq. (7). This generally results in a much smoother time dependence of the control parameters while the additional numerical effort for solving the Poisson equation is negligible. As for an optimal control, both Eqs. (6) and (7) yield $\nabla_\lambda J = 0$.

Here we solve the optimal control equations using the Matlab toolbox OCTBEC [23]. The ground and desired states of the Gross-Pitaevskii equation are computed by using the optimal damping algorithm [23,35]. The control parameters are obtained iteratively by using either a conjugate gradient method (GRAPE grad), which only uses first-order information, or a quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) scheme [36] (GRAPE BFGS), which also takes into account second-order information via an approximated Hessian. In both cases, the optimization employs a line search to determine the optimal step size in the direction of a given gradient. The pulse update is calculated for all time points simultaneously, making the GRAPE schemes *concurrent*.

B. Krotov's method: Functional and optimization scheme

Krotov's method [16] provides an alternative optimal control implementation. The main idea is to add to Eq. (2) a vanishing term [16,17,24], which is chosen such that the minimum of the new function is also a minimum of J. However, for nonoptimal $\lambda(t)$ one can devise a scheme that always gives a new control corresponding to a lower cost function. Thus, Krotov's method leads to a monotonically convergent optimization algorithm that is expected to exhibit much faster convergence.

Our implementation closely follows Refs. [17,20,24]. Specifically, the cost reads

$$J(\psi(T),\lambda) = J_T(\psi(T)) + \int_0^T \frac{[\lambda(t) - \lambda_{\mathrm{ref}}(t)]^2}{S(t)} dt, \qquad (8)$$

where the reference field $\lambda_{\mathrm{ref}}(t)$ is typically chosen to be the control from the previous iteration [37]. The second term in Eq. (8) penalizes changes in the control from one iteration to the next and ensures that, as an optimum is approached, the value of the functional is increasingly determined by only J_T. $S(t) = ks(t)$ is a shape function that controls the turning on and off of the control fields, k is a step-size parameter, and $s(t) \in [0,1]$ is bound between 0 and 1.

Let $\psi^{(i)}(t)$ and $\lambda^{(i)}(t)$ denote the wave function and control parameter, respectively, in the ith iteration of the optimal control loop. To get started, we first solve for an initial guess $\lambda^{(0)}(t)$ the Gross-Pitaevskii equation (1) and the adjoint equation (4) for the costate $p(t)$, which is backward propagated in time with the same terminal condition as in GRAPE in order to obtain $\psi^{(0)}(t)$ and $p^{(0)}(t)$. In the next step, we solve the Gross-Pitaevskii equation *simultaneously* with the equation for the new control field

$$\begin{aligned}\lambda^{(i+1)}(t) &= \lambda^{(i)}(t) + S(t)\mathrm{Re}\langle p^{(i)}(t)|\left[\frac{\partial V}{\partial\lambda}\right]_{\lambda^{(i+1)}(t)}|\psi^{(i+1)}(t)\rangle \\ &\quad + \mathrm{Re}\frac{\sigma(t)}{2i}\langle\Delta\psi(t)|\left[\frac{\partial V}{\partial\lambda}\right]_{\lambda^{(i+1)}(t)}|\psi^{(i+1)}(t)\rangle, \qquad (9)\end{aligned}$$

where $\psi^{(i+1)}(t)$ is obtained by propagating $\psi(t=0)$ forward in time using the updated pulse.[1] The fact that $\psi^{(i+1)}(t)$ appears on the right-hand side of the update equation implies that the update at a given time t depends on the updates at all earlier times, making Krotov's method *sequential*. This type of update makes it nonstraightforward to include a cost term on the derivative of the control as in Eq. (3), since the derivative at a given time t requires knowledge of past *and* future values of $\psi(t)$.

The last term in Eq. (9) with $\Delta\psi(t) = \psi^{(i+1)}(t) - \psi^{(i)}(t)$ is generally needed to ensure convergence in presence of the nonlinear mean-field term $\kappa|\psi(t)|^2$ of the Gross-Pitaevskii equation. Convergence is achieved through a proper choice of $\sigma(t)$ [17,24]. In this work we neglect this additional contribution for simplicity, as it is of only minor importance for the moderate κ values of our present concern.

The derivative $\partial V/\partial\lambda$ in Eq. (9) has to be computed for $\lambda^{(i+1)}(t)$, thus leading to an implicit equation for $\lambda^{(i+1)}(t)$. When k is chosen sufficiently small, such that the control parameter varies only moderately from one iteration to the next, one can obtain the new control fields approximately from

$$\lambda^{(i+1)}(t) \approx \lambda^{(i)}(t) + S(t)\mathrm{Re}\langle p^{(i)}(t)|\left[\frac{\partial V}{\partial\lambda}\right]_{\lambda^{(i)}(t)}|\psi^{(i+1)}(t)\rangle. \qquad (10)$$

Otherwise one can employ an iterative Newton scheme for the calculation of $\lambda^{(i+1)}(t)$, as briefly described in Appendix A. In all our simulations we found Eq. (10) to provide sufficiently accurate results. Once the new wave functions $\psi^{(i+1)}(t)$ and control parameters $\lambda^{(i+1)}(t)$ are computed, we get the adjoint variables $p^{(i+1)}(t)$ through the solution of Eq. (4) and continue with the Krotov optimization loop until the cost function J is small enough or a certain number of iterations is exceeded.

As a variant, we also use a combination of Krotov's method with the BFGS method (KBFGS) [20]. It includes an approximated Hessian via the Krotov gradient as an additional term in the update equation (9). However, for technical reasons and differently from the GRAPE BFGS algorithm, no line search is employed.

IV. RESULTS

In this paper, we consider two control problems. The first one is *condensate splitting*, where the condensate initially resides in one well which is subsequently split into a double well. In our simulations we employ the confinement potential of Lesanovsky *et al.* [38] where the control parameter $\lambda(t)$ is associated with a radio frequency magnetic field [22]. The objective is to bring at the terminal time T the condensate wave function to the ground state of the double-well potential.

---

[1] The costates of this work and of Ref. [24] are related through $p = i\chi$. With this definition the adjoint equation (4) and the terminal condition $p(T)$ are the same for GRAPE and Krotov. As consequence, the scalar products on the right-hand side of Eq. (9) involve the real part rather than the imaginary part.

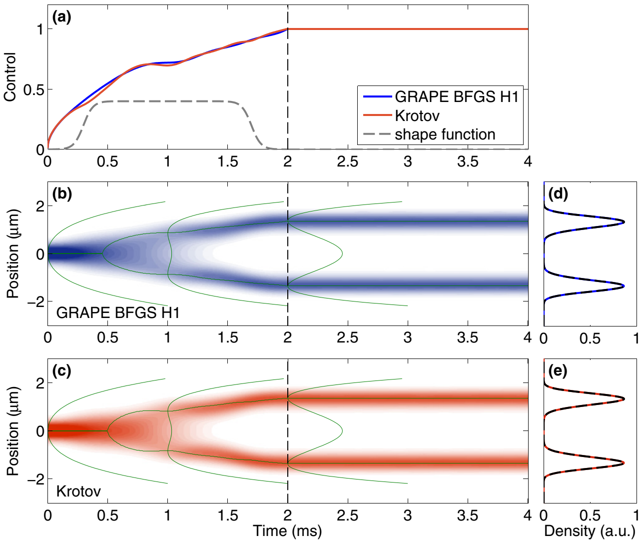

FIG. 1. (Color online) Wave-function splitting through the transformation of the confinement potential from a single to a double well. (a) The solid lines report the control parameters $\lambda(t)$ for the GRAPE and Krotov optimizations, respectively. The potential is held constant after the terminal time $T = 2$ ms. The dashed line shows the shape function $s(t)$ of Eq. (8) used in our version of Krotov's method, scaled by a factor of 0.4 for better visibility. (b), (c) Density plots of the condensate density $n(x,t) = |\psi(x,t)|^2$ during the splitting. The solid lines show the confinement potentials at three selected times and the time variation of the potential minima. (d), (e) Terminal (solid lines) and desired (dashed lines) densities, which are indistinguishable. In the optimization we set $\gamma = 10^{-6}$ and $k = 10^{-3}$.

In the second control problem the condensate wave function is excited from the ground to the first-excited state of a single-well potential. The confinement potential is an anharmonic single-well potential; details and a parametrized form of $V(x)$ can be found in Refs. [2,23,29]. The shakeup is achieved by displacing the potential origin according to $V(x - \lambda(t))$, where $\lambda(t)$ now corresponds to the position of the potential minimum, i.e., through *wave-function shaking*. Experimental realizations of such shaking protocols have been reported in Refs. [2,3,29].

In our simulations, GRAPE and Krotov start with the same initial guess. The terminal time is set to $T = 2$ ms throughout. Unless stated differently, we use a nonlinearity $\kappa/\hbar = 2\pi \times 250$ Hz ($\kappa = \pi/2$ for units with $\hbar = 1$ and time measured in milliseconds, as used in our simulations [23]).

A. Splitting vs shaking

Figure 1(a) shows the controls obtained from our GRAPE and Krotov optimizations for condensate splitting, together with [Figs. 1(b) and 1(c)] the density maps of the condensate wave function. The potential is held constant after the terminal time $T = 2$ ms of the control process. Figures 1(d) and 1(e) show the square moduli of the terminal (solid lines) and desired (dashed lines) condensate wave functions, which are almost indistinguishable, thus demonstrating the success of both control protocols. This can be also seen from the density maps which show no time variations at later times, when the potential is held constant, in accordance to the fact that the terminal wave function is the ground state of the double-well trap.

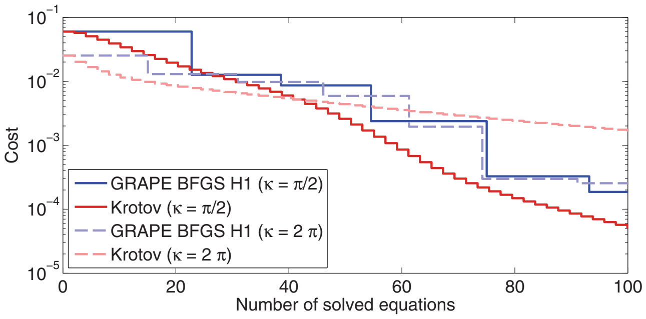

FIG. 2. (Color online) Cost function versus number of solved equations (either Gross-Pitaevskii or adjoint equation) for GRAPE and Krotov. For GRAPE, one optimization iteration consists of numerous solutions of the Gross-Pitaevskii equation (1) during a line search, which are followed by a solution of the adjoint equation (4) once a minimum is found, to obtain a new search direction $\nabla_\lambda J$. For Krotov one optimization iteration consists of a Gross-Pitaevskii solution, subject to Eq. (10), and a subsequent solution of the adjoint equation (4). In our simulations we use $k = 10^{-3}$. The dashed lines report results of simulations with a larger nonlinearity $\kappa/\hbar = 2\pi \times 1000$ Hz. In the legend we report the κ values in units used in our simulations, with $\hbar = 1$ and time measured in milliseconds [23].

Figure 2 compares the efficiency of the GRAPE and Krotov optimizations. We plot the cost function J_T versus the number n of equations solved during optimization. For both GRAPE and Krotov, n counts the solutions of either the Gross-Pitaevskii or the adjoint equation. The actual computer run times depend on the details of the numerical implementation but are comparable for both schemes. As can be seen in Fig. 2, in the GRAPE optimization the cost function decreases in large steps after a given number of solved equations, whereas in the Krotov optimization J_T decreases continuously. The cost evolution of GRAPE can be attributed to the BFGS search algorithm, where a line search is performed along a given search direction. Once the minimum is found, the step is accepted (J_T drops) and a new search direction is obtained through the solution of the adjoint equation. In contrast, the Krotov algorithm is constructed such that J_T decreases monotonically in each iteration step. Altogether, GRAPE and Krotov optimizations perform equally well.

In comparison to condensate splitting, the shakeup process is a considerably more complicated control problem. Figure 3 shows the optimized control parameters as well as the time evolution of the condensate densities. Both GRAPE and Krotov succeed comparably well. Regarding the control fields, the GRAPE one is smoother than the Krotov one, due to the penalty term on $\dot{\lambda}(t)$ in Eq. (3). From Fig. 4 we observe that a much higher number of optimization iterations is needed, in comparison to wave-function splitting, for both optimization methods to significantly reduce J_T. Initially, J_T decreases more rapidly for the Krotov optimization, but after a larger number n of solved equations, say around $n \sim 600$, GRAPE performs better.

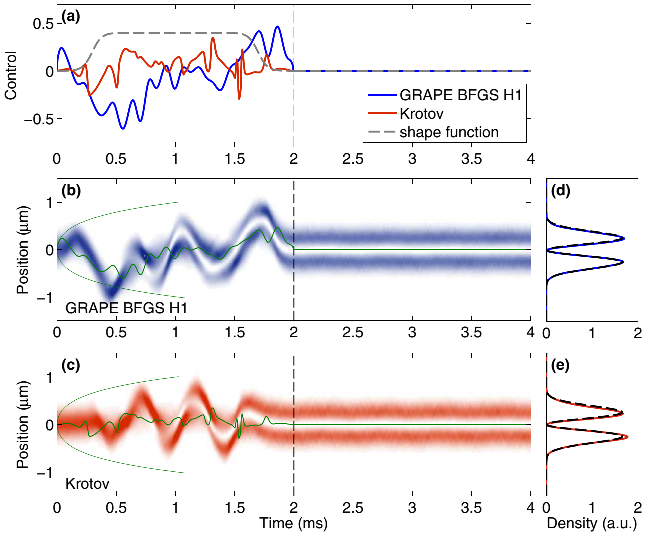

FIG. 3. (Color online) Same as Fig. 1 but for shaking process.

B. Influence of nonlinearity

We investigate the influence of the nonlinear atom-atom interaction on the convergence of the optimization loop. The dashed lines in Fig. 2 report results for splitting simulations with a larger nonlinearity $\kappa/\hbar = 2\pi \times 1000$ Hz. While the GRAPE convergence depends only weakly on κ, Krotov converges significantly slower for larger κ values.

Things are different for the shaking shown in Fig. 4. While the GRAPE performance again depends only weakly on κ, Krotov converges *faster* with increasing κ. Because of the lack of a line search in the Krotov algorithm, the convergence behavior is far more dependent on specific features of the control landscape which depend strongly on κ.

C. Convergence behavior

Next, we inquire into the details of the convergence properties for the optimization of the shakeup process. By comparing GRAPE with Krotov, we will identify the advantages and disadvantages of the respective optimization methods.

Figure 5(a) shows the terminal cost function J_T versus the number of solved equations of motion n for the different GRAPE schemes. It is evident that the conjugate gradient solutions reach a plateau after a certain number of iterations. In contrast, the BFGS solutions decrease significantly even at later stages of the optimization. We attribute this behavior to the use of the second-order-derivative information. The GRAPE BFGS scheme, which estimates the Hessian of J in addition to $\nabla_\lambda J$, can take larger steps to cross flat regions of J, contrary to the (first-order) GRAPE gradient scheme, which gets stuck.

Figure 5(b) shows the control fields for the GRAPE BFGS schemes. Although both optimization strategies perform equally well, the solutions obtained with the H^1 norm are smoother and probably better suited for experimental implementation.

Figures 6(a) and 6(b) present J_T versus n and the control parameters for the Krotov optimization, respectively. The solid line with $k = 0.005$ in Fig. 6(a) is identical to the one shown in Fig. 4. When we increase k (black line), the cost function drops more rapidly. However, we found that larger k values can lead to sharp variations in $\lambda(t)$ which might be problematic for experimental implementations, as is discussed in more detail below.

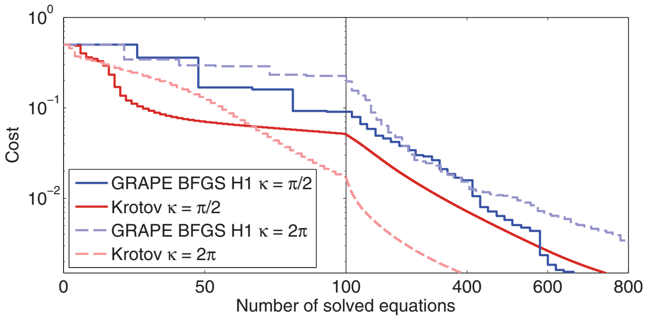

FIG. 4. (Color online) Same as Fig. 2 but for shaking process. We use $k = 5 \times 10^{-3}$.

In Fig. 6 we additionally display results for a simulation using a combination of Krotov's method with the BFGS scheme (KBFGS) [20]. The performance of KBFGS is similar to the simpler optimization procedure of Eq. (10), a finding in accordance with Ref. [20]. We attribute this to the fact that within the Krotov scheme only a small portion of the control landscape is explored, because the monotonic convergence enforces small control updates, in contrast to GRAPE where larger regions are scanned by the line search. As consequence, the improvement in the Krotov search direction via the Hessian is minimal.

Finally, the dashed line for adaptive k shows results for an optimization that starts with a small k value, which subsequently increases in each iteration until the cost decreases by a desired amount (here 2.5%) within one iteration. This k value is then kept constant for the rest of the optimization. The idea behind this strategy is that the choice of k is crucial for convergence, but the optimal value is different for each problem. Generally, finding a suitable value for k requires some trial and error.

D. Features of the control

For many experimental implementations it is indispensable to use smooth control parameters. In the following we investigate the smoothness of the optimal controls obtained by the different optimization methods.

Figure 7(a) shows for GRAPE BFGS H1 the evolution of the $\lambda(t)$ values during optimization. One observes that, during the first few iterations, the characteristic features of $\lambda(t)$ emerge, which then become refined in the course of further iterations. Figure 7(b) reports the power spectra (square moduli of Fourier transforms) of the $\lambda(t)$ history during optimization. During the first, say, 20 iterations the Fourier-transformed control parameter $\tilde{\lambda}(\nu)$ spectrally broadens, indicating the emergence of sharp features during optimization. With increasing iterations the spectral width of $\tilde{\lambda}(\nu)$ remains approximately constant.

Results of the GRAPE BFGS L2 optimization are shown in Figs. 7(c) and 7(d). We observe that, in contrast to the H1 results, $\lambda(t)$ acquires sharp features during optimization,

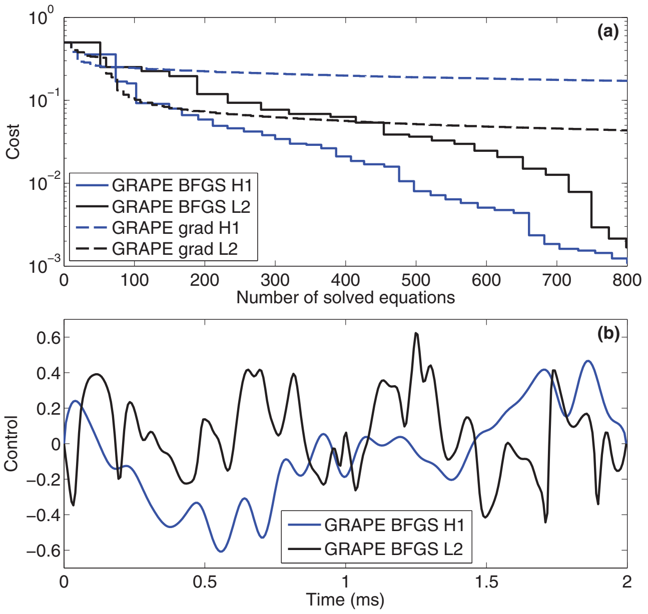

FIG. 5. (Color online) (a) Cost function versus number of solved equations for conjugate gradient (grad) and BFGS optimization schemes, and for search directions obtained from Eqs. (6) and (7) with H^1 or L^2 norm, respectively. (b) Optimal control parameters $\lambda(t)$ for BFGS solutions.

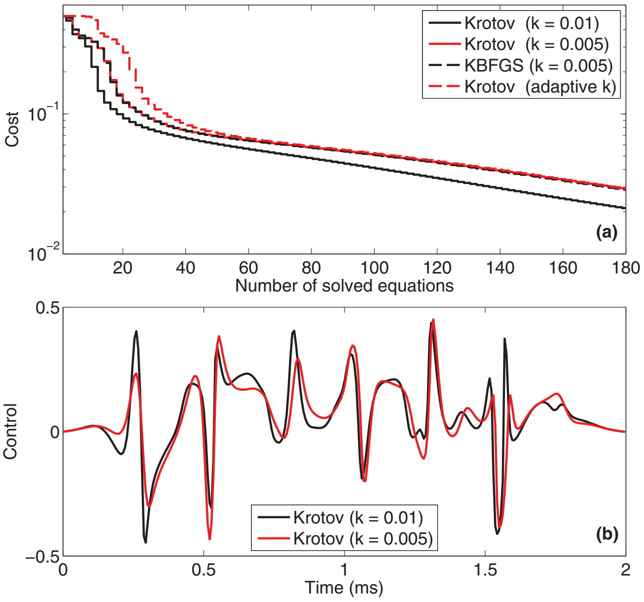

FIG. 6. (Color online) Same as Fig. 5 but for Krotov optimization. We investigate the influence of different mixing parameters k between the old and new control fields [see Eq. (10)], as well as a scheme with an adaptive k choice (see text for details). KBFGS reports the optimization result for a combination of Krotov's method with BFGS [20].

as also reflected by the broad power spectrum. This is because initially the gradient $\nabla_\lambda J$, which determines the search direction for improved control parameters, exhibits strong variations. These variations are washed out in the H1 optimization through the solution of the Poisson equation [see Eq. (7)], leading to significantly smoother control parameters.

In GRAPE, the user must additionally provide the weighting factor γ of Eq. (3) that determines the relative importance of terminal cost and control smoothness. For the problems under study, we found that the performance of GRAPE does not depend sensitively on the value of γ, and we usually use a small value such that the cost is dominated by the terminal cost.

Figures 8(a) and 8(c) show the $\lambda(t)$ history during a Krotov optimization for different step sizes k, and Figs. 8(b) and 8(d) report the corresponding power spectra. In comparison to the GRAPE BFGS H1 optimization, the power spectra are significantly broader, in particular for the larger k values. This is due to the fact that, in the functional used for the Krotov optimization, there is no penalty term that enforces smoothness of the control (and thus a narrow spectrum).

The choice of the step size k is rather critical for the Krotov performance. With increasing k the cost function decreases more rapidly during optimization. However, values of k that are too large can lead to numerical instabilities. These instabilities result from the discretization of the update equation; mathematically, Krotov is only guaranteed to converge monotonically for any value of k if the control problem is continuous.

One might wonder whether a combination of both approaches would give the best of two worlds. In Fig. 9 we present results for simulations where we start with a Krotov optimization and switch to GRAPE after a given number of iterations. As can be seen, the performance of this combined optimization does not offer a particular advantage over genuine GRAPE or Krotov optimizations. This is probably due to differences between the optimal control fields $\lambda(t)$ obtained by the two approaches, such that $\lambda(t)$ needs to be significantly modified when changing from one scheme to the other. In addition, the BFGS search algorithm of GRAPE uses the information of previous iterations in order to estimate the Hessian of the control space, and this information is missing when changing schemes.

V. CONCLUSIONS AND OUTLOOK

Based on the two examples investigated in the previous section; namely, wave-function splitting and shaking in a magnetic microtrap, we now set out to analyze the advantages and disadvantages of the GRAPE and Krotov optimization methods which are tied to the functional that is minimized in each case.

First, when the optimization converges fast to an optimal solution, such as for wave-function splitting investigated in Sec. IV A, both optimization algorithms perform equally well, even without carefully tuning the free parameters γ or k. For such problems, the choice of algorithm is a matter of personal preference. On the other hand, for optimization problems with slow convergence, such as wave-function shaking, more care has to be taken. Specifically, there are significant differences between the two algorithms in terms of free parameters vs speed of convergence as well as possible cost functionals vs features of the obtained optimal control.

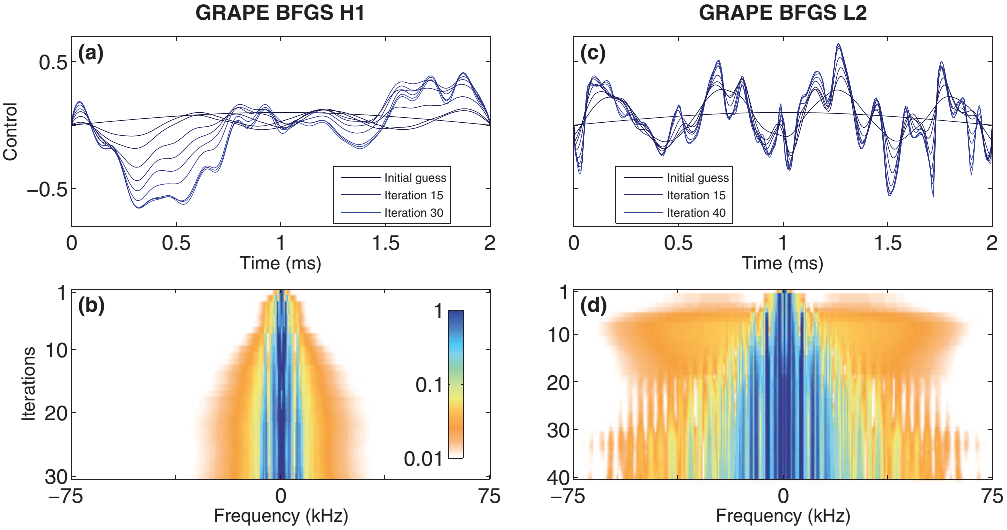

FIG. 7. (Color online) (a), (c) Evolution of control parameters during the optimization process for GRAPE. (b), (d) Density plot of power spectra of the control parameters displayed in panels (a) and (c). We use a logarithmic color scale. In panels (b) and (d) the numbers of iterations are chosen such that the final cost function J_T becomes approximately 10^{-2}, the numbers of solved equations (see Figs. 2 and 4) are approximately (b) 500 and (d) 700.

While GRAPE BFGS utilizes a line search to ensure monotonic convergence and to obtain the optimal step size in each iteration, the speed of convergence in Krotov's method is mainly determined by the free parameter k. On the one hand this means that GRAPE BFGS works better "out of the box" since it automatically determines the best step size in each step. On the other hand, the convergence is slowed down due to the necessity of a line search.

It is also evident from our results that both algorithms yield controls with features that can be understood in terms of additional costs introduced in the functional. For GRAPE we use a cost that penalizes a large derivative of the control which results in smooth controls in the end. For Krotov's method we employ a penalty on changes in the amplitude of the control in each iteration. Correspondingly, this leads to controls that have a smaller integrated intensity and come at the cost of a less smooth control.

In principle it is conceivable to modify the Krotov algorithm to take into account an additional cost term on the derivative of the control. While we conjecture that this will lead to controls that are comparable with those obtained in the GRAPE framework, the necessary modification of the Krotov algorithm is beyond the scope of the current work.

In the context of controlling Bose-Einstein condensates with experimentally smooth controls, the optimization with the GRAPE BFGS method, a functional enforcing smoothness, and use of the H^1 norm appears to be the method of choice. It is a black-box scheme with practically no problem-dependent

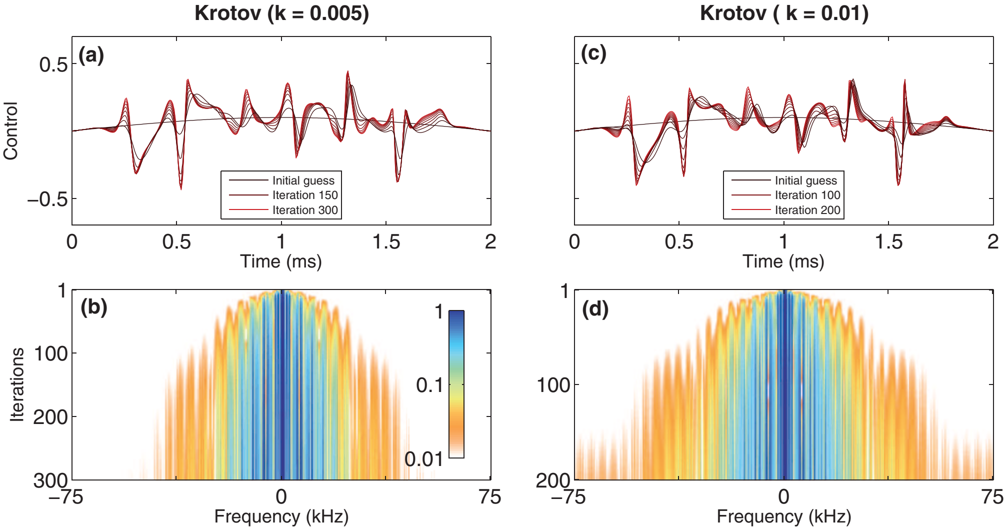

FIG. 8. (Color online) Same as Fig. 7 but for Krotov optimization. The numbers of solved equations are (b) 600 and (d) 400.

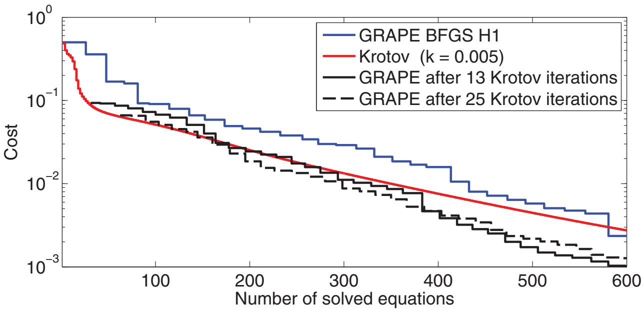

FIG. 9. (Color online) Same as Fig. 4, but for a combined GRAPE-Krotov scheme where one initially starts with the Krotov method and switches to GRAPE after a given number of iterations.

parameters, it gives the desired smooth control fields, and it works for various nonlinearity parameters κ.

In contrast, the Krotov optimization without an appropriate penalty term in the functional can converge faster but usually also leads to sharp features in the control. A sensitive choice of the step size k is indispensable to achieve a compromise between fast convergence and smoothness. If smoothness is not an issue or extremely fast convergence is needed, the Krotov method is preferable.

A combination of GRAPE and Krotov in the sense of switching from one method to the other during the optimization did not result in any significant gain. This is explained by the different control solutions that are found by the different methods which do not easily facilitate a transition between them. It points to the fact that many control solutions exist, and which solution is identified by the optimization depends strongly on the additional constraints [39] as well as the optimization method.

ACKNOWLEDGMENTS

This work has been supported in part by the Austrian science fund FWF under project P24248, by NAWI Graz, and by the European Union under Grant No. 297861 (QUAINT).

APPENDIX

In this appendix we show briefly how to numerically solve the equation

$$\lambda^{(i+1)}(t) = \lambda^{(i)}(t) + S(t)\mathrm{Re}\langle p^{(i)}(t)|\left[\frac{\partial V}{\partial \lambda}\bigg|_{\lambda^{(i+1)}(t)}\right]|\psi^{(i+1)}(t)\rangle, \tag{A1}$$

which differs from Eq. (10) in that the potential derivative is evaluated for $\lambda^{(i+1)}(t)$. Things can be easily generalized for the additional $\sigma(t)$ term of Eq. (9). Let $\lambda_0(t)$ denote an initial guess for the solution of Eq. (A1), e.g., the solution of Eq. (10). We now set $\lambda^{(i+1)}(t) = \lambda_0(t) + \delta\lambda(t)$, where $\delta\lambda(t)$ is assumed to be a small quantity. Thus, we can expand the second term on the right-hand side of Eq. (A1) in lowest order of $\delta\lambda(t)$ to obtain

$$\lambda_0(t) + \delta\lambda(t) \approx \lambda^{(i)}(t) + S(t)\mathrm{Re}\langle p^{(i)}(t)|\left[\frac{\partial V}{\partial \lambda}\bigg|_{\lambda_0(t)} + \frac{\partial^2 V}{\partial \lambda^2}\bigg|_{\lambda_0(t)}\delta\lambda(t)\right]|\psi^{(i+1)}(t)\rangle. \tag{A2}$$

Separating the contributions of $\delta\lambda$ from the rest, we get

$$\left(1 - S(t)\mathrm{Re}\langle p^{(i)}(t)|\left[\frac{\partial^2 V}{\partial \lambda^2}\bigg|_{\lambda_0(t)}\right]|\psi^{(i+1)}(t)\rangle\right)\delta\lambda(t) \approx -[\lambda_0(t) - \lambda^{(i)}(t)] + S(t)\langle p^{(i)}(t)|\left[\frac{\partial V}{\partial \lambda}\bigg|_{\lambda_0(t)}\right]|\psi^{(i+1)}(t)\rangle, \tag{A3}$$

which can be solved for $\delta\lambda(t)$. If $|\delta\lambda(t)| < \varepsilon$ is smaller than some small tolerance ε, we set $\lambda^{(i+1)}(t) \to \lambda_0(t) + \delta\lambda(t)$. Otherwise we set $\lambda_0(t) \to \lambda_0(t) + \delta\lambda(t)$ and repeat the Newton iteration until convergence. Typically only few iterations are needed to reach tolerances of the order of $\varepsilon = 10^{-6}$.

---

[1] C. Sayrin, I. Dotsenko, X. Zhou, B. Peaudecerf, T. Rybarczyk, S. Gleyzes, P. Rouchon, M. Mirrahimi, H. Amini, M. Brune *et al.*, Nature (London) **477**, 73 (2011).

[2] R. Bücker, J. Grond, S. Manz, T. Berrada, T. Betz, C. Koller, U. Hohenester, T. Schumm, A. Perrin, and J. Schmiedmayer, Nat. Phys. **7**, 508 (2011).

[3] S. van Frank, A. Negretti, T. Berrada, R. Bücker, S. Montangero, J. F. Schaff, T. Schumm, T. Calarco, and J. Schmiedmayer, Nat. Commun. **5**, 4009 (2014).

[4] M. Lapert, Y. Zhang, M. A. Janich, S. J. Glaser, and D. Sugny, Sci. Rep. **2**, 589 (2012).

[5] T. Häberle, D. Schmid-Lorch, K. Karrai, F. Reinhard, and J. Wrachtrup, Phys. Rev. Lett. **111**, 170801 (2013).

[6] L. Rybak, S. Amaran, L. Levin, M. Tomza, R. Moszynski, R. Kosloff, C. P. Koch, and Z. Amitay, Phys. Rev. Lett. **107**, 273001 (2011).

[7] R. González-Férez and C. P. Koch, Phys. Rev. A **86**, 063420 (2012).

[8] S. A. Rice and M. Zhao, *Optical Control of Molecular Dynamics* (Wiley and Sons, New York, 2000).

[9] P. Brumer and M. Shapiro, *Principles and Applications of the Quantum Control of Molecular Processes* (Wiley Interscience, Weinheim, 2003).

[10] D. Tannor, V. Kazakov, and V. Orlov, in *Time-Dependent Quantum Molecular Dynamics*, edited by J. Broeckhove and L. Lathouwers (Plenum, New York, 1992), pp. 347–360.

[11] J. Werschnik and E. K. U. Gross, J. Phys. B: At., Mol. Opt. Phys. **40**, R175 (2007).

[12] J. P. Palao and R. Kosloff, Phys. Rev. Lett. **89**, 188301 (2002).

[13] P. Doria, T. Calarco, and S. Montangero, Phys. Rev. Lett. **106**, 190501 (2011).

[14] A. Kaiser and V. May, J. Chem. Phys. **121**, 2528 (2004).

[15] T. Caneva, T. Calarco, and S. Montangero, Phys. Rev. A **84**, 022326 (2011).

[16] A. Konnov and V. Krotov, Autom. Remote Control (Engl. Transl.) **60**, 1427 (1999).

[17] S. E. Sklarz and D. J. Tannor, Phys. Rev. A **66**, 053619 (2002).

[18] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, J. Magn. Reson. **172**, 296 (2005).

[19] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, and T. Schulte-Herbrüggen, Phys. Rev. A **84**, 022305 (2011).

[20] R. Eitan, M. Mundt, and D. J. Tannor, Phys. Rev. A **83**, 053426 (2011).

[21] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro, Phys. Rev. Lett. **103**, 240501 (2009).

[22] U. Hohenester, P. K. Rekdal, A. Borzi, and J. Schmiedmayer, Phys. Rev. A **75**, 023602 (2007).

[23] U. Hohenester, Comput. Phys. Commun. **185**, 194 (2014).

[24] D. M. Reich, M. Ndong, and C. P. Koch, J. Chem. Phys. **136**, 104103 (2012).

[25] Y. Shin, M. Saba, T. A. Pasquini, W. Ketterle, D. E. Pritchard, and A. E. Leanhardt, Phys. Rev. Lett. **92**, 050405 (2004).

[26] T. Schumm, S. Hofferberth, L. M. Andersson, S. Wildermuth, S. Groth, I. Bar-Joseph, J. Schmiedmayer, and P. Krüger, Nat. Phys. **1**, 57 (2005).

[27] J. Grond, U. Hohenester, I. Mazets, and J. Schmiedmayer, New J. Phys. **12**, 065036 (2010).

[28] A. P. Peirce, M. A. Dahleh, and H. Rabitz, Phys. Rev. A **37**, 4950 (1988).

[29] R. Bücker, T. Berrada, S. van Frank, T. Schumm, J. F. Schaff, J. Schmiedmayer, G. Jäger, J. Grond, and U. Hohenester, J. Phys. B: At., Mol. Opt. Phys. **46**, 104012 (2013).

[30] A. Leggett, Rev. Mod. Phys. **73**, 307 (2001).

[31] J. Grond, G. von Winckel, J. Schmiedmayer, and U. Hohenester, Phys. Rev. A **80**, 053625 (2009).

[32] G. Jäger and U. Hohenester, Phys. Rev. A **88**, 035601 (2013).

[33] A. Borzi and U. Hohenester, SIAM J. Sci. Comput. **30**, 441 (2008).

[34] G. von Winckel and A. Borzi, Inverse Probl. **24**, 034007 (2008).

[35] C. M. Dion and E. Cances, Comput. Phys. Commun. **177**, 787 (2007).

[36] D. P. Bertsekas, *Nonlinear Programming* (Athena Scientific, Cambridge,1999).

[37] J. P. Palao and R. Kosloff, Phys. Rev. A **68**, 062308 (2003).

[38] I. Lesanovsky, T. Schumm, S. Hofferberth, L. M. Andersson, P. Krüger, and J. Schmiedmayer, Phys. Rev. A **73**, 033619 (2006).

[39] J. P. Palao, D. M. Reich, and C. P. Koch, Phys. Rev. A **88**, 053409 (2013).

7. Paper 4: Parametric squeezing amplification of Bose-Einstein condensates

Accepted to: Physical Review A (2015)

Parametric squeezing amplification of Bose-Einstein condensates

Georg Jäger,[1] Tarik Berrada,[2] Jörg Schmiedmayer,[2] Thorsten Schumm,[2] and Ulrich Hohenester[1]
[1] *Institut für Physik, Karl-Franzens-Universität Graz, Universitätsplatz 5, 8010 Graz, Austria*
[2] *Vienna Center for Quantum Science and Technology,*
Atominstitut, TU Wien, Stadionallee 2, 1020 Vienna, Austria

We theoretically investigate the creation of squeezed states of a Bose-Einstein Condensate (BEC) trapped in a magnetic double well potential. The number or phase squeezed states are created by modulating the tunnel coupling between the two wells periodically with twice the Josephson frequency, i.e., through parametric amplification. Simulations are performed with the multi configurational Hartree method for bosons (MCTDHB). We employ optimal control theory to bring the condensate to a complete halt at a final time, thus creating a highly squeezed state (squeezing factor of 0.12, $\xi_S^2 = -18$ dB) suitable for atom interferometry.

PACS numbers: 03.75.-b,39.20.+q,39.25.+k,02.60.Pn

I. INTRODUCTION

In atom chips, Bose-Einstein condensates (BECs) and ultracold atoms become trapped in the vicinity of a solid-state chip [1]. By changing the currents running through the wires mounted on the chip or modifying the strength of additional radio-frequency (rf) fields [2, 3], one can manipulate [3–5] and measure single quantum systems with extremely high precision. Possible applications range from atom interferometry [5–9], over quantum gates [10–12] and resonant condensate transport [13], to nonlinear atom optics [14–18].

In particular atom interferometry has attracted a lot of interest since atoms are massive objects sensitive to gravity. This opens new ways for measuring the gravitational constant [19], detection of gravitational waves, or the search for dark energy [20]. Using non-classical (squeezed) states brings the measurement sensitivity below the quantum noise limit [21]. Squeezed atom number states are typically created through condensate splitting and manipulation of the condensate around the point where the tunnel coupling strength becomes comparable with the nonlinear atom-atom interaction [22–24]. Possible routes towards squeezing are based on quasi-adiabatic splitting [25] or one-axis twisting [26].

It is often advantageous to seek for fast squeezing, for instance to achieve measurement series with high repetition rates or to suppress dephasing losses due to thermally excited atoms. In [27, 28] we demonstrated fast squeezing protocols that were obtained by using optimal control theory (OCT) [29, 30], a mathematical device allowing for optimization of certain control objectives. OCT protocols were succesfully implemented in atom chip experiments for twin-atom production [18] and interferometry [31].

In this paper we theoretically investigate the generation of squeezed states in a split BEC through parametric amplification. For a harmonic oscillator, parametric amplification can be achieved by modulating the spring constant with twice the resonance frequency, leading to an exponential increase of the oscillator's amplitude [32]. Similarly, modulating the tunnel coupling strength with twice the Josephson frequency leads to an exponential increase of number and phase fluctuations. To achieve fast squeezing, say on a time scale of 10 ms, one needs rather large tunnel coupling modulations which lead to additonal wavefunction oscillations of the split condensate, thus rendering the state useless for further interferometry once the wells become separated. We demonstrate that a final splitting stage of 2 ms, optimized with OCT, brings the condensate at a final time to halt and freezes the system in a highly squeezed state.

The motivation of this work lies in a direct experimental implementation. While the combined parametric amplification and splitting scheme investigated in this work leads to a slightly better squeezing compared to previous work [26–28], it is additionally simpler to implement and facilitates state tomography by releasing the condensate at different times and recording the time-of-flight images [18, 33].

We have organized our paper as follows. In Sec. II we discuss BEC interferometry and squeezing within a two-mode model and introduce a convenient Bloch sphere visualization for the many-body wavefunction. Squeezing through parametric amplification is discussed in Sec. III within the framework of the multi configurational Hartree method for bosons (MCTDHB) [34], which allows for the consideration of both wavefunction and atom number dynamics. We identify the pertinent parameters that lead to fast and efficient squeezing amplifications. In Sec. IV we employ the OCT framework to derive control ramps that freeze the condensate in a state with high number squeezing. Finally, Sec. V provides a short summary.

II. TWO-MODE MODEL

A. BECs in double wells

For the purpose of interferometry, we consider a 1d representation of a BEC in a double well trap, as depicted in

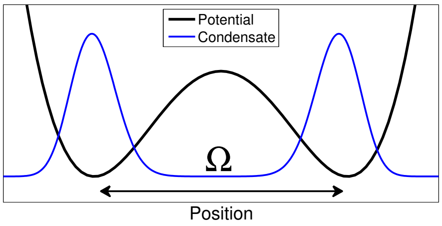

FIG. 1: (Color online) Schematic of a BEC wave function in a double well potential. Transforming a single well slowly into a double well produces a split BEC that can be used for interferometry. In the two-mode model atoms can reside in either the left or right well. Tunneling promotes atoms between the two wells, and the interwell distance allows control over the tunneling strength Ω.

Fig. 1. We assume that the trap is produced by the magnetic fields generated by an atom chip [1], which allow us to transform the potential from a single to a double well, thus creating a split BEC, and to change the distance between the two wells [35] in order to control the interwell coupling. Within the field of BEC interferometry, in the waiting phase the atoms in the two wells are decoupled and acquire different phases due to interactions with some external (classical) probe, such as gravity or magnetic fields. The phase shifts are finally read out through BEC interference.

The physics of double-well BECs is conveniently described in terms of a two-mode model, similar to Josephson junctions [36], where each atom can either reside in the left or the right well. With N atoms in the BEC, we can map the model to a spin $N/2$ system that captures many phenomena of double-well BECs. We introduce the pseudo spin operators [37]

$$J_x = \frac{1}{2}\left(a_l^\dagger a_r + a_r^\dagger a_l\right) \quad (1)$$

$$J_y = \frac{i}{2}\left(a_l^\dagger a_r - a_r^\dagger a_l\right) \quad (2)$$

$$J_z = \frac{1}{2}\left(a_l^\dagger a_l - a_r^\dagger a_r\right) \,, \quad (3)$$

with $a_{l/r}$ and $a_{l/r}^\dagger$ being the annihilation and creation operators for an atom in the left/right well, respectively. These operators have the following physical interpretations: J_x exchanges an atom between the left and the right well, and J_y and J_z measure the phase difference and atom number imbalance between the two wells, respectively. With these operators we can write down a model Hamiltonian in the form [38, 39]

$$H = -\Omega\, J_x + 2\kappa\, J_z^2 \,, \quad (4)$$

where Ω is the tunneling energy, accounting for the interwell tunneling, and κ is the charging energy describing the nonlinear interaction between atoms. For the interwell distances of our present concern, Ω can be assumed to be approximately proportional to the distance of the two wells (see Fig. 1), while κ has in general a more complicated behavior. Both quantities can be computed within the Gross-Pitaevskii framework [40].

States of a two-level quantum mechanical system (qubit) are conveniently depicted on the Bloch sphere [41, 42]. Such visualization is also possible for the two-mode model with a rather intuitive interpretation: A state where all the atoms are in the left or right well corresponds to a Bloch state on the north or south pole. We introduce $n = (n_l - n_r)/2$ for the atom number imbalance, with $n_{l/r}$ being the number of atoms in the left/right well. States where the atom number is exactly balanced, $n = 0$, are on the equator, and the angle φ describes the relative phase between the two wells (see Fig. 2). In addition to the mean values, also the atom number and phase uncertainties Δn and $\Delta\phi$ can be seen on the Bloch sphere: Δn corresponds to the height and $\Delta\phi$ to the width of the distribution, as shown in Fig. 2(a).

For any interferometry experiment, the important observable to be measured in the end is either the relative phase or number imbalance between the wells. Both measurements are subject to (shot) noise limiting the measurement precision, and thus render states with large Δn and $\Delta\phi$ fluctuations unfavorable. On the other hand, reduction of Δn and $\Delta\phi$ is possible but bound by the important relation [43]

$$\Delta n\, \Delta\phi \gtrsim \frac{1}{2} \,, \quad (5)$$

stating that we can in principle decrease one of the variances, however, at the cost of increasing the other one. For a binomial state we have $\Delta n = \sqrt{N}/2$ and $\Delta\phi = 1/\sqrt{N}$, leading to standard quantum shot noise [23]. In contrast, for states with smaller Δn or $\Delta\phi$ values, the so-called *squeezed states*, we can achieve a measurement precision below standard quantum shot noise [24].

In order to quantify how much a state is squeezed several factors have been used in the literature. The so-called number squeezing and phase squeezing factors $\xi_n = \Delta n/(\sqrt{N}/2)$ and $\xi_\phi = \Delta\phi/(1/\sqrt{N})$, respectively, provide information about how much a given state is squeezed in comparison to a binomial one. Both factors equal one for a binomial state. However, these factors completely neglect the coherence of the split condensate, a quantity of paramount importance for interferometry. Coherence is additionally considered by the factor [43, 44]

$$\alpha = \frac{2}{N}\sqrt{\langle J_x\rangle^2 + \langle J_y\rangle^2} = \langle \cos\phi\rangle \,, \quad (6)$$

where we have used the fact that $\langle \sin\phi\rangle = 0$ at equilibrium to arrive at the last expression on the right hand side. It is now convenient to introduce the so-called coherent spin squeezing factor [43]

$$\xi_S = \frac{\Delta n}{(\sqrt{N}/2)\alpha} = \frac{\xi_n}{\alpha} \,, \quad (7)$$

which is a direct measure of the useful number squeezing in the context of interferometry. In the following we

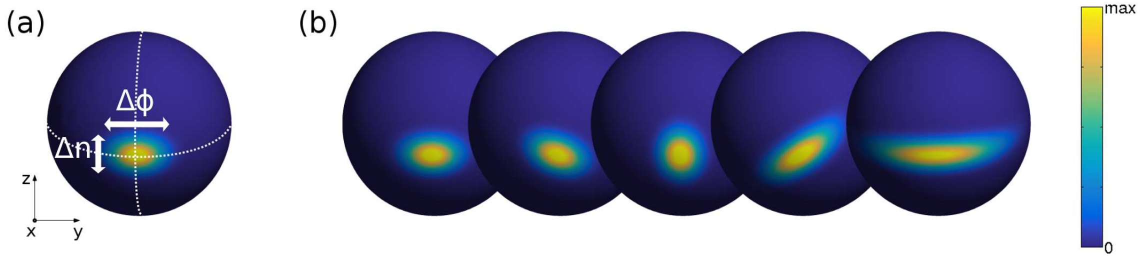

FIG. 2: (Color online) (a) State of a double-well BEC depicted on the Bloch sphere. For an equal number of atoms in the two wells, the distribution is centered around the equator, the height of the distribution corresponding to number fluctuations and the width to phase fluctuations. (b) Schematic view of parametric amplification on the Bloch sphere. During amplification the distribution rotates around the x-axis (driven by modulations of the tunneling strength Ω) and becomes more and more elongated under the influence of the atom-atom nonlinearity. Parametric amplification leads to an alternation between number and phase squeezed states, and the overall squeezing increases over time.

will refer to states with low squeezing factors as "highly squeezed states".

Squeezed states cannot only be used for measurements with precisions beyond the standard quantum limit [24], but also have other interesting properties. For instance, number squeezed states are very robust against dephasing effects [45]. It is therefore important to find ways of producing number or phase squeezed states, ideally on short time scales. A possible route towards number squeezing is to simply increase the distance between the two wells quasi-adiabatically [25]: this reduces tunneling and, in turn, Δn, since the number squeezing in the groundstate of Eq. (4) increases with decreasing Ω. Additionally $\Delta\phi$ increases. However, this process is relatively slow.

In this paper we will explore a different approach towards highly phase or number squeezed states on short time scales, which relies on parametric amplification through a periodic modulation of the tunnel coupling with twice the resonance frequency. In the following we briefly recall the mechanism underlying parametric amplification. We start with the Hamiltonian of Eq. (4) and rewrite it using the particle imbalance n and the relative phase ϕ [46],

$$H = -\Omega\,\cos\phi + 2\kappa\, n^2\,. \tag{8}$$

In the coupled regime the relative phase is assumed to be very small, so we can approximate $\cos\phi \approx 1 - \phi^2/2$. Apart from an irrelevant constant energy shift, this expansion leads to

$$H = \frac{\Omega}{2}\,\phi^2 + 2\kappa\, n^2\,. \tag{9}$$

From the commutation relations of the spin operators one observes that ϕ and n are canonically conjugate quantities, obeying $[\phi, n] = i$ [46]. The Hamiltonian of Eq. (9) thus has exactly the same form as the Hamiltonian of a harmonic oscillator, with phase and number playing the role of momentum and position, and the "mass" of the oscillator given by $1/\Omega$. Starting with a small amplitude at time zero, parametric amplification for the harmonic oscillator occurs for a time-dependent Ω which is modulated with twice the Josephson frequency $\omega_J = 2\sqrt{\kappa\Omega}$. This leads to an exponential increase of the oscillator's amplitude, in case of Eq. (9) the density n. We emphasize that parametric amplification is also possible for higher ϕ values without performing a Taylor expansion of $\cos\phi$, as discussed in more detail in Ref. [47].

Since for a split BEC we can modulate the tunneling parameter Ω by changing the distance between the wells, we can use parametric amplification in order to increase squeezing. We first start with a slightly number squeezed groundstate of a split but still tunnel-coupled BEC in a double well. In a next step, we modulate Ω with twice the Josephson frequency to get parametric squeezing amplification. In contrast to the above example of the harmonic oscillator, the Ω modulation leads to an amplification of the fluctuations rather than the mean values.

III. PARAMETRIC SQUEEZING AMPLIFICATION OF A BEC

In this section we show how to achieve parametric amplification for a BEC in a double well with realistic experimental parameters, in order to achieve high number or phase squeezing. To describe the BEC correctly, a simple two mode model with static orbitals is not sufficient, as will be discussed below, and one has to resort to a description scheme that accounts for both the atom number and wavefunction dynamics. In this work we employ the multi-configurational time-dependent Hartree method for bosons (MCTDHB) [34] using our recently developed Matlab toolbox OCTBEC [49].

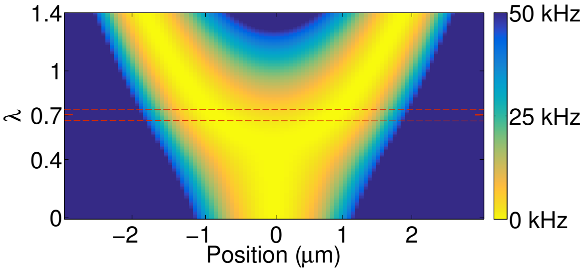

FIG. 3: (Color online) Lesanovsky-type potential [35], as used in our simulations, which allows to change from a single to a double well by modifying the control parameter λ associated with rf fields of an atom chip. The dashed lines indicate the 5% modulations used in our simulations. λ primarily controls the distance between the two wells, but additionally also modifies the barrier height.

A. Simulation details

We simulate parametric amplification of a BEC consisting of 1000 atoms in a double well potential within MCTDHB(2) [34], which expands the BEC wavefunction in two orbitals. The trap is a Lesanovsky type potential [35], giving rise to elongated, cigar shaped condensates. In this potential, the relevant splitting and amplification dynamics occurs in the radial direction, which allows us to introduce a 1d description scheme. The Lesanovsky potential has a single parameter λ, associated with the amplitude of a radio frequency field, that can turn the trap from a single into a double well, as depicted in Fig. 3.

B. Parametric amplification

In our simulations, we start with a BEC corresponding to the split ground state of the double well (we use $\lambda = 0.7$). It has been demonstrated experimentally that this state can be realized through adiabatic splitting of an elongated single atom trap, without generating significant heating or a noticable thermal fraction [9]. Finite temperature effects might lead to a slight broadening of $\Delta\phi$ but will not significantly influence the spin squeezing factor ξ_S [23]. In experiment, this might reduce the time that is avalaible for the parametric amplification.

In the split ground state the spin squeezing factor is initially already smaller than one ($\xi_S = 0.65$ for $\lambda = 0.7$). Starting at time zero, the distance between the double well is modulated with twice the Josephson frequency $\omega_J/(2\pi) = 220$ Hz, giving rise to a parametric amplification of squeezing. The squeezing value mainly depends on the amplitude of the modulation. Figs. 4 and 5 show the spin squeezing factor ξ_S and the atomic density during parametric amplification for λ-modulations with amplitudes of 1% and 5%, respectively. The density oscillates periodically for an amplitude of 1%, while strong excitation and non-periodic features can be observed for 5%.

Parametric amplification with an amplitude of 1% only produces a squeezing factor of $\xi_S \approx 0.4$ ($\xi_S^2 \approx -8$ dB), while the modulation with 5% leads to a much better squeezing of $\xi_S \approx 0.1$ ($\xi_S^2 \approx -20$ dB). However, squeezing becomes worse after roughly 10 ms. We attribute this degradation to dephasing effects: as depicted in Fig. 6, at later times the number distribution becomes curled around the x-axis of the Bloch sphere, indicating the partial occupation of states where all atoms reside in the left or right well, leading to complicated ensuing number dy-

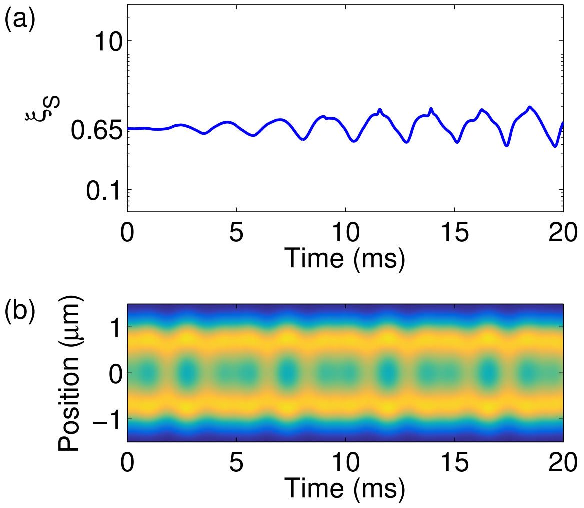

FIG. 4: (Color online) Parametric amplification with an amplitude of the control parameter λ of 1%: (a) Coherent spin squeezing factor ξ_S and (b) BEC density. Same color bar as in Fig. 2.

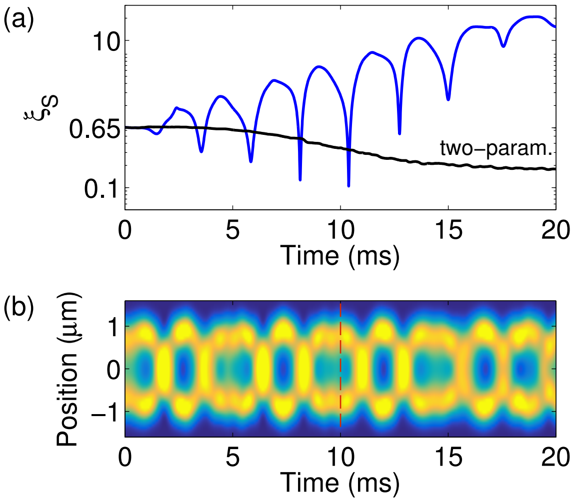

FIG. 5: (Color online) Same as Fig. 4 but for an amplitude of 5%. The black line reports for comparison results for a two parameter optimization [48].

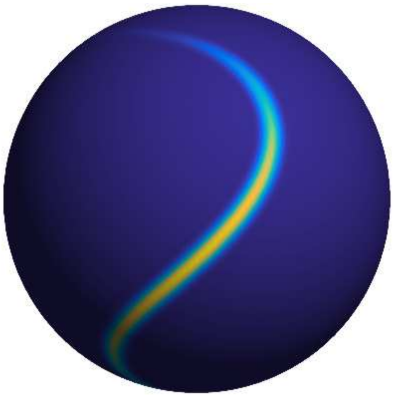

FIG. 6: (Color online) Snap shot of state that suffers dephasing, as obtained from the simulation with a 5% modulation at time $t = 12$ ms (see Fig. 5). Through parametric amplification wavefunction components with all atoms in the left or right well become populated, leading to a time evolution where the distribution “curls” around the Bloch sphere and squeezing is diminished. See Fig. 2 for color bar.

namics with a net effect reminiscent of dephasing.

A key requirement for interferometry on atom chips is a reliable and fast generation of squeezed states. To boost parametric amplification on short time scales, one has to use sufficiently high tunnel coupling modulations, which, in turn, lead to excitations of the condensate wave function. In this context, the consideration of both the atom number and wavefunction dynamics becomes mandatory in a simulation approach, thus calling for realistic many-body simulation approaches such as the MCTDHB framework of this work. Additionally, following the parametric amplification one has to modify the trap potential in such a way that the orbitals are brought to a halt. This step will be discussed in the next section. The main advantage of parametric squeezing amplification compared to other routes towards number squeezing [27, 28] is that the whole parametric amplification process can be implemented experimentally very easily, and only the final trapping stage requires some fine-tuning of the atom chip potentials. For comparison, the black line in Fig. 5 reports results for a two parameter optimization [48], whose final ξ_S value is also comparable to genuine OCT protocols for the optimization of number squeezing [27, 28]. Note that in comparison to these optimized protocols our simple parametric amplification scheme already leads to higher squeezing.

IV. CONDENSATE TRAPPING

A. Optimal control theory

To make parametric amplification useful in the context of squeezing generation, we should be able to trap the BEC after amplification in a highly number squeezed state and separate the two wells far enough to inhibit interwell tunneling. We will refer to the stage where the condensate is brought to a halt as “ trapping”, not to be confused with the atom trapping in order to produce a BEC on the atom chip. Trapping is shown in Fig. 7 (for details see discussion below) and is accomplished within the framework of OCT.

The general goal of OCT is to solve the following inverse problem: suppose that the state of a system Ψ_0 is known at some initial time t_0, and we are seeking for a desired state Ψ_d at some final time T. In order to bring the system from Ψ_0 to Ψ_d, we can tune some external control paramaters, such as the λ-parameter for the Lesanovsky potential. In general, the time dependence of the control parameter that brings the system from the initial to the desired state is unknown. OCT allows to find an optimal control in an iterative process. Many variants of OCT implementations exist, such as CRAB [50], Krotov’s method [51, 52], or a gradient ascent pulse engineering (GRAPE) scheme [27, 53]. In this work we employ the GRAPE algorithm implemented in the OCTBEC toolbox [49].

The OCT ingredients are the initial state of our system Ψ_0, a dynamic equation for the system’s time evolution (here MCTDHB), and a cost function that rates the success for a given control field $\lambda(t)$. For Ψ_0 we use the system’s state after an initial parametric amplification stage. As for the terminal cost, we are seeking for highly number squeezed states and for condensates at rest. This can be accomplished through a cost function of the form

$$\mathcal{J}_T = \langle\Psi|J_z^2|\Psi\rangle + \frac{\gamma}{N}\langle\Psi|H|\Psi\rangle\,, \tag{10}$$

which consists of two parts: the first one favors strongly number-squeezed states, the second one minimal energy and thus a condensate at rest. γ is a parameter that weights the relative importance for these two optimization goals.

A slight complication arises for the squeezing term in Eq. (10), as J_z is defined in the left-right basis whereas the natural MCTDHB basis is a gerade-ungerade basis [27]. To switch between the two bases, we use

$$\phi_l = \frac{1}{\sqrt{2}}(\phi_g + \tilde{f}\phi_u) \tag{11}$$

$$\phi_r = \frac{1}{\sqrt{2}}(\phi_g - \tilde{f}\phi_u)\,, \tag{12}$$

where $\tilde{f} = f/|f|$ is the relative phase between the orbitals, which is obtained from the wave function overlap for $x > 0$ (θ denotes the Heaviside step function)

$$f = \int \theta(x)\ \phi_g^*(x)\phi_u(x)\,dx\,. \tag{13}$$

The constraint that the BEC dynamics is governed by the MCTDHB equations of motion is included within a Lagrangian framework, and the full Lagrangian contains an additional cost term that favors control fields where

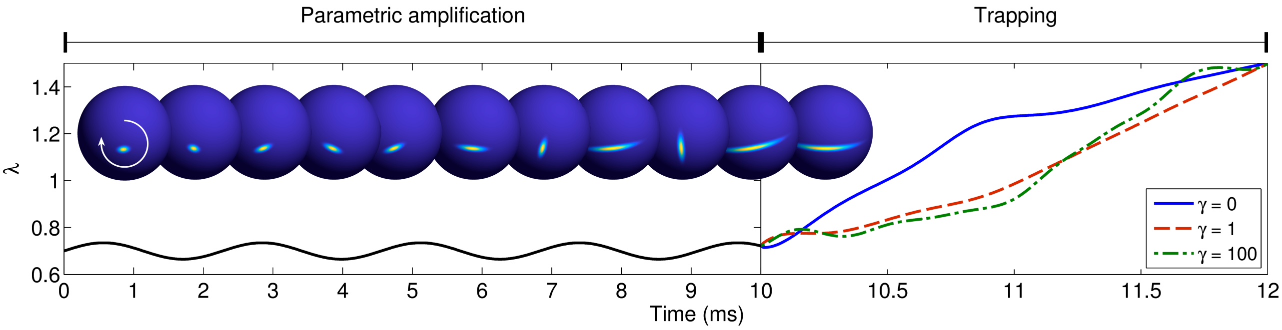

FIG. 7: (Color online) Parametric squeezing amplification (for times below 10 ms) and trapping (for times later than 10 ms, note different time axes). We show the time evolution of the control parameter λ for trapping ramps obtained for different weighting factors γ. The inset shows Bloch sphere representations of states which rotate clockwise around the x-axis and become continuously squeezed. See Fig. 2 for color bar.

the control parameter changes smoothly [49],

$$\mathcal{L} = \mathcal{J}_T + \frac{\nu}{2}\int_0^T [\dot\lambda]^2\,dt + \mathrm{Re}\langle\tilde\Psi, i\dot\Psi - F(\Psi,\lambda)\rangle\,. \tag{14}$$

Here F is a short-hand notation for the equations of motion, ν is a weighting parameter, and $\langle a,b\rangle$ is a short-hand notation for $\int_0^T dt \int dx\, a^*(x,t)b(x,t)$.

We next derive from this Lagrangian the optimality system that is needed for OCT. With exception of the cost function, the pertinent equations for the MCTDHB approach can be found in Ref. [49], and we thus only comment on the functional derivatives of the terminal cost function $\mathcal{J}_T$. Because of the relative phase $\tilde f$, see Eq. (13), appearing in the operator J_z, these derivatives are somewhat involved. After some calculations, which are briefly sketched in appendix A, we arrive at

$$\frac{\partial\mathcal{J}_T}{\partial C^*} = J_z^2|C\rangle + \frac{\gamma}{N}\,H|C\rangle \tag{15a}$$

$$\frac{\partial\mathcal{J}_T}{\partial\phi_g^*} = \langle C|J_z\frac{\partial J_z}{\partial\phi_g^*} + \frac{\partial J_z}{\partial\phi_g^*}J_z|C\rangle \tag{15b}$$

$$\frac{\partial\mathcal{J}_T}{\partial\phi_u^*} = \langle C|J_z\frac{\partial J_z}{\partial\phi_u^*} + \frac{\partial J_z}{\partial\phi_u^*}J_z|C\rangle\,, \tag{15c}$$

with

$$\frac{\partial J_z}{\partial\phi_g^*} = \frac{\theta(x)\phi_u}{4}\left(\frac{a_g^\dagger a_u}{|f|} - a_u^\dagger a_g\frac{(f^*)^2}{|f|^3}\right) \tag{16a}$$

$$\frac{\partial J_z}{\partial\phi_u^*} = \frac{\theta(x)\phi_g}{4}\left(\frac{a_u^\dagger a_g}{|f|} - a_g^\dagger a_u\frac{(f)^2}{|f|^3}\right)\,. \tag{16b}$$

For the optimizations we employ the Matlab toolbox OCTBEC [49]. See also Refs. [53, 54] for a detailed description of our OCT implementation.

B. Trapping

In our OCT simulations we first perform a parametric amplification with an amplitude of 5% and $t_0 = 10$ ms,[1] as shown in Figs. 5 and 7. The system's state at this terminal time is then used for Ψ_0 in our OCT algorithm. For the initial guess of the splitting and trapping ramp we use a linear ramp for λ and a time interval of $T - t_0 = 2$ ms. The initial guess was then optimized with the scheme described in Sec. IV A and for different weighting parameters γ.

Fig. 7 shows the resulting ramps for γ-factors of 0, 1, and 100. For the additional cost penalization term in Eq. (14) we use a small value of $\nu = 10^{-6}$ such that the control selection is only governed by $\mathcal{J}_T$ of Eq. (10). The difference between these ramps is attributed to the impact of the γ-factor that weights between the different optimization objectives of squeezing and trapping. Fig. 8 depicts the resulting spin squeezing factors ξ_S for the ramps shown in Fig. 7, and Fig. 9 the corresponding atom densities.

All three ramps produce squeezing values lower than 0.13, corresponding to $\xi_S^2 \approx -18$ dB. This is roughly 10 dB above the Heisenberg limit of -28 dB. As expected, the squeezing values are better for optimizations with smaller γ values, although the influence is not overly large. From the density maps shown in Fig. 9 we infer that the ramp with $\gamma = 0$ leads to an excited BEC, the ramp with $\gamma = 1$ produces an only weakly excited BEC, and the ramp with $\gamma = 100$ results in a BEC that is

---

[1] The success for optimizing squeezing and wavefunction trapping depends on the initial and terminal times t_0 and T, respectively. $t_0 = 10$ ms was obtained from a linesearch, where we used a linear λ ramp for trapping in order to find the "best" initial time in the interval $t_0 \in [9, 11]$ ms. Also the length of the trapping sequence (here 2 ms) was optimized through a similar linesearch.

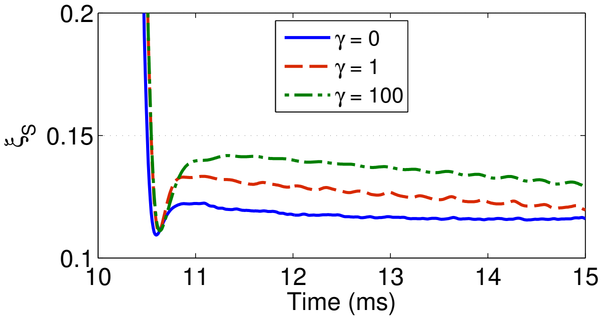

FIG. 8: (Color online) Coherent spin squeezing during and after trapping, for the ramps shown in Fig. 7. Smaller γ values lead to better squeezing.

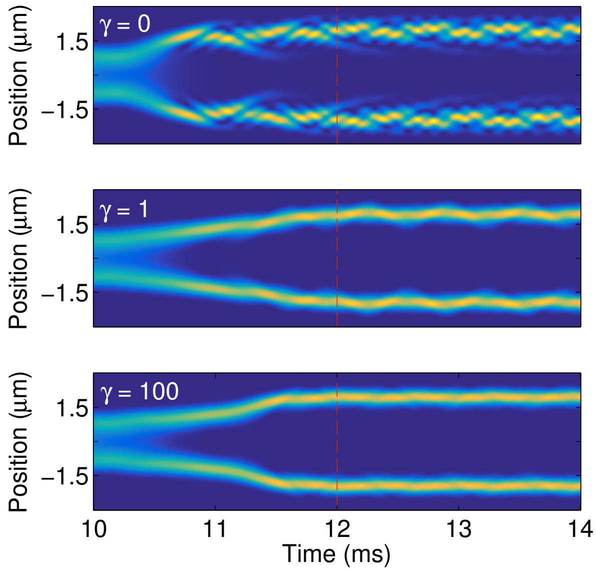

FIG. 9: (Color online) BEC density maps during and after trapping (dashed lines indicate end of trapping stage, see Fig. 2 for color bar) for the ramps shown in Fig. 7. Larger γ values lead to a less excited BEC.

almost at rest.

V. SUMMARY

We have discussed a parametric amplification scheme for creating and trapping a BEC in a highly squeezed state, with a squeezing value of $\xi_S^2 \approx -18$ dB. Squeezing amplification is achieved in a split BEC through modulation of the tunnel coupling with twice the Josephson frequency. To achieve high squeezing on short time scales, one has to use sufficiently large modulation amplitudes, which, in turn, lead to condensate oscillations. These oscillation can be brought to halt through a splitting ramp optimized within the OCT framework. Compared to other protocols for number squeezing [27, 28], the method presented here needs OCT only for the final trapping stage of the squeezed state.

Acknowledgements

This work has been supported in part by the Austrian science fund FWF under project P24248 and by NAWI Graz.

Appendix A

We describe the BEC dynamics within the framework of MCTDHB(2) [34]. In this method, the BEC wave function is expanded into a set of time dependent orbitals, which, for a spatially symmetric problem, can be classified according to their parity as gerade and ungerade, i.e., $\phi_g(x)$ and $\phi_u(x)$. In order to find the optimality system given in section IV A we have to calculate all the derivatives of the cost function of Eq. (10), namely $\frac{\partial \mathcal{J}_T}{\partial C^*}$, $\frac{\partial \mathcal{J}_T}{\partial \phi_g^*}$, and $\frac{\partial \mathcal{J}_T}{\partial \phi_u^*}$. The difficulty here is that the operator J_z depends explicitly on the orbitals,

$$J_z = \frac{1}{2}\left(\tilde{f} a_g^\dagger a_u + \tilde{f}^* a_u^\dagger a_g\right), \tag{A1}$$

namely through the factor $\tilde{f}$ that depends on ϕ_g and ϕ_u, see Eq. (13). Performing the functional derivative $\frac{\partial \mathcal{J}_T}{\partial C^*}$ is straightforward, and we arrive at

$$\frac{\partial}{\partial C^*}\left(\langle C|J_z^2|C\rangle + \frac{\gamma}{N}\,\langle C|H|C\rangle\right) = J_z^2|C\rangle + \frac{\gamma}{N}\,H|C\rangle\,. \tag{A2}$$

For $\frac{\partial \mathcal{J}_T}{\partial \phi_g^*}$ the second term of the cost function vanishes, since there is no dependence on the orbitals. We start by using the chain rule

$$\frac{\partial \mathcal{J}_T}{\partial \phi_g^*} = \frac{\partial \mathcal{J}_T}{\partial J_z}\frac{\partial J_z}{\partial \phi_g^*} = \langle C|J_z\frac{\partial J_z}{\partial \phi_g^*} + \frac{\partial J_z}{\partial \phi_g^*}J_z|C\rangle\,. \tag{A3}$$

To calculate $\frac{\partial J_z}{\partial \phi_g^*}$ we first use

$$\frac{\partial f}{\partial \phi_g^*} = \frac{1}{2|f|}\theta(x)\phi_u\,, \qquad \frac{\partial f^*}{\partial \phi_g^*} = -\frac{(f^*)^2}{2|f|^3}\theta(x)\phi_u\,,$$

and arrive at

$$\frac{\partial J_z}{\partial \phi_g} = \frac{\theta(x)\phi_u}{4}\left(\frac{a_g^\dagger a_u}{|f|} - a_u^\dagger a_g\frac{(f^*)^2}{|f|^3}\right)\,. \tag{A4}$$

The calculation of $\frac{\partial J_z}{\partial \phi_u^*}$ is very similar and we find

$$\frac{\partial J_z}{\partial \phi_u} = \frac{\theta(x)\phi_g}{4}\left(\frac{a_u^\dagger a_g}{|f|} - a_g^\dagger a_u\frac{(f)^2}{|f|^3}\right)\,. \tag{A5}$$

This leads us to our final result of Eq. (15).

---

[1] J. Reichel and V. Vuletic, eds., *Atom Chips* (Wiley, Weinheim, 2011).

[2] R. Folman, P. Krüger, J. Schmiedmayer, J. Denschlag, and C. Henkel, Adv. in Atom. Mol. and Opt. Phys. **48**, 263 (2002).

[3] S. Hofferberth, I. Lesanovsky, B. Fischer, J. Verdu, and J. Schmiedmayer, Nature Physics **2**, 710 (2006).

[4] P. Krüger, X. Luo, M. W. Klein, K. Brugger, A. Haase, S. Wildermuth, S. Groth, I. Bar-Joseph, R. Folman, and J. Schmiedmayer, Phys. Rev. Lett. **91**, 233201 (2003).

[5] T. Schumm, S. Hofferberth, L. M. Andersson, S. Wildermuth, S. Groth, I. Bar-Joseph, J. Schmiedmayer, and P. Krüger, Nature Phys. **1**, 57 (2005).

[6] W. Hänsel, J. Reichel, P. Hommelhoff, and T. W. Hänsch, Phys. Rev. A **64**, 063607 (2001).

[7] E. Andersson, T. Calarco, R. Folman, M. Andersson, B. Hessmo, and J. Schmiedmayer, Phys. Rev. Lett. **88**, 100401 (2002).

[8] Y.-J. Wang, D. Z. Anderson, V. M. Bright, E. A. Cornell, Q. Diot, T. Kishimoto, M. Prentiss, R. A. Saravanan, S. R. Segal, and S. Wu, Phys. Rev. Lett. **94**, 090405 (2005).

[9] T. Berrada, S. van Frank, R. Bücker, T. Schumm, J.-F. Schaff, and J. Schmiedmayer, Nat Commun **4** (2013).

[10] T. Calarco, E. A. Hinds, D. Jaksch, J. Schmiedmayer, J. I. Cirac, and P. Zoller, Phys. Rev. A **61**, 022304 (2000).

[11] E. Charron, M. Cirone, A. Negretti, J. Schmiedmayer, and T. Calarco, Phys. Rev. A **74**, 012308 (2006).

[12] P. Treutlein, T. W. Hänsch, J. Reichel, A. Negretti, M. A. Cirone, and T. Calarco, Phys. Rev. A **74**, 022312 (2006).

[13] T. Paul, K. Richter, and P. Schlagheck, Phys. Rev. Lett. **94**, 020404 (2005).

[14] L. Deng, E. W. Hagley, J. Wen, M. Trippenbach, Y. Band, P. S. Jiulienne, J. E. Simsarian, K. Helmerson, S. L. Rolston, and W. D. Phillips, Nature **398**, 218 (1999).

[15] C. Orzel, A. K. Tuchman, M. L. Fenselau, M. Yasuda, and M. A. Kasevich, Science **291**, 2386 (2001).

[16] G. K. Campbell, J. Mun, M. Boyd, E. W. Streed, W. Ketterle, and D. E. Pritchard, Phys. Rev. Lett. **96**, 020406 (2006).

[17] A. Perrin, H. Chang, V. Krachmalnikoff, M. Schellekens, D. Boiron, A. Aspect, and C. I. Westbrook, Phys. Rev. Lett. **99**, 150405 (2007).

[18] R. Bücker, J. Grond, S. Manz, T. Berrada, T. Betz, C. Koller, U. Hohenester, T. Schumm, A. Perrin, and J. Schmiedmayer, Nature Physics **7**, 608 (2011).

[19] G. Rosi, F. Sorrentino, L. Cacciapuoti, M. Prevedelli, and G. M. Tino, Nature **510** (2014).

[20] P. Hamilton, M. Jaffe, P. Haslinger, Q. Simmons, H. Müller, and J. Khoury, Science **349**, 849 (2015).

[21] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, Rev. Mod. Phys. **81**, 1051 (2009).

[22] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, Nature **464**, 1170 (2010).

[23] J. Esteve, C. Gross, A. Weller, S. Giovanazzi, and M. K. Oberthaler, Nature **455**, 1216 (2008).

[24] C. Gross, T. Zibold, E. Nicklas, J. Estève, and M. K. Oberthaler, Nature **464**, 1165 (2010).

[25] J. Javanainen and M. Y. Ivanov, Phys. Rev. A **60**, 2351 (1999).

[26] M. Kitagawa and M. Ueda, Phys. Rev. A **47**, 5138 (1993).

[27] J. Grond, J. Schmiedmayer, and U. Hohenester, Phys. Rev. A **79**, 021603 (2009).

[28] J. Grond, G. von Winckel, J. Schmiedmayer, and U. Hohenester, Phys. Rev. A **80**, 053625 (2009).

[29] A. P. Peirce, M. A. Dahleh, and H. Rabitz, Phys. Rev. A **37**, 4950 (1988).

[30] S. Glaser et al.,arXiv:1508.00442 (2015).

[31] S. van Frank, A. Negretti, T. Berrada, R. Bücker, S. Montangero, J.-F. Schaff, T. Schumm, T. Calarco, and J. Schmiedmayer, Nat Commun **5** (2014).

[32] L. D. Landau and E. M. Lifshitz, *Mechanics* (Pergamon Press, New York, 1991).

[33] R. Bücker, T. Berrada, S. van Frank, J.-F. Schaff, T. Schumm, J. Schmiedmayer, G. Jäger, J. Grond, and U. Hohenester, Journal of Physics B: Atomic, Molecular and Optical Physics **46**, 104012 (2013).

[34] O. E. Alon, A. I. Streltsov, and L. S. Cederbaum, Phys. Rev. A **77**, 033613 (2008).

[35] I. Lesanovsky, T. Schumm, S. Hofferberth, L. M. Andersson, P. Krüger, and J. Schmiedmayer, Phys. Rev. A **73**, 033619 (2006).

[36] A. Barone and G. Paternor, *Physics and Applications of the Josephson Effect* (John Wiley, 1982).

[37] G. J. Milburn, J. Corney, E. M. Wright, and D. F. Walls, Phys. Rev. A **55**, 4318 (1997).

[38] D. Ananikian and T. Bergeman, Phys. Rev. A **73**, 013604 (2006).

[39] G. J. Milburn, J. Corney, E. M. Wright, and D. F. Walls, Phys. Rev. A **55**, 4318 (1997).

[40] S. Raghavan, A. Smerzi, S. Fantoni, and S. R. Shenoy, Phys. Rev. A **59**, 620 (1999).

[41] F. Bloch, Phys. Rev. **70**, 460 (1946).

[42] F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, Phys. Rev. A **6**, 2211 (1972).

[43] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, Phys. Rev. A **50**, 67 (1994).

[44] L. Pitaevskii and S. Stringari, Phys. Rev. Lett. **87**, 180402 (2001).

[45] G.-B. Jo, Y. Shin, S. Will, T. A. Pasquini, M. Saba, W. Ketterle, D. E. Pritchard, M. Vengalattore, and M. Prentiss, Phys. Rev. Lett. **98**, 030407 (2007).

[46] C. Menotti, J. R. Anglin, J. I. Cirac, and P. Zoller, Phys. Rev. A **63**, 023601 (2001).

[47] A. Smerzi, S. Fantoni, S. Giovanazzi, and S. R. Shenoy, Phys. Rev. Lett. **79**, 4950 (1997).

[48] C. K. Law, H. T. Ng, and P. T. Leung, Phys. Rev. A **63**, 055601 (2001).

[49] U. Hohenester, Comp. Phys. Commun. **185**, 194 (2014).

[50] T. Caneva, T. Calarco, and S. Montangero, Phys. Rev. A **84**, 022326 (2011).

[51] D. M. Reich, M. Ndong, and C. P. Koch, The Journal of Chemical Physics **136** (2012).

[52] S. E. Sklarz and D. J. Tannor, Phys. Rev. A **66**, 053619 (2002).

[53] G. Jäger, D. M. Reich, M. H. Goerz, C. P. Koch, and U. Hohenester, Phys. Rev. A **90**, 033628 (2014).

[54] R. Bücker, T. Berrada, S. van Frank, J.-F. Schaff, T. Schumm, J. Schmiedmayer, G. Jäger, J. Grond, and U. Hohenester, J. Phys B **46**, 104012 (2013).

8. Conclusion and outlook

8.1. Comparison of different optimization methods

There are many different OCT techniques and algorithms, but not all of them are equally suitable to solve BEC problems. The suitability of each method also depends on the specific problem and how it is tackled. For most BEC problems the control describes a time dependent magnetic field. This field can be numerically treated through discretization in time or by describing it by a set of parameters. Discretization leads to a very high dimensional problem, while parametrizing usually poses too strong constraints on the possible control fields. For some algorithms it is necessary to parametrize, since they do not converge in a control space with thousands of dimensions. Especially genetic algorithms or CRAB rely heavily on such parametrization, which can be a problem for a typical BEC optimization. In general, the shape of the optimal control is completely unknown, so it is very difficult to find a working parametrization. Without any information about the final control, one has to expand in some base functions, like sines or polynomials, but for a finite number of parameters there is always the risk that the best controls are excluded from the control space. In that sense, CRAB and genetic algorithms are not very well suited to solve complex BEC optimization problems.

Nevertheless it is possible to use CRAB for BEC optimizations, if the shape of the control is roughly known beforehand. In [22] for example CRAB was used in order to find a control that produces a superposition of two trap eigenstates. This was possible, because the rough shape of the control was known and it was sufficient to optimize 60 Fourier components. For general problems, however, such a parametrization might not work, so GRAPE and Krotov's method have a big advantage there, since they can work with discretized controls and a high dimensional control space.

Especially GRAPE was very successful in the field of BEC optimization. It was possible to bring a BEC from the ground state of a single well trap into the ground state of a double well trap [64], as well as from its ground state to its first excited state (paper 1, [51]). A detailed comparison of GRAPE and Krotov's algorithm was performed in paper 3 [88] and showed that also Krotov's method is capable of such optimizations and performs equally well. Although both algorithms work, there are important differences between them that not only influence performance, but also the resulting controls.

One such difference is that the update in Krotov's method is sequential, while GRAPE's is concurrent. Using GRAPE, one first calculates the gradient for every point in time and then updates all points simultaneously. This has the big advantage that additional constraints on the control can be implemented very easily. Demanding a smooth control, which is favorable for most experimental implementations, can be done by simply adding a term to the cost functional that favors smooth controls. This is quite easy, since when calculating the gradient, the control is already known at all points in time, so properties like smoothness, which rely on that information, are accessible. Krotov's method on the other hand is sequential. Each point in time is updated individually. This has the advantage that the most recent information about the control is used for the update, as illustrated in Fig. 8.1. The downside of this approach is that the information about the shape of the control is incomplete most of the time, making it very difficult to include terms in the cost function that require information about the control at each point in time.

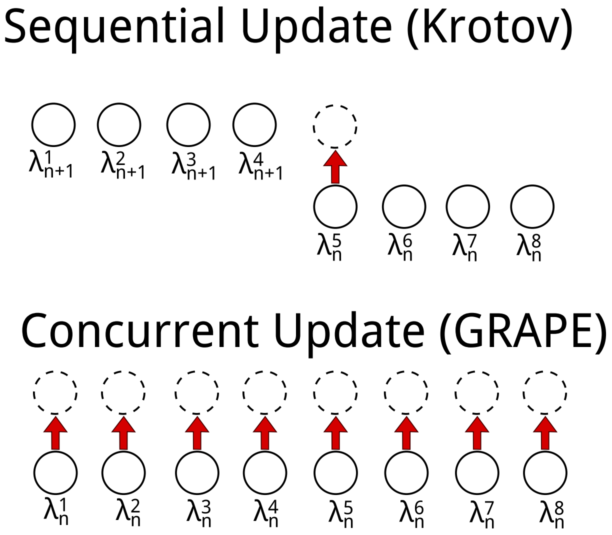

Figure 8.1.: The difference between a sequential and concurrent update: For sequential updates each point in time is updated separately, for concurrent updates the update is performed simultaneously for all points in time.

The fact that it is very difficult to include a smoothing term in Krotov's method leads to very different results for GRAPE and Krotov regarding the final control. Both controls have an equally low cost, the shape however is very different as can be seen in Fig. 8.2.

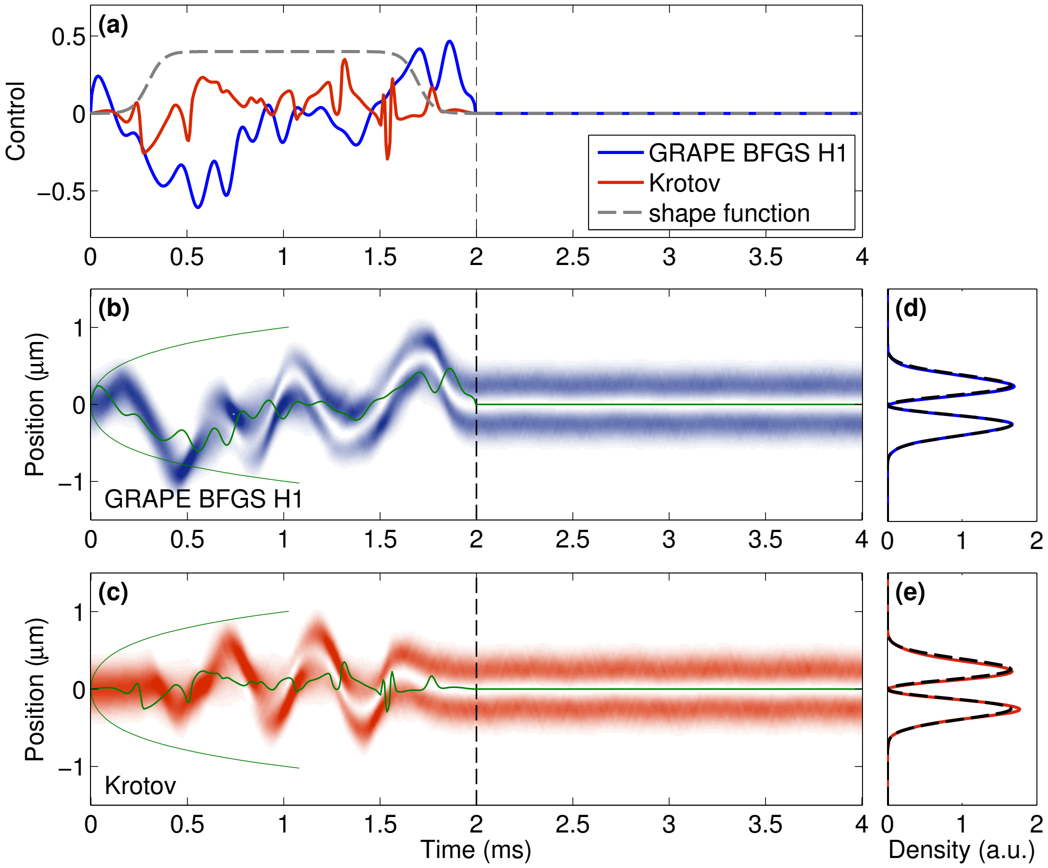

Figure 8.2.: Comparison of GRAPE and Krotov Optimization: (a) The control produced by GRAPE (blue) is much smoother than the one produced by Krotov's method (red). (b) and (c): Time development of the BEC density for both GRAPE (blue) and Krotov's method (red) (d) and (e): Both controls produce a final state that matches the target state. From paper 3 [88]

While the GRAPE control is relatively smooth, the Krotov control has very sharp features. In simulations these features do not pose a problem, but for an experimental realization they could be difficult to implement or might lead to unexpected behavior. If it is important that the resulting control is very smooth or one has other, more complicated constraints, GRAPE might be the method of choice.

Also concerning input parameters the two algorithms differ a lot. GRAPE works more or less like a black box. The only value that can be chosen is the cost tolerance at which the algorithm stops. However, this parameter does not need to be finetuned, it simply corresponds to the precision one is interested in. Krotov's method on the other hand has many parameters that can be chosen more or less arbitrarily. First and foremost the success of Krotov's method depends heavily on the choice of the Krotov parameters α, β and γ. In order to get fast convergence, all these parameters must be finetuned. There are ways of calculating upper bounds for them [83], but using these upper bounds often results in extremely slow convergence. Using smaller values can speed up convergence, but there is also the possibility that the algorithm does not converge at all, if one uses too small parameters. Choosing suitable values requires intuition and experience, and often it is convenient to simply determine them by trial and error, i.e. starting with very small parameters and slowly increasing them until the cost begins to decrease [82]. There is a similar problem with choosing the step size of Krotov's method. Too small step sizes lead to slow convergence, while too large step sizes inhibit convergence altogether. Unfortunately a line search in order to find the optimal step size is not possible for Krotov's method, since this would be computationally too expensive. The existence of many parameters that can be changed can be an advantage, but also a disadvantage. On the one hand finetuning all those parameters can be cumbersome, on the other hand it can potentially improve convergence significantly.

In terms of computational effort both algorithms are comparable, although the types of equations that need to be solved are very different. For Krotov's method it is necessary to solve the adjoint equation, followed by a simultaneous solution of the control equation and the Gross-Pitaevskii equation, which is computationally expensive. Additionally these equations need to be solved with very high precision, because the update is extremely sensitive to errors. The equations for GRAPE are entirely different. First one has to solve the Gross-Pitaevskii equation, followed by the adjoint equation. Most of the time, however, is spent on the following line search. In order to find an optimal step size the Gross-Pitaevskii equation needs to be solved 10 or maybe even 100 times in each iteration step. However, these calculations are usually fast.

In terms of the number of calculations that need to be solved GRAPE and Krotov's method perform equally well, as seen in Fig. 8.3. However, in terms of actual computation time the situation is different. Since the equations that need to be solved for GRAPE can be solved in a fraction of the time it takes to solve the Krotov equations, GRAPE usually outperforms Krotov's method.

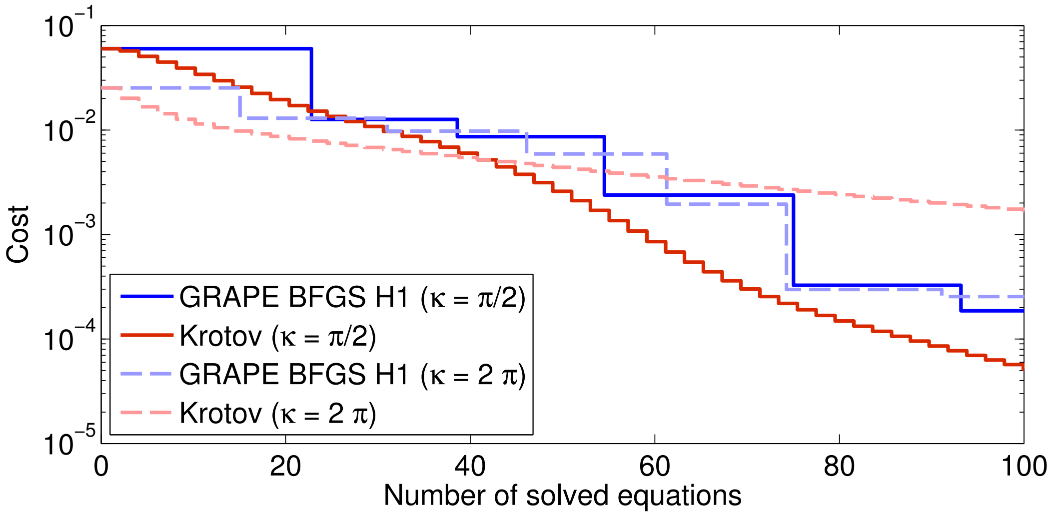

Figure 8.3.: Cost development of GRAPE and Krotov's method: Both algorithms produce controls with similar cost after the same amount of solved equations. The equations for GRAPE can however be solved much quicker than those for Krotov's method. From paper 3 [88]

The nonlinearity parameter κ plays an important role in these optimizations. For vanishing or very small κ, both algorithms converge very fast and there is little difference between them in terms of performance. High κ on the other hand slows down the convergence for GRAPE as well as Krotov's method. While GRAPE can in principle deal with large nonlinearities, Krotov's method needs to be adapted, in particular the Krotov term σ must be increased in order to ensure convergence.

To sum up, both GRAPE and Krotov's method are well suited to solve many different problems in the field of controlling BECs. However, GRAPE is very easy to use without any parameters that need to be finetuned, it has a simple way of favoring smooth controls, which is important for experiments, and it is computationally faster. This makes it the method of choice for most BEC optimizations.

8.2. Optimized generation of condensate wave functions

The main goal of most BEC optimizations is to bring the BEC from a target state Ψ_0 to a desired state Ψ_{des}. The actual shape of these states depends on the problem at hand.

Wave function translation

The simplest manipulation of a BEC one can perform is a simple translation. Here the BEC resides in a single well and the control translates the whole potential equally. The goal is to transfer the BEC a given distance without exciting it. Successful optimizations of such kind have been demonstrated in [64] and are depicted schematically in Fig. 8.4.

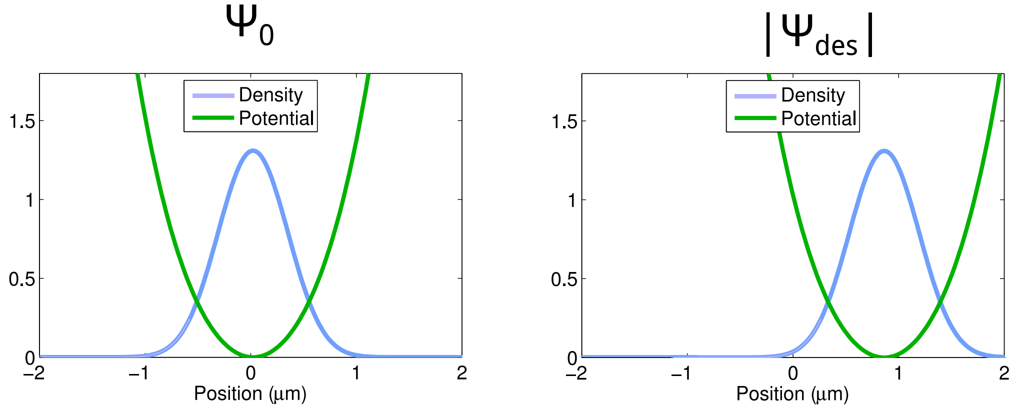

Figure 8.4.: Translating a BEC: Inital state (left) and target state (right)

This example demonstrates the principles of OCT, but is not overly interesting for experimental implementation.

Wave function splitting

A more relevant example of BEC optimization is wave function splitting, also demonstrated in [64]. Here the goal is to split the BEC into two coherent BECs that can be used for interferometry or other purposes. This process is more complicated and requires more than a simple translation of the potential. Instead a Lesanovsky type potential is used, as depicted in Fig. 8.5. This potential can be changed from a single well into a double well by a radio frequency field. The strength of this field is the control parameter λ that needs to be optimized.

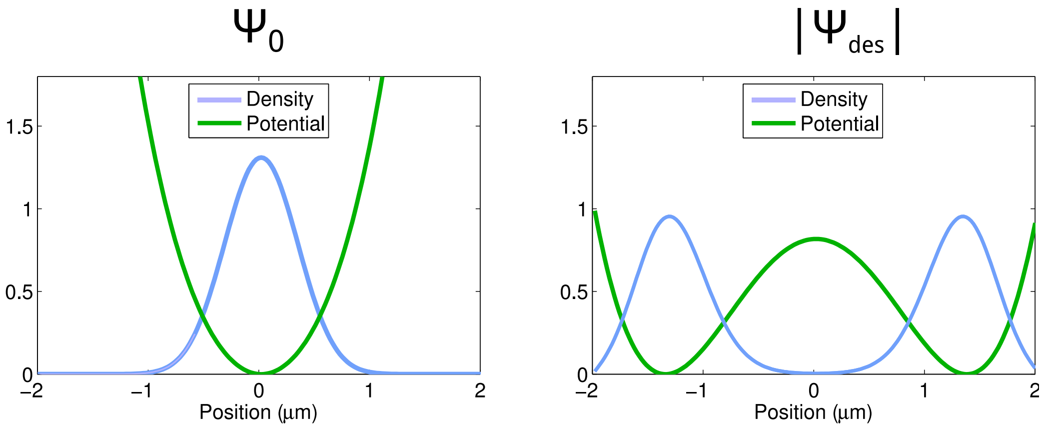

Figure 8.5.: Lesanovsky potential: Single well for $\lambda = 0$ (left), double well for $\lambda = 1$ (right). The optimization objective is to bring the BEC from the ground state of the unsplit trap to the ground state of the split trap.

Using GRAPE, this problem can be solved within a few iterations, typically < 10, corresponding to a computation time of a few minutes on a single core. While a linear, unoptimized control typically excites the BEC, especially for short transition times, the optimized control leads to a BEC that is almost exactly in its ground state. This state can then be used for interferometry or other applications.

Wave function shaking

A problem with even more interesting applications is called "wave function shaking". Again, the initial state is a BEC in the ground state of a single well trap. The target state is the first excited state of the same potential, depicted in Fig. 8.6.

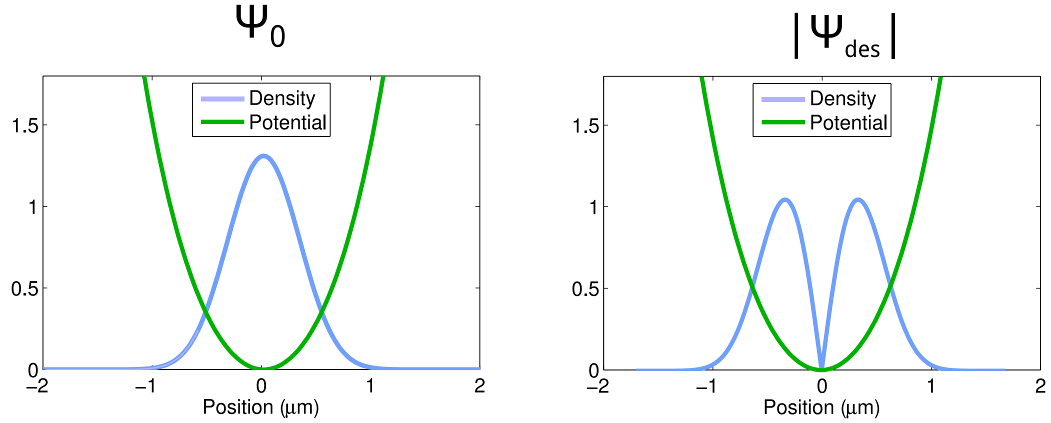

Figure 8.6.: Wave function shaking. The optimization objective is to bring the BEC from the ground state of the trap to the first excited state by “shaking”.

The density of the target state looks similar to the one for wave function splitting, but the underlying process is entirely different. For wave function shaking the target state is not the ground state of a split trap, but the first excited state of a single well trap. Such a state cannot be reached by controlling the radio-frequency field of a Lesanovsky type potential. Instead, the whole potential is moved in space, similar to wave function translation. “Shaking” the trap in that way brings additional energy into the system and can promote the BEC into the first excited state. The actual shape of the control is hereby of utmost importance and can not be guessed or calculated analytically in any way. For simpler problems, like wave function translation or wave function splitting, even a linear control can produce relatively low costs. For wave function shaking this is not the case, which makes this problem all the more interesting. The target is very specific: all the atoms should be in the first excited state, none of them are allowed in the ground state or in a higher excited state. With a guessed control this goal can simply not be reached in a satisfying way.

Such a state inversion from the ground to the first excited state was experimentally shown in paper 1, [51]. The optimization was performed with GRAPE (BFGS) and lead to a control that was used in an experiment in cooperation with the group of Jörg Schmiedmayer at the TU Wien. Theoretical simulation and experimental results match very well as can be seen in Fig. 8.7.

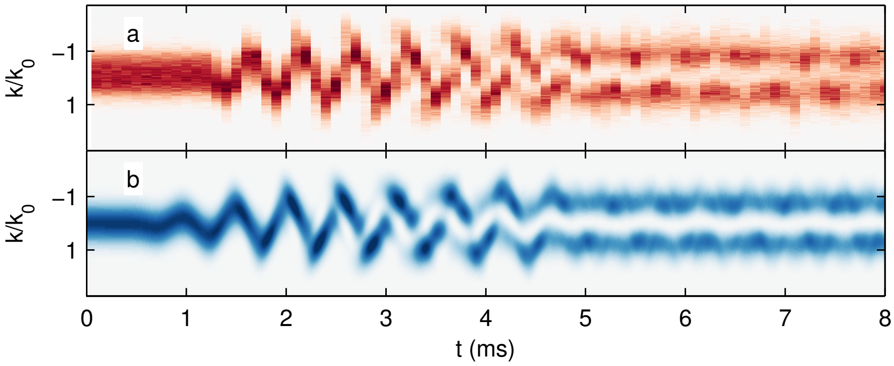

Figure 8.7.: An optimized control leads to state inversion of a BEC. BEC density for the experiment (a) and the simulation (b) as observed in TOF experiments. From paper 1 [51]

The main source of the small discrepancy between theory and experiment is the experimental setup that works as a filter for the control. The optimized control gets changed and distorted, leading to a suboptimal result. This problem has been tackled in paper 2, [89], by including the experimental filter in the OCT calculations. With this adaption an arbitrary filter function can be included in the optimization process, which helps to close the gap between theory and experiment.

Performing a state inversion makes many other interesting experiments possible. One prominent example is the generation of twin atom beams [90], the matter wave equivalent of twin photon beams. Similar to a laser, there is an excited active medium that produces a beam, here with atoms. This technique can be used for generating nonlocal correlations and entanglement as required for precision metrology and quantum communication and may path the way for future BEC applications.

8.3. Matter wave interferometry

One of the most promising applications where BEC control is needed is matter wave interferometry. In order to surpass the classical limit in terms of precision it is necessary to produce squeezed states, see section 2.3. In paper 4 [91] a technique called parametric squeezing amplification (see section 2.3.2) is used to generate such states on short timescales.

The main idea of this technique is to trap a BEC in a double well and then modulate the tunnel coupling at twice the Josephson frequency. We simulate this process within MCTDHB(2) for different amplitudes and find very high squeezing for an amplitude of 5%. However, it is also evident that this squeezing becomes worse after a certain time of amplification due to dephasing and other effects of the BEC dynamics. This raises the question if it is possible to trap the BEC in a highly squeezed state using OCT.

This optimization problem is fundamentally different from the ones presented above.

Previously we used the Gross-Pitaevskii equation to describe the BEC in a mean field approximation. Here we are interested in number squeezing, a property that can only be calculated within a many-body model. We use MCTDHB(2) (see section 1.2.4) in order to describe both orbital and number dynamics.

Although this makes the equations of motion and the optimality system more complicated, the steps for formulating the OCT problem remain the same. First we have to define the cost functional. We do not have a desired state for this problem, the goal is to find a state that is maximally squeezed. The exact shape of this state is unknown. Nevertheless, we can formulate a cost functional by finding an operator whose expectation value needs to be minimized. In this case we can use the square of the pseudospin operator J_z. Additionally we have a second requirement for the state at time T: Since parametric squeezing amplification can excite the BEC, which is unfavorable for interferometry, we include a term in the cost that favors states close to the ground state. The cost $\mathcal{J}_T$ then reads:

$$\mathcal{J}_T = \langle\Psi|J_z^2|\Psi\rangle + \frac{\gamma}{N}\langle\Psi|H|\Psi\rangle \tag{8.1}$$

γ is a weighting parameter that determines the relative importance of the two terms. For convenience it is normalized by the number of atoms N.

The next step is finding the initial state of the system Ψ_0. For this problem this is not the ground state, but a state that occurs after a certain time of parametric amplification. We optimize this amplification time by a line search algorithm and find a value of 10 ms. Ψ_0 is then calculated by simulating 10 ms of parametric amplification within MCTDHB(2) and it servers as a starting point for the OCT calculation.

In order to find an optimal control for this problem we use the MATLAB toolbox OCTBEC [92]. For different weighting parameters γ we find different controls, shown in Fig. 8.8.

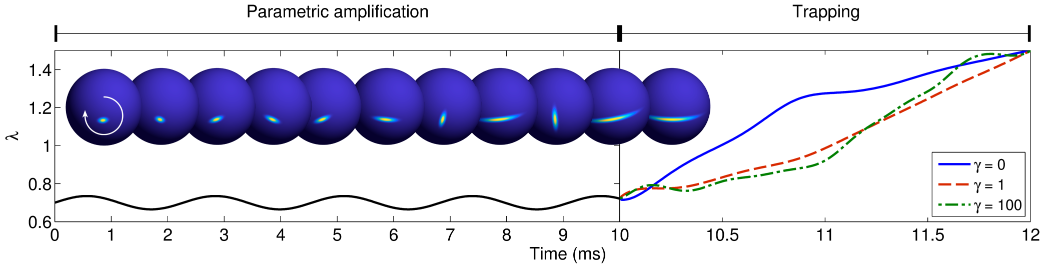

Figure 8.8.: Control fields that produce highly squeezed states: The parametric amplification stage (0-10 ms) is the same for all three controls, the trapping stage (10-12 ms) differs depending on the weighting parameter γ. The inset shows the parametric squeezing amplification on the Bloch sphere.
From paper 4, [91].

All controls produce very highly squeezed states with a squeezing factor of $\xi_S < 0.13$ ($\xi_s^2 = -18$ dB). Using such a state for an interferometer drastically improves its precision: While a binomial state leads to an error of $1/\sqrt{N}$, using a squeezed state reduces the error to $\xi_S/\sqrt{N}$. In other words the error is reduced to 13% of the classical value.

It has been shown experimentally that non-classical states can improve the precision of a matter wave interferometer [22], in principle down to the Heisenberg limit. The experimental realization of parametric squeezing amplification will be an important next step towards that goal.

8.4. Outlook

The field of BECs is very active with rapid advancements in recent years and one expects many interesting results and applications in the near future. Matter wave interferometry is especially promising. Using non-classical states to increase a matter wave interferometer's precision has been experimentally demonstrated in [22] and methods of producing very highly squeezed states have been theoretically investigated [44, 60, 61]. Especially parametric squeezing amplification (paper 4, [91]) is a promising candidate for improving the precision of a matter wave interferometer. If these advancements continue, it is to be expected that interferometers using squeezed states will outperform conventional ones in terms of precision for a lot of applications. Additionally, many (proposed) applications for matter wave interferometry are in areas where light wave interferometry does not work. Since BECs are sensitive to gravity, matter waves can for example be used to search for gravity waves [19]. They also present an entirely new method to measure the gravitational constant on a microscopic scale [21]. BEC experiments can also help in the search for dark energy [20], one of the most widely investigated problems in astrophysics.

Furthermore the manipulation of BECs and the creation of specific many body states has other popular applications, beyond interferometry. Quantum simulation is a rapidly growing field, where quantum systems, like BECs, are investigated as models for more complicated systems. That way one can extract information about problems that are far too complex to be tackled numerically. BECs are very versatile in this area, since a magnitude of different states can be produced, depending on the used atom chip and trap potential, so that many different systems can by analyzed. In the near future better control of BECs may even increase their versatility and make quantum simulations possible, that are currently beyond our reach.

Both quantum cryptography and quantum computing rely heavily on the production of highly entangled states. BECs are a perfect tool for creating hundreds of entangled atoms and are therefore a promising candidate for applications in both fields. The high controllability of BECs and the possibility to work with a large number of atoms simultaneously may lead to big advances in these areas.

BECs are fascinating objects and path the way to new, exciting experiments and technologies. Studying them and learning to control them gives us insight in fundamental physical principles, while it also enables us to utilize them in many different areas.

A. Calculation of the optimality system

The starting point of the calculation of the optimality system is the cost function (3.11). If we explicitly write down the Lagrangian, which ensures that the Gross-Pitaevskii equation is satisfied, it reads:

$$\mathcal{L} = \frac{1}{2}\left(1 - |\langle\Psi(T)|\Psi_{\text{des}}\rangle|^2 + \nu \int_0^T \dot{\lambda}(t)^2 dt\right) + \operatorname{Re}\left\langle p\,, i\dot{\Psi} - \left(-\frac{1}{2}\nabla^2 + V_\lambda + \kappa\,|\Psi|^2\right)\Psi\right\rangle \tag{A.1}$$

The goal is now to find the three derivatives $\partial\mathcal{L}/\partial\Psi^*, \partial\mathcal{L}/\partial\lambda$ and $\partial\mathcal{L}/\partial p^*$.

We will begin by calculating $\partial\mathcal{L}/\partial p^*$. The first step is to separate $\mathcal{L}$ into terms that depend on p^* and terms that do not depend on p^*. Now we can denote all the terms that are independent of p^* as “const”, since they will simply vanish during the derivative. Writing the real part of the second term of $\mathcal{L}$ as $\operatorname{Re}(A) = \frac{1}{2}(A + A^*)$ gives a factor $1/2$ and the complex conjugate of the term, which does not depend on p^* and can be included in the constant “const”:

$$\mathcal{L} = \frac{1}{2}\int dt\ p^*\left(i\dot{\Psi} - \left(-\frac{1}{2}\nabla^2 + V_\lambda + \kappa\,|\Psi|^2\right)\Psi\right) + \text{const}$$

The functional derivative $\partial\mathcal{L}/\partial p^*$ can now be calculated quite easily, as $\mathcal{L}$ is linear in p^*.

$$\frac{\partial\mathcal{L}}{\partial p^*} = \frac{1}{2}\left(i\dot{\Psi} - \left(-\frac{1}{2}\nabla^2 + V_\lambda + \kappa\,|\Psi|^2\right)\Psi\right)$$

To find the minimum of $\mathcal{L}$ we set $\partial\mathcal{L}/\partial p^* = 0$ and arrive at the first equation of the optimality system.

$$i\dot{\Psi} = \left(-\frac{1}{2}\nabla^2 + V_\lambda + \kappa\,|\Psi|^2\right)\Psi$$

Next we calculate $\partial\mathcal{L}/\partial\Psi^*$. Again, we separate into constant and Ψ^*-dependent terms. We find three terms that depend on Ψ^* and write them in a way that the actual dependency on Ψ^* is evident.

$$\mathcal{L} = \frac{1}{2}\int dt\left(p^*(-\kappa)\Psi^*\Psi\Psi + (i\dot{\Psi})^* p - \Psi^*\left(-\frac{1}{2}\nabla^2 + V_\lambda + \kappa\Psi^*\Psi\right)p\right) + \text{const}$$

We will now investigate each of these terms separately. The first one is linear in Ψ^*, and we find

$$\frac{\partial\mathcal{L}}{\partial\Psi^*}\int dt\ p^*(-\kappa)\Psi^*\Psi\Psi = p^*(-\kappa)\Psi\Psi$$

For the second term we perform a partial integration and again exploit that the derivatives of the constant terms vanish.

$$\frac{\partial\mathcal{L}}{\partial\Psi^*}\int dt\ (-i)\dot{\Psi}^*p = \frac{\partial\mathcal{L}}{\partial\Psi^*}\left[(-i)\Psi^*p\right]_0^T - \frac{\partial\mathcal{L}}{\partial\Psi^*}\int dt\ (-i)\dot{p}\Psi^* = i\dot{p}$$

Here we get an additional term $p(T)$ from the integration by parts that gives rise to the boundary condition for $p(T)$. The third term consists of a linear, and a quadratic part and the derivative yields:

$$\begin{aligned}
&\frac{\partial\mathcal{L}}{\partial\Psi^*}\int dt\ \ -\Psi^*\left(-\frac{1}{2}\nabla^2+V_\lambda+\kappa\Psi^*\Psi\right)p = \\
&= \frac{\partial\mathcal{L}}{\partial\Psi^*}\int dt\ \ -\Psi^*\left(-\frac{1}{2}\nabla^2+V_\lambda\right)p-\kappa\Psi^*\Psi^*\Psi p = \\
&= \left(-\frac{1}{2}\nabla^2+V_\lambda\right)p-2\kappa\Psi^*\Psi p
\end{aligned}$$

Now we add those three terms together and get $\partial\mathcal{L}/\partial\Psi^*$. To find the minimum of $\mathcal{L}$ we set the derivative to zero. This gives us the second equation of the optimality system, the so-called adjoint equation.

$$\dot{p} = \left(-\frac{1}{2}\nabla^2+V_\lambda+2\kappa|\Psi|^2\right)p+\kappa\Psi^2p^*$$

The last derivative we need to calculate is $\partial\mathcal{L}/\partial\lambda$. Rewriting the Lagrangian gives

$$\mathcal{L} = \int dt\ \frac{\nu}{2}\dot{\lambda}\dot{\lambda}+\mathrm{Re}\left\langle p, i\dot{\Psi}-\left(-\frac{1}{2}\nabla^2+V_\lambda+\kappa\ ,\Psi|^2\right)\Psi\right\rangle+\mathrm{const.}$$

We begin with the first term. Again we use partial integration and utilize that constant terms, like λ evaluated at a certain time, vanish.

$$\begin{aligned}
\frac{\nu}{2}\tfrac{\delta}{\delta\lambda}\int dt\ \dot{\lambda}\dot{\lambda} &= \frac{\nu}{2}\tfrac{\delta}{\delta\lambda}\left[\ddot{\lambda}\lambda\right]_0^T-\frac{\gamma}{2}\tfrac{\delta}{\delta\lambda}\int dt\ \lambda\ddot{\lambda} = \\
&= \frac{\nu}{2}\bigg(\int dt\ \tfrac{\delta}{\delta\lambda}\lambda\ddot{\lambda}-\int dt\ \lambda\tfrac{\delta}{\delta\lambda}\ddot{\lambda}\bigg) = \\
&= \frac{-\nu}{2}\bigg(\ddot{\lambda}+\int dt\ \lambda\tfrac{\delta}{\delta\lambda}\ddot{\lambda}\bigg) = \\
&= \frac{-\nu}{2}\bigg(\ddot{\lambda}+\left[\lambda\tfrac{\delta}{\delta\lambda}\dot{\lambda}\right]_0^T-\int dt\ \dot{\lambda}\tfrac{\delta}{\delta\lambda}\dot{\lambda}\bigg) = \\
&= \frac{-\nu}{2}\bigg(\ddot{\lambda}+\left[\lambda\tfrac{\delta}{\delta\lambda}\dot{\lambda}\right]_0^T-\left[\ddot{\lambda}\tfrac{\delta}{\delta\lambda}\dot{\lambda}\right]_0^T+\int dt\ \ddot{\lambda}\tfrac{\delta}{\delta\lambda}\lambda\bigg) = \\
&= \frac{-\nu}{2}\left(\ddot{\lambda}+\ddot{\lambda}\right) = -\nu\ddot{\lambda}
\end{aligned}$$

In the second term the only dependency on λ appears in V_λ.

$$\frac{\partial}{\partial\lambda}\mathrm{Re}\left\langle p, i\dot{\Psi}-\left(-\frac{1}{2}\nabla^2+V_\lambda+\kappa\left|\Psi\right|^2\right)\Psi\right\rangle = \mathrm{Re}\left\langle p\left|\frac{\delta V_\lambda}{\delta\lambda}\right|\Psi\right\rangle$$

We add those two terms and find $\partial\mathcal{L}/\partial\lambda$. Demanding that this derivative also equals zero gives the last equation of the optimality system, the control equation:

$$\gamma\ddot{\lambda} = -\operatorname{Re}\left\langle p \left| \frac{\delta V_\lambda}{\delta\lambda} \right| \Psi \right\rangle$$

We have now derived the three equations of the optimality system, namely the Gross-Pitaevskii equation, the adjoint equation and the control equation, from the three functional derivatives $\partial\mathcal{L}/\partial\Psi^*, \partial\mathcal{L}/\partial\lambda$ and $\partial\mathcal{L}/\partial p^*$. To complete the optimality system we also need to find the boundary conditions of those equations. The wave function Ψ at time 0 is Ψ_0. For the control λ we know both the value at the initial time (λ_0) and the value at the final time T ($\lambda(T)$). The last boundary condition comes from the derivative of the cost functional $\partial\mathcal{J}/\partial\Psi^*$.

The complete optimal control system, including boundary conditions, is:

$$i\dot{\Psi} = \left(-\frac{1}{2}\nabla^2 + V_\lambda + \kappa\left|\Psi\right|^2\right)\Psi \tag{A.2a}$$

$$\dot{p} = \left(-\frac{1}{2}\nabla^2 + V_\lambda + 2\kappa|\Psi|^2\right)p + \kappa\Psi^2 p^* \tag{A.2b}$$

$$\gamma\ddot{\lambda} = -\operatorname{Re}\left\langle p \left| \frac{\delta V_\lambda}{\delta\lambda} \right| \Psi \right\rangle \tag{A.2c}$$

$$\lambda(0) = \lambda_0 \tag{A.2d}$$

$$\lambda(T) = \lambda_T \tag{A.2e}$$

$$\Psi(0) = \Psi_0 \tag{A.2f}$$

$$ip(T) = -\left\langle \Psi_{\text{des}} \left| \Psi(T) \right.\right\rangle \Psi_{\text{des}} \tag{A.2g}$$

B. Calculation of the optimality system for a filtered control

The derivation of the optimality system for a filtered control is similar to the calculations done in appendix A. There are, however, some important differences.

We start again with the Lagrangian (3.11), but we have to replace the control λ with the filtered control $\lambda^\star$:

$$\lambda^\star(t) = \int_0^t h(\tau)\lambda(t-\tau)d\tau\,, \tag{B.1}$$

where h is the filter function which is determined by the experimental parameters. We find a new Lagrangian:

$$\mathcal{L} = \frac{1}{2}\left(1 - |\langle\Psi(T)|\Psi_{\mathrm{des}}\rangle|^2 + \nu\int_0^T \dot{\lambda}(t)^2 dt\right) + \mathrm{Re}\left\langle p\,, i\dot{\Psi} - \left(-\frac{1}{2}\nabla^2 + V_{\lambda^\star} + \kappa\,|\Psi|^2\right)\Psi\right\rangle \tag{B.2}$$

Like in appendix A we now have to calculate the functional derivatives $\partial\mathcal{L}/\partial\Psi^*$, $\partial\mathcal{L}/\partial\lambda$ and $\partial\mathcal{L}/\partial p^*$.

The derivation of $\partial\mathcal{L}/\partial p^*$ is very similar to the one in appendix A, we only have to replace V_λ with $V_{\lambda^\star}$. We arrive at the first equation of the optimality system.

$$i\dot{\Psi} = \left(-\frac{1}{2}\nabla^2 + V_{\lambda^\star} + \kappa\,|\Psi|^2\right)\Psi$$

Also $\partial\mathcal{L}/\partial\Psi^*$ does not change much due to the introduction of a filtered control. Again it is enough to simply replace V_λ with $V_{\lambda^\star}$, so that the second equation of the optimality system reads:

$$\dot{p} = \left(-\frac{1}{2}\nabla^2 + V_{\lambda^\star} + 2\kappa|\Psi|^2\right)p + \kappa\Psi^2 p^*$$

The last derivative, $\partial\mathcal{L}/\partial\lambda$, is however more complicated once the control is replaced by a filtered control. We start by writing the Lagrangian in a way that makes the actual dependency on λ apparent. Note that we still search for the derivative with respect to λ, although the potential depends on $\lambda^\star$.

$$\mathcal{L} = \int dt\ \frac{\nu}{2}\dot{\lambda}\dot{\lambda} + \mathrm{Re}\left\langle p\,, i\dot{\Psi} - \left(-\frac{1}{2}\nabla^2 + V_{\lambda^\star} + \kappa\,|\Psi|^2\right)\Psi\right\rangle$$

The first term is calculated in the same way we did earlier for the Lagrangian without the filter:

$$
\begin{aligned}
\frac{\nu}{2}\tfrac{\delta}{\delta\lambda}\int dt\ \dot{\lambda}\dot{\lambda} &= \frac{\nu}{2}\tfrac{\delta}{\delta\lambda}\left[\ddot{\lambda}\lambda\right]_0^T - \frac{\nu}{2}\tfrac{\delta}{\delta\lambda}\int dt\ \lambda\ddot{\lambda} = \\
&= \frac{\nu}{2}\Big(\int dt\ \tfrac{\delta}{\delta\lambda}\lambda\ddot{\lambda} - \int dt\ \lambda\tfrac{\delta}{\delta\lambda}\ddot{\lambda}\Big) = \\
&= \frac{-\nu}{2}\Big(\ddot{\lambda} + \int dt\ \lambda\tfrac{\delta}{\delta\lambda}\ddot{\lambda}\Big) = \\
&= \frac{-\nu}{2}\Big(\ddot{\lambda} + \left[\lambda\tfrac{\delta}{\delta\lambda}\dot{\lambda}\right]_0^T - \int dt\ \dot{\lambda}\tfrac{\delta}{\delta\lambda}\dot{\lambda}\Big) = \\
&= \frac{-\nu}{2}\Big(\ddot{\lambda} + \left[\lambda\tfrac{\delta}{\delta\lambda}\dot{\lambda}\right]_0^T - \left[\ddot{\lambda}\tfrac{\delta}{\delta\lambda}\lambda\right]_0^T + \int dt\ \ddot{\lambda}\tfrac{\delta}{\delta\lambda}\lambda\Big) = \\
&= \frac{-\nu}{2}\left(\ddot{\lambda} + \ddot{\lambda}\right) = -\nu\ddot{\lambda}
\end{aligned}
$$

Before we derive the last term, we first calculate $\frac{\delta}{\delta\lambda(t)}\lambda^\star(s)$, since we are going to need it in the next step.

$$
\begin{aligned}
\lambda^\star(s) &= \int_0^s h(\tau)\lambda(s-\tau)d\tau \\
\frac{\delta\lambda^\star(s)}{\delta\lambda(t)} &= \int_0^s h(\tau)\delta(s-\tau-t)d\tau
\end{aligned}
$$

With that result we can now calculate the derivative of the second term:

$$
\begin{aligned}
&\tfrac{\partial}{\partial\lambda(t)}\,\mathrm{Re}\left\langle p\ , i\dot{\Psi} - \Big(-\frac{1}{2}\nabla^2 + V_{\lambda^\star} + \kappa\left|\Psi\right|^2\Big)\Psi\right\rangle = \\
&= \mathrm{Re}\int_0^T \tfrac{\delta}{\delta\lambda(t)}\left\langle p(s)\left|V_{\lambda^\star(s)}\right|\Psi(s)\right\rangle ds = \\
&= \mathrm{Re}\int_0^T \tfrac{\delta\lambda^\star(s)}{\delta\lambda(t)}\tfrac{\delta}{\delta\lambda^\star(s)}\left\langle p(s)\left|V_{\lambda^\star(s)}\right|\Psi(s)\right\rangle ds = \\
&= \mathrm{Re}\int_0^T ds\int_0^s d\tau h(\tau)\delta(s-\tau-t)\left\langle p(s)\left|\frac{\delta V_{\lambda^\star(s)}}{\delta\lambda^\star(s)}\right|\Psi(s)\right\rangle = \\
&= \mathrm{Re}\int_t^T h(s-t)\left\langle p(s)\left|\frac{\delta V_{\lambda^\star(s)}}{\delta\lambda^\star(s)}\right|\Psi(s)\right\rangle ds
\end{aligned}
$$

This leads to the third equation of the optimality system for a filtered control:

$$
\gamma\ddot{\lambda} = -\,\mathrm{Re}\int_t^T h(s-t)\left\langle p(s)\left|\frac{\delta V_{\lambda^\star(s)}}{\delta\lambda^\star(s)}\right|\Psi(s)\right\rangle ds
$$

Additional to the boundary condition we calculated in appendix A, the filter gives us two new boundary conditions: The filtered control is $\lambda_0^\star$ at time 0 and $\lambda_T^\star$ at time T. This completes the optimality system for a filtered control:

$$i\dot{\Psi} = \left(-\frac{1}{2}\nabla^2 + V_{\lambda^\star} + \kappa \left|\Psi\right|^2\right)\Psi \tag{B.3a}$$

$$\dot{p} = \left(-\frac{1}{2}\nabla^2 + V_\lambda^\star + 2\kappa|\Psi|^2\right)p + \kappa\Psi^2 p^* \tag{B.3b}$$

$$\gamma\ddot{\lambda} = -\operatorname{Re}\int_t^T h(s-t)\left\langle p(s)\left|\frac{\delta V_{\lambda^\star(s)}}{\delta\lambda^\star(s)}\right|\Psi(s)\right\rangle ds \tag{B.3c}$$

$$\lambda(0) = \lambda_0 \tag{B.3d}$$

$$\lambda(T) = \lambda_T \tag{B.3e}$$

$$\lambda^\star(0) = \lambda_0^\star \tag{B.3f}$$

$$\lambda^\star(T) = \lambda_T^\star \tag{B.3g}$$

$$\Psi(0) = \Psi_0 \tag{B.3h}$$

$$ip(T) = -\left\langle \Psi_d \left| \Psi(T) \right.\right\rangle \Psi_d \tag{B.3i}$$

Bibliography

[1] S. Bose, “Plancks Gesetz und Lichtquantenhypothese,” *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 26, pp. 178–181, 1924.

[2] A. Einstein, “Quantentheorie des einatomigen idealen Gases,” *Sitzungber. Kgl. Akad. Wiss.*, no. 261, 1924.

[3] A. Einstein, “Quantentheorie des einatomigen idealen Gases: Zweite Abhandlung,” *Sitzungber. Kgl. Akad. Wiss.*, no. 3, 1925.

[4] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, “Observation of Bose-Einstein condensation in a dilute atomic vapor,” *Science*, vol. 269, no. 5221, pp. 198–201, 1995.

[5] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, “Bose-Einstein condensation in a gas of sodium atoms,” *Phys. Rev. Lett.*, vol. 75, pp. 3969–3973, Nov 1995.

[6] M. R. Andrews, C. G. Townsend, H.-J. Miesner, D. S. Durfee, D. M. Kurn, and W. Ketterle, “Observation of interference between two bose condensates,” *Science*, vol. 275, no. 5300, pp. 637–641, 1997.

[7] T. Schumm, P. Krüger, S. Hofferberth, I. Lesanovsky, S. Wildermuth, S. Groth, I. Bar-Joseph, L. Andersson, and J. Schmiedmayer, “A double well interferometer on an atom chip,” *Quantum Information Processing*, vol. 5, pp. 537–558, 2006.

[8] R. Folman, P. Krüger, D. Cassettari, B. Hessmo, T. Maier, and J. Schmiedmayer, “Controlling cold atoms using nanofabricated surfaces: Atom chips,” *Phys. Rev. Lett.*, vol. 84, pp. 4749–4752, May 2000.

[9] W. Hansel, P. Hommelhoff, T. W. Hansch, and J. Reichel, “Bose-Einstein condensation on a microelectronic chip,” *Nature*, vol. 413, pp. 498–501, Oct 2001.

[10] R. Folman, P. Krüger, J. Schmiedmayer, J. Denschlag, and C. Henkel, “Microscopic atom optics: From wires to an atom chip,” vol. 48 of *Advances In Atomic, Molecular, and Optical Physics*, pp. 263 – 356, Academic Press, 2002.

[11] T. Bravo, C. Sabín, and I. Fuentes, “Analog quantum simulation of gravitational waves in a Bose-Einstein condensate,” *EPJ Quant. Technol.*, vol. 2, p. 3, 2015.

[12] T. Schumm, S. Hofferberth, L. M. Andersson, S. Wildermuth, S. Groth, I. Bar-Joseph, J. Schmiedmayer, and P. Kruger, “Matter-wave interferometry in a double well on an atom chip,” *Nat Phys*, vol. 1, 2005.

[13] W. Hänsel, J. Reichel, P. Hommelhoff, and T. W. Hänsch, "Trapped-atom interferometer in a magnetic microtrap," *Phys. Rev. A*, vol. 64, p. 063607, Nov 2001.

[14] G.-B. Jo, Y. Shin, S. Will, T. A. Pasquini, M. Saba, W. Ketterle, D. E. Pritchard, M. Vengalattore, and M. Prentiss, "Long phase coherence time and number squeezing of two Bose-Einstein condensates on an atom chip," *Phys. Rev. Lett.*, vol. 98, p. 030407, 2007.

[15] T. Berrada, S. van Frank, R. Bücker, T. Schumm, J.-F. Schaff, and J. Schmiedmayer, "Integrated Mach-Zehnder interferometer for Bose-Einstein condensates," *Nat Commun*, vol. 4, 2013.

[16] J. Grond, U. Hohenester, J. Schmiedmayer, and A. Smerzi, "Mach-Zehnder interferometry with interacting trapped Bose-Einstein condensates," *Phys. Rev. A*, vol. 84, p. 023619, Aug 2011.

[17] Y.-J. Wang, D. Z. Anderson, V. M. Bright, E. A. Cornell, Q. Diot, T. Kishimoto, M. Prentiss, R. A. Saravanan, S. R. Segal, and S. Wu, "Atom michelson interferometer on a chip using a Bose-Einstein condensate," *Phys. Rev. Lett.*, vol. 94, p. 090405, Mar 2005.

[18] E. Andersson, T. Calarco, R. Folman, M. Andersson, B. Hessmo, and J. Schmiedmayer, "Multimode interferometer for guided matter waves," *Phys. Rev. Lett.*, vol. 88, p. 100401, Feb 2002.

[19] A. Einstein, "Über Gravitationswellen," *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften Berlin*, 1918.

[20] P. Hamilton, M. Jaffe, P. Haslinger, Q. Simmons, H. Müller, and J. Khoury, "Atom-interferometry constraints on dark energy," *Science*, vol. 349, no. 6250, pp. 849–851, 2015.

[21] G. Rosi, F. Sorrentino, L. Cacciapuoti, M. Prevedelli, and G. M. Tino, "Precision measurement of the Newtonian gravitational constant using cold atoms," *Nature*, vol. 510, 2014.

[22] S. van Frank, A. Negretti, T. Berrada, R. Bücker, S. Montangero, J.-F. Schaff, T. Schumm, T. Calarco, and J. Schmiedmayer, "Interferometry with non-classical motional states of a Bose-Einstein condensate," *Nat Commun*, vol. 5, 2014.

[23] A. Görlitz, J. M. Vogels, A. E. Leanhardt, C. Raman, T. L. Gustavson, J. R. Abo-Shaeer, A. P. Chikkatur, S. Gupta, S. Inouye, T. Rosenband, and W. Ketterle, "Realization of Bose-Einstein condensates in lower dimensions," *Phys. Rev. Lett.*, vol. 87, p. 130402, Sep 2001.

[24] D. S. Petrov, G. V. Shlyapnikov, and J. T. M. Walraven, "Regimes of quantum degeneracy in trapped 1d gases," *Phys. Rev. Lett.*, vol. 85, pp. 3745–3749, Oct 2000.

[25] E. H. Lieb, R. Seiringer, and J. Yngvason, “One-dimensional bosons in three-dimensional traps,” *Phys. Rev. Lett.*, vol. 91, p. 150401, Oct 2003.

[26] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, “Theory of Bose-Einstein condensation in trapped gases,” *Rev. Mod. Phys.*, vol. 71, pp. 463–512, Apr 1999.

[27] A. J. Leggett, “Bose-Einstein condensation in the alkali gases: Some fundamental concepts,” *Rev. Mod. Phys.*, vol. 73, pp. 307–356, Apr 2001.

[28] C. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases.* Cambridge University Press, 2002.

[29] N. N. Bogoliubov, “On the theory of superfluidity,” *J. Phys. (USSR)*, vol. 11, no. 23, 1947.

[30] H. M. J. M. Boesten, C. C. Tsai, J. R. Gardner, D. J. Heinzen, and B. J. Verhaar, “Observation of a shape resonance in the collision of two cold ^{87}Rb atoms,” *Phys. Rev. A*, vol. 55, pp. 636–640, Jan 1997.

[31] K. Sakmann, A. I. Streltsov, O. E. Alon, and L. S. Cederbaum, “Reduced density matrices and coherence of trapped interacting bosons,” *Phys. Rev. A*, vol. 78, p. 023615, Aug 2008.

[32] O. Penrose and L. Onsager, “Bose-Einstein condensation and liquid helium,” *Phys. Rev.*, vol. 104, pp. 576–584, Nov 1956.

[33] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, “Theory of Bose-Einstein condensation in trapped gases,” *Rev. Mod. Phys.*, vol. 71, pp. 463–512, Apr 1999.

[34] L. P. Pitaevsk, “Vortex lines in an imperfect bose gas,” *Soviet Physics JETP-USSR*, vol. 13, no. 2, 1961.

[35] E. P. Gross, “Structure of a quantized vortex in boson systems.,” *Nuovo Cimento*, vol. 20, pp. 454–477, Oct 1960.

[36] S. Raghavan, A. Smerzi, S. Fantoni, and S. R. Shenoy, “Coherent oscillations between two weakly coupled Bose-Einstein condensates: Josephson effects, π oscillations, and macroscopic quantum self-trapping,” *Phys. Rev. A*, vol. 59, pp. 620–633, Jan 1999.

[37] A. Barone and G. Paternor, “Physics and applications of the josephson effect,” *John Wiley and Sons Inc.*, 1982.

[38] G. J. Milburn, J. Corney, E. M. Wright, and D. F. Walls, “Quantum dynamics of an atomic Bose-Einstein condensate in a double-well potential,” *Phys. Rev. A*, vol. 55, pp. 4318–4324, Jun 1997.

[39] G. J. Milburn, J. Corney, E. M. Wright, and D. F. Walls, “Quantum dynamics of an atomic Bose-Einstein condensate in a double-well potential,” *Phys. Rev. A*, vol. 55, pp. 4318–4324, Jun 1997.

[40] D. Ananikian and T. Bergeman, "Gross-pitaevskii equation for bose particles in a double-well potential: Two-mode models and beyond," *Phys. Rev. A*, vol. 73, p. 013604, Jan 2006.

[41] G. J. Milburn, J. Corney, E. M. Wright, and D. F. Walls, "Quantum dynamics of an atomic Bose-Einstein condensate in a double-well potential," *Phys. Rev. A*, vol. 55, pp. 4318–4324, Jun 1997.

[42] F. Bloch, "Nuclear induction," *Phys. Rev.*, vol. 70, pp. 460–474, Oct 1946.

[43] F. T. Arecchi, E. Courtens, R. Gilmore, and H. Thomas, "Atomic coherent states in quantum optics," *Phys. Rev. A*, vol. 6, pp. 2211–2237, Dec 1972.

[44] J. Grond, J. Schmiedmayer, and U. Hohenester, "Optimizing number squeezing when splitting a mesoscopic condensate," *Phys. Rev. A*, vol. 79, p. 021603, Feb 2009.

[45] O. E. Alon, A. I. Streltsov, and L. S. Cederbaum, "Multiconfigurational time-dependent hartree method for bosons: Many-body dynamics of bosonic systems," *Phys. Rev. A*, vol. 77, p. 033613, Mar 2008.

[46] H.-D. Meyer, F. Gatti, and G. A. Worth, "Multidimensional quantum dynamics," 2009.

[47] A. I. Streltsov, O. E. Alon, and L. S. Cederbaum, "Role of excited states in the splitting of a trapped interacting Bose-Einstein condensate by a time-dependent barrier," *Phys. Rev. Lett.*, vol. 99, p. 030402, Jul 2007.

[48] H.-D. Meyer, U. Manthe, and L. Cederbaum, "The multi-configurational time-dependent hartree approach," *Chemical Physics Letters*, vol. 165, no. 1, 1990.

[49] Physikalisch-Technische Reichsanstalt (Germany), *Zeitschrift für Instrumentenkunde.* No. v. 11, Springer, 1891.

[50] Physikalisch-Technische Reichsanstalt (Germany), *Zeitschrift für Instrumentenkunde.* No. v. 12, Springer, 1892.

[51] R. Bücker, T. Berrada, S. van Frank, J.-F. Schaff, T. Schumm, J. Schmiedmayer, G. Jäger, J. Grond, and U. Hohenester, "Vibrational state inversion of a Bose-Einstein condensate: optimal control and state tomography," *J. Phys B*, vol. 46, p. 104012, 2013.

[52] S. Hofferberth, I. Lesanovsky, B. Fischer, J. Verdu, and J. Schmiedmayer, "Radiofrequency-dressed-state potentials for neutral atoms," *Nat Phys*, vol. 2, 2006.

[53] G.-B. Jo, Y. Shin, S. Will, T. A. Pasquini, M. Saba, W. Ketterle, D. E. Pritchard, M. Vengalattore, and M. Prentiss, "Long phase coherence time and number squeezing of two Bose-Einstein condensates on an atom chip," *Phys. Rev. Lett.*, vol. 98, p. 030407, Jan 2007.

[54] L. Pezzé, L. A. Collins, A. Smerzi, G. P. Berman, and A. R. Bishop, "Sub-shot-noise phase sensitivity with a Bose-Einstein condensate Mach-Zehnder interferometer," *Phys. Rev. A*, vol. 72, p. 043612, Oct 2005.

[55] J. Grond, U. Hohenester, J. Schmiedmayer, and A. Smerzi, "Mach-Zehnder interferometry with interacting trapped Bose-Einstein condensates," *Phys. Rev. A*, vol. 84, p. 023619, Aug 2011.

[56] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, "Squeezed atomic states and projection noise in spectroscopy," *Phys. Rev. A*, vol. 50, pp. 67–88, Jul 1994.

[57] J. Esteve, C. Gross, A. Weller, S. Giovanazzi, and M. K. Oberthaler, "Squeezing and entanglement in a Bose-Einstein condensate," *Nature*, vol. 455, pp. 1216–1219, Oct 2008.

[58] L. Pitaevskii and S. Stringari, "Thermal vs quantum decoherence in double well trapped Bose-Einstein condensates," *Phys. Rev. Lett.*, vol. 87, p. 180402, 2001.

[59] C. Gross, T. Zibold, E. Nicklas, J. Estève, and M. K. Oberthaler, "Nonlinear atom interferometer surpasses classical precision limit," *Nature*, vol. 464, pp. 1165–1169, Apr 2010.

[60] J. Grond, J. Schmiedmayer, and U. Hohenester, "Shaking the condensates: Optimal number squeezing in the dynamic splitting of a Bose-Einstein condensate," *Physica E: Low-dimensional Systems and Nanostructures*, vol. 42, no. 3, pp. 432 – 435, 2010.

[61] J. Grond, G. von Winckel, J. Schmiedmayer, and U. Hohenester, "Optimal control of number squeezing in trapped Bose-Einstein condensates," *Phys. Rev. A*, vol. 80, p. 053625, Nov 2009.

[62] G. R. Jin and C. K. Law, "Relationship between spin squeezing and single-particle coherence in two-component Bose-Einstein condensates with josephson coupling," *Phys. Rev. A*, vol. 78, p. 063620, Dec 2008.

[63] C. Menotti, J. R. Anglin, J. I. Cirac, and P. Zoller, "Dynamic splitting of a Bose-Einstein condensate," *Phys. Rev. A*, vol. 63, p. 023601, Jan 2001.

[64] U. Hohenester, P. K. Rekdal, A. Borzì, and J. Schmiedmayer, "Optimal quantum control of Bose-Einstein condensates in magnetic microtraps," *Phys. Rev. A*, vol. 75, p. 023602, Feb 2007.

[65] A. Borzì, G. Stadler, and U. Hohenester, "Optimal quantum control in nanostructures: Theory and application to a generic three-level system," *Phys. Rev. A*, vol. 66, p. 053811, Nov 2002.

[66] M. A. Daleh, A. P. Peirce, and H. Rabitz, "Optimal control of quantum-mechanical systems: Existence, Numerical approximations, and applications," *Phys. Rev. A*, vol. 37, pp. 4950–4964, 1988.

[67] U. Hohenester and G. Stadler, "Quantum control of electron-phonon scatterings in artificial atoms," *Phys. Rev. Lett.*, vol. 92, p. 196801, May 2004.

[68] T. Calarco, U. Dorner, P. S. Julienne, C. J. Williams, and P. Zoller, "Quantum computations with atoms in optical lattices: Marker qubits and molecular interactions," *Phys. Rev. A*, vol. 70, p. 012306, Jul 2004.

[69] C. P. Koch, J. P. Palao, R. Kosloff, and F. m. c. Masnou-Seeuws, "Stabilization of ultracold molecules using optimal control theory," *Phys. Rev. A*, vol. 70, p. 013402, Jul 2004.

[70] C. Brif, R. Chakrabarti, and H. Rabitz, "Control of quantum phenomena: past, present and future," *New Journal of Physics*, vol. 12, no. 7, p. 075008, 2010.

[71] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, "Whither the future of controlling quantum phenomena?," *Science*, vol. 288, no. 5467, pp. 824–828, 2000.

[72] C. M. Tesch and R. de Vivie-Riedle, "Quantum computation with vibrationally excited molecules," *Phys. Rev. Lett.*, vol. 89, p. 157901, 2002.

[73] K. W. Moore and H. Rabitz, "Exploring quantum control landscapes: Topology, features, and optimization scaling," *Phys. Rev. A*, vol. 84, p. 012109, Jul 2011.

[74] M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving Linear Systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952.

[75] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms 1. general considerations," *IMA Journal of Applied Mathematics*, vol. 6, no. 1, 1970.

[76] R. Fletcher, "A new approach to variable metric algorithms," *The Computer Journal*, vol. 13, no. 3, 1970.

[77] D. Goldfarb, "A family of variable-metric methods derived by variational means," *Math. Comp.*, vol. 24, 1970.

[78] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization," *Math. Comp.*, vol. 24, 1970.

[79] D. F. Shanno and P. C. Kettler, "Optimal conditioning of quasi-Newton methods," *Math. Comp.*, vol. 24, 1970.

[80] D. P. Bertsekas and D. P. Bertsekas, *Nonlinear Programming.* Athena Scientific, 2nd ed., 1999.

[81] V. F. K. A. I. Konnov, "On global methods for the successive improvement of control processes," *Autom. Remote Control (Engl. Transl.)*, vol. 60, 1999.

[82] S. E. Sklarz and D. J. Tannor, "Loading a Bose-Einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear Schrödinger equation," *Phys. Rev. A*, vol. 66, p. 053619, Nov 2002.

[83] D. M. Reich, M. Ndong, and C. P. Koch, "Monotonically convergent optimization in quantum control using krotov's method," *The Journal of Chemical Physics*, vol. 136, no. 10, 2012.

[84] T. Caneva, T. Calarco, and S. Montangero, "Chopped random-basis quantum optimization," *Phys. Rev. A*, vol. 84, p. 022326, Aug 2011.

[85] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. New York, NY, USA: Cambridge University Press, 1988.

[86] M. Mitchell, *An introduction to genetic algorithms*. Cambridge, Mass: MIT Press, 1996.

[87] G. von Winckel and A. Borzì, "Computational techniques for a quantum control problem with H 1 -cost," *Inverse Problems*, vol. 24, no. 3, p. 034007, 2008.

[88] G. Jäger, D. M. Reich, M. H. Goerz, C. P. Koch, and U. Hohenester, "Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Comparison of gradient-ascent-pulse-engineering and Krotov optimization schemes," *Phys. Rev. A*, vol. 90, p. 033628, Sep 2014.

[89] G. Jäger and U. Hohenester, "Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Consideration of filter effects," *Phys. Rev. A*, vol. 88, p. 035601, Sep 2013.

[90] R. Bücker, J. Grond, S. Manz, T. Berrada, T. Betz, C. Koller, U. Hohenester, T. Schumm, A. Perrin, and J. Schmiedmayer, "Twin-atom beams," *Nature Physics*, vol. 7, pp. 608–611, 2011.

[91] G. Jäger, T. Berrada, J. Schmiedmayer, T. Schumm, and U. Hohenester, "Parametric-squeezing amplification of Bose-Einstein condensates," *Phys. Rev. A (accepted Nov. 2015).*

[92] U. Hohenester, "Octbec a matlab toolbox for optimal quantum control of Bose-Einstein condensates," *Computer Physics Communications*, vol. 185, no. 1, pp. 194 – 216, 2014.

Acknowledgements

I would like to thank my advisor Ulrich Hohenester for his help and guidance. He always pointed me in the right direction and over the years I was able to learn a lot from him. I also want to offer my thanks to Jörg Schmiedmayer and his whole group. Collaborating with them was a great experience and stimulating discussions with Tarik Berrada and Jean-François Schaff provided vital input for this work. I am grateful for the warm hospitality I experienced during my visit in Kassel. Working with Christiane Koch, Daniel Reich and Michael Goerz was not only fruitful, but also very enjoyable. Furthermore, I want to thank Judith for her support, her patience and her love.