We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK
CITATION
INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Stochastic Optimization of Bose-Einstein Condensation Using a Genetic Algorithm

W. Rohringer, D. Fischer, M. Trupke, J. Schmiedmayer, T. Schumm
Vienna Center for Quantum Science and Technology, Atominstitut,
TU Wien, 1020 Vienna
Austria

1. Introduction

In 1924/25, Satyendranath Bose and Albert Einstein predicted that particles of integer spin (today called bosons) should undergo a quantum statistical phase transition when cooled to temperatures very close to absolute zero (Bose, 1924; Einstein, 1925). This phenomenon, nowadays called Bose-Einstein condensation, has been experimentally observed seventy years later using ultracold (nanokelvin temperatures) gases of bosonic neutral atoms (Anderson et al., 1995; Bradley et al., 1995; Davis et al., 1995c). This spectacular experimental achievement was awarded the Nobel price in 1995 and has triggered a new research direction "ultracold quantum gases" with more than 200 groups worldwide today. [1]

Bose-Einstein condensates (BEC) of ultracold atomic gases represent a fascinating exotic state of matter with properties entirely determined by the laws of quantum mechanics. Although they consisting of several thousands up to millions of atoms, the quantum state can in most cases be described by a single collective wave function with a common quantum phase. This wave function usually has a spatial extent ranging between 1-100 microns, allowing us to observe quantum mechanics essentially by eye using basic magnification optics.

The Bose-Einstein condensate can be described as a coherent matter wave in close analogy to the optical field emitted by (or inside) a laser. This analogy has brought about the field of "quantum-atom-optics" which aims to implement standard elements and experiments known from laser optics using matter waves. As heavy rest mass particles, such as atoms, are very sensitive to gravity or acceleration/rotations, matter-wave interferometers promise orders of magnitude in sensitivity gain compared to their photonic counterparts (Berman, 1996).

Over the last decade, experimental tools for the creation and manipulation of ultracold quantum matter have reached an impressive degree of sophistication. This allows the construction of tailored Hamiltonians for the system and the "quantum simulation" of more general physical situations, not connected to atomic physics alone. Optical lattice potentials allow to mimic solid state physics with a high degree of control over essentially all experimental parameters. Particle interactions can be tuned via Feshbach resonances, allowing the study of superconductivity, superfluidity and the formation of ultracold

---

[1] (n.d.). see atom traps world wide: http://www.uibk.ac.at/exphys/ultracold/atomtraps.html and links therein.

molecules. One-dimensional and two-dimensional quantum systems have been realized with ultracold gases in constrained geometries. Experiments have also been extended to fermionic atoms which follow completely different quantum statistics at low temperatures and will one day allow to simulate electrons.

The creation of an ultracold gas of atoms is a complex and delicate procedure which involves many steps like laser cooling, conservative atom trapping, evaporative cooling etc. The fundamental steps common to most experimental approaches will briefly be outlined in section 2. Together with the actual experiment to be performed and the detection, a whole experimental sequence takes between several tens of seconds and a few minutes. Some operations within this sequence take place (and have to be timed) on a microsecond timescale, hence a computer based experimental control is inevitable. As the detection of the system is almost always destructive, the experimental cycle has to be repeated many times to accumulate statistics or vary experimental parameters. Complex optimizations or multi-parameter scans can require several days of continuous operation.

In most setups working with ultracold atoms the result of an experiment is an image of the atomic density distribution (see section 2.2 for details). These images are acquired using computer controlled CCD cameras, yielding a graphical file for immediate data processing. With modern computers and efficient algorithms, the analysis of a single result image takes a few seconds, usually much faster than the entire experimental cycle. Hence effectively real-time analysis for various feedback schemes is available, which is the basis for the stochastic optimization methods described herein.

In our work, we close the loop between computer-based experimental control and equally computer-based real-time analysis to automatically optimize various experimental tasks using a genetic algorithm (GA). These tasks range from optimizing specific parts of the experimental sequence for optimal result parameters (atom number, temperature, phase space density) to complex ramp shapes that produce quantum gases in specific external or internal states. To our knowledge, there are very few implementations of stochastic optimization to physical systems. Aside from our experiment (Rohringer et al., 2008; Wilzbach et al., 2009), some examples are the optimization of the temporal shape of laser pulses (Baumert et al., 1997), or the tailoring of pulse shapes to control chemical reactions (Assion et al., 1998).

We wish to underline that none of the methods described here is specific to our experimental setup or physical system. This approach is applicable in a large variety of fields of experimental and industrial research.

The following chapter will start with a brief and general introduction to experiments with ultracold atoms in section 2. In section 3 we describe the implementation and internal structure of the GA. Examples of stochastic optimizations on various levels are given in section 4. For comparison, a brief overview of purely computer-based optimization tasks is given in section 5. In section 6 we close with a summary and outlook on further developments.

2. Experiments with ultracold quantum gases

This section will give a brief introduction to experiments with ultracold atoms. Since the first realization of Bose-Einstein condensation in 1995, experimental techniques have evolved and diversified. We will focus on the major steps which are still common to most approaches and which are necessary to understand the optimizations performed and presented in section 4. For simplicity, we divide an experimental cycle into two main phases: first the preparation of the ultracold gas or Bose-Einstein condensate, which in itself consists of several stages. The second phase concerns the detection of the sample after manipulation, and the acquisition

of significant physical quantities. These allow an evaluation of the experiment run and the chosen experimental parameters. This analysis is the starting point for the genetic optimization routines described in sections 3 and 4. For a more comprehensive review on the creation and characterization of ultracold Bose and Fermi gases see (Ketterle et al., 1998) and (Ketterle & Zwierlein, 2008).

2.1 Preparation of ultracold atomic gases

In the following, we characterize a gas of neutral atoms by its temperature T, and by the corresponding de Broglie wavelength $\lambda_{dB} = (2\pi\hbar^2/mk_BT)^{1/2}$, where m is the mass of the atom and k_B is Boltzmanns constant.

The de Broglie wavelength can be regarded as the size of a quantum mechanical wave function of an individual atom of the gas. It increases as the gas gets colder. The gas density n is related to the average distance d between atoms through $n = d^{-1/3}$. The quantum phase transition to a Bose-Einstein condensate takes place when bosonic atoms are cooled to a point where the atomic wavepackets start to overlap, more precisely at a critical temperature T_c where the phase space density $n\lambda_{dB}^3 \approx 2.612$. This temperature T_c is typically between $100\,\mathrm{nK}$ and $1\,\mu\mathrm{K}$, the atomic density is between $10^{13}\,\mathrm{cm}^{-3}$ and $10^{14}\,\mathrm{cm}^{-3}$ (compare figure 1).

The starting point for experiments with ultracold quantum gases is usually a thermally activated source of neutral atoms, providing particles at temperatures around $500\,\mathrm{K}$ and densities of $10^8\,\mathrm{cm}^{-3}$. [2] Therefore to reach Bose-Einstein condensation the temperature has to be reduced by 9 orders of magnitude, while the atomic density needs to be increased by up to 6 orders of magnitude. This enormous cooling power is attained by using a combination of extremely efficient techniques which will be briefly outlined in the following. A schematic trajectory through phase space on the path towards Bose-Einstein condensation is depicted in figure 1.
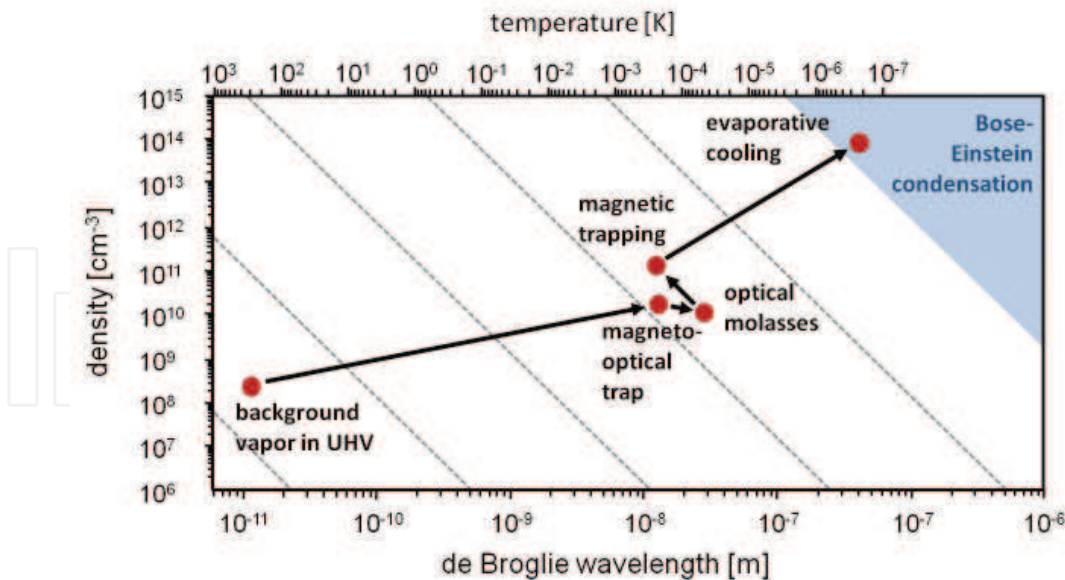

Fig. 1. Schematic representation of a typical trajectory through phase space in a Bose-Einstein condensation experiment. The various steps are explained in the text.

---

[2] The densities at the starting point of the experiment vary significantly depending on the specific approach, they may reach $10^{14}\,\mathrm{cm}^{-3}$ in high flux Zeeman slowers.

2.1.1 Laser cooling

Laser cooling and trapping relies on light forces, emerging when a (near-) resonant laser interacts with atomic transitions (see (Metcalf, 1999) for a detailed description). When an atom absorbs a photon of energy $h\nu_{atom}$, its momentum changes by $\hbar k = 2\pi\nu_{atom}/c$ where k is the wave vector of the laser. When (spontaneously) emitting the photon again, the atom momentum changes again by $\hbar k$. While the momentum "kick" in emission is in random direction and averages out over many absorption-emission cycles, the momentum transfer in absorption is directive, pointing in the direction of the laser (light pressure). Hence, on average, one recoil momentum of $2\pi\nu_{atom}/c$ is transferred to the atom per cycle. The interaction with the laser changes the momentum of an atom, which is the action of a "light force". This force is only determined by the frequency ν_{atom} and the scattering rate R of the atom: $F = \dot{p} = \hbar k R$ with

$$R = \frac{\Gamma}{2} \frac{s_0}{1 + s_0 + (2\Delta/\Gamma)^2} \tag{1}$$

where $\Gamma = 1/\tau$ is the transition linewidth, $s_0 = I/I_{sat}$ is the saturation parameter, I_{sat} is the saturation intensity of the atomic transition and $\Delta = \nu_{atom} - \nu_{laser}$ the laser detuning with respect to the atomic transtion frequency ν_{atom}. The maximum light force amounts to $F_{max} = \Gamma\hbar k/2$ on resonance and is hence mainly determined by the linewidth of the used atomic transition. For the 780 nm D2 transition of ^{87}Rubidium used in the experiments presented here, the acceleration of an atom at rest by a resonant laser is $a \approx 10^5$ m/s^2, four orders of magnitudes higher than gravity!

Under the influence of the light force, the atom changes its velocity v quickly, giving rise to the Doppler effect. The laser now interacts with an effective detuning $\Delta_{eff} = \Delta + kv$. After a series of absorption-emission cycles, (≈ 800 in the case of ^{87}Rubidium), the effective detuning is so large that no further interaction with the laser takes place. Turning this argument around, choosing the detuning of the laser allows to selectively address a specific velocity class of atoms, making the light force velocity-dependent. If the laser frequency is tuned below the atomic resonance frequency ν_{atom} ("red" detuning), the laser preferably interacts with atoms moving towards the laser, slowing them down. This method is routinely employed to slow down atomic beams coming from a hot thermal source (Metcalf, 1999).

If two counterpropagating laser beams of equal intensity and frequency are used on the atoms, the resulting light force has a dispersion-like shape. Around zero velocity, the force can be approximated to

$$F = \hbar k^2 \frac{8s_0\Delta/\Gamma}{(1 + (2\Delta/\Gamma)^2)^2} v. \tag{2}$$

The light force takes the form of a friction or "molasses" force (optical molasses, see below). The cooling strength can be adjusted again by adjusting the detuning Δ. A smaller (red) detuning leads to colder temperatures. However a more narrow velocity class is then addressed by the laser, leading to a lower number of cooled atoms. Therefore a compromise between number and temperature has to be found experimentally. An optimization of laser cooling parameters using a stochastic GA can be found in section 4.

The above description takes place entirely in momentum/velocity space, so far no spatial dependence and hence no trapping is introduced. To render the optical cooling force spatially dependent, the magnetic Zeemann effect is employed. Using a quadrupole magnetic field (e.g. generated by coils in anti-Helmholtz configuration) the atomic transition is shifted, depending on the position. With the right laser detuning (and polarization) and the right quadrupole field

the effective cooling force can be designed so that it points towards zero magnetic field, where the atoms will accumulate. The scheme can easily be extended to three dimensions, realizing a true 3d trapping of neutral atoms in free space.

The combination of optical and magnetic fields to at the same time cool and spatially trap atoms is called magneto-optical-trap (MOT) and has become a standard tool in atomic physics. Its development, together with a thorough theoretical explanation of the relevant effects, led to the award of a Nobel price in 1997 (Chu, 1998; Cohen-Tannoudji, 1998; Phillips, 1998).

A MOT usually catches atoms from the low-velocity tail ($\approx 10\,\mathrm{m/s}$) of the thermal distribution of a background gas at room temperature or from a slowed atomic beam. Typical total atom numbers (^{87}Rubidium) are $10^8 - 10^9$ with a density of $10^{11}\,\mathrm{cm}^{-3}$ and a temperature of $\approx 200\,\mu$K. This constitutes a significant step towards Bose-Einstein condensation, as illustrated in figure 1. Almost all experiments with ultracold gases start with a phase of magneto-optical trapping.

2.1.2 Optical molasses

As described above, the experimental settings used in a magneto-optical trap are usually optimized for trapping high number of atoms, rather than for obtaining the lowest possible temperatures. Before loading atoms from a MOT into a conservative trapping scheme, a short phase of "optical molasses" (1-100 ms) is employed to further lower the temperature of the gas. Here, the magnetic fields are quickly switched off, extinguishing the spatial trapping. The lasers are adjusted to different detunings (usually significantly further from the atomic resonance) and amplitudes to provide an optimal optical molasses. Atoms hence expand in the laser field, but reduce their kinetic energy (and hence the temperature of the sample, once recaptured in a conservative trap). As the atoms fulfill a damped Brownian motion, they will ultimately diffuse out of the volume that can be captured by a conservative trap. Again a compromise has to be found between temperature (favoring long molasses times) and atom number (favoring short molasses times) to be transferred.

Several different processes contribute to the enhanced cooling in the optical molasses, such as Sisyphus cooling or dark state effects which go beyond the simple model of Doppler cooling described above. For a comprehensive overview see (Metcalf, 1999).

2.1.3 Conservative atom trapping

Even though laser cooling and optical molasses allow an enormous gain in phase space density, there are fundamental limits. As all optical forces rely on absorption and successive emission of photons, the random momentum transfer in spontaneous emission induces a heating mechanism which limits the temperatures that can be achieved. A lower bound, termed the recoil limit, can be obtained by calculating the energy associated with a single photon recoil:

$$E_{recoil} = k_B T_{recoil} = \frac{\hbar^2 k^2}{2m}. \tag{3}$$

For ^{87}Rubidium atoms and the standard $780\,\mathrm{nm}$ D2 transition, this corresponds to a temperature of $0.4\,\mu$K.

Therefore to confine and further cool the gas, another trapping scheme has to be employed which does not rely on photon exchange. Such traps can be constructed using the interaction of an electric[3] or magnetic dipole moment with external electric or magnetic fields. These

---

[3] Note that most experiments with ultracold atoms are performed with alkali atoms, where the electric dipole moment is only induced in the presence of en external electric field.
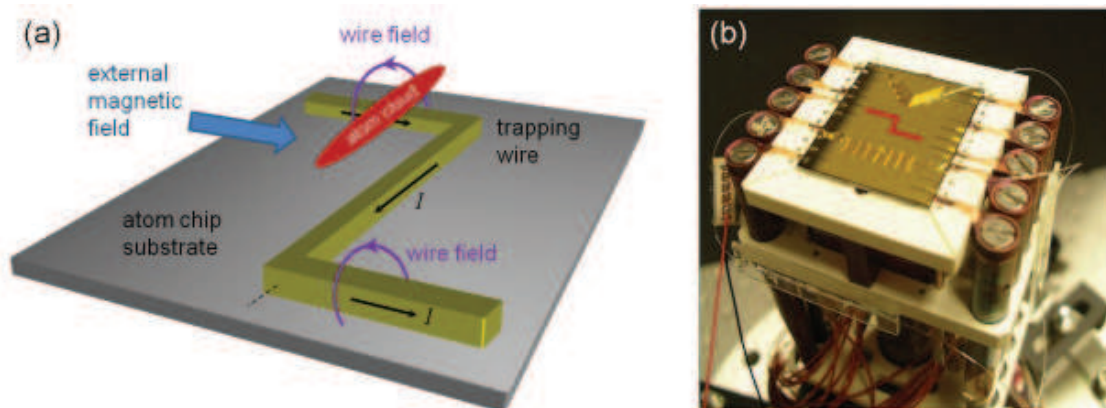
Fig. 2. (a) Schematic representation of a magnetic wire trap. The combination of a homogeneous external magnetic field and the magnetic field of a current carrying wire gives rise to a three-dimensional potential minimum above the wire, that can be used to trap and manipulate neutral atoms. (b) Experimental implementation of a wire trap using an atom chip. The position of the central trapping wire is indicated in red. Optical fibres used for on-chip fluorescence detection can also be seen.

fields give rise to a shift of the internal atomic energy via the Stark or Zeemann effect which acts as an effective potential for the atoms. As the experiments described in the following only employ magnetic interactions, we will concentrate on magnetic trapping. A comprehensive review on optical trapping relying on the electric interaction can be found in (Grimm et al., 2000).

The interaction of an atomic magnetic moment $\vec{\mu}$ with an external, inhomogeneous magnetic field $\vec{B}(\vec{r})$ gives rise to the potential $V(\vec{r}) = -\vec{\mu}\vec{B}(\vec{r})$. Solving the Zeeman Hamiltonian within an adiabatic approximation gives rise to the magnetic quantum numbers m_F and we can write the potential $V(\vec{r}) = m_F g_F \mu_B |B(\vec{r})|$, where g_F is the Landé factor and μ_B the Bohr magneton. For atomic states where $m_F g_F > 0$ atoms are attracted to a spatial minimum of the magnetic field ("low-field seekers") whereas for $m_F g_F < 0$ atoms are attracted to a magnetic field maximum ("high-field seekers")[4]. From Maxwell's equations, one can derive that only a minimum of magnetic field can be created in free space, hence only low-field seekers can be trapped magnetically (Wing, 1984). All experiments described in the following are in the $|F = 2, m_F = 2 >$ state of ^{87}Rubidium, where $g_F = 1/2$.

To give an order of magnitude: a magnetic field of 1 Gauss (10^{-4} Tesla) leads to a potential energy of $k_B \times 67\,\mu$K. As magnetic traps usually work with tens of Gauss, atomic clouds prepared by laser cooling and optical molasses ($\approx 50\,\mu$K) can easily be captured. However, directly catching from room temperature background gas would require hundreds of Tesla.

A plethora of magnetic field configurations has been developed over time to trap atoms and reviewing them here is beyond the scope of this chapter. An overview can be found in (Ketterle et al., 1998). A simple and elegant way to produce magnetic traps with strong spatial confinement are wire traps as used in the experiments described below (Folman et al., 2002; Fortagh & Zimmermann, 2007). In brief, combining the magnetic field of a current carrying wire and homogeneous fields produced by external coils creates a magnetic trap following the geometry of the wire (compare figure 2). Integrating these trapping wires by using techniques from electronic circuit lithography ("atom chips") allows the creation of potential landscapes and provides a high degree of spatial control over the atomic gas, with

---

[4] Obviously, atoms in the $m_F = 0$ state are insensitive to magnetic fields

high temporal resolution. Bose-Einstein condensation on a chip was first demonstrated in 2001 (Hänsel et al., 2001; Ott et al., 2001; Schneider et al., 2003), and atom chips have since then become a standard tool in ultracold atom research.

2.1.4 Evaporative cooling

Magnetic trapping as described above provides a means to confine neutral atoms of a specific temperature in free space. However, as magnetic (and also electric) potentials are conservative, no cooling takes place. The temperature of the gas can be changed by (adiabatically) changing the atomic confinement, however, phase space density is maintained and hence no progress towards Bose-Einstein condensation can be made.
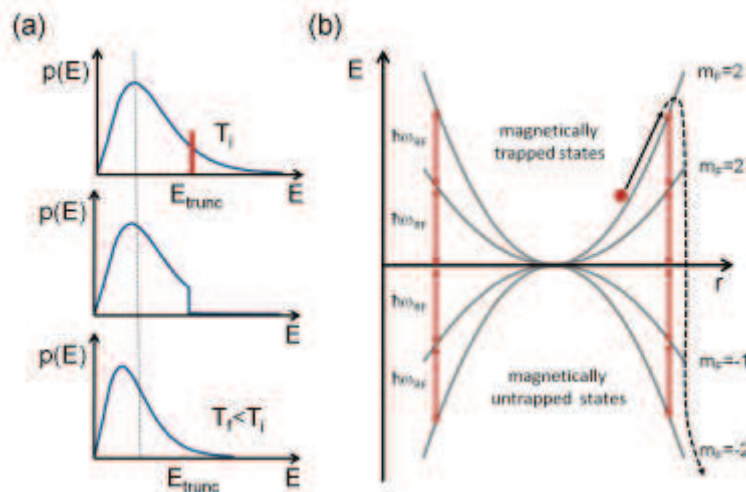

Fig. 3. (a) Principle of evaporative cooling. A thermal Maxwell-Boltzmann distribution characterized by a temperature T_i is truncated at energy E_{trunc} with $E_{trunc} > k_B T_i$. The truncated system relaxes to thermal equilibrium at a lower temperature T_f. (b) Selective removal of hot atoms by adjusting the frequency ω_{RF} driving spin flip transitions to untrapped states.

To gain in phase space density and decrease the gas temperature in a steady trap, an additional cooling mechanism termed "evaporative cooling" is employed. Similar to blowing onto a coffee cup, evaporative cooling relies on the selective removal of energetic (hot) particles. The system successively relaxes back to thermodynamic equilibrium (via particle collisions) at a lower temperature (see figure 3).

To selectively remove hot atoms from the magnetic trap, spin-flip transitions between trapped and untrapped Zeeman states are induced using radio frequency (RF) fields. By choosing the RF-field's frequency, the transitions occur at distinct regions in space (magnetic equipotential surfaces). The "hottest" atoms with highest kinetic energy explore the outermost regions of the magnetic trap, so these can be removed selectively ("RF knife"). As the system re-thermalizes at lower temperature, the frequency of the RF field has to be adjusted. This leads to dynamic forced evaporative cooling with time (Davis et al., 1995b; Luiten et al., 1996). As illustrated in figure 1, evaporative cooling allows to reduce the temperature of a gas by another two orders of magnitude and has enabled the creation of Bose-Einstein condensates (Davis et al., 1995a). However, the atom number is reduced by a similar factor. So far, no cooling method able to achieve such low temperatures while maintaining the initial atom number has been found. The details of the evaporative cooling process crucially depend on the details of the experimental implementation, in particular on the lifetime of

the magnetically trapped atoms and the collision and hence re-thermalization rate. The optimization of evaporative cooling using a stochastic GA is described in section 4.

2.2 Detection

Most information in ultracold atom experiments is gained through the interaction of atoms with light. Although several non-optical methods for specific applications or unique atom species (Santos et al., 2002) are in use, optical imaging is currently the main detection method for cold atomic gases. As far as imaging is concerned, atom-light interactions can be divided into three processes: absorption, re-emission and phase alteration of incident light, giving rise to three detection methods, two of which - absorption and fluorescence imaging - are used in our experiment[5]. Both methods result in a picture of the atom cloud, destroying it in the process. Quantities characterizing the state of the cloud can be extracted either from a single picture or from a series of images taken while varying one of the experimental parameters from shot to shot. The following section will describe these quantities and how to measure them. Suitable light sources for excitation of atoms are lasers tunable in the frequency range near an atomic transition, while CCD - cameras are ideal detectors for light (or lack thereof) whenever time resolution is not a critical factor.[6] In order to reach the resolutions needed for the examination of structures like interference fringes or vortices inside cold atomic clouds, it is neccessary to put special emphasis on the imaging optics, which usually has to be custom-built for each experiment.

2.2.1 Absorption imaging

This method consists in recording the shadow which an atom cloud casts onto a detector due to the absorption of a certain fraction of photons when irradiated with laser light. By comparison with the intensity of the beam in absence of the atomic cloud, the atoms' density distribution and a series of other parameters discussed in section 2.2.3 can be calculated. The detection beam can be absorbed almost completely if the density of the atom cloud becomes sufficiently high. These *optically dense* clouds make a quantitative analysis of an absorption image difficult. In order to compensate for this, it is possible to detune the laser light from resonance, lowering the absorption cross-section. The former introduces diffraction as well as a phase shift of the transmitted light, and going to high detunings as well as filtering out the unscattered transmitted light components leads to dispersive imaging. Absorption imaging introduces heating: Since each absorbed photon transfers a momentum $\hbar k$ to the atom, the cloud is literally blown away by the imaging light, making absorption imaging a destructive technique.

2.2.2 Fluorescence imaging

In fluorescence imaging, the atom cloud is also irradiated with a laser beam. However now it is not the transmitted intensity that is measured, but the number of photons scattered into the solid angle Ω covered by the detector. If Ω were equal to 4π, fluorescence imaging would just collect all the photons missing from an absorption picture. However, since the detector usually has a coverage factor $f_c = \frac{\Omega}{4\pi}$ of a few percent, in comparison the fluorescence signal is about a factor 100 weaker. Yet, fluorescence imaging has two advantages: In situations where the

---

[5] For a review of dispersive imaging methods, see for example (Ketterle et al., 1998) and references therein.

[6] For fast detection, photomultipliers or avalanche photodiodes are required, trading their advantages for lower detection efficiency and, in most cases, lower spatial resolution.

dissipative light force is used to trap the atoms - the MOT and optical molasses phase in our experiment - fluorescence photons come "for free" and allow non-destructive measurements. Additionally, fluorescence imaging has favourable properties for imaging dilute atom clouds. A comparison of the signal to noise ratio (SNR) - neglecting all noise sources except atomic shot noise - for absorption and fluorescence imaging yields:

$$\frac{SNR_f}{SNR_a} = \sqrt{\frac{f_c}{OD}}. \tag{4}$$

Thus, if the cloud's optical density OD drops below the coverage factor, fluorescence imaging becomes better in terms of SNR, at least as long as other noise sources deliver a comparable contribution for both methods. Hence, fluorescence imaging is the preferred technique for detecting extremely dilute clouds. In contrast to absorption imaging, fluorescence detectors need not be exposed to the high light intensities of the source laser illuminating the atoms. As a consequence, highly sensitive detectors like EMCCD cameras and avalanche photodiode - based single photon counting modules can be employed. Several techniques based on fluorescence, like lightsheet imaging (Bücker et al., 2009; Rottmann, 2006) or fiber-based detection methods allow to detect atomic clouds with single atom sensitivity.

2.2.3 Evaluating absorption images

The intensity of a monochromatic light beam travelling in y - direction through an opaque medium is attenuated with

$$\frac{dI}{dy} = -\sigma I n \tag{5}$$

where σ denotes the scattering cross-section and n the density of the atomic cloud. As long as the cross-section is a constant with respect to intensity, i.e. in the case of linear optics defined by low intensities[7], this equation can simply be integrated to yield *Lambert - Beer's law* as result:

$$I = I_0 e^{-\sigma n}. \tag{6}$$

If we allow a density distribution in the (x, z) direction, this gives $I(x, z) = I_0(x, z)e^{-\sigma n_c(x,z)}$, with column density $n_c(x, z) = \int dy\, n(x, y, z)$. The scattering cross-section depends on the detuning Δ, with $\sigma(\Delta) = \sigma_0 / \left(1 + (2\Delta/\Gamma)^2\right)$. The column density can then be expressed as:

$$n_c(x, z) = \sigma(\Delta) \ln\left(\frac{I_0(x, z)}{I_t(x, z)}\right). \tag{7}$$

This measured column density allows to deduce important basic properties of the dilute atomic cloud:

- **Total atom number**
  In the continuous case, the total atom number can be obtained by integrating the column density:

$$N = \int n_c(x, z)\, dx\, dz. \tag{8}$$

  In the experiment, we have to consider that x and z are discrete, their step size defined by the area A imaged onto a single CCD pixel, with the magnification M. Therefore, for

---

[7] For high intensities, stimulated emission begins to play a role, leading to an enhanced forward scattering rate

a square region of interest containing p pixels, equation 8 becomes the discrete sum over these pixels:

$$N = A \sum_p n_p(x,y) = \frac{\Delta x_{Pixel} \Delta y_{Pixel}}{M^2} \sum_p n_p(x,y). \tag{9}$$

- **Thermal Atom Density in the Trap**

  In order to derive temperature or temperature-dependent quantities like phase space density, or to gain information about the trapping potential, we need a model which describes the density distribution of the atoms in the trap, as well as its evolution after releasing the cloud from the trap. Since thermal atom clouds and Bose - Einstein condensates represent different thermodynamic phases, where density is closely linked to the phase transition's order parameter, one expects different behaviour in the two regimes. Trapping potentials created by static magnetic fields are harmonic around the field minimum:

$$V(x,y,z) = \frac{m}{2} \left(\omega_x^2 x^2 + \omega_y^2 y^2 + \omega_z^2 z^2 \right). \tag{10}$$

For a thermal cloud of bosons, the density distribution for high temperatures can be expressed as (Ketterle et al., 1998; Reichl, 1998)

$$n(\mathbf{r}) = \left(\frac{2\pi\hbar^2}{mk_BT} \right)^{3/2} g_{3/2} \left(z \left(\mathbf{r} \right) \right) \tag{11}$$

with $z = e^{(\mu - V(x,y,z))/k_BT}$. Here, $g_j(z) = \sum_i \frac{z^i}{i^j}$ is the *Bose function*, introducing *Bose enhancement*, which means increased density compared to the classical case, where the distribution would be Gaussian. For high temperatures or low densities, we can neglect Bose enhancement, and with the halfwidths $w_i = \sqrt{\frac{2k_BT}{m\omega_i^2}}, i = x, y, z$ and zero chemical potential, we recover a Gaussian distribution:

$$n(x,y,z) = n_0 e^{-\left(\frac{x^2}{w_x^2} + \frac{y^2}{w_y^2} + \frac{z^2}{w_z^2} \right)}. \tag{12}$$

Experimentally, we only have access to column densities:

$$n_c(x,z) = \int dy \, n(x,y,z) = n_0 \sqrt{\pi} w_y e^{-\left(\frac{x^2}{w_x^2} + \frac{z^2}{w_z^2} \right)} = \tilde{n}_0 e^{-\left(\frac{x^2}{w_x^2} + \frac{z^2}{w_z^2} \right)}. \tag{13}$$

We can determine $\tilde{n}_0$ by normalization with respect to the atom number:

$$N = \int dx \, dz \, n_c(x,z) \Rightarrow \tilde{n}_0 = \frac{N}{\pi w_x w_z}. \tag{14}$$

By comparison with equation 13 we can calculate the *peak density* of the thermal cloud:

$$n_0 = \frac{\tilde{n}_0}{\sqrt{\pi} w_y}. \tag{15}$$

Equivalently, n_0 can be determined directly from normalization of equation 12. Since the picture is integrated along y, we need an assumption about the density distribution on this axis. Our magnetic traps yield cigar-shaped atom clouds, therefore it usually holds that $w_y = w_z$.

- **Temperature**
  A thermal atom cloud released from a trap by suddenly (non-adiabatically) switching off the trapping potential expands isotropically, the intitial isotropic velocity distribution being conserved. The evolution of the cloud's half-width is given by

$$w_{r_i}(t)^2 = \frac{2k_BT}{m\omega_i^2} + \frac{2k_BT}{m}t^2. \tag{16}$$

By repeatedly measuring $w_{r_i}^2$ at different times of flight t^2 and plotting the two quantities against each other, the temperature can be determined by the resulting line's slope $\frac{2k_BT^2}{m}$. For $t = 0$, an estimation of the trap frequency in direction r_i is possible.

- **Phase Space Density**
  The important quantity which has to reach a treshold value of 2.612 in order to achieve Bose–Einstein condensation is the phase space density of the atomic cloud:

$$D = n_0 \, \lambda_{dB}^3 \tag{17}$$

comprising the peak density n_0 and the thermal de Broglie wavelength $\lambda_{dB} = \sqrt{\frac{2\pi\hbar^2}{mk_BT}}$.

3. Stochastic optimization in an ultracold atom experiment

3.1 Setup of the feedback loop

The goal of our experiment is the investigation of ultracold ^{87}Rubidium clouds in chip - based magnetic traps, employing all steps involving preparation, manipulation and detection of the atoms described in the previous sections. The focus lies on the application of fiber optics integrated directly on the chip as a tool for the detection of ultracold atoms (Heine et al., 2010).

Figure 4 illustrates our hardware feedback loop. A real-time control system governs the behavior of the experimental apparatus via 30 output channels providing analog control voltage signals, as well as 28 digital TTL channels, with a time resolution of 25 μs. The control channels allow us to manipulate practically all aformentioned experimental quantities, including laser detunings and intensities, magnetic field strengths and radio frequency fields, both in timing and magnitude. User input is collected by an interface software written in MATLAB.

Our absorption images of atomic clouds are taken with a Pixelfly QE interline CCD camera, read out and evaluated by a MATLAB program on a dedicated computer. The algorithm providing feedback between acquisition and control software is also implemented in MATLAB as part of the acquisition software, and communication between acquisition and control hardware is established via a UDP interface provided by MATLAB.

Briefly, our experimental sequence consists of a magneto-optical trap followed by a molasses stage. The atoms are then loaded into a magnetic trap created using a macroscopic wire structure located behind the atom chip, and subsequently cooled using RF-induced evaporation. All data discussed in section 4 was obtained by absorption imaging of atomic

clouds released from the magnetic trap at this experimental stage. This point in the sequence is crucial as it determines the atomic phase space density available for experiments using the chip structures, including condensation in the chip trap and transport to the fiber detector.
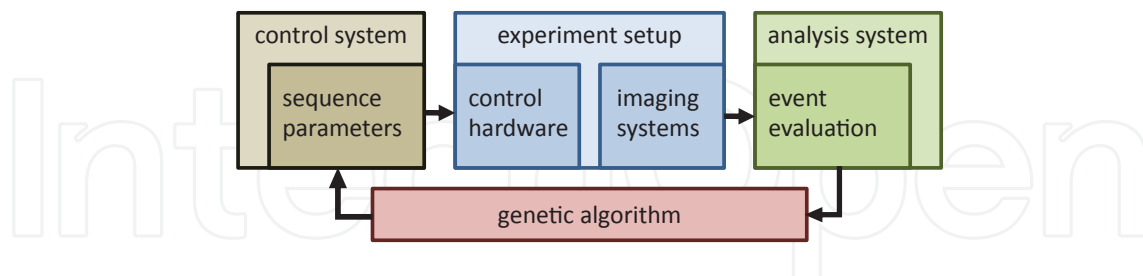

Fig. 4. Scheme of the hardware feedback loop.

In order to perform an experiment, the right set of input parameters for all the devices integrated within the experiment must be found. While for some devices optimal values can be directly calculated or gained through simulations, usually it is only possible to constrain the range of useful values to a certain extent. Within this value range, optimization is necessary. The typical approach is to scan one variable while fixing all others, and repeating this procedure for all variables, in some cases iteratively, until satisfactory conditions for the experiment are met. This is a time consuming task, and inefficient if subsets of these parameters are coupled.

Instead of performing this task manually, we implement an automatic optimization scheme. With the availability of a control system allowing relevant parameters to be set via a program running on a PC, evaluation software capable of automatically acquiring measurements as well as extracting all interesting information and with the possibility of communication between these two programs, the technical requirements for the implementation of an automated optimization scheme directly into a hardware feedback loop are met.

The big number of different optimization problems arising in our setup, depending on the choice of parameters to optimize, makes the implementation of a deterministic algorithm unfeasible. On the other hand, stochastic optimization algorithms have been successfully used in many applications.

3.2 Choice of algorithm

The question can be raised whether it is possible to build an optimal algorithm, outperforming all the others on all possible optimization problems. However, according to a 'No Free Lunch' - theorem for search and optimization (Wolpert & Macready, 1995; 1997) there is no such intrinsically optimal algorithm. This can be expressed as follows:

All algorithms that search for an extremum of an objective function perform exactly the same, when averaged over all possible objective functions. In particular, if algorithm A outperforms algorithm B on some objective functions, then loosely speaking there must exist exactly as many other functions where B outperforms A. (Wolpert & Macready, 1995)

In order to define a performance measure for an algorithm a, let $P(d_m | f, m, a)$ be the conditional probability of obtaining a particular sample d_m by iterating an algorithm a m times on an objective function f. For a finite problem space and a finite space of objective function values, it can be shown (Wolpert & Macready, 1997) that for any two algorithms a_1 and a_2 it

applies that

$$\sum_f P\left(d_m|f,m,a_1\right) = \sum_f P\left(d_m|f,m,a_2\right).$$ (18)

Therefore in order to deliver optimal performance, optimization algorithms have to be matched or tailored to specific problems. The main point in this context is the balance between what is called 'exploration versus exploitation'. Any random search mechanisms within an algorithm contribute to exploration, while gradient search - based elements exploit the parameter space structure in order to find the optimum. As a consequence, picking an algorithm and choosing its parameters to fit the problem at hand is as valid an approach as choosing a specific algorithm.

Out of the many different methods available, we choose to implement a real coded genetic algorithm (RCGA). Belonging to the first stochastic optimization methods developed, the convergence behavior of genetic algorithms is well documented for different classes of problem spaces. They belong to the class of global optimization algorithms, exploiting information from different parts of the problem space in parallel as opposed to local algorithms like e.g. stochastic hill climbing or simulated annealing. This property reduces the probability of premature convergence towards local optima. RCGAs allow to define states directly from real valued optimization parameters, without any intermediate encoding, which is simple and intuitive. While early literature suggests that binary encoding is key to the convergence properties of genetic algorithms, more recent studies, backed up by an increasing number of applications, have shown that real coded algorithms suffer from no general disadvantages as compared to other encoding schemes, and even perform better for many applications.

3.3 Implementation

The basic concept of the algorithm, as depicted in figure 5, is the same as applies to the initial canonical genetic algorithm and most subsequent implementations. After generation of a random starting population of states, real valued parameter vectors, the experiment is performed and evaluated for each state with respect to a measured value representing the objective function of the optimization problem. Based on the measurement results, the states are ranked and fitness values are assigned accordingly. The fitness values determine the probability for each state to be selected as a parent for a recombination procedure providing the next generation of states. After recombination, each state undergoes mutation, a stochastic alteration of one of its parameter values, with a preset probability. Subsequently, the next iteration begins by evaluating the resulting parameter vectors by experiment.

The time consuming process in this setup is evaluating the objective function, which means performing the experiment, with a duration of 35 seconds. Even more so than in purely computational applications of GAs, it is crucial to minimize the number of iterations before finding the optimum. Since runtime crucially depends on the number of states within each generation, the population size, the design goal is to keep this number as low as possible while preventing premature convergence due to rapidly decreasing diversity of the states. Simulations, backed up by our experiments, indicate that with proper adjustment of the genetic operators, as described in the next section, a population size of 20 states can ensure reliable convergence behavior for realistic problem spaces.
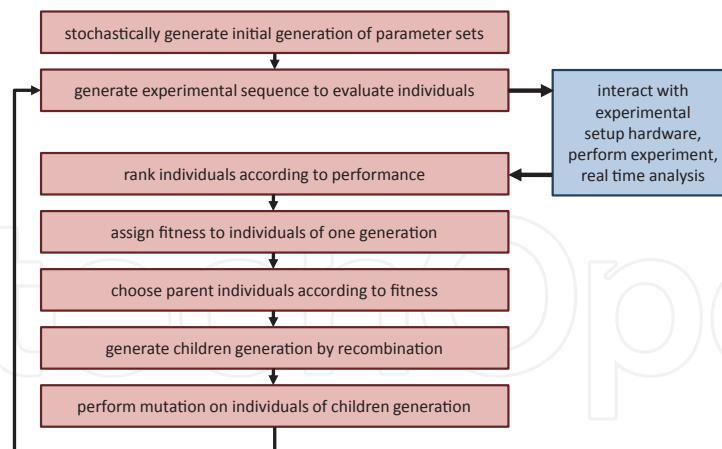
Fig. 5. Basic concept of the genetic algorithm.

3.3.1 Fitness assignment

Most fitness assignment schemes developed for genetic algorithms can be devided into two different classes: ranking based fitness assignment and proportional fitness assignment. The latter class has been shown to suffer from two problems, *premature convergence* and *stagnation* (Herrera et al., 1998; Pohlheim, 1999; Weicker, 2002, 2.Auflage 2007).

Ranking - based fitness assignment (Baker, 1985) avoids the two stated problems by distributing fitness values independently from the actual objective function values. A simple implementation chosen in our algorithm is linear ranking. Here, the sorted population members S_i, $i \in \{1...N_p\}$ are assigned a fittness given by

$$F(S_i) = \frac{2}{N_p} \left(1 - \frac{i-1}{N_p - 1} \right), \tag{19}$$

where N_p is the population size. This redistributes the fitness values linearly between 0 and 1. Several nonlinear ranking methods can be applied to shift recombination probability towards good or bad states.

3.3.2 Selection

Out of several available selection schemes, we chose *Stochastic Universal Sampling* (SUS), a variant of roulette selection. Given that each state has a fitness value $F(S_i)$ between 0 and 1 with a total fitness of 1, one can interpret the total fitness as the area of a roulette wheel devided into N_p sections with area $F(S_i)$. Creating a random number between 0 and 1 and selecting the state occupying the area including the random number is equivalent to spinning a roulette wheel and waiting for the pointer to stop in one of the sections. In this picture, SUS represents a roulette wheel which is spun once with n pointers equally partitioned between 0 and 1. SUS results in zero bias and minimum spread and is a widely used selection method. In order to speed up convergence, the best 20 percent of the parent states are taken over into the next generation, providing an *elitist* selection scheme.

3.3.3 Recombination

Generally, recombination is a genetic operator using two parent states to generate a new state. For binary strings, this means breaking each of the two strings at specific points and creating a new string by concatenating fragments stemming from different parents. The only option to

vary in this case is the number of fragments the parents are broken in, leading to *single-point*, *multi-point* or *uniform* crossover. The latter represents simply an extreme case of multi-point crossover, where the m - bit parent states are decomposed into n fragments.

For RCGAs, the state is represented by a string or vector of real numbers. Translating the idea of crossover to this situation gives what is called **discrete recombination**. One method to implement this is to decide for each vector component v_j of the child state $S_i^C(v_j)$ which parent (P_1 or P_2) contributes the variable value:

$$v_j^C = v_j^{P_1} a_j + v_j^{P_2} \left(1 - a_j\right).$$ (20)

Here, a_j is randomly chosen to be 0 or 1 for each v_j^C separately and $j \in \{1, ..., m\}$, where m denotes the total number of variables.

With discrete recombination, only variable values already realized in the starting population can be reached. In order to gain access to new value regions, real number represented states allow interpolation between two values. The most general recombination scheme to be derived this way is **intermediary recombination**. It can be implemented with equation 20, but an a_j picked from the interval $[-d, 1 + d]$ with uniformly distributed probability for each variable separately. This operator is also called BLX - α (*blend crossover*), where α equals $2d$. The hypercuboid of possible new values has a volume of

$$V_{PS}^C = (1 + 2d) \prod_{j=1}^{m} l_j$$ (21)

with l_j as length of the value region spanned by the variables $v_j^{P1, P2}$ and a total of m variables. For $d = 0$, the hypercuboid containing the possible children values is as big as the one spanned by the parent variables. Since the probability for a child value to lie inside the cuboid is higher than the probability to lie on its bounds, the cuboid volume will decrease with a growing number of iterations in this case, restricting the accessible part of the problem space without any influence of selection. By stretching the children's value space by the factor $(1 + 2d)$ one can compensate for this effect. Empirically, a value of $d = 0.25$ (BLX-0.5) has proven to conserve the cuboid volume in the limit of a large iteration number.

If a_j is not chosen for each variable separately, but rather once in the beginning of the recombination phase and kept the same for all variables, the **linear recombination** can be recovered as special case of intermediary recombination.

Since intermediary recombination with $d = 0.25$ (BLX-0.5) gives optimal convergence behaviour in many computer experiments Herrera et al. (1998); Pohlheim (1999), it has been chosen for our algorithm. Since it delivers real numbers as variable values while our stepsizes imposes a whole number representation, the routine's results are rounded to match the nearest allowed variable value.

3.3.4 Mutation

Mutation in an RGCA means randomly changing values in the state vector. A commonly used mutation routine has been presented in Mühlenbein & Schlierkamp-Voosen (1993) and Mühlenbein & Schlierkamp-Voosen (1995), and can be described as follows:

$$v_j^{mut} = v_j + s_j \cdot r D_j \cdot 2^{-u\kappa_m}.$$ (22)

Here, v_j^{mut} and v_j denote the mutated and source states respectively, while s_j randomly chooses the sign of the mutation step, r defines the *mutation range* as fraction of the variable *definition domain* D_j. The last term designates the used distribution characterized by the random number u which is uniformly distributed in the interval $[-1, 1]$ and the *mutation precision* κ_m. The latter defines a lower limit of $\frac{1}{2}^{-\kappa_m}$ for the mutation step size. Favoring small mutation steps over big ones, nonuniform mutation operators like this have shown to be advantageous for RCGAs in computer experiments.

From runs on test problems, our algorithm with population sizes between 20 and 30 has given good results with $\kappa_m = 10$, $r = 0.2$ and an overall mutation rate of 10 percent. This is consistent with literature suggesting that optimal mutation rates are inversely proportional to population size Haupt (2000).

4. Examples of stochastic optimization

4.1 Optimization of an optical molasses

The measurement presented here allows a simple comparison of a grid scan to the genetic-algorithm approach. The optical molasses has already been briefly described in section 2.1.2. In this phase, it is possible to reduce the temperature of ^{87}Rubidium atoms from the magneto-optical trap by an order of magnitude to ensure that a large fraction of the cloud has low enough energy to be trapped in a conservative magnetic trapping potential. This phase relies purely on the interaction of atoms with laser light. In our measurement, the optimized quantity is the atom number within the magnetic trap after the molasses phase, and the optimized parameters are molasses duration and laser detuning. Variations in the experimental conditions due for example to environmental noise, lead to a statistical uncertainty in the measured value. We therefore average over multiple experimental runs to reduce this uncertainty. The successful optimization of this experimental stage is shown in figure 6. In the 2d grid scan, we changed the detuning in steps of 5 MHz and the molasses duration in steps of 0.2 ms, and computed the average of 4 atom-number measurements at each pair of values, resulting in a scan duration of approximately 17 hours. For the GA optimization, we used a population size of 20 states and recorded the convergence over 16 generations. In this case, we only averaged over 3 atom number measurements. This optimization approach led to a reduction of the runtime to approximately 9 hours. The set of surviving parameters has already clustered near the optimum after this time, demonstrating the efficiency of the approach.

4.2 Optimization of evaporative cooling in a magnetic trap

The goal of evaporative cooling in a magnetic trap is to increase phase space density (see sections 2.1.4 and 2.2) of the atomic cloud as efficiently as possible. Efficiency means maximizing the amount of removed energy per removed atom. Technically, evaporative cooling as described in 2.1.4 is implemented with the help of an RF source that is tunable in frequency and therefore can create *RF - cooling ramps* by lowering the frequency of the applied field on timescales ranging from milliseconds to seconds. Frequencies from 10 MHz down to the 100 kHz range with sub - kHz stability and corresponding resolution have to be provided. In an ideal system, efficiency grows with ramp duration. The slower the ramp, the more time the system has to stochastically produce atoms with high energy that are removed by the RF field, while fast cooling means removing more atoms with lower energy. In reality, constraints for the steepness of these ramps are imposed mainly by two mechanisms: On the one hand,
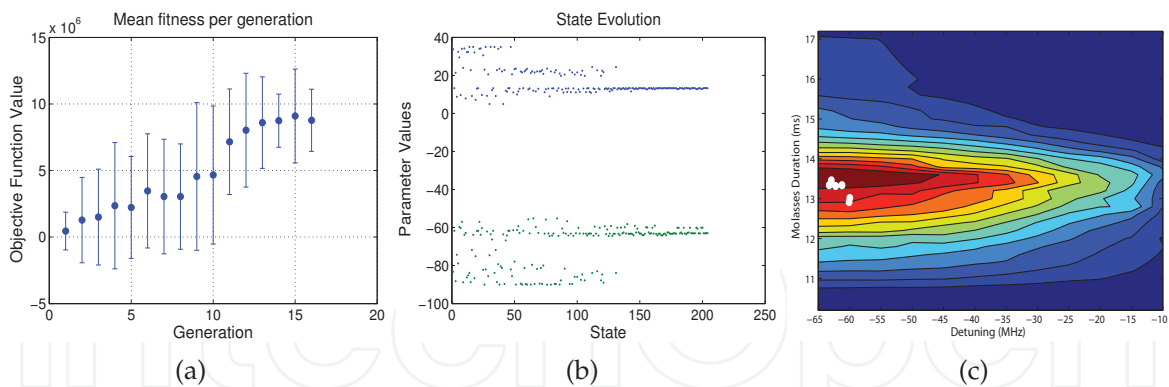
Fig. 6. Results of the optical molasses optimization. (a) Plot of the mean fitness per generation. (b) State Evolution. Blue points correspond to times given in milliseconds, whereas green points represent laser detunings, in MHz. (c) The plot depicts atom number as a function of duration and laser detuning. The points indicate the set of surviving parameters found by the algorithm.

evaporative cooling competes with different loss or heating mechanisms leading to decreasing phase space density by removal of cold atoms, providing an upper limit for the duration of useful cooling ramps. On the other hand, it is necessary to ensure that the atomic cloud has sufficient time to thermalize by interatomic collisions, which provides a lower bound for ramp duration. The optimum depends critically on technical details and usually has to be found empirically.

We optimize our RF-cooling ramp to yield a maximum in a dimensionless parameter $PSD \propto N^{1/3}/T$, which therefore requires a simultaneous measurement of the atom number N and the cloud temperature T. The temperature is determined from the expansion of the atom cloud as it is released from the trap 2. Ours is a cloud with a normal density distribution in the x-direction, described by

$$n_x \propto Exp\left[-\left(\frac{x}{\sigma_N(x)}\right)^2\right].$$ (23)

Here σ is the $1/e-$radius of the cloud. The temporal evolution of its spatial extent is given by

$$\sigma_N(x,t) = \sqrt{\sigma_N^2(x,0) + (\sigma(v_x) \times t)^2}.$$ (24)

The velocity distribution is related to the temperature with $\sigma^2(v_x) = 2k_B\overline{T}_x/m$. The same relationships hold for expansion along the other directions. However, because of the anisotropy of the trap, the initial cloud size as well as the velocity distribution will be different along each axis. The trap shape is near-identical along the x- and z-directions, but rather more elongated along the y-axis. To measure the expansion, we record four absorption images of the cloud at 4, 8, 12 and 16 ms after release from the trap. These two-dimensional images allow us to extract the expansion rate for the x- and $y-$ axes, and give us four measurements of the atom number for a particular set of GA parameters. The fitness of a given parameter set is determined by averaging over the measured temperatures in the x- and $y-$ axes, as well as averaging over the atom numbers from the four images.

The cooling ramp under consideration consists of two linear segments determined by three RF - frequencies and two times (see fig. 7 a). We present two measurements:

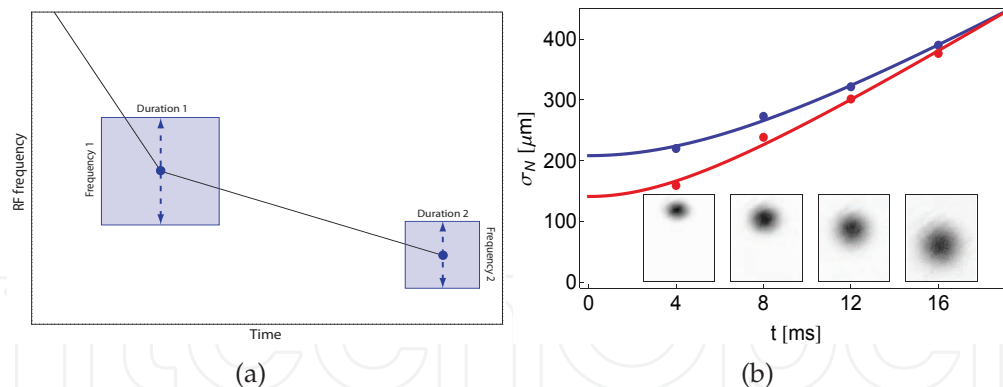(a)   (b)

Fig. 7. Example of data for one population member. a) Schematic of an RF - ramp, depicting the used optimization parameters. Arrows indicate the RF frequencies modified in the 2d - optimization run. In the 4d scan, the cooling ramp was optimized over the rectangular areas spanning duration and frequency of the two stages. b) Fits to equation (24) for the x- and $y-$ axes in red and blue, respectively. The inset panels show absorption images of the atom cloud as it falls and expands after release from the magnetic trap, at the corresponding times. Each panel shows a region of approximately $1.5\,\text{mm} \times 1.8\,\text{mm}$.

- A **2d** measurement, where the algorithm adjusts the intermediate and final value of the RF frequency. The values ranges from 0 to 10 and 0 to 1 MHz, respectively.

- A **4d** measurement, where the algorithm adjusts both field values from the 2d measurement as well as both times. RF values range from 2 to 6 and 0.3 to 0.6 MHz respectively, in steps of 0.01 MHz. The times for the first and second ramp segment can take values each from 500 to 2000 ms (duration 1) and 500 to 3500 ms (duration 2), in steps of 1 ms.

The results are summarized in figure 8. Unlike the molasses optimization described above, we have not compared this measurement to a grid scan. In the case of the 4d-optimization this is simply not feasible, given that the explored parameter space contains of over 10^{11} individual points corresponding to over 10000 years of experiment runtime. Even by reducing the time resolution to 10 ms as well and changing the frequency steps to 0.05 MHz, the measurement duration remains large, at roughly 50 years. Unless an optimization approach like ours is implemented, only physical arguments and a certain degree of parameter separability can be used to find useful working points for ramp parameters for this and similar problems, leading to a labor-intensive manual search.

For the 2d measurement, the algorithm finds values of 3 MHz and 0.65 MHz for the RF - ramp frequencies. Although we cannot make a quantitative statement about the quality of this solution without knowledge of the parameter space, these values are similar to the cooling ramp values which have been successfully used before GA - optimization as well as the corresponding phase space densities. Note that the algorithm has not fully converged for the intermediate RF value; a competing subpopulation with 2 MHz RF - value is still present in the last generation.

For the 4d measurement, the algorithm also finds an intermediate value of 3 MHz, but a lower end frequency of 0.47 MHz. Objective function values, especially those of the best performing states within the run, are only slightly inferior to those in the last generations of the 2d measurement. It is interesting to see however that performance increases towards the end of the run, at the same time when the competing subpopulation for short times of duration 1

(a) 2d optimization run. (b) 4d optimization run I. (c) 4d optimization run II.
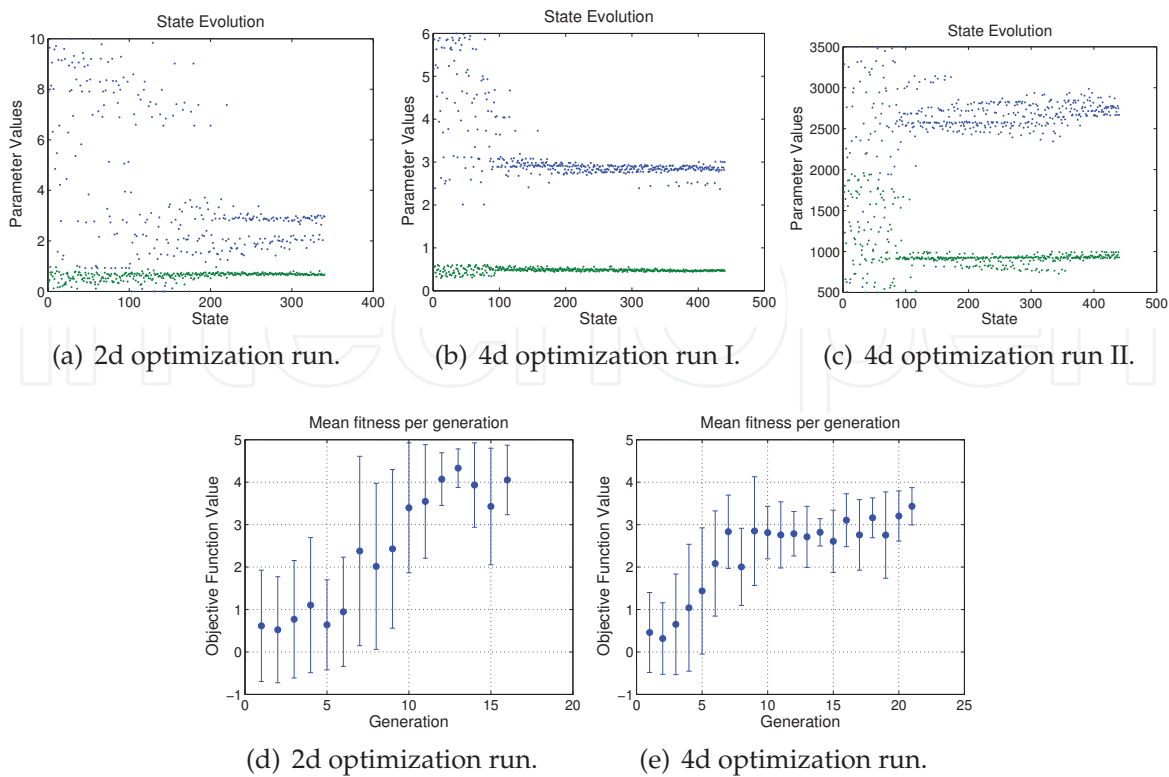
(d) 2d optimization run. (e) 4d optimization run.

Fig. 8. Results of the RF - cooling ramp optimization. The parameters in (a) and (b) are frequencies of the RF field during the ramp, given in MHz. In (c), parameters are times in milliseconds. The parameters represented in (b) and (c) belong to states consisting of two times and two frequencies, but are presented in distinct graphs due to the different unit scales. (e), (f). Mean fitness per generation for the 2d and 4d optimization runs.

(green points) vanishes and duration 2 (blue points) begins to develop a trend towards bigger values. A noteworthy point, however, is that overall ramp duration for these results with 3.7 seconds is significantly shorter than the preset 4.5 seconds, with only marginally worse phase space densities.

In summary the algorithm has found useful working points in both optimization runs, reproducing optimal values found with the help of other experiments in one case, and significantly shortening the cooling ramp with only a small tradeoff in phase space density in the second measurement.

5. Computer experiments

As stated above, algorithm runtime as part of the hardware feedback loop is on the order of a few hours due to the duration of an experimental sequence of 35 seconds. As a consequence, in order to characterize the algorithms performance and for parameter tuning, computer experiments on several test problems have been carried out, with runtimes of seconds to minutes and full knowledge of the parameter space.

The tests go from simple, unimodal problems in two dimensions to complicated, multidimensional functions commonly used as test functions for stochastic optimization problems. In each case, the algorithms task was to find the global maximum or minimum of the respective test function.

Representative for all optimization problems, figure 9 illustrates the algorithm's walk through the parameter space of the 2d Rastrigin function (see section 5.2) through a subset of stages out of a total of 48 generations.
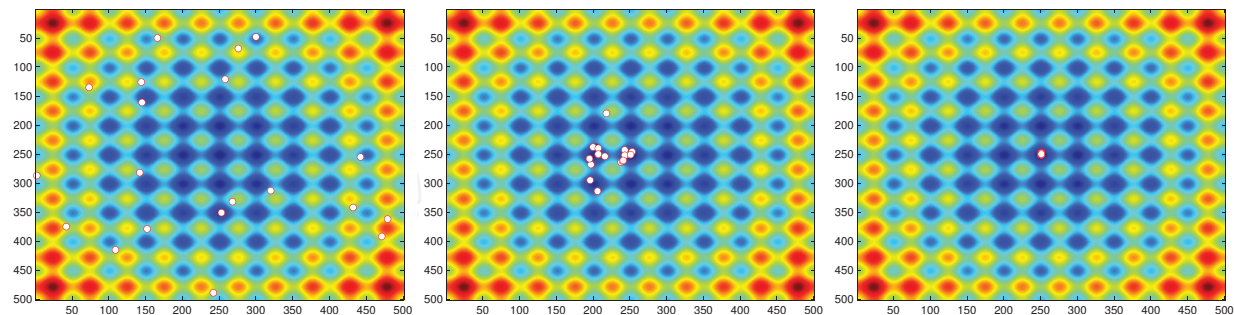

Fig. 9. Population in parameter space in generations 1, 32 and 48.

5.1 Unimodal test functions

A first, very simple function is

$$f(x_i) = \sum_{i=1}^{n} x_i^2 \tag{25}$$

which is known as *De Jong's first function* as benchmark for optimization algorithms. It is continuous, convex and unimodal. As second problem, we used a bivariate normal distribution, creating a peaked structure in an otherwise flat parameter space as depicted in figure 10. This problem is slightly more difficult since the algorithm can find an exploitable gradient only in the vicinity of the optimum.

(a) Peaked 2d distribution   (b) De Jong F1 - Function
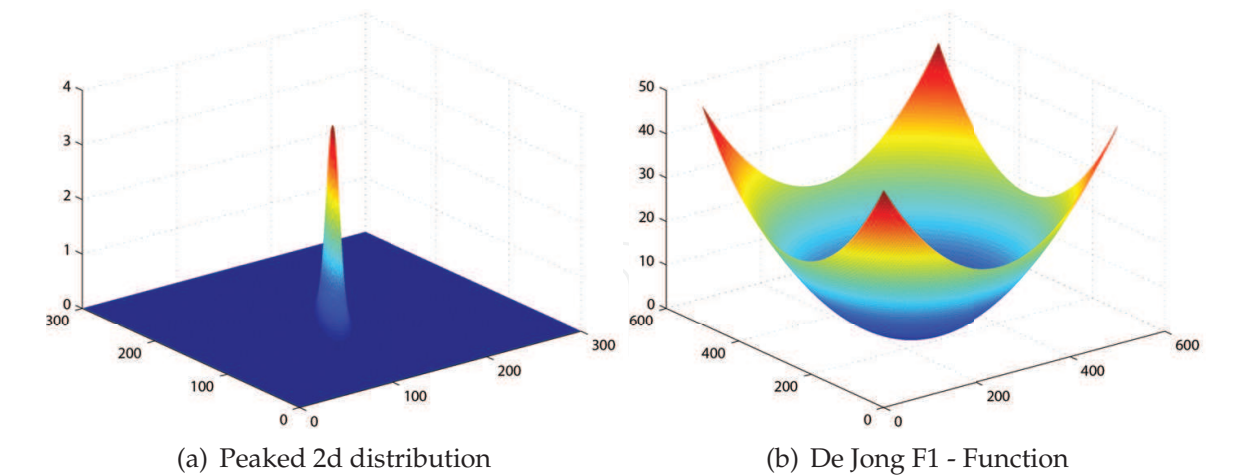
Fig. 10. 2d representations of the used unimodal test functions.

5.1.1 Results

In this benchmark, the algorithm was supposed to minimize De Jong's function in three and five dimensions as well as the bivariate Gaussian distribution. Each task was repeated 50 times in order to gain statistics about the convergence behavior.

As an example, the first two panels in figure 11 show the properties of one typical optimization run.

In the different fitness graphs, the value of the objective function is plotted as fitness measure. With exception of the bivariate normal distribution, all functions had to be minimized, thus smaller values correspond to better states.
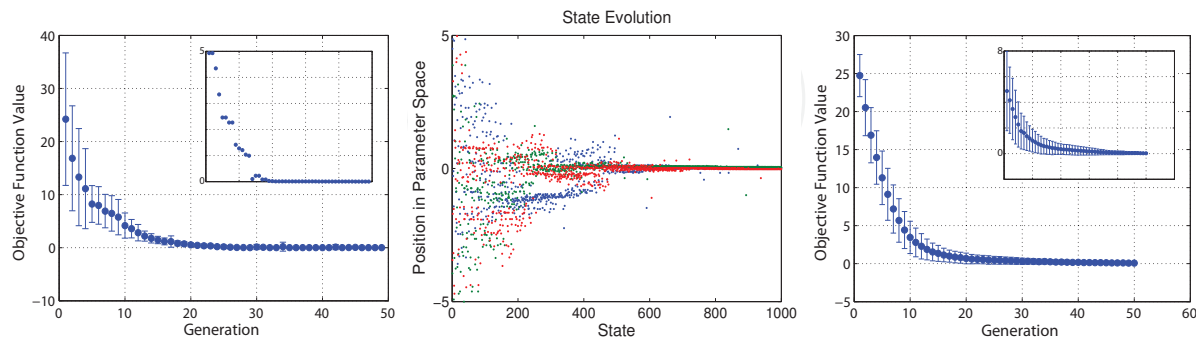

Fig. 11. Results for De Jong's function in three dimensions. Panel 1 contains the mean fitness of each generation within one typical optimization run. Error bars show the standard deviation, corresponding to the spread of different states within the generation, with the fitness of each generation's best state depicted in the inset. The evolution of states is given in Panel 3. Different colors correspond to different variables, i. e. components of the state vectors. Panel 3 contain the fitness of each generation averaged over the total number of 50 runs and the fitness of each generation's best state averaged over the total number of 50 runs in the inset. The spread in expected convergence time is related to the fitness variance expressed through the bars.

The results demonstrate the algorithm's ability to converge towards the global optimum located at the center of the parameter space on a scale of 20 to 30 generations, while finding the optimum takes 10 - 20 generations.

Figure 12 summarizes convergence behavior for the same function, but in 5 dimensions. For the data presented in figure 12 a), the population size was set to 20 states per generation, as opposed to 30 states per generation for b). The additional degrees of freedom cause bigger fitness spreads and hence slower convergence for both settings, although the effect is not as pronounced for the runs with larger population size. The results underline the tradeoff between population size and convergence time in terms of generations one is confronted with in large parameter spaces.

To conclude, the results on the bivariate Gaussian distribution are presented in figure 12 c). For this problem, the algorithm had to find a sharp maximum with an objective function value of 3.5. The peaked structure also represents itself in a fitness spread, since even states close to the minimum achieve a significantly lower objective function value than maximally possible.

5.2 Multimodal test functions
The first multimodal benchmark used is known as *Rastrigin's function*:

$$f(x_i) = 10 \cdot n + \sum_{i=1}^{n} \left[x_i^2 - 10 \cdot cos(2\pi x_i) \right]. \tag{26}$$

Both this form and the 2d representation given in figure 13 show that this is an overall convex function with a sinusoidal modulation, creating a large number of local optima.
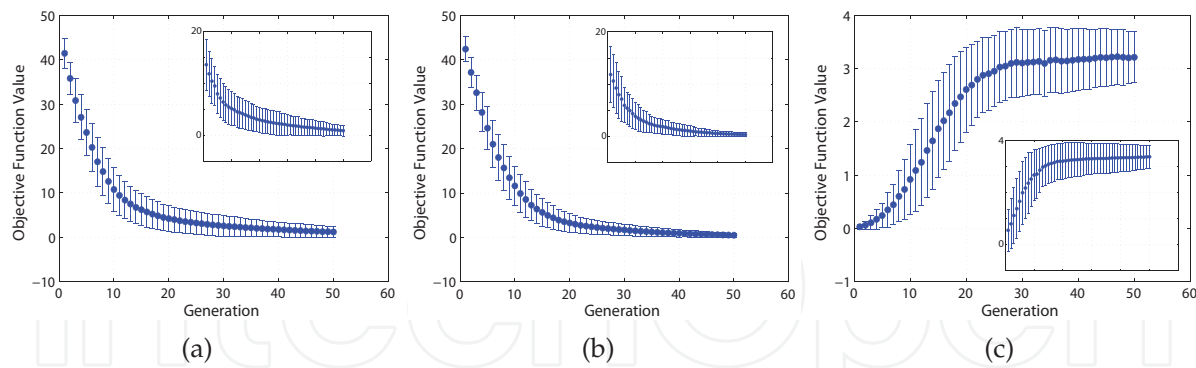
(a)   (b)   (c)

Fig. 12. Results for De Jong's function in five dimensions. Figures (a) and (b) represent results gained with a population size of 20 states and 30 states respectively. (c) Results for the optimization of a 2d bivariate Gaussian function.

As a second benchmark, we use the function

$$f(x_i) = \sum_{i=1}^{n} \left[-x_i \cdot sin\left(\sqrt{|x_i|} \right) \right],$$ (27)

also known as *Schwefel's function*. It does not feature as many local extrema as Rastrigin's function within the search space under consideration, however it is a *deceptive function* in that the global minimum is distant in parameter space from the next best local minima.

(a) Rastrigin's function   (b) Schwefel's F6 function
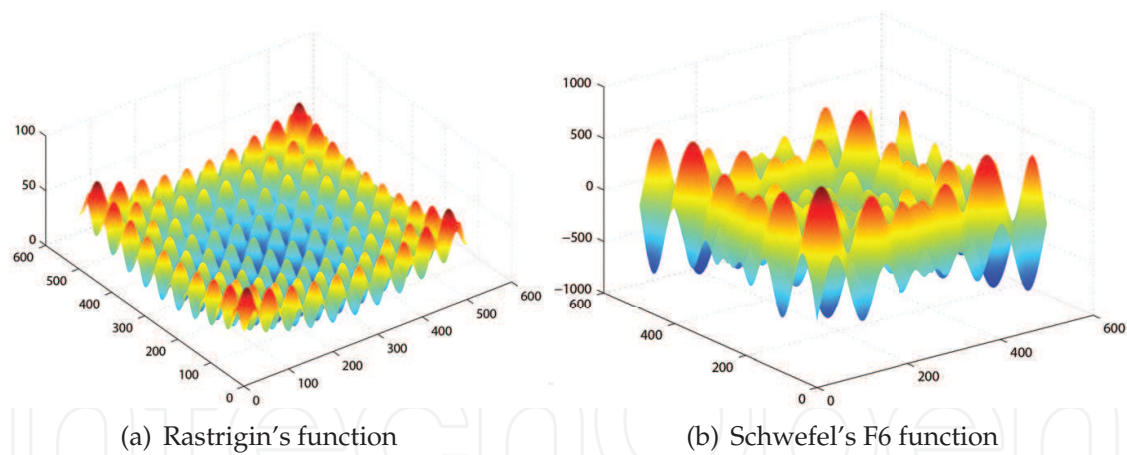
Fig. 13. 2d representations of the used multimodal test functions.

5.2.1 Results

Results on the 2d and 5d Rastrigin function are summarized in figure 14. While the algorithm can locate the global optimum of this complicated test problem in 2d, for the 5d problem with a population size of 20 states in most of the cases the optimization gets stuck in a local minimum. Raising population size to 50 states shows a clear improvement. Evolution of the fitness per generation as well as each generation's best state's fitness for these cases are shown in figure 14 a, b and c and the insets, respectively.

The last test problem presented here is optimization of Schwefel's function. The global minimum has a function value of $-418 \cdot d$, where d represents the number of dimensions.
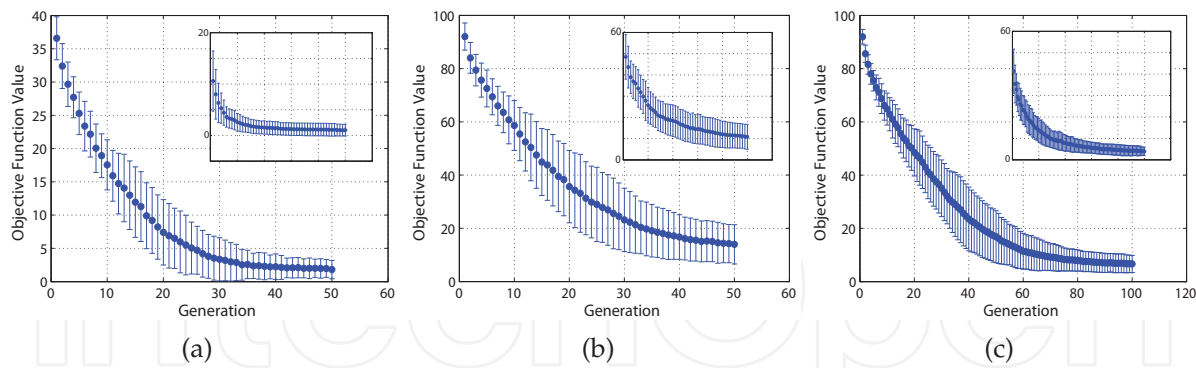
Fig. 14. Convergence behavior for Rastrigin's function (see text).

Similarly to the case of Rastrigin's function, a population of 20 states per generation still locates the global minimum in 60 % of runs for 2d (figure 15 a), while in five dimensions (figure 15 b), the algorithm usually converges towards a local optimum.
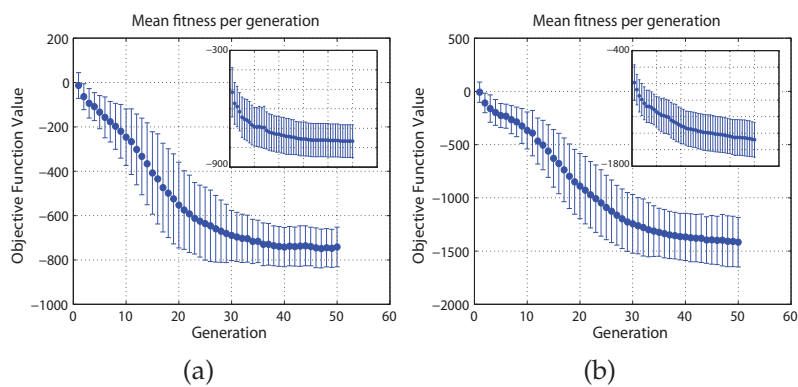

Fig. 15. Evolution of the fitness per generation as well as each generation's best state's fitness in 2d (a), and 5d (b).

In summary, the computer experiments suggest that for moderately complex optimization problems, our simple RCGA can perform optimization tasks even with comparatively small population sizes. In complicated problems with many local optima, bigger population sizes cannot be avoided.

6. Summary and outlook

In this chapter, we have described the implementation of shot-to-shot real-time stochastic optimization of a physics experiment. Our approach is broadly applicable and can be implemented for all computer-controlled parameters of any given physical apparatus. As with most implementations of GA optimization, the approach is particularly useful for multidimensional parameter spaces with multiple local optima and little or no quantitative predictions of their coordinates. This is a situation which arises often in atomic physics as well as in other branches of experimental research, and is usually tackled by making use of intuition, physical arguments and a certain degree of parameter separability to restrict the initial parameter space. This approach often amounts to a human-controlled, semi-stochastic search which is usually time-consuming and has no guarantee of yielding a global optimum.

Using the genetic-algorithm approach, we have seen rapid convergence to optimal parameters in 2- and 4-dimensional parameter spaces, and that the approach is robust even in the presence of local optima and experimental noise. We envisage a number of possible augmentations for future implementations. Among these are the weighting of the fitness of a population member according to its experimental uncertainty, and the inclusion of qualitative physical predictions in some implementations. These predictions can then be progressively quantified with each generation and used to steer mutations to speed up the search convergence. In future implementations, this method may also be extended to perform "optimal experimental control" that automatically finds the best experimental sequence to produce a defined target quantum state. In conclusion, optimization using a genetic algorithm can be an efficient tool to improve the performance of a 'real-life' apparatus.

7. References

(n.d.). see atom traps world wide: http://www.uibk.ac.at/exphys/ultracold/atomtraps.html and links therein.

Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. (1995). Observation of Bose-Einstein condensation in a dilute atomic vapor, *Science* 269: 198.

Assion, A., Baumert, T., Bergt, M., Brixner, T., Kiefer, B., Seyfried, V., Strehle, M. & Gerber, G. (1998). Control of Chemical Reactions by Feedback-Optimized Phase-Shaped Femtosecond Laser Pulses, *Science* 282(5390): 919–922.
URL: *http://www.sciencemag.org/cgi/content/abstract/282/5390/919*

Baker, J. E. (1985). Adaptive selection methods for genetic algorithms, *Proceedings of the 1st International Conference on Genetic Algorithms*, Lawrence Erlbaum Associates, Inc., Mahwah, NJ, USA, pp. 101–111.

Baumert, T., Brixner, T., Seyfried, V., Strehle, M. & Gerber, G. (1997). Femtosecond pulse shaping by an evolutionary algorithm with feedback, *Applied Physics B: Lasers and Optics* 65: 779–782. 10.1007/s003400050346.
URL: *http://dx.doi.org/10.1007/s003400050346*

Berman, P. R. (1996). *Atom Interferometry*, Academic Press, New York.

Bose, S. N. (1924). Plancks gesetz und lichtquantenhypothese tex, *Z. Phys.* 26: 178.

Bradley, C. C., Sackett, C. A., Tollett, J. J. & Hulet, R. G. (1995). Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions, *Phys. Rev. Lett.* 75: 1687.

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

Chu, S. (1998). The manipulation of neutral particles, *Rev. Mod. Phys.* 70: 685.

Cohen-Tannoudji, C. N. (1998). Manipulating atoms with photons, *Rev. Mod. Phys.* 70: 707.

Davis, K. B., Mewes, M.-O., Andrews, M. R., van Druten, N. J., Durfee, D. S., Kurn, D. M. & Ketterle, W. (1995). Bose einstein condensation in a gas of sodium atoms, *Phys. Rev. Lett.* 75: 3969.

Davis, K. B., Mewes, M.-O., Ioffe, M. A., Andrews, M. R. & Ketterle, W. (1995). Evaporative cooling of sodium atoms, *Phys. Rev. Lett.* 74: 5202.

Davis, K., Mewes, M.-O. & Ketterle, W. (1995). An analytical model for evaporative cooling of atoms, *Appl. Phys. B* 60: 155.

Einstein, A. (1925). Quantentheorie des einatomigen idealen gases. ii, *Sitzungsber. Preuss. Akad. Wiss.* Bericht 1: 3–14.

Folman, R., Krüger, P., Schmiedmayer, J., Denschlag, J. & Henkel, C. (2002). Microscopic atom optics: from wires to an atom chip, *Adv. At. Mol. Opt. Phys.* 48: 263–356.

Fortagh, J. & Zimmermann, C. (2007). Magnetic microtraps for ultracold atoms, *Rev. Mod. Phys.* 79: 235.

Grimm, R., Weidemüller, M. & Ovchinnikov, Y. B. (2000). Optical dipole traps for neutral atoms, *Adv. At. Mol. Opt. Phys.* 42: 95.

Hänsel, W., Hommelhoff, P., Hänsch, T. W. & Reichel, J. (2001). Bose-Einstein condensation on a microelectronic chip, *Nature* 413: 498.

Haupt, R. (2000). Optimum population size and mutation rate for a simple real genetic algorithm that optimizes array factors, *Antennas and Propagation Society International Symposium, 2000. IEEE* 2: 1034–1037 vol.2.

Heine, D., Rohringer, W., Fischer, D., Wilzbach, M., Raub, T., Loziczky, S., Liu, X., Groth, S., Hessmo, B. & Schmiedmayer, J. (2010). A single-atom detector integrated on an atom chip: fabrication, characterization and application, *New Journal of Physics* 12(9): 095005.
URL: *http://stacks.iop.org/1367-2630/12/i=9/a=095005*

Herrera, F., Lozano, M. & Verdegay, J. L. (1998). Tackling real-coded genetic algorithms operators and tools for behavioural analysis, *Artificial Intelligence Review* 12(4): 265–319.
URL: *citeseer.ist.psu.eduherrera98tackling.html*

Ketterle, W., Durfree, D. S. & Stamper-Kurn, D. M. (1998). Making, probing and understanding bose-einstein condensates, Contribution to the proceedings of the 1998 Enrico Fermi summer school on Bose-Einstein condensation in Varenna, Italy, Academic Press, p. 1. and references therein.

Ketterle, W. & Zwierlein, M. W. (2008). Making, probing and understanding ultracold fermi gases, Contribution to the proceedings of the 1998 Enrico Fermi summer school on ultracold Fermi gases condensation in Varenna, Italy, IOS Press, p. 1. and references therein.

Luiten, O. J., Reynolds, M. W. & Walraven, J. T. M. (1996). Kinetic theory of the evaporative cooling of a trapped gas, *Phys. Rev. A* 53: 381.

Metcalf, H. J. (1999). *Laser Cooling and Trapping*, Springer Verlag, Heidelberg Berlin New York.

Mühlenbein, H. & Schlierkamp-Voosen, D. (1993). Predictive models for the breeder genetic algorithm, i.: continuous parameter optimization, *Evol. Comput.* 1(1): 25–49.

Mühlenbein, H. & Schlierkamp-Voosen, D. (1995). Analysis of selection, mutation and recombination in genetic algorithms, *Evolution and Biocomputation, Computational Models of Evolution*, Springer-Verlag, London, UK, pp. 142–168.

Ott, H., Fortagh, J., Schlotterbeck, G., Grossmann, A. & Zimmermann, C. (2001). Bose-Einstein condensation in a surface microtrap, *Phys. Rev. Lett.* 87: 230401.

Phillips, W. D. (1998). Laser cooling and trapping of neutral atoms, *Rev. Mod. Phys.* 70: 721.

Pohlheim, H. (1999). *Evolutioni£¡re Algorithmen: Verfahren, Operatoren und Hinweise fi£¡r die Praxis*, Springer Verlag. ISBN 3-540-66413-0.

Reichl, L. E. (1998). *A Modern Course In Statistical Physics*, John Wiley & Sons.

Rohringer, W., Bücker, R., Manz, S., Betz, T., Koller, C., Göbel, M., Perrin, A., Schmiedmayer, J. & Schumm, T. (2008). Stochastic optimization of a cold atom experiment using a genetic algorithm, *Applied Physics Letters* 93(26): 264101.
URL: *http://link.aip.org/link/?APL/93/264101/1*

Rottmann, J. (2006). *Towards a Single Atom Camera*, Diploma thesis, University of Heidelberg.

Santos, F. P. D., Leonard, J., Wang, J., Barrelet, C., Perales, F., Rasel, E., Unnikrishnan, C., Leduc, M. & Cohen-Tannoudji, C. (2002). Production of a bose einstein condensate of metastable helium atoms, *Eur. Phys. J. D* 19: 103–109.

Schneider, S., Kasper, A., Hagen, C. V., Bartenstein, M., Engeser, B., Schumm, T., Bar-Joseph, I., Folman, R., Feenstra, L. & Schmiedmayer, J. (2003). Bose-Einstein condensation in a simple microtrap, *Phys. Rev. A* 67: 023612.

Weicker, K. (2002, 2.Auflage 2007). *Evolutionï£¡re Algorithmen*, B.G. Teubner Verlag / GWV Fachverlage GmbH.

Wilzbach, M., Heine, D., Groth, S., Liu, X., Raub, T., Hessmo, B. & Schmiedmayer, J. (2009). Simple integrated single-atom detector, *Opt. Lett.* 34(3): 259–261.
URL: *http://ol.osa.org/abstract.cfm?URI=ol-34-3-259*

Wing, W. H. (1984). On neutral particle trapping in quasistatic electromagnetic fields, *Prog. Quant. Electr.* 8: 181.

Wolpert, D. H. & Macready, W. G. (1995). No free lunch theorems for search, *Technical report*, Santa Fe Institute.

Wolpert, D. H. & Macready, W. G. (1997). No free lunch theorems for optimization, *IEEE Transactions on Evolutionary Computation* 1: 67.
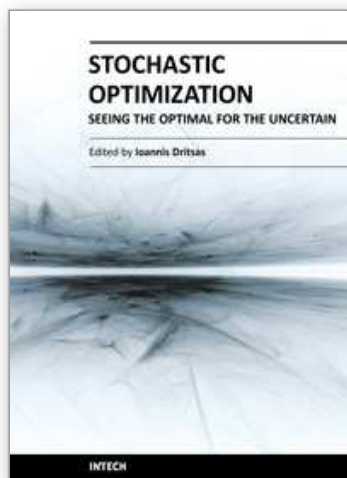
Stochastic Optimization - Seeing the Optimal for the Uncertain

Edited by Dr. Ioannis Dritsas

Stochastic Optimization Algorithms have become essential tools in solving a wide range of difficult and critical optimization problems. Such methods are able to find the optimum solution of a problem with uncertain elements or to algorithmically incorporate uncertainty to solve a deterministic problem. They even succeed in â€œfighting uncertainty with uncertaintyâ€. This book discusses theoretical aspects of many such algorithms and covers their application in various scientific fields.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

W. Rohringer, D. Fischer, M. Trupke, J. Schmiedmayer and T. Schumm (2011). Stochastic Optimization of Bose-Einstein Condensation Using a Genetic Algorithm, Stochastic Optimization - Seeing the Optimal for the Uncertain, Dr. Ioannis Dritsas (Ed.), ISBN: 978-953-307-829-8, InTech, Available from: http://www.intechopen.com/books/stochastic-optimization-seeing-the-optimal-for-the-uncertain/stochastic-optimization-of-bose-einstein-condensation-using-a-genetic-algorithm

INTECH
open science | open minds