Probing fluctuations in a lattice of mesoscopic atomic ensembles

Caspar F. Ockeloen

van der Waals-Zeeman Institute, University of Amsterdam
August 26, 2010

Master’s thesis in Physics
Supervisor
Dr. R. J. C. Spreeuw

Daily supervisor
Dr. S. Whitlock

University of Amsterdam
Abstract

We experimentally probe intrinsic atom number fluctuations in trapped ensembles of ultracold atoms on a magnetic lattice atom chip. Advanced post-processing and optimal analysis techniques allow the detection of intrinsic fluctuations in the number of atoms in small atomic ensembles with absorption imaging. A fringe removal algorithm reduces imaging noise to the fundamental photon shot noise limit and proves beneficial even in the absence of fringes. Maximum-likelihood estimation is used to optimally detect the number of atoms in a single realisation of the experiment. Combined, these techniques offer a factor 3 improvement in signal to noise, to a minimum resolvable population difference of 17 atoms/shot in a double well experiment. In a lattice of tightly confining magnetic traps, atoms undergo density dependent three-body loss. This loss process naturally reduces the shot-to-shot atom number fluctuations to below the Poisson level. We directly observe sub-Poissonian atom number fluctuations in our lattice, and find a Fano factor of $F = 0.53 \pm 0.22$ for ensembles comprising between 50 and 300 atoms each, in very good agreement with an analytic theory, which predicts $F = 3/5$. Our results provide the basis for improved readout of atom interferometers and for the study of quantum information science with mesoscopic ensembles.
Contents

Introduction

1 Magnetic lattice atom chip
   1.1 Magnetic lattice atom chip ................................................................. 9
   1.2 Experimental procedure ............................................................................ 10
   1.3 Manipulation of atoms in the lattice ....................................................... 11
   1.4 Internal state coherence ............................................................................ 12
   1.5 Quantum information processing with mesoscopic qubits .................... 14

2 Improved atom number detection in absorption imaging ...................... 15
   2.1 Absorption imaging .................................................................................. 15
   2.2 Fringe removal .......................................................................................... 16
   2.3 Optimal atom number estimation ............................................................ 20
   2.4 General detection limit for absorption imaging ....................................... 22
   2.5 Atom number fluctuations in double-well systems .................................. 23
   2.6 Experimental setup .................................................................................. 24
   2.7 Results ....................................................................................................... 27
   2.8 Conclusion ................................................................................................. 29

3 Sub-Poissonian atom number fluctuations by three-body loss ............... 31
   3.1 Three-body recombination ....................................................................... 31
   3.2 Experimental setup .................................................................................. 34
   3.3 Single trap analysis .................................................................................. 35
   3.4 Fluctuation correlation analysis ............................................................... 37
   3.5 Conclusion ................................................................................................. 39

Conclusion ......................................................................................................... 41

Appendices ......................................................................................................... 42
   A Propagation of variances ......................................................................... 43
   B Fringe removal implementation ................................................................. 45
   C Fluctuation-correlation analysis ............................................................... 47

Bibliography ....................................................................................................... 49

Nederlandse Samenvatting .............................................................................. 53
Introduction

Ultracold neutral atoms are a promising starting point for quantum metrology and quantum information science [123]. Due to their intrinsically weak interaction with the environment, trapped ensembles of neutral atoms can be largely robust to decoherence effects. In cold atom experiments dilute clouds of atoms, containing up to millions and down to only a few atoms, can be isolated, controlled and detected, opening a way to the study of many-body entanglement in a controllable environment.

A high level of control over the atomic motion can be achieved on atom chips, microstructured devices suspended in vacuum, where atoms can be trapped and manipulated near the chip surface using various structures on the chip [4]. Magnetic lattice atom chips, incorporating a patterned array of permanently magnetic material, enable arrays of hundreds of microtraps, with almost arbitrary lattice geometry and sizes [56]. Such a system provides a promising host structure for quantum information science, which could readily be scaled up to incorporate thousands of microtraps on a single device. Quantum bits (qubits) could for example be encoded in the internal state of single atoms, each trapped in a single microtrap, providing extremely long storage times.

Alternatively, arrays of small ensembles, containing only a few to a few hundred atoms each, could be used to create mesoscopic qubits. Here, collective excitation to high-lying Rydberg states would create multi-particle entangled qubit states, and provide a basis for two-qubit operations based on their long range interaction properties [78910]. As the excitation rate to such states depends on the number of particles in each trap, high-fidelity quantum operations would require well-known and well-defined numbers of atoms in each trap. Therefore, the study and control of atom number fluctuations in small ensembles is an important step towards quantum information science with mesoscopic qubits.

In this thesis, we study the atom number fluctuations in a two-dimensional lattice of hundreds of mesoscopic ensembles, created on a magnetic lattice atom chip. We employ advanced post-processing and optimal analysis of laser-illuminated absorption images, to optimise the sensitivity for detecting the atom number in a single realisation of the experiment. This allows us to detect intrinsic atom number fluctuations in small ensembles of only a few tens of atoms. We then make use of density-dependent atom loss due to three-body recombination, which naturally reduces the atom number fluctuations. In our experiment, we directly measure fluctuations below the standard atom shot noise or Poisson limit. Making use of our lattice geometry, we employ fluctuation correlation analysis to detect sub-Poissonian atom number fluctuations for a wide range of mean atom numbers.

The optimal analysis techniques presented in this thesis improve standard absorption imaging for the detection of small atom numbers, which we demonstrate by applying our techniques to the study of intrinsic atom number fluctuations in a double-well potential. This study impacts a wide variety of experiments, including atom interferometry [111213], atomic clocks [14151617], Josephson physics [181920] and experiments on number number/spin-squeezing and entanglement in atomic ensembles [2122232425]. Our analysis provides a basis to improve the readout of trapped
atom interferometers to the quantum limit and to better resolve number squeezing and entanglement with small atomic ensembles.

This thesis is laid out as follows. In chapter 1 our magnetic lattice atom chip experiment is introduced. We demonstrate site-resolved absorption imaging, enabling access to the full two-dimensional lattice distribution, and measurement of internal state coherence in a single magnetic trap. In chapter 2 we demonstrate improved atom number detection techniques using advanced post-processing to reduce noise due to optical fringes and the effect of photon shot noise, without affecting the atomic distribution. We derive the Cramér-Rao lower bound for atom number detection in our experiment, and the maximum likelihood estimator that attains the lower bound. We experimentally test our techniques in a double-well potential on our atom chip. In chapter 3 we study reduced atom number fluctuations due to three-body loss in our magnetic lattice of traps. We derive a stochastic model for the evolution of fluctuations due to multi-particle loss, and use both direct detection and fluctuation correlation analysis to measure sub-Poissonian atom number fluctuations in our lattice.
1 Magnetic lattice atom chip

Ensembles of neutral atoms trapped in micro-potentials are a promising starting point for quantum metrology and quantum information science. A fundamental difficulty in such technologies is posed by decoherence of the quantum states used as qubit or interferometer states. Ground state neutral atoms intrinsically exhibit weak interactions with the environment, enabling long coherence times of internal spin states, up to seconds or even one minute [14, 15].

While internal state coherence could be employed for quantum information storage, interactions between (ensembles of) atoms are required for entanglement-based technologies such as quantum metrology and information processing. Potential interactions for this include collisional interactions in a Bose-Einstein condensate, recently demonstrated in number/spin-squeezing experiments [21, 22, 23] and long-range electric dipole-dipole interactions of Rydberg states, further discussed later in this chapter.

1.1 Magnetic lattice atom chip

Trapped ensembles of ultra-cold atoms can be experimentally realised through laser cooling and optically or magnetically trapping of atoms [26]. For the purpose of quantum information science, it is desirable to prepare multiple ensembles, for example in multi-well or periodic lattice potentials, to facilitate scaling to many qubits.

A common technique for creating lattices of atomic clouds consists of a magnetic or optical trap modulated by a standing wave optical lattice, splitting the trap into wells, typically separated by half the wavelength of the lattice light. This way, lattices can be created in multiple dimensions, but optically resolving and addressing single lattice sites remains difficult.

Magnetic traps on atom chips provide an alternative route to trapping multiple ensembles. An atom chip is a micro-fabricated device, where the atoms are trapped and manipulated close to the surface by various structures that can be implemented on the chip, such as current-carrying wires, structured magnetic material and even optical fibers. In particular, large arrays of traps can be created with a layer of patterned magnetic material. These magnetic lattice atom chips provide tremendous design flexibility, with almost arbitrary lattice geometries and lattice spacings, ranging in principle from optically resolvable down to sub-μm scale [6]. Furthermore, trap frequencies can be varied independent from the lattice spacing, enabling tight localisation of the atoms compared to the inter-site distance.

We employ in our experiments a magnetic lattice atom chip [27, 5], shown schematically in figure 1.1. Here, a patterned layer of permanently magnetic FePt-film creates, in conjunction with a homogeneous external bias field, a vast two-dimensional lattice of magnetic potentials which we load with ultra-cold $^{87}$Rb atoms. The individual magnetic traps are separated by the lattice spacings of 22 μm and 36 μm in two directions, large enough to enable individual ensembles to be optically resolved and addressed. In total we load $\sim 500$ traps of our lattice, each typically $\approx 10 \text{ μm}$ away from the surface.
1.2 Experimental procedure

In a typical realisation of the experiment, we collect and laser-cool a cloud of atoms in in a mirror-magneto-optical trap. The atoms are then transferred to the $|F = 2, m_F = +2\rangle$ hyperfine ground state and loaded into a single magnetic trap, formed by the field of a Z-shaped current-carrying wire, located beneath the atom chip surface, and an external bias field. To load the magnetic lattice, this trap can be merged with the lattice traps by ramping off the Z-wire current. We can also perform experiments directly in the Z-wire trap, positioning the trap bottom with external magnetic fields. Both in the Z-wire trap and the lattice micro-traps, we apply radio frequency evaporative cooling to create cold thermal ensembles of atoms.

We detect the atoms using a sensitive absorption imaging technique, detailed in chapter 2. To access the full two-dimensional distribution in the magnetic lattice, it was necessary to develop a reflection imaging setup, shown in figure 1.2. Probe light, circularly polarised by a quarter wave plate, reflects perpendicularly off the chip surface, passing the atoms twice. After passing the quarter wave plate again, the remaining light
Figure 1.3: Absorption image of the loaded lattice, averaged over 28 realisations, showing $\sim 500$ traps loaded with 200-2500 atoms each.

is linearly polarised perpendicular to the incoming probe beam, allowing us to separate the two with a polarising beam splitter. The shadow cast by the atoms is then imaged onto a charge coupled device (CCD) camera.

Figure 1.3 shows an averaged absorption image of the loaded magnetic lattice. Here, $\sim 500$ traps are loaded with between 200 and 2500 atoms each. Individual traps are resolved, with an optical resolution of $\approx 7.5 \mu m$ (Rayleigh criterion), enabling individual detection of the number of atoms in each ensemble.

1.3 Manipulation of atoms in the lattice

In addition to readout of individual lattice site, we have implemented techniques to manipulate the atoms in the lattice. We incorporate optical addressing of individual lattice sites, opening a route to the manipulation of the quantum state of each ensemble independently. We are able to manipulate atoms in a single microtrap without affecting neighbouring lattice sites, clearly demonstrating that individual sites are optically resolved. Here we benefit from our lattice spacing compared to optical lattices, where atoms are typically trapped at sub-micrometre spacings. At these lattice spacings, single site resolution has recently been achieved but remains challenging [28, 29, 30]. We furthermore manipulate the lattice as a whole using an atomic shift register, where the entire lattice is shifted back and forth in one direction.

For addressing individual atomic ensembles we have incorporated in the imaging optics a second light path. By placing a single-mode optical fiber in a second imaging plane, created with a polarising beam splitter cube, we image the output facet of the fiber onto a single atomic ensemble, near the centre of the magnetic lattice (see fig. 1.2). Here, we estimate an Airy disk radius of $\sim 7 \mu m$, small compared to our lattice spacing to avoid influencing neighbouring traps, but large enough to provide a highly homogeneous intensity over the extent of the atomic cloud. The addressing light is reflected off the chip surface and partially imaged on the CCD sensor, allowing for alignment of the focal point. The optical addressing setup, including additional optics compensating for aberrations, is detailed in [31].

To demonstrate the addressability of single sites, we apply a pulse of optical pumping light during 1 ms, resonant with the $F = 2 \to F' = 2$ transition of the D2 line, resulting in a high probability of changing the spin of the atoms to an untrapped state and thus removing the atoms from the trap. Figures 1.4a,b show absorption images before and after removing a lattice site. We observe no loss or heating in the neighbouring ensembles, even for an addressing pulse of 10 ms. For quantum information processing on our atom chip, this single site addressing system could be used to manipulate the
internal state of individual ensembles, for example using a two-photon Raman transition between hyperfine ground state levels or coherent excitation to Rydberg states.

We employ an atomic shift register, detailed in [5], to shift the entire lattice in the direction corresponding to a lattice spacing of 36 μm, analogous to the electronic shift registers found in CCDs. Rotating the external bias fields in three dimensions moves the position of the field minima with respect to the magnetic pattern, without creating field zeros that would lead to loss of atoms. Figure 1.4 shows shifts of 2 lattice periods, using a site emptied by the addressing laser as a marker. During the shifting, the empty sites do not refill, demonstrating there is no interaction between neighbouring sites. Similar shift registers have been reported in [32, 33].

1.4 Internal state coherence

In order to store and manipulate quantum information in our magnetic lattice, internal electronic states could be used as ‘qubit states’. To enable long storage times, the hyperfine levels $|0\rangle = |F = 1, m_F = -1\rangle$ and $|1\rangle = |F = 2, m_F = 1\rangle$ of the $^{87}$Rb ground state are particularly promising. These states undergo identical linear Zeeman shift at a so-called ‘magic’ magnetic field of $\sim 3.23$ Gauss. As a result, the relative energy difference between the two states is nearly constant around the magic field, strongly reducing dephasing of atoms prepared in a superposition of the qubit states.

We investigate coherent manipulation of the qubit states on our atom chip with a two-photon microwave and radio-frequency (rf) transition. For this, we prepare a single magnetic Z-wire trap loaded with cold thermal atoms in state $|0\rangle$. We apply 1 MHz rf and microwaves around 6.8 GHz for a variable pulse time $\tau_p$. Both frequency sources are phase locked to a $^{87}$Rb standard, providing phase coherence between them and long term stability of the experiment. The rf is transmitted by a coil situated beneath the chip surface. The microwave signal is amplified up to several Watt and transmitted to the atoms through a sawed-off wave guide antenna.
Figure 1.5 shows Rabi oscillations for the two-photon transition. All atoms are prepared in state $|0\rangle$, before applying the pulse. Then, the population in each state is recorded through absorption imaging in separate runs of the experiment, and the experiment is repeated for varying pulse duration. The resulting Rabi oscillations have almost full contrast over the first cycles. For this data, we measure a $\pi/2$-pulse duration of $\approx 3$ ms, whereas the contrast decays during the pulse to $1/e$ in $\approx 60$ ms. The Rabi oscillations start $\sim 1$ ms after the pulse is switched on, due to slow switch on of the microwave pulse. This delay was fixed in further experiments.

**Frequency domain Ramsey spectroscopy**

To estimate the coherence time of superpositions of the qubit states, we perform frequency domain Ramsey spectroscopy. Here, $N$ atoms are prepared in state $|0\rangle$, and two microwave plus rf pulses with a duration of $\tau_p$ each are applied, with a hold time $\tau_R$ between them (see inset of figure 1.6). Here, $\tau_p$ is chosen to correspond to a $\pi/2$-pulse on resonance. When scanning the microwave frequency around the two-photon resonance, Ramsey fringes can be observed in the population $N_1$ of state $|1\rangle$, which can be described as

$$N_1 = \frac{1}{2} N (1 + c \cos(2\pi\tau_R\delta)) \sin^2(\delta/w)$$

(1.1)

Here, $\delta$ the microwave detuning with respect to the two-photon resonance and $w$ the envelope width depending on $\tau_p$. The contrast $c = \exp(-\tau_R/\tau_c)$ describes any decoherence and/or dephasing between the two pulses, with a combined coherence time $\tau_c$.

Figure 1.6 shows the population in state $|1\rangle$ for a Ramsey time of $\tau_R = 25$ ms. We fit equation (1.1) to the data, yielding a contrast of $c = 0.79 \pm 0.03$, resulting in a coherence time of $\tau_c = 110 \pm 20$ ms, two orders of magnitude larger than our typical pulse time for $\pi/2$-rotations. This provides a viable starting point for quantum information with neutral atoms. We expect the coherence time could be improved by further optimisation of the experiment, for example by compensating for spatially inhomogeneous energy shifts. In a similar setup, coherence times exceeding one second have been demonstrated [14].
1.5 Quantum information processing with mesoscopic qubits

The qubit states defined in the previous section provide, in principle, a basis for the storage of information in the magnetic lattice, which could be encoded in the superpositions of the qubit states in each micro-trap. One way to encode a qubit in each microtrap, would be to reduce the density in the lattice to exactly one atom per site. However, the single-atom-per-site regime is technically difficult to achieve. Furthermore, loss of a single atom due to collision with background gas would directly lead to the loss of a qubit, drastically reducing the coherence time of a quantum computer.

A promising alternative route towards quantum information processing employs mesoscopic ensembles in each trap, consisting of only a few tens of atoms. Here, qubits could be encoded in recently observed collective excitations produced via Rydberg dipole blockade [7, 8, 9, 10, 34, 35, 36]. In the dipole blockade effect, the strong electric dipole moment of an atom excited to a Rydberg state shifts the energy levels of surrounding atoms, prohibiting the excitation of more than one atom to the Rydberg state.

A simple encoding scheme could be as follows. An ensemble is prepared with each atom in state $|0\rangle$, where we denote the state of the ensemble as $|\tilde{0}\rangle = |000\ldots\rangle$. A Rabi $\pi/2$-pulse to a Rydberg level $|R\rangle$ is applied. Due to the dipole blockade effect, only one excitation to $|R\rangle$ is possible, shared among all atoms in the ensemble: $|\tilde{R}\rangle = (|R00\ldots\rangle + |0R0\ldots\rangle + \ldots)/\sqrt{N}$. Such a state could provide robustness against loss of particles [37]. For robustness against interaction with the environment, a second pulse could be used to transfer $|R\rangle$ back to the other ground state level $|1\rangle$, creating the second qubit state $|\tilde{1}\rangle$.

Depending on the chosen spacing from site to site, the dipole blockade effect may extend to neighbouring lattice sites. If so, the scheme can be readily extended to a CNOT type operation between two neighbouring ensembles, providing a basis for quantum information processing in the lattice.

In collective excitation schemes, an ensemble containing $N$ atoms would benefit from a $\sqrt{N}$ enhancement in Rabi frequency, increasing the number of operations achievable within the coherence time. However, the $N$-dependence of the Rabi frequency adversely affects the fidelity of operations, if the number of atoms fluctuates from shot to shot and/or from site to site. Intrinsic atom number fluctuations would typically lead to a Poisson distribution at best, with fluctuations particularly significant for small ensembles. Suppressed atom number fluctuations would provide robustness against these errors, especially if combined with composite pulse techniques that reduce the effect of fluctuations on fidelity [8, 38]. Thus, quantum information science with collective excitations require small and well-defined atom numbers in the lattice.
2 Improved atom number detection in absorption imaging

Essential to the study of atom number fluctuations, is the ability to accurately detect the number of atoms in small ensembles in a single realisation of the experiment. A standard technique for the detection of trapped atoms is laser-illuminated absorption imaging, but detection sensitivity is limited by intrinsic photon shot noise and technical noise sources such as interference fringes in the images.

In this chapter, we demonstrate improved detection of trapped ensembles of ultra-cold atoms using advanced post-processing and optimal analysis techniques, significantly lowering the detection limit of standard absorption imaging and allowing us to better measure the intrinsic atom number fluctuations in atomic clouds. First, we employ a fringe removal algorithm that strongly suppresses interference fringes and at the same time significantly reduces the effect of photon shot noise in the images. Then, we establish the ultimate limit for measuring the populations based on the Cramér-Rao bound and derive maximum-likelihood estimators to attain that limit [39].

The relevance of these optimal analysis techniques extends beyond the study of fluctuations in our magnetic lattice atom chip. In particular, the precise measurement of the relative population difference in a double-well potential is of interest to new technologies such as atom interferometry [11, 12, 13], atomic clocks [14, 15, 16, 17] and Josephson physics [18, 19, 20]. Recent experiments on number number/spin-squeezing and entanglement in atomic ensembles exemplify the importance of accurate atom number detection in double-well systems [21, 22, 23, 24, 25].

Our analysis provides a basis to improve the readout of trapped atom interferometers to the quantum limit and to better resolve number squeezing and entanglement with small atomic ensembles. To demonstrate this, we apply our techniques to the measurement of intrinsic atom number fluctuations in a double-well potential, created at the edge of our magnetic lattice. The results of this chapter have also been published in reference [40].

2.1 Absorption imaging

In absorption imaging, atoms are briefly exposed to nearly homogeneous probe laser light, typically tuned to resonance with a cycling atomic transition. The atoms absorb part of the probe light, and the resulting absorption profile is imaged onto a CCD camera, creating an absorption image $A$. After removing the atoms, the local probe intensity is recorded in a reference image $R$. To remove any stray light and CCD dark counts, a third frame may be recorded with no probe light on and subtracted from both $A$ and $R$. From these, the atomic density image can be calculated, as shown in the following paragraphs.
**Atomic density**

Consider a thin section of the atomic distribution with thickness $dz$ and homogeneous three-dimensional density $n_{3D}$. When the atoms are illuminated by a resonant probe laser along the $z$-direction, the light intensity $I(z)$ satisfies Beer’s law, including the effect of saturation \[26\],

$$\frac{\partial I(z)}{\partial z} = -n_{3D} \sigma_0 \frac{1}{\alpha + I(z)}/(\alpha I_0) I(z).$$

(2.1)

Here, $\sigma_0$ is the resonant atomic cross-section, $I_0^s$ the saturation intensity ($1.67\text{mW/cm}^2$ for our transition) and $\alpha$ a dimensionless parameter, depending on the polarisation of the probe light and the quantisation axis. We solve eq. \(2.1\) for $I(z)$ between $I(z)$ and $I(z + dz)$, and solve the result for $n_{3D}$. Integrating along $z$, we find the atomic column density $n = \int n_{3D} dz$, in agreement with reference \[41\],

$$n = -\frac{\alpha}{\sigma_0} \log \left( \frac{I_A}{I_R} \right) + \frac{\alpha}{\sigma_0} s \left( 1 - \frac{I_A}{I_R} \right),$$

(2.2)

where $I_R$ is the probe intensity, $I_A$ the intensity after the cloud, and $s = I_R/\alpha I_0^s$ the effective saturation of the atomic transition, measured at the position of the atoms. To calculate the density image $n$ from the absorption and reference images, equation \(2.2\) is applied pixel by pixel, using $I_A/I_R = A/R$ and a spatially dependent saturation $s$ calculated from $R$ and the imaging and camera parameters.

The measured $A$ and $R$ are, apart from possible technical noise sources, corrupted by intrinsic photon shot noise for which $\text{var}(A) = A/g$ and $\text{var}(R) = R/g$, with $g$ the camera gain in counts per electron detected on the CCD. Using equation \(A.3\), and assuming small absorption, the expected photon shot noise in $n$ is

$$\sigma_n = \frac{\alpha \Delta}{\sigma_0} (1 + s) \sqrt{\frac{2}{gR}},$$

(2.3)

measured in atoms/pixel. Here $\Delta$ is the pixel area and the factor $2$ arises from summing the photon shot noise in $A$ and $R$.

For small absorption, the $z$-dependence of the saturation correction in equation \(2.1\) can be neglected. In that case, the measured atomic density can be approximated as $n \approx \alpha (1 + s) O/\sigma_0$, with $O$ the (widely used) optical density, defined as

$$O = -\log(A/R).$$

(2.4)

In the experiments described in this chapter, we do not use this approximation.

### 2.2 Fringe removal

In addition to photon shot noise, absorption images are in practice often troubled by fringes, due to diffraction and interference of the probe beam by optical elements in the imaging path. In principle these fringes would be divided out in calculating $n$; however, small vibrations of optical elements give rise to fluctuating fringe patterns between absorption and reference image, resulting in imperfect normalisation and thus fringe noise in $n$.

We employ a fringe removal algorithm that greatly reduces fringe noise by constructing an optimal reference image for each absorption image \[42, 43\]. In contrast to filtering techniques such as Fourier filtering, the absorption image itself is not altered, and thus no assumptions are made about the atomic distribution. Furthermore, the technique
reduces imaging noise even when no fringes are present, by strongly suppressing the photon shot noise in the reference image.

Our algorithm is similar to modern facial recognition algorithms [44], and closely related to a technique recently used in astronomical image analysis for detecting extrasolar planets [45].

**Fringe removal algorithm**

A typical data set consists of many absorption images $A_x$ and reference images $R_x$, where $x$ indexes pixels in the image. The algorithm aims to find for each absorption image a corresponding optimal reference image $Q_x$ as a linear combination of all reference images in the set, $Q_x = \sum_k c_k R_k$, where $k$ indexes realisations of the experiment. To find the optimal values for $c_k$, we minimise the least squares difference $\delta = \sum_x m_x (A_x - Q_x)^2$. Here, the mask $m_x$ selects a background region ($m_x = 1$), excluding the signal region ($m_x = 0$). Setting the derivatives with respect to each $c_j$ to zero,

$$\frac{\partial \delta}{\partial c_j} = 2 \sum_x m_x R_{x,j} \left( A_x - \sum_k c_k R_{x,k} \right) = 0.$$  (2.5)

Rearranging the terms, we obtain the system of linear equations

$$\sum_k c_k B_{j,k} = \sum_x m_x R_{x,j} A_x,$$  (2.6)

with the square matrix $B$ defined by

$$B_{j,k} = \sum_x m_x R_{x,j} R_{x,k}.$$  (2.7)

Equation (2.6) can be readily solved for $c_k$.

As a typical data set consists of hundreds of images, it is useful to solve equation (2.6) efficiently. Therefore, we decompose $B$ once using LU decomposition or singular value decomposition (SVD) and substitute the result for each $c_k$. MATLAB code for our implementation is included in appendix B. The algorithm is sufficiently fast to decompose $B$ and process new absorption images in $\lesssim 1$ s, allowing for live processing between experimental cycles.

The optimal reference image $Q_x$ is a weighted average of many reference images, which can strongly reduce photon shot noise in $Q_x$ compared to a single reference image. Noise in the resulting density image $n$ is therefore reduced even in the absence of fringes. For a sufficiently large dataset where many reference images are contributing to each $Q_x$, this results in a new, ultimate photon shot noise limit for $n$,

$$\sigma_n = \frac{\alpha \Delta}{\sigma_0} (1 + s) \sqrt{\frac{\gamma}{g R}},$$  (2.8)

a factor $\sqrt{2}$ lower compared to standard absorption imaging.

**Experimental results**

To benchmark the fringe removal algorithm, we apply our algorithm to a set of images recorded with our lattice imaging setup, shown in figure 1.2. Here, the probe beam is aligned perpendicular to the atom chip, and reflects off the lattice structure. The height profile of the patterned magnetic structure gives rise to additional fringes, with the same periodicity as the lattice of atoms itself. Removing these fringes with fourier
Figure 2.1: Optical density image of the magnetic lattice without (left) and with (right) fringe removal. The areas enclosed in red dashed lines are used as the fringe removal background region. The areas in blue boxes are used to characterise the residual background noise. In these regions, the variance from pixel to pixel was reduced by a factor 5.

Figure 2.2: Average fourier transform amplitude in a background region (blue boxes in fig. 2.1) of 250 optical density images (a) without and (b) with fringe removal. The peak amplitude of the most prominent peaks in (a) was reduced by a factor $\sim 10$ in (b). The spectrum of the data region, shown in part (c), has the same frequency components.
filtering or similar techniques would be impossible without significantly affecting the atomic distribution.

Figure 2.1 shows the effect of fringe removal for an optical density image of the magnetic lattice, selected to have significant fringes in the raw image. For this particular image, the pixel-to-pixel variance in a background region was reduced by a factor 5 by fringe-removal with a basis of 250 reference images. To further quantify the effect of fringe removal, we compute for all 250 images in the set the fourier transform in a background region, and average the spectra. The result is shown in figure 2.2. Without fringe removal, the average spectrum has similar characteristics to a spectrum calculated from the central region of the lattice, significantly hindering the detection of the atom number in each individual trap. After applying fringe removal, the amplitudes of the most prominent frequency components are strongly reduced, by a factor $\sim 10$.

We expect the fringe removal algorithm to also reduce photon shot noise. To verify this, we record a set of images for varying intensity using the imaging setup shown in figure 2.4. In this setup we observe no fringes due to reflection off the chip surface, and use a short time of 30 ms between recording the absorption and reference image. As a result, the recorded images show less significant fringes to begin with, providing an independent measure of the reduction in uncorrelated noise. We apply the fringe removal algorithm with a basis of $\sim 250$ images.

Figure 2.3 shows the measured imaging noise with and without fringe removal, $\text{var}(A/Q)$ and $\text{var}(A/R)$ respectively, as a function of the mean probe intensity $\langle R \rangle$, expressed in CCD counts per pixel. The variances are measured in a signal free region, separate from the fringe removal background region. Without fringe removal, our data is in good agreement with the expected photon shot noise, $2/\langle g(R) \rangle$, plus camera readout noise $\sigma^2_{rd}$. Application of the fringe removal algorithm reduces the measured variances by a factor of $1.9 \pm 0.3$ over the full rage of intensities (inset figure 2.3), and the resulting images are close to the ultimate photon shot noise limit of $1/\langle g(R) \rangle$.  

Figure 2.3: Imaging noise for various probe intensities with and without fringe removal, $\text{var}(A/Q)$ (blue circles) and $\text{var}(A/R)$ (red squares) respectively. The predicted photon shot noise plus camera readout noise is shown as a solid line. The dashed line shows the ultimate photon shot noise limit.
2.3 Optimal atom number estimation

In a typical cold atom experiment, the experimental signal is encoded in the number of atoms remaining at the end of an experimental cycle, either in a single cloud or in the multiple atomic ensembles. Of particular interest are double-well experiments and atom interferometers including atomic clocks and number/spin-squeezing experiments. In an atom interferometer, the interferometric phase can be mapped onto a population difference, where the individual populations are imaged onto different regions of a CCD camera.

In order to establish a limit on the signal-to-noise ratio (SNR) for total and difference population that can be obtained from such images, we derive the Cramér-Rao bound (CRB) for estimating both the atom number in a single cloud and the sum and difference populations in a double-well system. The CRB is a powerful tool, providing a minimum variance attainable for any (unbiased) parameter estimate, independent of the method used to extract the information [39]. To find an optimal method, we derive for both situations the maximum-likelihood estimator (MLE), and show it attains the CRB. Finally, we derive a general detection limit for the analysis of a single cloud, based on the CRB and imaging parameters, and show single atom sensitivity is attainable with realistic imaging parameters.

Single cloud analysis

We model an absorption image of a single atomic cloud (measured in atoms/pixel) consisting of $N$ atoms as

$$n_i(x) = Np(x) + d_i(x), \quad (2.9)$$

where $i$ labels realisations of the experiment and $x$ labels pixels within the image. Here, $p(x)$ is the normalised atomic distribution in the image ($\sum_x p(x) = 1$) depending on the cloud shape and optical resolution. The additive noise $d_i(x)$ is considered to be uncorrelated gaussian noise originating from photon shot noise, with standard deviation $\sigma_n$. For each pixel in the image, we write the probability distribution function (PDF)

$$P_x(n_i; x) = \frac{1}{\sqrt{2\pi\sigma_n}} \exp \left( -\frac{[n_i(x) - Np(x)]^2}{2\sigma_n^2} \right). \quad (2.10)$$

As the noise is uncorrelated between pixels, the PDF for a full image is given by $P(n_i; N) = \prod_{x=1}^X P_x(n_i; x)$, where $X$ is the total number of pixels in the image. $P(n_i; N)$ describes the probability of measuring a particular absorption image $n_i(x)$ for a given atom number $N$; however, it can alternatively be interpreted as the likelihood of the atom number being $N$ for a given image $n_i(x)$. By taking the natural logarithm, we then obtain the log-likelihood function for the problem,

$$l(N; n_i) = -X \log(\sqrt{2\pi\sigma_n}) - \sum_{x=0}^X \frac{[n_i(x) - Np(x)]^2}{2\sigma_n^2}. \quad (2.11)$$

The Fisher information is defined as $I(N) = -E[\partial^2 l(N; n_i)/\partial N^2]$, where $E[...]$ denotes the expectation value based on $P(n_i; N)$. The CRB is given by the inverse of the Fisher information and is written as

$$\text{var}(\hat{N}) \geq \sigma_n^2 \left( \sum_{x=1}^X p(x)^2 \right)^{-1}, \quad (2.12)$$

where the term in brackets can be interpreted as the effective number of pixels occupied by the atomic distribution.
Equation (2.12) gives the lowest possible detection noise for any unbiased estimator of \( N \). To find an estimator that attains the CRB we calculate the maximum-likelihood estimator, found by maximising the log-likelihood function and then solving for \( N \). We find

\[
\hat{N}_i = \frac{\sum_{x=1}^{X} n_i(x)p(x)}{\sum_{x=1}^{X} p(x)^2},
\]

which can be interpreted as a weighted sum over the density distribution \( n_i(x) \), where the weight is the local signal-to-noise ratio. It is worth noting that the CRB can also be attained by a least squares fit of the amplitude of the same distribution function \( p(x) \).

In the present analysis, the mode function \( p(x) \) is assumed to be a known function. In general, \( p(x) \) can be obtained in a model independent way by averaging many images, thus reducing noise.

**Sum and difference population in a double-well potential**

The Cramér-Rao bound and corresponding maximum likelihood estimator can be generalised to a multi-parameter problem. Of particular interest for interferometry and measurement of atom number fluctuations is the estimation of sum and difference population in a system with two atomic clouds. For this, we write the model density image

\[
N_+ = p(x) + q(x) + d_i(x),
\]

where \( N_p \) and \( N_q \) are the populations in the two atomic ensembles, with mode functions \( p(x) \) and \( q(x) \), respectively, normalised to \( \sum_x p(x) = \sum_x q(x) = 1 \). We write the log-likelihood function in terms of the sum and difference populations, \( \hat{N}_+ = N_p + N_q \),

\[
l(N_+; n_i) = -\sum_{x=0}^{X} \left\{ n_i(x) \right\} - \frac{1}{2} \left\{ \left[ (N_+ + N_-)p(x) + (N_+ - N_-)q(x) \right]^2 \right\} + \text{const.} \quad (2.15)
\]

The Fisher information matrix is now the matrix of second derivatives of \( l \) with respect to \( N_+ \) and \( N_- \),

\[
\mathcal{I}(N_+, N_-) = -E \left[ \left( \begin{array}{cc}
\frac{\partial^2}{\partial N_+^2} & \frac{\partial^2}{\partial N_-^2} \\
\frac{\partial^2}{\partial N_+ \partial N_-} & \frac{\partial^2}{\partial N_- \partial N_+}
\end{array} \right) l(N_+; n_i) \right]. \quad (2.16)
\]

The multi-parameter CRB is then the matrix inverse of \( \mathcal{I} \) and provides a bound on the covariance matrix,

\[
\text{Cov}(\hat{N}_+, \hat{N}_-) \geq C = \frac{4\sigma_n^2}{(u^+u^-) - v^2} \left( \begin{array}{cc}
u^- & -v \\
-v & u^+
\end{array} \right), \quad (2.17)
\]

with the parameters \( u^\pm = \sum_x [p(x) \pm q(x)]^2 \) and \( v = \sum_x [p(x)^2 - q(x)^2] \). For non-overlapping mode functions, \( 4/u^\pm \) can be interpreted as the effective number of pixels occupied by \( p + q \). The parameter \( v \) is related to the difference in occupied area between \( p \) and \( q \).

The matrix inequality \( X \geq C \) in eq. (2.17) denotes that the matrix \( X - C \) is positive semi-definite, which for the diagonal elements implies \( X_{i,i} \geq C_{i,i} \). The off-diagonal elements of \( C \) can be used in establishing the CRB for derived quantities, using (A.5).

Of particular interest is the relative population difference \( N_-/N_+ \), for which we obtain

\[
\text{var} \left( \frac{N_-}{N_+} \right) \geq \frac{\langle N_- \rangle^2 C_{1,1} + \langle N_+ \rangle^2 C_{2,2} - 2\langle N_- \rangle \langle N_+ \rangle C_{1,2}}{\langle N_+ \rangle^4}. \quad (2.18)
\]
The maximum likelihood estimator for $N^\pm$ is again found by maximising the log-likelihood function and solving for $N^+$ and $N^-$, yielding

$$\hat{N}^\pm_i = \frac{2u^\mp}{(u^+u^- - v^2)} \sum_x n_{i,x} \left( [p(x) \mp q(x)] - \frac{v[p(x) \mp q(x)]}{u^\mp} \right). \quad (2.19)$$

For non-overlapping clouds, the mode functions $p$ and $q$ can be found by averaging over many reference images. If $p$ and $q$ are not spatially separated, one could record a set of images where each ensemble is individually populated, or apply independent component analysis to isolate the signal components [46].

### 2.4 General detection limit for absorption imaging

To establish a general limit for detecting the number of atoms in an atomic ensemble with absorption imaging, we write the single parameter CRB (eq. (2.12)) for a 2D Gaussian distribution with $e^{-1/2}$-radius $w << \sqrt{\Delta}$, where $\Delta$ is the pixel area,

$$\text{var}(\hat{N}) \geq \sigma_n^2 4\pi w^2 / \Delta \quad (2.20).$$

Assuming photon shot noise limited detection, the noise per pixel is given by equation (2.8). After substituting in experimental parameters, we define the lowest detectable atom number by setting $\text{SNR} = 1$,

$$N_{\text{min}} = \frac{\sqrt{8\pi s(1 + s)w}}{Q_e \sigma_0 \Gamma^3} \quad (2.21).$$

Here $\tau$ is the exposure time, $Q_e$ the CCD quantum efficiency, $s$ the relative saturation and $\Gamma$ the natural line width of the imaging transition.

Assuming the atomic cloud is localised, the cloud width is initially limited by the optical resolution $w_0$ ($e^{-1/2}$-radius). During the probe pulse, however, the cloud broadens due to the random nature of photon scattering and the transferred photon momentum. This recoil blurring or heating effect, characterised by a blurring width $w_r$, adds to the observed width as $w^2 = w_0^2 + w_r^2$. To estimate $w_r$, we assume free space imaging, and integrate the velocity $\Delta v = \sqrt{N_{\text{sc}}/3v_r}$ acquired during the probe pulse with a constant scattering rate $N_{\text{sc}}$ (neglecting doppler shifts due to the acquired velocity), where $v_r$ is the velocity acquired by one photon recoil. We obtain

$$w_r^2 = 2 \frac{\Gamma s}{27} \left( \frac{5v_r^2 w_0^2}{\Gamma v_r^2} \right)^{1/3}. \quad (2.22)$$

By minimising $N_{\text{min}}$ with respect to $s$ and $\tau$ we obtain optimal imaging parameters $s_{\text{opt}} = 2/3$ and

$$\tau_{\text{opt}} = \frac{3}{2} \left( \frac{5v_r^2 w_0^2}{\Gamma v_r^2} \right)^{1/3}. \quad (2.23)$$

For these parameters, the Cramér-Rao bound is

$$\text{var}(\hat{N}) \geq \frac{20\pi}{3} \frac{\alpha}{Q_e \sigma_0} \left( \frac{5v_r w_0^2}{\Gamma} \right)^{2/3}. \quad (2.24)$$

To establish a realistic limit for absorption imaging, we consider atomic parameters for $^{87}\text{Rb}$ and realistic imaging parameters, consisting of $Q_e = 0.9$, $\alpha = 1$ and an optical resolution of $1.2\mu \text{m}$ (Rayleigh criterion, equivalent to $w_0 = 0.50\mu \text{m}$ and based on an objective lens with a numerical aperture of 0.4 [47]). For these parameters, we predict a detection limit of $N_{\text{min}} = 0.5$ atoms/shot at $\tau_{\text{opt}} = 13\mu \text{s}$. 22
The model presented here overestimates the effect of recoil blurring for several reasons. Most prominently, the width in the imaged is taken to be determined by the final width of the distribution, whereas in reality the image contains the integrated atomic distribution during the pulse. Secondly, any Doppler shifts during the probe pulse are ignored. Finally, for in-trap imaging the recoil blurring would be better described as heating of the atoms in the trap, also resulting in smaller broadening effects.

2.5 Atom number fluctuations in double-well systems

In a double well experiment, the measured populations will fluctuate from shot to shot due to different noise sources. Detection noise, for which the minimum contribution is derived in the previous section, is one such noise source. Another important contribution is intrinsic atom shot noise, which arises due to the stochastic nature of loss from the trap or, in the case of an atom interferometer, the projection of a coherent superposition state onto the two states of the double well. There can also be technical common-mode noise sources, for example due to magnetic field fluctuations from shot to shot, influencing the trap bottom, loading efficiency or other steps in the experimental cycle.

We investigate the expected effect of these noise sources on the measured populations from a basic model including the expected scaling of each type of noise. We are primarily interested in the relative population difference, $N^−/N^+$, which is the relevant quantity for the readout of an atom interferometer, and which we show to be robust against common-mode fluctuations of the atom number.

Relative population difference

We model the measured atom numbers in the two wells of a double well system to fluctuate about their mean as

$$
N_p = (1 + \gamma)\langle N_p \rangle + \delta_p \sqrt{1 + \gamma \sqrt{\langle N_p \rangle}} + \epsilon_p,
$$

$$
N_q = (1 + \gamma)\langle N_q \rangle + \delta_q \sqrt{1 + \gamma \sqrt{\langle N_p \rangle}} + \epsilon_q.
$$

(2.25)

Here, we include several noise sources, uncorrelated with each other and each with a mean of zero. Atom shot noise, for which $\text{var}(N_{p,q}) \propto \langle N_{p,q} \rangle$, is represented by $\delta_p$ and $\delta_q$, with $\text{var}(\delta) = 1$ for Poissonian fluctuations. The parameters $\epsilon_p$ and $\epsilon_q$ represent detection noise, independent of the mean atom number. Technical fluctuations in the total atom number generally scale with $\text{var}(N_{p,q}) \propto \langle N_{p,q} \rangle^2$ and are modelled by $\gamma$. All fluctuating parameters are defined with a mean of zero.

Using equation (A.3), the expected variance in the relative population difference $N^−/N^+$ is independent of $\gamma$,

$$
\text{var}\left(\frac{N^-}{N^+}\right) = \frac{\langle N^+ \rangle^2 - \langle N^- \rangle^2}{\langle N^+ \rangle^4} \text{var}(\delta) + \frac{\langle N^+ \rangle^2 + \langle N^- \rangle^2}{\langle N^+ \rangle^4} 2\text{var}(\epsilon).
$$

(2.26)

Here we assumed that the atom shot noise and detection noise have the same amplitude in both clouds. For equal populations ($\langle N^- \rangle = 0$) and $\text{var}(\delta) = 1$, equation (2.26) reduces to

$$
\text{var}\left(\frac{N^-}{N^+}\right) = \frac{\langle N^+ \rangle + 2\text{var}(\epsilon)}{\langle N^+ \rangle^2},
$$

(2.27)

which can be interpreted as atom shot noise plus the detection noise originating from both clouds (hence the factor 2). As equation (2.26) does not depend on $\text{var}(\gamma)$, the measured $\text{var}(N^−/N^+)$ is robust against technical fluctuations. For low atom number, detection noise dominates the expected variance, whereas for $\langle N^+ \rangle \gg 2\text{var}(\epsilon)$ the variance is dominated by atom shot noise.
Mirror images

For the purpose of experimentally measuring atom number fluctuations, a mirror imaging setup can be employed. In such an imaging system, the probe laser reflects off a mirror surface close to the atomic distribution, passing the atoms twice and creating two independent images of the atomic distribution on separate regions of the CCD camera [18]. The mirror images then provide an additional means to separate atom number fluctuations from detection noise, as detection noise is expected to be uncorrelated between them, whereas actual atom number fluctuations will be identical in both images.

We model the mirror image with a second set of measured atom numbers, $N'_p$ and $N'_q$ (and correspondingly $N'^{\pm} = N'_p \pm N'_q$), which differ from $N_p$ and $N_q$ only by having independent detection noise parameters $\epsilon'_p$ and $\epsilon'_q$. We consider $\text{var}(\epsilon)$ to be identical for both images.

To measure the atom number fluctuations without including detection noise, we measure the covariance of the relative difference populations, which, using equation (A.2), for our model is

$$\text{cov} \left( \frac{N^-}{N^+}, \frac{N'^-}{N'^+} \right) = \frac{\langle N^+ \rangle^2 - \langle N^- \rangle^2}{\langle N^+ \rangle^4} \text{var}(\delta),$$

(2.28)

thus providing a means to measure atom shot noise robust against both technical noise and detection noise.

2.6 Experimental setup

We experimentally test our optimal analysis techniques by investigating atom shot noise in two wells of a multi-well trap. We load up to a total of $4 \times 10^3$ $^{87}$Rb atoms in the trap, created at the edge of our magnetic lattice. Here, edge effects of the lattice structure combine with a Z-shaped current-carrying wire in the chip and homogeneous external fields to create a four-well potential (two deep central wells and two shallow outer wells). We use radio frequency evaporative cooling in this potential to lower the temperature to $\sim 1\mu$K, thereby creating independent atomic ensembles. We vary the mean number of atoms by holding the atoms in the trap for $0-4$ s and by reducing the amount of Rb dispensed during loading of the trap. While holding the atoms, a fixed temperature is ensured by a radio frequency ‘knife’ cutting off the tail of the thermal distribution at a fixed energy.

Imaging system

We image the atoms in-situ onto a back-illuminated deep-depletion CCD camera (Andor iKon-M934 BR-DD) with a quantum efficiency of $Q_e = 0.9$ electrons/photon at our probe wavelength of $\lambda = 780$ nm. We directly measure the camera conversion gain $Q_{e,g}$ by exposing the camera to light from an aperture, and comparing the integrated number of counts in the image to the expected number of photons based on exposure time and measured probe power after the aperture. We measure $Q_{e,g} = 0.87 \pm 0.05$ counts/photon and a readout noise level of $\sigma_{rd} \approx 13$ counts. The imaging system, depicted in figure 2.4, has a magnification factor of 4 using a 75 mm objective lens, achieving an optical resolution of 9.6 $\mu$m (Rayleigh criterion) with a pixel area of $\Delta = (3.3 \pm 0.1 \mu m)^2$ in the object plane. The probe beam is slightly inclined with respect to the gold-coated chip surface, creating two mirror images of the same atomic distribution. This reflection imaging technique (fig. 2.4) provides an additional means to separate atom number fluctuations from detection noise, through correlating the two mirror images.

In calculating the atomic density distribution using equation (2.22), we determine the local saturation parameter pixel-by-pixel from the reference image, $s = 2RE_{ph}/(Q_{e,g}I_0^c \Delta \tau)$. 

24
Figure 2.4: (a) Schematic depiction of the mirror imaging setup, showing the atom chip (A), probe beam (P), lens system (O, S) and CCD camera. The angle of the probe beam has been exaggerated to illustrate the reflection setup. (b) Close-up around the atom chip surface. Part of the probe light is reflected before and part after passing through the atoms, depicted by solid and dashed lines, respectively. The two separate shadows cast by the atoms are then imaged onto separate regions of the CCD camera.

Figure 2.5: Typical absorption image of the atomic distribution (90 × 71 pixels) for a hold time of 1.6 s. Four independent atomic ensembles are visible (horizontally distributed), with the two mirror images vertically separated. The two central clouds contain ≈ 400 atoms each. Dashed boxes indicate the background region for the fringe removal algorithm and the signal regions are indicated by solid lines.

Here $E_{\text{ph}}$ is the photon energy. The factor 2 is specific to reflection imaging, as the probe beam passes the atoms twice.

Figure 2.5 shows a typical absorption image of the atomic distribution, for a hold time of 1.6 s. The four potential wells are horizontally distributed and well resolved by our optical resolution, and the two mirror images are vertically separated. The image contains ≈ 400 atoms in each central well.

**Optimised and calibrated imaging parameters**

The signal-to-noise ratio (SNR) for atom number detection can be optimised by carefully choosing the probe intensity, or more specifically the saturation parameter $s$, and exposure time $\tau$. The photon shot noise per pixel $\sigma_n$ decreases with increasing exposure
Figure 2.6: Atom number signal-to-noise ratio (colour scale, arbitrary units) as function of probe duration $\tau$ and saturation parameter $s$. (a) Experimental SNR, obtained from measured cloud area and photon shot noise. Each coloured patch corresponds to a separate realisation of the experiment. The white area (top right) has been excluded due to saturation of the CCD sensor. The black circle indicates the optimal parameters used in further experiments. (b) Theoretical SNR based on equations (2.21) and (2.22) for our optical resolution. Black lines are constant SNR contours, and the black circle indicates the theoretical optimal parameters.

We record a series of absorption images for constant atom number as a function of both $s$ and $\tau$, and measure the cloud area $a$, increasing the amount of imaging noise integrated when measuring the atom number. The optimal parameters derived in section 2.4 are based on a simplified model. To obtain model independent optimal parameters, we directly optimise the SNR experimentally.

For the chosen parameters, we calibrate $\alpha$ by comparing the integrated absorption signal with a second set of images, recorded after free expansion using a weak probe ($s \approx 0.02$) and well-defined quantisation axis, such that we can assume $\alpha = 1$ for these images. The comparison yields $\alpha = 3.0 \pm 0.2$ for in-situ imaging.

To verify the linearity of our imaging system, we compare the integrated absorption in the inner wells with the outer wells. For independent ensembles dominated by one-body loss, we expect a fixed ratio between inner and outer atom numbers. Figure 2.7 shows the inner versus outer atom number for a wide range of atom numbers. We observe a ratio of 3.4 over the full range of atom numbers, confirming detection linearity.

Cramér-Rao bound

After applying the fringe removal algorithm, the noise from pixel to pixel in our images is $\sigma_n = 1.3 \pm 0.3$ atoms, close to the ultimate limit of 1.1 atoms due to photon shot noise in $A$. From the ultimate photon shot-noise limit and the average distribution of our atoms ($1/u^k \approx 30$ pixels, $v \approx 0$), we obtain a CRB of $\text{var}(N^k) \geq 200 \text{ atoms}^2$. 
corresponding to a minimum resolvable populations difference of 14 atoms/shot. The single-cloud CRB for our data corresponds to a detection limit of 10 atoms/trap/shot.

Apart from uncorrelated noise, we also observe in our images a small residual correlated noise component on a length scale comparable with our cloud size, with an rms amplitude of $\sigma = 0.06 \pm 0.01$ atoms.

### 2.7 Results

We have measured the relative population difference $N^-/N^+$ for varying total atom number to resolve the intrinsic atom number fluctuations and demonstrate the improvements gained using the maximum-likelihood estimator. We record data for 40 hold times from $0 - 4$ s, varying the total atom number $\langle N^+ \rangle$ from $\sim 4000$ down to $\sim 20$ atoms. The experiment is repeated 40 times for each hold time. To correct for small variations in the positions of the atoms, we apply a sub-pixel registration algorithm to align each image with the average [49].

We measure the sum and difference populations $N^\pm$ for each realisation in two ways, restricting our analysis to the lower mirror image. Firstly, we directly integrate a region of the density image for each cloud, with the regions shown in figure 2.5. The regions are chosen to include practically all the atomic signal, but not too large to avoid needlessly integrating extra pixel noise. Secondly, we apply the maximum likelihood estimator given in equation (2.19). The mode functions $p \pm q$ are determined by averaging all the data to reduce noise, and segmenting the result to the same regions as used for integration. Finally, we compute for both methods the variance $\text{var}(N^-/N^+)$ by applying equation (A.3).

Figure 2.8 shows the measured variance for both methods as a function of $\langle N^+ \rangle$, together with the expected variance for CRB limited detection and atom shot noise, given by equation (2.20). For large $\langle N^+ \rangle$, the result for both methods is atom shot-noise limited. For straight integration, the measured variance is detection noise limited for $\langle N^+ \rangle$ below $\sim 1000$ atoms. The MLE significantly improves the detection limit, reaching atom shot-noise limited detection down to $\sim 300$ atoms.

To quantify the results, we fit to our data the model $\text{var}(N^-/N^+) = \tilde{C}_{2,2}/\langle N^+ \rangle^2 + 1/\langle N^+ \rangle$, where we use equation (2.18) that for our data $\langle N^- \rangle \approx 0$. For the integration method we find $\tilde{C}_{2,2} = (1.3 \pm 0.3) \times 10^3$ atoms$^2$, whereas for the MLE $\tilde{C}_{2,2} =$
Figure 2.8: Relative variance as a function of total atom number measured by maximum likelihood estimation (squares) and straight integration (circles). The dash-dotted line shows the expected fluctuations combining atom shot noise (dashed line) and the minimum detection noise contribution given by the CRB (dotted line). Solid lines indicate fits for the detection noise contributions.

Figure 2.9: Covariance between $N^-/N^+$ in the upper and lower mirror image (squares) along with the predicted atom shot noise (dashed line).

270 ± 40 atoms$^2$. The latter corresponds to a minimum resolvable population difference of 17 atoms in a single shot.

The measured noise is slightly larger than the minimum value given by the CRB, but is consistent with the expected noise when taking into account the added effect of the residual correlated noise $\sigma_c$. Although $\sigma_c$ has a small amplitude in each pixel, the correlated noise has a significant effect on the measured atom number, as summing over $a$ pixels results in an integrated variance of $a^2\sigma_n^2$, compared to $a\sigma_n^2$ for the uncorrelated noise.

We can further distinguish atom shot noise from detection noise in our data by com-
paring the measured relative population numbers in the upper and lower mirror images. For the covariance \( \text{cov}(N^-/N^+, N'^-/N'^+) \), we expect only atom number fluctuations to play a role. Figure 2.9 shows the measured covariance using the MLE, along with the expected result for atom shot-noise given by equation (2.28). Our data is consistent with the model for the full range of atom numbers, demonstrating a robust method for observing intrinsic atom number fluctuations.

2.8 Conclusion

In this chapter, we have described and demonstrated improved detection of small atomic ensembles through advanced post-processing and optimal image analysis. A fringe removal algorithm is used to minimise fringe noise and the effect of photon shot noise. We maximise the signal-to-noise obtainable from absorption images by deriving the Cramér-Rao bound for atom number detection, and the maximum-likelihood estimator that attains that bound.

The combined application of fringe removal and MLE to our experiment offers a factor 3 improvement in signal-to-noise ratio, allowing us to resolve a minimum population difference of 17 atoms in a single shot in a double well system. Correspondingly, we are able to measure atom shot-noise limited detection of ensembles comprising of as few as 270 atoms, a factor 9 lower than without our analysis.

Averaging the measurements from both mirror images would provide a further factor \( \sqrt{2} \) improvement in SNR. Better SNR can also be obtained by localising the atoms in three dimensions, and correspondingly improving the optical resolution. For well-localised atoms and realistic imaging parameters, we show single atom sensitivity in a single cloud is feasible with absorption imaging.

The optimal analysis techniques are applicable to many cold atom experiments, and provide a basis to improve the readout of trapped atom interferometers to the quantum limit or better resolve number squeezing and entanglement in small atomic ensembles.
3 Sub-Poissonian atom number fluctuations by three-body loss

Arrays of small atomic ensembles, each with a well-defined atom number, provide a starting point for quantum information processing with mesoscopic qubits. Intrinsic to these mesoscopic ensembles, however, are shot-to-shot atom number fluctuations due to the stochastic nature of the processes involved in creating the atomic clouds. When there is no interaction between particles, this leads to the standard shot noise limit, where the number of atoms from realisation to realisation follows a Poisson distribution, with a width of $\sqrt{N}$ for a mean of $\langle N \rangle$ particles. In experiments, technical noise sources may further broaden the atom number distribution.

In this chapter, we show that three-body loss, where particles leave the trap in groups of three, naturally reduces the shot-to-shot fluctuations of the total atom number in the trap to sub-Poissonian levels. Three-body loss is itself a random process, and may therefore intuitively be expected to increase fluctuations. Due to its density dependent nature, however, three-body loss can rapidly reduce atom number fluctuations to below the Poisson level. We show this theoretically by deriving a stochastic model for general loss processes involving multiple particles.

We experimentally measure sub-Poissonian atom number distributions in our magnetic lattice atom chip. Sensitive absorption imaging enables the detection of the atom number fluctuations in each trap individually, providing a direct measure of reduced fluctuations. We further reduce the effect of residual imaging noise by applying spatial correlation analysis, enabling us to separate atom number fluctuations from detection noise and find sub-Poissonian statistics for between 50 and 300 atoms per trap. The results over this range are in very good agreement with the stochastic model. The results of this chapter have also been published in reference [50].

3.1 Three-body recombination

For trapped ultracold gases, inelastic density-dependent decay is the dominant loss process. In $^{87}$Rb this is typically due to three-body recombination [51, 52, 53], whereby all three atoms are lost from the trap. The loss rate of atoms depends on the probability of finding three atoms together. Measurement of the evolution of the mean atom number due to three-body loss has been used as a sensitive probe for density correlations and fluctuations within degenerate and nondegenerate Bose gases [51, 55].

We are primarily interested in the evolution of fluctuations due to the loss of particles. As the probability to lose atoms from the trap depends on density, and thus for a thermal cloud on the number of atoms in the trap, we expect a trap loaded with more than the average number of atoms to decay faster than average. Therefore, we expect density-dependent loss to reduce the shot-to-shot fluctuations.
Stochastic model for multi-body loss

Following reference [56], we write a stochastic model for multiple loss processes, where for each process loss of particles occurs in groups of \( \rho \) bodies, characterised by loss rates \( k_\rho \). We write the master equation for the probability \( P(N,t) \) of having \( N \) particles remaining at time \( t \),

\[
\frac{dP(N,t)}{dt} = \sum_\rho \frac{k_\rho}{\rho N_0^{\rho-1}} (E^\rho - 1) \left[ \prod_{r=0}^{\rho-1} (N-r) \right] P(N,t). \tag{3.1}
\]

Here, \( N_0 \) is the initial number of particles and \( E \) is the step operator, defined by its effect on an arbitrary function \( f(N) \),

\[
E^\rho f(N) = f(N + \rho).
\tag{3.2}
\]

Equation (3.1) is a set of coupled differential equations, one for each \( N \). For small initial atom numbers, a numerical solution can be obtained by limiting the number of equations to some cut-off value for \( N \) and choosing specific values of \( \rho \) to include.

Here, we obtain an approximate analytical solution, valid for \( N_0 \gg 1 \). We expect the solution to be a peaked function around the mean \( \langle N \rangle \), with a width on the order \( \sqrt{\langle N \rangle} \). Therefore, we change variables to the macroscopic evolution of the mean atom number plus a fluctuation term,

\[
N = N_0 \eta(t) + \sqrt{N_0} \xi,
\tag{3.3}
\]

where \( \eta = \langle N \rangle / N_0 \) is the mean fractional atom number, such that \( \langle \xi \rangle = 0 \). The evolution of \( \eta \) is described by

\[
\frac{d\eta(t)}{dt} = - \sum_\rho k_\rho \eta^\rho. \tag{3.4}
\]

The probability distribution of \( \xi \) is written \( P(\xi,t) \), and its time derivative is expanded as

\[
\frac{dP(\xi,t)}{dt} = \frac{\partial P(\xi,t)}{\partial \xi} \frac{d\xi}{dt} + \frac{\partial P(\xi,t)}{\partial N} \frac{dN}{dt} = \frac{\partial P(N,t)}{\partial t} + \sqrt{N_0} \frac{d\eta(t)}{dt} \frac{\partial P(\xi,t)}{\partial \xi}. \tag{3.5}
\]

Substituting in equation (3.1), we obtain the master equation for the evolution of \( \xi \),

\[
\frac{dP(\xi,t)}{dt} = \sum_\rho \frac{k_\rho}{\rho N_0^{\rho-1}} (E^\rho - 1) \left[ \prod_{r=0}^{\rho-1} (N_0 \eta(t) + \sqrt{N_0} \xi - r) \right] P(\xi,t) + \sqrt{N_0} \frac{d\eta(t)}{dt} \frac{dP(\xi,t)}{d\xi}. \tag{3.6}
\]

For \( N_0 \gg 1 \), the step operator can be expanded as [56]

\[
E^\rho \approx 1 + \frac{\rho}{\sqrt{N_0}} \frac{d}{d\xi} + \frac{\rho^2}{2 N_0} \frac{d^2}{d\xi^2}. \tag{3.7}
\]

For \( N_0 \gg m \), we can also ignore the \( r \)-term of the product in equation (3.6). Substituting in both approximations, and keeping only terms to zeroth order in \( N_0 \), we obtain

\[
\frac{dP}{dt} = \sum_\rho k_\rho \left( \sqrt{N_0} \eta^\rho \frac{dP}{d\xi} + \rho \eta^\rho \frac{d(P)}{d\xi} + \frac{\rho}{2} \eta^{\rho+1} \frac{d^2 P}{d\xi^2} \right) + \sqrt{N_0} \frac{d\eta}{dt} \frac{dP}{d\xi}. \tag{3.8}
\]

Substituting in equation (3.4), we obtain a linear Fokker-Planck equation,

\[
\frac{dP}{dt} = \sum_\rho k_\rho \left( \rho \eta^\rho \frac{d(P)}{d\xi} + \frac{\rho}{2} \eta^{\rho+1} \frac{d^2 P}{d\xi^2} \right). \tag{3.9}
\]
Multiplying both sides by $\xi^2$ and integrating over $\xi$, we obtain an equation for the second moment $\langle \xi^2 \rangle = \int \xi^2 P d\xi$,

$$\frac{d\langle \xi^2 \rangle}{dt} = -2 \sum_{\rho} \rho k_\rho \nu^{\rho-1} \langle \xi^2 \rangle + \sum_{\rho} \rho k_\rho \nu^\rho. \quad (3.10)$$

Here, we assume $P(\xi, t)$ is a peaked function in $\xi$, decaying sufficiently fast such that $\xi P = 0$ and $\xi dP/d\xi = 0$ for $\xi \to \pm \infty$.

It is convenient to write the fluctuations in $N$ in terms of the Fano factor, or relative variance, defined as

$$F \equiv \frac{\langle N^2 \rangle - \langle N \rangle^2}{\langle N \rangle} = \frac{\langle \xi^2 \rangle}{\eta}, \quad (3.11)$$

such that $F = 1$ for a Poisson distribution and $F < 1$ for sub-Poissonian fluctuations.

Combining equations (3.4) and (3.10), the evolution of the Fano factor depends only on the fractional atom number $\eta$,

$$\frac{dF}{d\eta} = \sum_{\rho} k_\rho \nu^{\rho-1} [(2\rho - 1)F - \rho] \sum_{\rho} k_\rho \nu^\rho. \quad (3.12)$$

For a single loss process involving $\rho$ bodies, the Fano factor decays as

$$F(\eta) = \frac{\rho}{2\rho - 1} + \eta^{2\rho-1} \left( F_0 - \frac{\rho}{2\rho - 1} \right), \quad (3.13)$$

where $F_0$ is the initial Fano factor at $t = 0$ where $\eta = 1$. It reaches an asymptotic value of $F \to \rho/(2\rho - 1)$ for $\eta \to 0$, indicating the lowest Fano factor possible through $\rho$-body loss is $F = 1/2$. For one-body loss, the mean atom number decays exponentially as

$$\eta(t) = \exp(-k_1 t). \quad (3.14)$$

**Three-body loss**

In our experiment, the dominating loss processes are one-body loss and three-body loss. For this case, the evolution of the mean is

$$\eta(t) = \frac{\exp(-k_3 t)}{\sqrt{1 + (k_3/k_1)[1 - \exp(-2k_1 t)]}}, \quad (3.15)$$

or, when dominated by three-body loss,

$$\eta(t) = \frac{1}{\sqrt{1 + 2k_3 t}}. \quad (3.16)$$

The Fano factor for three-body loss, given by

$$F(\eta) = \frac{3}{5} + \eta^{5} \left( F_0 - \frac{3}{5} \right), \quad (3.17)$$

decays as $\eta^5$ to a steady state value of $F = 3/5$, significantly below the Poisson noise level of $F = 1$. Including one-body loss, the Fano factor can be written in terms of $k = k_1/k_3$,

$$F(n) = \frac{n(n^2 + k)^2}{4k^{5/2}} \left[ \sqrt{k} \left( \frac{3n^4 + 5n^2k + 4k^2}{n(n^2 + k)^2} + \frac{4k^2(F_0 - 1) - 5k - 3}{(k + 1)^2} \right) + 3 \arctan \left( \frac{\sqrt{k}(n - 1)}{k + n} \right) \right]. \quad (3.18)$$
Figure 3.1: Predicted decay of the Fano factor and mean atom number for one- and three-body loss. (a) Fano factor as function of $\eta$, with an artificially large initial Fano factor of $F_0 = 5$. For one-body loss (dashed red line), $F$ decays linearly, to $F = 1$ (solid black line) when all atoms are lost. For three-body loss (solid blue line), the fluctuations decay rapidly as $F \propto \eta^5$, reaching sub-Poissonian fluctuations before half the atoms are lost and finally reaching its asymptotic value of $F = 3/5$ (dashed black line) (b) Decay of the mean atom number $\eta$ as function of time, for typical values of $k_1$ and $k_3$ in our experiment. The dashed line shows exponential decay with the same decay constant $k_1$, corresponding to one-body loss only.

Figure 3.1 shows the theoretical results for one-body and three-body loss. In figure 3.1a, the Fano factor is plotted as function of $\eta$ for both one-body and three-body loss, assuming an initial super-Poissonian distribution with large Fano factor. For one-body loss, $F$ decays to its asymptotic value of 1 (Poisson noise), but only linear in $\eta$. As a result, the distribution is not reduced to Poisson noise until all atoms are lost from the trap. For three-body loss, the Fano factor decays rapidly, reaching sub-Poissonian levels before half of the atoms are lost, and finally reaching the asymptotic value of $F = 3/5$. Figure 3.1b shows the decay of combined one- and three-body loss for typical parameters for our experiment ($k_1 = 0.5 \text{ s}^{-1}$ and $k_3 = 10 \text{ s}^{-1}$), along with one-body loss only. The non-exponential behaviour of three-body loss provides a means to distinguish the processes in experimental data.

3.2 Experimental setup

To probe atom number fluctuations in our experiment, we load $\sim 250$ traps of our lattice with a few hundred atoms each, evaporatively cooled in the lattice to $\sim 3 \mu\text{K}$. The tight confinement (mean trap frequency $\Omega = 10.0 \pm 0.5 \text{ kHz}$) in each trap results in a high atomic density ($\approx 2 \times 10^{14} \text{ cm}^{-3}$), for which three-body decay is the dominant loss process. We counteract any heating that may occur due to three-body decay by limiting the trap depth with an rf ‘knife’ to $\sim 35 \mu\text{K}$.

We detect the atom number in each trap, by imaging the atomic distribution in situ with our reflection imaging setup described in section 1.2. We use an exposure time of $0.15 \text{ ms}$ and the saturation parameter is $s = 0.6$. We apply our fringe removal algorithm to the recorded images, and compute for each run of the experiment an optical density (OD) image. We align each image to the average distribution using sub-pixel registration, to correct for small shifts from shot to shot due to movement of the atom chip. We independently calibrate the effective absorption cross section $\sigma_{\text{abs}} = \sigma_0/\alpha(1+s)$ by comparing for each trap individually the measured cloud temperature, measured through radio frequency spectroscopy, with the three-body loss rate [5]. We find $\alpha(1+s) = 3.1 \pm 0.5$, consistent with the expected value of $\alpha(1+s) = 3.2$, based on polarisation and the saturation parameter. The maximum optical density in our measurements is
Figure 3.2: Subsection of an optical density image near the centre of the magnetic lattice. The image is taken after a hold time of 25 ms, with \( \sim 250 \) atoms per trap. The black rectangle highlights the trap labelled \( m = 138 \).

\[ \sim 0.1 \] Figure 3.2 shows a subsection of an optical density image containing \( \sim 250 \) atoms per trap.

We hold the atoms for a variable time after evaporative cooling, before taking an absorption image of the distribution. We have recorded two sets of data. The first spans from \( t = 0 \) ms to \( t = 880 \) ms in 40 intervals, and the experiment is repeated 15 times for each interval. The second set spans from \( t = 21 \) ms to \( t = 2.5 \) s, with 40 intervals and 19 repeats. The intervals were selected on a power-law scale for both data sets.

3.3 Single trap analysis

We extract for each micro-trap the evolution of the mean atom number and the variance. For this, we obtain an approximate shape function for each lattice site by decomposing the mean shape of each site into a sum of Gaussian profiles. We then extract the atom number in each trap by fitting the amplitude of the corresponding shape function with a least squares fit. To correctly include small overlaps of the shape functions, we include the eight neighbouring traps in each fit.

Figure 3.3 shows the evolution of the mean atom number \( \langle N \rangle \), obtained from averaging over 19 realisations, for a selected trap. A fit to equation (3.15) yields one- and three-body decay rates \( k_1 = 0.52 \pm 0.03 \) s\(^{-1}\) and \( k_3 = 10.4 \pm 0.4 \) s\(^{-1}\) and initial atom number \( N_0 = 354 \pm 4 \). The measured decay curve is in excellent agreement with the theory, to the level of \( \pm 3 \) atoms over the full range of atom numbers.

From the measured atom numbers we can infer the fluctuations from shot to shot in each trap individually. For this, we use the vast number of traps measured in each run of the experiment to normalise out noise in the total atom number, for our data mainly due to fluctuations of the probe frequency.

Figure 3.4 shows histograms of the number of atoms in a selected trap, denoted \( m = 138 \), over 19 averages for two different hold times. In figure 3.4a, a histogram for a hold time of 25 ms is shown, with \( \langle N_m \rangle = 280 \pm 3 \). To measure the variance, we assume a Gaussian distribution and fit an error function to the cumulative distribution function, obtained by sorting the 19 measurements by \( N_m \). The fitted width corresponds to \( \text{var}(N_m) = 140 \pm 50 \), consistent with calculating the variance directly. The result is significantly below the Poisson level of \( \text{var}(N_m) = \langle N_m \rangle \), indicated by the dashed line. Thus, our data directly shows sub-Poissonian fluctuations, even including the added presence of detection noise. Figure 3.4b shows a histogram after a long hold time of...
Figure 3.3: Decay of the mean atom number $\langle N \rangle$ as function of time for the trap highlighted in figure 3.2. The blue line is a fit to equation (3.15). The dashed line shows the corresponding one-body loss contribution. The shaded region indicates the range of atom numbers probed in our data for all traps. The lower section shows residuals of the fit and standard errors on the measurement of $\langle N \rangle$, demonstrating agreement on the level of ±3 atoms.

Figure 3.4: Histograms of the atom number in trap $m = 138$ for hold times of (a) 25 ms and (b) 1.98 s, along with the expected distribution for Poisson noise (dashed lines) and Poisson noise combined with detection noise (dash-dotted lines). Gaussian fits to the distribution (solid lines) indicate directly resolved sub-Poissonian fluctuations in (a), whereas in (b) the fit is in excellent agreement with detection noise.

1.98 s, where the mean atom number has decayed to $\approx 30$ atoms. Here, the measured fluctuations are limited by detection noise.

The same analysis is performed for each lattice site and each hold time, resulting in $250 \times 40$ observations. Figure 3.5 shows the measured Fano factor as function of mean atom number for all traps (points), with trap $m = 138$ highlighted (open circles). The observed fluctuations have two contributions: atom shot noise with a constant $F$ (Poisson noise is indicated by a dashed line) and a detection noise contribution corresponding to a fixed variance of 64 atoms$^2$/trap/shot (dotted line). For trap $m = 138$, we obtain a fitted constant Fano factor of $F = 0.57$ (solid line). For $N \gtrsim 100$ we find the vast majority of data points are below the combined Poisson fluctuations and detection noise, indicating $F < 1$ for this range. For smaller atom numbers, the measured fluctuations are dominated by detection noise and we cannot distinguish the atom number fluctuations directly.

The measured atom number fluctuations are significantly below the Poisson noise level, most notably for large $\langle N \rangle$ (corresponding to short hold times), where the detection
Figure 3.5: Measured Fano factor for all traps (points) for all hold times, with trap \( m = 138 \) highlighted (open circles), as function of mean atom number. For high \( \langle N \rangle \), our data is directly below the Poisson level (dashed line). For \( \langle N \rangle \gtrsim 100 \), the vast majority of data points is below the combined Poisson noise plus detection noise (dash-dotted line). Fitting a constant \( F \) with detection noise (dotted line) to trap \( m = 138 \) yields \( F = 0.57 \).

Figure 3.6: (a) Typical optical density image of our lattice, for peak optical density comparable to the noise amplitude. (b) Autocorrelation function of (a), clearly revealing the lattice structure. The uncorrelated imaging noise is only present in the central pixel.

3.4 Fluctuation correlation analysis

The measured fluctuations in each trap individually are, for small atom numbers, dominated by detection noise. To account for detection noise and investigate the sub-Poissonian noise over the full range of atom numbers in our data, we perform spatial correlation analysis of the individual images. Here, we benefit from our lattice geometry and separate different noise contributions based on their respective correlation length scales. This way, we can isolate atom number fluctuations from uncorrelated detection noise.

Figure 3.6 shows a typical optical density image of the lattice with low signal-to-noise-ratio, and the autocorrelation function of the same image. Where in the real
Figure 3.7: Central section of a fluctuation correlation image for $t = 25$ ms, averaged over 19 realisations. The central (red) pixel contains the uncorrelated imaging noise. The broader peak surrounding it represents shot-to-shot fluctuations within each cloud, averaged over all clouds. The smaller peaks arranged in a lattice in the rest of the image represent common-mode shot-to-shot fluctuations due to fluctuating probe frequency.

space image the signal is hardly distinguishable from the noise, in the correlation image different correlation length scales can be clearly identified. A central peak arises from the average shape of the imaged traps, and is surrounded by peaks arising from the lattice structure. The imaging noise is uncorrelated, and therefore only contributes to the zeroth pixel.

**Fluctuation correlation function**

The autocorrelation images of the data primarily contain information about the mean number of atoms and the lattice structure. To measure the fluctuations, we first subtract for each individual image the mean image, averaged over all realisations of the experiment with the same hold time. We take the autocorrelation function of the resulting fluctuation image, to create fluctuation correlation images. These contain information about the fluctuations separated by correlation length scale, and can be averaged together while retaining information about the noise.

Figure 3.7 shows the central section of a fluctuation correlation image, averaged over 19 realisations with a hold time of $t = 25$ ms. The image shows a sharp peak at the centre, corresponding to the uncorrelated imaging noise. This sits on top of a broader peak representing the fluctuations correlated over the length scale of a single cloud, which corresponds to shot-to-shot fluctuations of the number of atoms in each trap. An array of neighbouring peaks, spaced at the lattice period, represents the correlated noise across traps, which we attribute to fluctuations of the probe detuning.

As derived in appendix C, the mean fluctuation correlation images can be used to extract the lattice-averaged Fano factor $\bar{F}$. We calculate mean fluctuation correlation images for both of our data sets, and subtract the autocorrelation in a background region to eliminate residual correlated imaging noise, due to diffraction of the probe light off our lattice structure. To obtain the contributions of the central and surrounding peaks, we fit Gaussian distributions to the data and use the fitted volumes as a robust measure for calculating the Fano factor. Similarly, we obtain the volume of the central peak in the pre-averaged autocorrelation image for $P_0$ in equation (C.10). The overlap term this equation is $< 0.1$ for our data.

Figure 3.8 shows the extracted Fano factor for both data sets, analysed separately. Horizontal lines indicate Poisson noise ($F = 1$, dashed line) and the expected limit for three-body loss ($F = 3/5$, dotted line). The data show sub-Poissonian atom-number fluctuations for $\langle N \rangle \gtrsim 50$ up to 300 atoms per trap. A fit over this range indicates
$F = 0.53 \pm 0.08$, where the uncertainty indicates the standard deviation. We independently estimate a systematic uncertainty of $\pm 0.2$ incorporating uncertainties in the absolute atom number calibration, background noise contribution, and the overlap between traps. The data are clearly below the Poisson level, and are in good agreement with the expectation of $F = 3/5$. For $\langle \hat{N} \rangle \lesssim 50$ we expect one-body loss to dominate, increasing the fluctuations to the Poissonian level. The solid line shows equation (3.18) for one- and three-body loss.

### 3.5 Conclusion

We have shown theoretically and experimentally that random three-body loss naturally leads to reduced fluctuations in the remaining atom number, to below the standard shot noise limit. For the range of 50 to 300 atoms, we measure clearly sub-Poissonian fluctuations, with a Fano factor in good agreement with our theoretical steady-state prediction of $F = 3/5$.

Three-body decay is a simple method to prepare small and well-defined atom numbers in ensembles of ultracold atoms containing tens to a few hundred of atoms. We expect this to be an ideal system for the study of collective excitations, for example via laser-excited Rydberg states. In our optically accessible lattice of microtraps, this would be a solid basis for quantum information processing with neutral atoms.
Conclusion

We have demonstrated sub-Poissonian atom number fluctuations in a lattice of mesoscopic ensembles on a magnetic lattice atom chip. Reduced fluctuations due to density-dependent loss allow us to create a lattice of small atomic ensembles with a well-defined atom number in each ensemble, opening the way to quantum information science with collective excitations in mesoscopic qubits. We have described and demonstrated improved atom number detection based on advanced post-processing and optimal analysis techniques, enabling the study of atom number fluctuations for small ensembles and providing a basis for improved readout of trapped atom interferometers.

Using a fringe removal algorithm, we strongly suppress the effect of interference fringes commonly present in absorption imaging and reduce the effect of photon shot noise by a factor 1.9 in variance, without affecting the atomic distribution. We derive the Cramér-Rao lower bound for the detection of total and relative atom numbers, and the corresponding maximum likelihood estimator to optimally extract information from our image data. We apply these techniques to the measurement of the relative atom number in a double well system, created on our magnetic lattice atom chip. Here, we show atom shot noise limited detection down to a total atom number of 270, a factor 9 lower than would be possible for the same data without our analysis. This corresponds to a minimum resolvable population difference of 17 atoms/shot. The techniques are directly applicable to atom interferometry experiments, where a population difference can be used to extract the interferometric phase. There, we expect our analysis would provide similar improvements, enabling atom shot noise limited readout and the study of squeezing and entanglement in smaller ensembles. For realistic experimental conditions, we expect a detection sensitivity of less than one atom per shot in a single trap to be feasible with absorption imaging.

We use our improved detection techniques to probe atom number fluctuations in a lattice of atomic ensembles, created on a magnetic lattice atom chip. In the tightly confining micro-traps, density-dependent three-body loss naturally reduces fluctuations to below the Poisson level. We develop a stochastic model for atom loss due to multi-body processes, providing the evolution of both the mean atom number and the fluctuations. We directly measure sub-Poissonian fluctuations in individual traps, and employ fluctuation-correlation analysis to measure a Fano factor of $F = 0.53 \pm 0.08 \pm 0.2$ (statistical, systematic uncertainty), in good agreement with the expected $F = 3/5$ for three-body loss.

Our lattice of microtraps exhibiting suppressed atom number fluctuations provides a basis for the study of quantum information science with mesoscopic ensembles. In a separate work, our experiment has been used to probe Rydberg-surface interaction through electromagnetically induced transparency using coherent excitation to Rydberg states [57]. In further studies, coherent coupling to Rydberg states could be used to study collective excitations in a lattice of traps, and to investigate coherent coupling between otherwise independent lattice sites using the long-range Rydberg dipole-dipole interaction. In future improvements of the experimental apparatus, we intend to use magnetic lattices with square and triangular symmetries with potentials optimised for
strong confinement \cite{6}. Combining smaller lattice periods with high resolution imaging through an in-vacuum objective lens, we expect to reach a regime of strong nearest neighbour interactions with Rydberg states while maintaining optical resolution of single sites and improving the single trap detection limit to \(\sim 1\) atom per shot.
\section{Propagation of variances}

Here, we investigate variance of a derived quantity, which is a function of several parameters. We first consider the general covariance between two derived quantities, and then write the special case of variance. We apply the same result to obtain the Cramér-Rao bound for derived quantities.

Consider the functions $f(a_1, a_2, \ldots)$ and $g(b_1, b_2, \ldots)$, where $a_i$ and $b_j$ are measured quantities with means $\langle a_i \rangle$ and $\langle b_j \rangle$ respectively. Assuming a sufficiently large sample size, we can approximate $f$ by a linear expansion around its mean,

$$f(a_1, a_2, \ldots) \approx \langle f \rangle + \sum_i \frac{\partial f}{\partial a_i} (a_i - \langle a_i \rangle), \quad (A.1)$$

The covariance $\text{cov}(f, g) = \langle fg \rangle - \langle f \rangle \langle g \rangle$ is then

$$\text{cov}(f, g) = \sum_i \sum_j \frac{\partial f}{\partial a_i} \frac{\partial g}{\partial b_j} \text{cov}(a_i, b_j). \quad (A.2)$$

The variance of a single derived quantity can be found using $\text{var}(f) = \text{cov}(f, f)$. For example, for $f(a_1, a_2)$,

$$\text{var}(f) = \left( \frac{\partial f}{\partial a_1} \right)^2 \text{var}(a_1) + \left( \frac{\partial f}{\partial a_2} \right)^2 \text{var}(a_2) + 2 \frac{\partial f}{\partial a_1} \frac{\partial f}{\partial a_2} \text{cov}(a_1, a_2). \quad (A.3)$$

The Cramér-Rao bound for a derived quantity can be found using the same logic. Given a multi-parameter CRB $C$ for which the limit on the covariance matrix is written as

$$\text{Cov}(a_1, a_2, \ldots) \geq C, \quad (A.4)$$

the CRB for a function $f(a_1, a_2, \ldots)$ is

$$\text{var}(f) \geq \sum_i \sum_j \frac{\partial f}{\partial a_i} \frac{\partial f}{\partial a_j} C_{i,j}, \quad (A.5)$$

provided that $f$ itself is an efficient estimate, at least for large enough data sets.
B Fringe removal implementation

The following MATLAB code implements the fringe removal algorithm, using either LU decomposition or singular value decomposition.

function [optrefimages] = fringeremoval(absimages,refimages,bgmask,method)
% FRINGEREMOVAL - Fringe noise removal from absorption images.
% Creates an optimal reference image for each absorption image in a set as
% a linear combination of reference images with coefficients chosen to
% minimize least-squares residuals between absorption image and optimal
% reference image.
% The coefficients are obtained by solving a linear set of equations using
% matrix inverse by LU decomposition or singular value decomposition (SVD).
% Algorithm inspired by J. Kronjaeger PhD thesis (Hamburg-2007).
% Syntax:
% [optrefimages] = fringeremoval(absimages,refimages,bgmask,method);
% Required inputs:
% absimages - Cell array of absorption images
% refimages - Cell array of raw reference images
% The number of refimages can differ from the number of absimages.
% Optional inputs:
% bgmask - Array specifying background region used,
% 1=background, 0=data. Defaults to all ones.
% method - Method used for solving linear equations. String
% containing 'LU' (default) or 'SVD'.
% Outputs:
% optrefimages - Cell array or of optimal reference images,
% equal in length to absimages.
% Dependencies: none
% Authors: Shannon Whitlock, Caspar Ockeloen
% Reference: C. F. Ockeloen, A. F. Tauschinsky, R. J. C. Spreeuw, and
% S. Whitlock, Improved detection of small atom numbers through
% image processing, arXiv:1007.2136
% Email:
% May 2009; Last revision: 14 July 2010
% Process inputs
svd = strcmpi(method,'svd'); % Default to LU
nimgs = length(absimages);
nimgsR = length(refimages);
xdim = size(absimages{1},2);
ydim = size(absimages{1},1);
if not(exist('bgmask','var')); bgmask=ones(ydim,xdim); end
k = find(bgmask(:)==1); % Index k specifying background region

% Flatten absorption and reference images
R = double(reshape(cat(3,refimages{:}),xdim*ydim,nimgsR));
A = double(reshape(cat(3,absimages{:}),xdim*ydim,nimgs));

% Ensure there are no duplicate reference images
% R=unique(R','rows')'; % comment this line if you run out of memory

% Invert B = R*R' with the chosen method
if svd
    Binv = pinv(R(k,:)'*R(k,:)); % SVD through PINV
else
    [L,U,p] = lu(R(k,:)'*R(k,:),',vector'); % LU decomposition
end

optrefimages=cell(nimgs,1);
for j=1:nimgs
    b=R(k,:)'*A(k,j);

    % Obtain coefficients c which minimise least-square residuals
    if svd
        c = Binv*b;
    else
        lower.LT = true; upper.UT = true;
        c = linsolve(U,linsolve(L,b(p,:),lower),upper);
    end

    % Compute optimised reference image
    optrefimages{j}=reshape(R*c,[ydim xdim]);
end
C Fluctuation-correlation analysis

The lattice-averaged Fano factor can be obtained from the fluctuation correlation function, defined by

\[
\chi_i(\delta) = \int [\tilde{n}_i(x) - \langle \tilde{n}_i(x) \rangle] [\tilde{n}_i(x - \delta) - \langle \tilde{n}_i(x - \delta) \rangle] \, dx. \tag{C.1}
\]

Here, \(\tilde{n}_i(x)\) is a measured optical density image of the lattice, where \(x\) is the spatial coordinate, \(i\) labels a realisation of the experiment, and \(\langle \ldots \rangle\) denotes averaging over all realisations with the same experimental parameters.

Model

We model our data by taking an infinite lattice of traps with identical normalised point spread functions \(p(x)\), corrupted by additive Gaussian noise \(d_i(x)\), uncorrelated from pixel to pixel and with a mean of 0. The model distribution is

\[
\tilde{n}_i(x) = c_i \sum_m N_{i,m} p_m(x) + d_i(x), \tag{C.2}
\]

where \(c_i\) accounts for correlated noise (for example due to probe frequency noise), \(N_{i,m}\) is the number of atoms in trap \(m\) and \(p_m(x) = p(x - m\lambda)\) with \(\lambda\) the trap spacing. \(N_{i,m}\) is uncorrelated in \(i\) and \(m\). We write for the mean distribution

\[
\langle \tilde{n}_i(x) \rangle = \langle c \rangle \sum_m \langle N_m \rangle p_m(x). \tag{C.3}
\]

Mean fluctuation correlation image

To reduce noise, we average fluctuation correlation images over realisations. In \(\langle \chi(\delta) \rangle\), we expect three distinct features. At \(\delta = 0\), a delta-function peak will occur, due to the uncorrelated photon shot noise. It is surrounded by a broader peak, with a width determined by the point spread function, containing the correlations of each trap with itself. At spacings of \(m\lambda\), there will be peaks due to correlations between neighbouring traps. Taking the mean of (C.1), we have

\[
\langle \chi(\delta) \rangle = \int \langle \tilde{n}(x) \tilde{n}(x - \delta) \rangle dx - \int \langle \tilde{n}(x) \rangle \langle \tilde{n}(x - \delta) \rangle dx. \tag{C.4}
\]

We are not interested in the uncorrelated imaging noise, and exclude the point \(\delta = 0\) from our analysis. For \(\delta \neq 0\), we substitute the model to find

\[
\langle \chi(\delta) \rangle = \int \langle c^2 \rangle \left( \sum_m N_{i,m} p_m(x) \sum_l N_{i,l} p_l(x - \delta) \right) dx - \int \langle c \rangle^2 \left( \sum_m N_{i,m} p_m(x) \right) \left( \sum_l N_{i,l} p_l(x - \delta) \right) dx. \tag{C.5}
\]
Here, we used that $\langle d(x) \rangle = 0$ and that $\int d_i(x)d_i(x - \delta)dx = 0$ for $\delta \neq 0$.

The amplitude of the central peak, excluding the photon shot noise, is found by taking $\langle \chi_0 \rangle = \langle \chi(\delta \to 0) \rangle$. Changing the order of sums, integrals and averages, and separating the terms where $l = m$ in the double sum, we obtain

$$\langle \chi_0 \rangle = \left( \langle c^2 \rangle - \langle c \rangle^2 \sum_m \langle N_m \rangle \right) \int p^2(x)dx + \left( \langle c^2 \rangle - \langle c \rangle^2 \right) \Xi, \quad (C.6)$$

where the integral $\int p^2(x)dx$ is independent of $m$ and $\Xi$ accounts for overlap between neighbouring traps,

$$\Xi = \sum_m \sum_{l \neq m} \langle N_m \rangle^2 \int p_m(x)p_n(x)dx. \quad (C.7)$$

All other peaks in $\langle \chi(\delta) \rangle$ are essentially identical, assuming the overlap between neighbouring sites is small. We find their amplitude by taking $\delta = \lambda$ in equation $\langle C.5 \rangle$.

Because of the small overlap, we can assume the two sums in the first term to be independent, and write

$$\langle \chi_\lambda \rangle = \left( \langle c^2 \rangle - \langle c \rangle^2 \right) \int \left( \sum_m \langle N_m \rangle p_m(x) \right)^2 dx$$

$$= \left( \langle c^2 \rangle - \langle c \rangle^2 \right) \left( \sum_m \langle N_m \rangle^2 \int p^2(x)dx + \Xi \right). \quad (C.8)$$

**Lattice-averaged Fano factor**

We define the lattice-averaged Fano factor $\overline{F}$, where the average is weighted by the mean number of atoms per cite $\langle N_m \rangle$, as

$$\langle F \rangle = \frac{\sum_m (\langle F_m \rangle \langle N_m \rangle)}{\sum_m \langle N_m \rangle} = \frac{\sum_m \langle N_m^2 \rangle - \sum_m \langle N_m \rangle^2}{\sum_m \langle N_m \rangle}. \quad (C.9)$$

The lattice-averaged Fano factor can be obtained directly from the fluctuation correlation images. We divide equations $\langle C.6 \rangle$ and $\langle C.8 \rangle$ both by $\sum_m N_m$, and substitute in equation $\langle C.9 \rangle$ and $c = \langle c^2 \rangle - \langle c \rangle^2$. Solving the resulting system of equations for $\langle F \rangle$ and $c$, we find

$$\overline{F} = \frac{\langle \chi_0 \rangle - \langle \chi_\lambda \rangle}{\langle \chi_0 \rangle + P_0} \left( N \right) \left( 1 + \frac{\sum_m \sum_{l \neq m} \int p_m(x)p_n(x)dx}{\int p^2(x)dx \sum_m \langle N_m \rangle} \right), \quad (C.10)$$

where $\langle N \rangle = \sum_m \langle N_m \rangle^2 / \sum_m \langle N_m \rangle$ is the weighted average atom number and

$$P_0 = \langle c \rangle^2 \int \sum_m \sum_l \langle N_m \rangle \langle N_l \rangle p_m(x)p_l(x)dx, \quad (C.11)$$

which can be measured directly from the data in the autocorrelation function of $\langle \tilde{n} \rangle$, by measuring the amplitude of the broad central peak, in the same way as for $\langle \chi_0 \rangle$. The overlap correction term can be obtained from estimated individual trap point spread functions.
Bibliography


Nederlandse samenvatting

Ultrakoude neutrale atomen, deeltjes die zijn afgekoeld tot vlak bij het absolute nulpunt, bieden een veelbelovend uitgangspunt voor moderne quantumtechnieken, zoals atoomklokken, nauwkeurige metingen van natuurconstanten en in de toekomst mogelijk quantumcomputers, waarmee bepaalde berekeningen kunnen worden gedaan in een fractie van het aantal stappen nodig op een klassieke computer. Hiervoor is het nodig de atomen te lokaliseren, gecontroleerd interacties tussen quantumtoestanden aan te kunnen sturen en nauwkeurig de quantumtoestand van de deeltjes te meten.

Wij gebruiken een atoomchip om atomen te vangen in een rooster van koude gaswolkjes. De chip bestaat uit een metalen plaatje, opgehangen in vacuum, met op de chip stroomdraden en een magnetisch laagje met een roosterstructuur (figuur 1a). De magneetvelden die hierdoor worden opgewekt maken het mogelijk om in het vacuum een rooster van honderden magneetvallen te maken, op 10 μm van het oppervlak, en met ongeveer 20 μm tussen de vallen. In de vallen worden wolkjes Rubidiumatomen gevangen en afgekoeld tot ∼1 μK, een miljoenste graad boven het absolute nulpunt, waarbij er slechts enkele tientallen tot honderden atomen per wolkje zijn. Deze kleine gaswolkjes bieden een mogelijk uitgangspunt voor een quantumcomputer, door gebruik te maken van Rydbergtoestanden. Dit zijn quantumtoestanden waarbij het elektron in een hoge schil wordt gebracht, waardoor er sterke elektrische interacties tussen atomen ontstaan.

Centraal bij kleine aantallen deeltjes staan echter altijd fluctuaties in het aantal deeltjes, tussen verschillende vallen en tussen herhalingen van het experiment. Het vangen en afkoelen van de atomen bestaat uit statistische processen, waardoor er een kansverdeling ontstaat voor het aantal deeltjes per val. Als bij deze processen alle atomen onafhankelijk elkaar zijn, zoals bij lage dichtheid het geval is, dan wordt de verdeling van atoomaantallen op zijn best (d.w.z. voor een technisch ideaal experiment) een Poissonverdeling, waarbij bij N deeltjes per val de fluctuaties een standaardafwijking van √N hebben. De relatieve fluctuaties zijn daardoor groter bij kleine atoomaantallen.

Figuur 1: Atoomchip met magnetisch rooster. (a) Schets van het centrale deel van de de atoomchip. Wolkjes atomen worden gevangen boven de randen van het magnetische patroon. (b) Absorptie-afbeelding van het rooster, geladen met atomen. Ongeveer 500 wolkjes zijn geladen, met in het centrale deel ongeveer 2500 atomen per wolkje.
Voor een quantumcomputer met kleine atoomwolkjes, zouden gereduceerde fluctuaties een beter uitgangspunt zijn.

In dit onderzoek worden de fluctuaties in aantal atomen onderzocht bij kleine atoomwolkjes. Om fluctuaties te kunnen meten, is het nodig het aantal atomen in elk wolkje nauwkeurig te kunnen meten. We meten de atomen door absorptie van laserlicht af te beelden op een camera, met voldoende resolutie om de afzonderlijke wolkjes te kunnen onderscheiden (figuur 1b). We verbeteren deze standaardtechniek door optimale data-analyse, waarmee we alle informatie uit elke absorptieafbeelding halen. Hierdoor kunnen we meten met een onzekerheid van ongeveer 10 atomen per wolkje, een factor 3 beter dan zonder onze analyse. Deze technieken zijn toepasbaar in een breed scala aan experimenten waarbij met kleine aantallen atomen wordt gewerkt.

In ons rooster gebruiken we drie-deeltjesverlies om fluctuaties in het aantal atomen per wolkje te reduceren tot minder dan een Poissonverdeling. Voor de relatief hoge dichtheden in onze magneetvallen is het belangrijkste proces waarmee atomen de trap verlaten een drie-deeltjesreactie, waarbij een molecuul gevormd wordt en alle drie deeltjes tegelijk de val verlaten. De kans dat dit gebeurt hangt af van de kans dat drie deeltjes bij elkaar in de buurt komen, en dus van de dichtheid. Een val die aan het begin van het experiment meer atomen dan gemiddeld bevat, heeft een hogere dichtheid en zal dus sneller atomen verliezen. Op deze manier wordt het aantal atomen in elke val dichter bij het gemiddelde gebracht, en worden de fluctuaties teruggedrongen tot minder dan een Poissonverdeling. In ons rooster meten we een verdeling met een standaardafwijking van \( \sqrt{0.53 \pm 0.22} \) bij gemiddeld \( N \) atomen, significant minder dan een Poissonverdeling en in overeenkomst met een theoretische verwachting van \( \sqrt{0.6N} \).

Ons rooster biedt, dankzij gereduceerde fluctuaties, een goed uitgangspunt voor een quantumcomputer met kleine atoomwolkjes. Dankzij geavanceerde analyse van absorptieafbeeldingen kunnen we kleine atoomaantallen meten in elk wolkje. Met toekomstige verbeteringen aan het experiment verwachten we de detectie verder te kunnen optimaliseren, tot een nauwkeurigheid van beter dan één atoom per wolkje. Deze detectietechnieken kunnen gebruikt worden voor het uitlezen van quantumtoestanden van de wolkjes, en bieden de mogelijkheid om bijvoorbeeld atoomklokken met gevangen atomen beter uit te lezen.