Sub-atom shot noise Faraday imaging of ultracold atom clouds

M A Kristensen¹, M Gajdacz¹, P L Pedersen¹, C Klempt², J F Sherson¹, J J Arlt¹ and A J Hilliard¹

¹ Institut for Fysik og Astronomi, Aarhus Universitet, Ny Munkegade 120, DK-8000 Aarhus C, Denmark
² Institut für Quantenoptik, Leibniz Universität Hannover, Welfengarten 1, D-30167 Hannover, Germany

E-mail: mick@phys.au.dk

Received 24 August 2016
Accepted for publication 15 November 2016
Published 17 January 2017

Abstract
We demonstrate that a dispersive imaging technique based on the Faraday effect can measure the atom number in a large, ultracold atom cloud with a precision below the atom shot noise level. The minimally destructive character of the technique allows us to take multiple images of the same cloud, which enables sub-atom shot noise measurement precision of the atom number and allows for an in situ determination of the measurement precision. We have developed a noise model that quantitatively describes the noise contributions due to photon shot noise in the detected light and the noise associated with single atom loss. This model contains no free parameters and is calculated through an analysis of the fluctuations in the acquired images. For clouds containing ~\(N\approx 10^6\) atoms, we achieve a precision more than a factor of two below the atom shot noise level.

Keywords: non-destructive imaging, sub atom shot noise detection, ultracold atom clouds, Faraday rotation, dispersive imaging

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

1. Introduction

Dispersive imaging is the standard method to perform spatially resolved in situ measurements on trapped ultracold atom clouds. The atomic density distribution gives rise to a spatially varying refractive index for off-resonant imaging light and the dispersive imaging technique converts the acquired phase shift into an intensity modulation that can be measured on a scientific camera. Imaging with off-resonant light permits a sample to be probed several times due to the low rate of spontaneous emission. This minimally destructive character of the imaging is typically used to measure the cloud’s evolution as it undergoes some dynamic process, such as the formation of a Bose–Einstein condensate (BEC) [1–3]. It has also been used to measure the speed of sound [4] and to directly observe vortices [5] in a Bose gas. Most recently, dispersive imaging was used to observe the quantum enhancement of the refractive index in a BEC [6].

Here, we investigate the use of repeated dispersive images to precisely determine the number of atoms in an ultracold thermal cloud. We make use of up to 100 images of a single cloud to enhance the measurement of the atom number. Another significant benefit of this approach comes from analyzing the statistics of the repeated measurements, which provides an ‘in-measurement’ characterization of the imaging noise. There are two fundamental noise sources that limit the precision in determining the atom number: the photon shot noise on the detected light and the noise associated with atom loss from spontaneous photon scattering induced by the imaging light. We develop a noise model that quantitatively describes these noise contributions. This model contains no free parameters and is based on an analysis of the fluctuations in the acquired images. For clouds containing ~\(N\approx 5\times 10^6\) atoms, a single dispersive image achieves a precision of approximately four times the atom shot noise level \(\sqrt{N}\). When averaging over only 10 images of the same cloud, our imaging achieves sub-atom shot noise precision.

The dispersive imaging technique we employ is based on the Faraday effect, whereby the linear polarization plane of an off-resonant imaging beam is rotated due to the vector part of the atomic polarizability. In earlier work, we compared the signal-to-noise ratio for four common dispersive imaging
Techniques [7]: phase contrast [4] and dark field imaging [1], which typically make use of the scalar part of the atomic polarization, and two variants of Faraday imaging [8, 9].

However, our experiments and analysis investigate a fundamental feature of any dispersive imaging method: the interplay between the acquired phase shift and the concomitant spontaneous photon scattering that must occur [10]. In this sense, dispersive imaging of a non-degenerate gas occupies an interesting position between the paradigms of classical and quantum measurements: the trapped cloud initially contains exactly \( N_0 \) atoms, but the imaging induces loss in the cloud through spontaneous photon scattering, causing a decrease in \( N \) and an increase in the atom number uncertainty \( \Delta N \) due to the stochastic character of the loss process.

The paper is structured as follows. In the following section, our realization of Faraday imaging is outlined, and in section 3 we give details on the experimental setup. Section 4 describes the processing we perform on the raw data to achieve high quality Faraday images. We also discuss the steps we take to obtain accurate atom numbers through a cross-calibration of Faraday imaging with the more common absorption imaging. In section 5, we analyze the achieved precision in the determination of the atom number in a single Faraday image. We then extend this analysis to the case of averaging over several images. We offer conclusions on our work in section 6.

2. Dark field Faraday imaging

Faraday imaging exploits circular birefringence in spin-polarized atomic samples: a linearly polarised off-resonant beam of light may be decomposed into an equal superposition of left- and right-hand circular polarizations; as the beam passes through the atomic sample, the light–matter interaction leads to a differential phase shift of the two circular components of the light leading to a rotation of the linear polarization plane. For an ultracold cloud of \(^{87}\)Rb in the \( |F = 2, m_F = 2 \rangle \) state that is spin polarised along the \( z \) axis and probed with light close to the D2 transition, the Faraday rotation angle is given by

\[
\theta(x, y) = \frac{\hbar m_F \Gamma^2}{16\pi \Delta_{\text{eff}}} \int n(x, y, z) dz \equiv c_F(\Delta_{\text{eff}}) \tilde{n}(x, y),
\]

where \( n(x, y, z) \) is the atomic density distribution, \( \lambda = 780.241 \text{ nm} \) is the wavelength and \( \Gamma = 2\pi \times 6.067 \text{ MHz} \) is the natural linewidth. The effective detuning of the imaging beam from the three possible optical transitions \( F = 2 \rightarrow F' = 1, 2, 3 \) weighted by their transition strengths is given by \( \Delta_{\text{eff}} = (28/\Delta_{2,2} - 5/\Delta_{2,3} - 3/\Delta_{2,1})/20 \). The far right-hand side of equation (1) gives the rotation angle as a product of the Faraday coefficient \( c_F \) and atomic column density \( \tilde{n}(x, y) \).

Figure 1 shows the key elements of the imaging system we use to measure this rotation angle. An off-resonant imaging beam initially linearly polarised along the \( y \) axis propagates along the quantization axis defined by a bias magnetic field oriented along the \( z \) axis. The spatial distribution of the atomic cloud is mapped onto the polarization plane of polarization. A polarizing beam splitter removes the unscattered light from the imaging path, and the transmitted light is imaged on a scientific camera. An additional imaging telescope between the PBS and camera is not shown.

3. Experimental setup

The atomic sample is produced by standard techniques [11, 12]. We first generate a large cloud containing \( \sim 10^9 \) \(^{87}\)Rb atoms in a vapour pressure magneto-optical trap (MOT), whereupon the atoms are transferred into a magnetic quadrupole trap. The quadrupole trap is transported on a translation stage to a region of the vacuum chamber that is held at a lower pressure, where the atoms are transferred to a Ioffe–Pritchard (IP) magnetic trap [12, 13]. The IP trap has axial and radial trapping frequencies \( \omega_z = 2\pi \times 17.1 \text{ Hz} \) and \( \omega_r = 2\pi \times 296 \text{ Hz} \), respectively, with an axial bias field of 330 mG. Forced radio-frequency (RF) evaporative cooling is used to produce ultracold clouds. For the characterization of Faraday imaging presented here, we produce clouds in the \( |F = 2, m_F = 2 \rangle \) state containing on average \( 5.8 \times 10^6 \) atoms at a temperature of 18 \( \mu \text{K} \).

The Faraday imaging light is off-resonant and applied in pulses, with an image acquired for each imaging pulse. It is generated by an external cavity diode laser, which is locked with a detuning \( \Delta_{2,3} = 2\pi \times (1200 \pm 1) \text{ MHz} \) to the blue of the \( F = 2 \rightarrow F' = 3 \) transition. The detuning is achieved by
offset-locking the imaging laser to a master laser, which is locked to a resonance in $^{87}$Rb by standard saturated absorption spectroscopy [14]. At the position of the atoms, the intensity distribution of the imaging light is Gaussian with a waist of 1.4 mm, leading to an approximately uniform intensity over the cloud. For the experiments presented here, we use an intensity of 0.5 mW cm$^{-2}$. The imaging pulses have rectangular temporal envelopes, which are generated by an acousto-optic modulator (AOM). The optical power of each imaging pulse is monitored on a photodiode (PD) on the reflecting port of the PBS (see figure 1). This signal is used to regulate the amplitude of the pulses through feedback to the AOM to achieve a relative error in the number of photons per pulse of approximately $4 \times 10^{-3}$. The Faraday rotated light that is transmitted through the PBS is imaged on an electron multiplying charge coupled device (EMCCD) camera. The imaging system has a magnification of 4.85.

Figure 2 shows the experimental sequence we use to characterize the precision of the Faraday imaging. After initial evaporative cooling, the RF frequency is held constant, producing a trap with an effective depth of 64.3 $\mu$K. A series of 100 Faraday images is taken with 7 ms between the start of each imaging pulse. The duration of each pulse can be varied to investigate the imaging precision as a function of the total time over which the atoms are probed. After the 100 Faraday images have been taken, the magnetic trap is extinguished and an absorption image is acquired after a time of flight.

4. Faraday image processing and accuracy

In this section, we describe the processing steps we perform to obtain high quality Faraday images and to determine the atom number. In principle, equation (1) provides a link between rotation angles and the column density of the atom cloud. However, even for the thermal clouds in this work, density dependent collective light scattering effects lead to deviations from this light scattering model [9]. For the imaging parameters used here and consistent with our earlier work [7], the atom number inferred from Faraday imaging is approximately 36% lower than the one from absorption imaging.

Of course, absorption imaging itself requires careful calibration to achieve accurate measured atom numbers. We calibrate the absorption imaging by the procedure described in [15], whereby a scaling factor for the effective saturation intensity of the imaging transition is empirically determined. An independent method to verify this calibration relies on measuring the scaling of the imaging noise as the atom number is varied [16].

Such calibration methods invariably make use of Poissonian statistics in some aspect of the imaging process. In our case, we use the fact that single particle loss from the trap is a Poissonian process [17], such that the atom number variance equals the mean number of atoms; i.e., $\Delta N^2 = N$. In recent work, we used Faraday imaging to probe the increase in atom number noise as a function of applied single particle loss [18]. Here, we use the single particle loss induced by the imaging itself as a way to link the mean atom number with the variance. Making use of the fact that imaging accuracy and precision are linked by the requirement of self-consistency is a generic approach [19].

To summarize, we achieve accurate atom numbers from Faraday imaging by calibrating the Faraday images with the results from absorption imaging [7, 15]; we then check this calibration for self-consistency by verifying the atom noise scaling of the Faraday images.

The light detected on the EMCCD camera scales with the Faraday 'signal' defined as

$$S \equiv \frac{I(\theta)}{I_0} = \sin^2(\theta).$$

Due to the finite extinction ratio of the PBS, (cube suppression) $CS = 1.5 \times 10^{-3}$, a small fraction of the non-rotated light leaks through the cube. This leads to an intensity after the PBS given by: $I(\theta) = I_0[\sin^2(\theta) + CS\cos^2(\theta)][1 + CS]^{-1}$ [7]. Experimentally, we calculate the signal in a region-of-interest (ROI) that encompasses the cloud as

$$S = \left(\frac{I(\theta)}{I_{\text{ref}}} - 1\right)\left(\frac{CS}{1 - CS}\right),$$

where $I_{\text{ref}} \equiv CS \cdot I_0$ is the non-rotated light leaking through the cube; i.e., we use $I_{\text{ref}}$ as a measure of $I_0$ in a given image so that we avoid having to take an additional picture without the atoms. We obtain $I_{\text{ref}}$ by taking the mean pixel count in a ROI outside the area covered by the atomic cloud. Both $I(\theta)$ and $I_{\text{ref}}$ are corrected for the baseline level of the camera, which is the mean pixel count evaluated in a masked area of the camera. Figure 3(a) illustrates the regions of interest where we measure the signal intensity $I(\theta)$, reference intensity $I_{\text{ref}}$ and baseline. Figure 3(d) shows the resulting evaluated signal in its ROI.

To evaluate the Faraday images, we can work directly with the signal $S$ or evaluate the distribution of rotation angles $\theta$ induced by the cloud.
An alternative approach is to work only with the total signal sum $\Sigma_S = \Sigma \Sigma_S$, where the sum runs over all pixels in the signal ROI. Since $\Sigma_S$ is a scalar function of $N$ and $T$, this processing method does not provide independent measures of the cloud’s parameters. In practice, however, $N$ and $T$ are highly correlated in an experimental run because the evaporative cooling relies on elastic collisions between atoms [22]. This means that $\Sigma_S$ can be calibrated through its correlation with $N$ and $T$ obtained from absorption imaging [18], and the calibration in figure 4(a), which is linear in the atom number. Despite these complications, we show in section 5 that working directly with the signal sum $\Sigma_S$ yields superior imaging precision over the angle sum $\Sigma_\varphi$ because we avoid the noise induced by model fitting. Additionally, this simple analysis is well-suited for implementation on a field programmable gate array (FPGA), which has allowed us to obtain real-time information on the cloud properties during individual experimental runs [18].

Figure 4(b) shows the signal sum $\Sigma_S$ in the 100th Faraday image as a function of the atom number extracted from the subsequent absorption image, for the same datasets as in figure 4(a). The data has been fitted by a quadratic function, which is motivated by the scaling of the signal sum in the limit of small Faraday rotation angles: for the thermal clouds employed in this work and our choice of detuning, the largest rotation angle (at the center of the cloud) is small, i.e., $\theta < 15^\circ$, such that the dependence of $\Sigma_S$ on atom number and temperature can be approximated by the first order series expansion of equation (2): $\Sigma_S \approx \theta^2 = (\kappa \bar n)^2$. Inserting a 2D Gaussian distribution for the column density yields the small angle dependence of the signal sum on the atom number and temperature: $\Sigma_S \propto N^2 / T$. The exact form of the quadratic fit function contains additional parameters used to model technical details in the imaging. The function is given by: $\Sigma_S = b_1(N - b_2) + b_3$. Here $b_2$ again models a lower bound in the detection of small rotation angles. The parameter $b_3$ models imaging beam inhomogeneity, which gives rise to an offset in the detected signal sum.

This approach is based on the simplifying approximation that the temperature of the sample is unchanged due to the imaging. For the range of imaging pulse durations we investigate here, we observe that the cloud temperature variation is less than 1.5%, as determined from the absorption

---

6 The Faraday rotation angles can also be reconstructed from the evaluated signal through $\theta = \arcsin(\sqrt{S})$ [21], but this approach can be problematic. First, since the baseline is subtracted, the pixel values where the signal is low can become negative, which complicates the evaluation of the square root. Secondly, the steep derivative of the square root for small arguments amplifies the noise in pixels with low signal.

---

#### Figure 3.
(a) Full frame of a Faraday Image. The ROIs for signal (innermost square), reference (outermost square, excluding a safety margin) and baseline (bottom square) are indicated. The greyed areas are discarded safety margins. (b), (c) Cross-sections of evaluated signal (blue points) with corresponding cross-sections of the fit (red line). (d) Evaluated signal from a single image.

---

#### Figure 4.
Accurate atom numbers can be obtained through cross-calibration of Faraday imaging with absorption imaging. (a) Angle sum as a function of the atom number obtained from absorption images (blue points) and fitted first order polynomial (red line). (b) Signal sum as a function of the atom number obtained from absorption images (blue points) and fitted second order polynomial (red line). (c) Ratio between atom number obtained from the angle sum using the calibration in (a) and the atom number obtained from the signal sum using the calibration in (b) as a function of imaging pulse number. The ratio is 1 for the 100th image, since both the signal sum and the angle sum have been calibrated with the atom number obtained from an absorption image taken directly after the final Faraday image.

If we work with the rotation angles, the cloud parameters are obtained by fitting with $S_{\theta} \sim \sin^2(\theta_{\theta})$, where $\theta_{\theta}$ is a 2D Gaussian distribution. This corresponds to the intrap column density distribution $\bar n$ for a thermal cloud characterized by an atom number $N$ and temperature $T$ [20], scaled by the Faraday coefficient $F$ (see equation (1)). Figures 3(b) and (c) show cross-sections of $\bar n$ and this fit. The area integral of the fitted distribution $\theta_{\theta}$ thus provides a measure of the total atom number, which we denote $^6 \Sigma_\varphi$. The exact form of the quadratic fit function contains additional parameters used to model technical details in the imaging. The function is given by: $\Sigma_S = b_1(N - b_2) + b_3$. Here $b_2$ again models a lower bound in the detection of small rotation angles. The parameter $b_3$ models imaging beam inhomogeneity, which gives rise to an offset in the detected signal sum.
imaging. In appendix A, we show that the heating induced by photon scattering from the imaging beam is negligible compared to the temperature of the sample, and that the dominant loss mechanism is spontaneous decay into states that are not trapped by the magnetic potential.

While the imaging does not influence the temperature of the sample, the cloud cools down during imaging due to free evaporation. Estimating the atom number based on the signal sum calibration in figure 4(b) alone, leads to a systematic underestimation of the atom number in the earlier images because the cloud is hotter there than in later images. This occurs due to the coupling of $N$ and $T$ which may be approximately understood from the small angle dependence of the signal sum.

Figure 4(c) shows the ratio of the atom number obtained from the signal sum using the fitted quadratic calibration function in figure 4(b) and the atom number evaluated from the angle sum using the fitted linear calibration function in figure 4(a). For the first image, the signal sum underestimates the atom number by 8% decreasing to zero for the 100th image. The atom number obtained from $S_\beta$ in the final image is accurate due to the back-to-back comparison with absorption imaging. The data has been fitted by a quadratic function, allowing us to correct for the underestimation and thus obtain an accurate atom number for each image from the signal sum alone.

To summarize this section, we employ two main methods to obtain simple measures of the atom number from the imaging light. In addition, the images suffer from technical noise in the camera and the imaging beam. The technical noise source, given that the imaging pulses are stochastic, is readout noise, we take advantage of the camera’s electron multiplying (EM) feature to amplify the detected signal above the readout noise. The EM amplifier consists of a multi-stage avalanche-type amplification, and leads to a multiplicative noise term increasing the photon shot noise by a factor approaching $\sqrt{2}$ [23, 24]. By using large ROIs to estimate $I_{\text{ref}}$ and the baseline, we ensure that the imaging precision is limited by the noise in the signal region. The application of EM-gain reduces the size requirement of the reference ROI.

5.1. Single image detection

Figure 5(a) shows the atom number measured with an imaging pulse duration of 0.66 ms as a function of imaging pulse number. The atom number has been obtained from the signal sum using the fitted quadratic function in figure 4(b), and corrected for the variation in temperature using the fitted quadratic correction in figure 4(c). The fluctuations of the atom number about the mean decay come from the light shot noise, the stochastic noise arising from atom loss, and potential technical noise. Given that the mean atom loss is a deterministic process, we can remove the decay by subtracting or dividing $N$ by the appropriate exponential curve. For simplicity, we normalize the atom number by its mean over several experimental runs and shift it to be centered on zero, yielding the ‘normalized atom number’

$$N_{ij} = \frac{N_{ij}}{\langle N_{ij}\rangle_{\text{runs}}} - 1,$$

where $i$ denotes the run number and $j$ denotes the image number in a given run. $\langle N_{ij}\rangle_{\text{runs}}$ is the mean atom number in

![Figure 5](image-url)
with this choice of scaling, the Poissonian is the number of electrons per $1/\sqrt{N}$ shows the same initial; continues to decrease as $\tau$ increases. For large $\tau$, the mean decay can be eliminated, such that $\sigma_N$ continues to decrease as the imaging pulse duration is increased; this noise approaches a constant offset (dotted line) for large $\tau$, which we attribute to technical noise in the imaging light. The total noise model (blue line) for the normalized atom number is comprised of the photon shot noise and technical noise contributions added in quadrature. The corresponding curve (red dash-dash-dot line) for the atom number is comprised of the photon shot noise, technical noise and mean atom loss noise contributions added in quadrature. The decay per image is so small that the stochastic atom loss noise is negligible. Error bars represent $1\sigma$ (standard deviation) statistical uncertainties.

the $j$th image, averaged over all runs in the dataset. Figure 5(b) shows the atom number traces from 5(a) in this normalized form, where the fluctuations about each trace give the relative uncertainty in $N$.

One of the most attractive experimental features of taking multiple images of the same cloud is the ability to use the statistics of the imaging series to provide an in situ measure of the imaging precision. To characterize the measurement noise, we evaluate the two-sample variance for all pairs of images along the individual traces. For the atom number, this is given by

$$\sigma_N^2 = \frac{1}{2} \langle (N_j-1 - N_j)^2 \rangle. \quad (5)$$

The two-sample variance can be evaluated for the atom number and normalized atom number traces shown in figures 5(a) and (b). The two-sample variance is insensitive to sufficiently slow trends or drifts in the measurements, but captures the relevant noise.

Figure 6 shows the relative two-sample deviation, for the atom number and normalized atom number as a function of imaging pulse duration $t$. The two-sample deviation is defined as the square root of the two-sample variance, and it is plotted relative to $\sqrt{N}$; with this choice of scaling, the Poissonian noise level is achieved at $\sigma/\sqrt{N} = 1$. Initially, the two-sample deviation of the atom number decreases as the imaging pulse duration is increased, but at 0.5 ms it reaches a minimum and then increases approximately linearly with $t$. The two-sample deviation of $N$ shows the same initial behavior, but does not increase for the range of $t$ we consider because the mean atom loss contribution to the noise has been removed through the normalization. This shows that in a single Faraday image, the imaging noise lies within a factor of 3 of the Poisson noise level. To gain insight into the contributions of the fundamental noise sources, we evaluate the photon shot noise and the stochastic noise due to loss of atoms: the detection of photons and the loss of atoms are examples of single particle events leading to Poisson statistics.

The photon shot noise can be evaluated directly from the Faraday images. The number of photoelectrons generated when an incident photon hits the CCD is Poisson distributed; i.e., the variance in the number of generated photoelectrons is equal to the mean number $N_{el}$. By error propagation, the uncertainty in the number of photoelectrons leads to an uncertainty in the detected signal given by [7]:

$$\sigma_{psn}^2 = \left( \frac{\partial N}{\partial \Sigma el} \right)^2 \sigma_{\Sigma el}^2 = \beta^2 \sigma_{el}^2,$$

where the error propagation coefficient is obtained from the calibrations in figures 4(b) and (c). It is evident from figure 6, that the photon shot noise is the dominant noise source for short pulses, but decreases as $1/\sqrt{t}$ as the pulse duration is increased.

Atom loss leads to an increase in the two-sample variance due to both the stochastic nature of the discrete particle loss and through the decay of the mean atom number. Assuming that only single-particle loss occurs, the two-sample variance due to loss is given by [25]

$$\sigma_{loss}^2 = \frac{N_0}{2\tau} + \frac{N_0^2}{2\tau^2}, \quad (6)$$

where $N_0$ is the atom number at the start of the imaging and $\tau$ is the lifetime during imaging, such that the mean atom number is given by $N = N_0 \exp(-t/\tau)$. The first term in equation (6) describes stochastic atom loss noise and the second describes the effect of mean decay on the two-sample variance. Atom loss due to residual thermalization and background losses gives rise to a small extra contribution of atom loss noise, leading to an offset in the mean atom loss noise curve in figure 6. The lifetime $\tau$ has been estimated by fitting the atom number traces. The second term can be eliminated by removing the mean decay from the traces as we do for the normalized atom number $N$. For all the pulse durations considered here, the stochastic noise in a single image is negligible.

Figure 6 shows that the photon shot noise $\sigma_{psn}$ and the atom loss noise $\sigma_{loss}$ plus a constant technical noise contribution $\sigma_{tech}$ account extremely well for the observed noise in single images. We emphasize that the photon shot noise and atom loss noise models include no fitted parameters: they are evaluated directly from the Faraday image traces. The dominant source of the technical noise is likely to be the
frequency stability of the detection laser: the laser has \( \sim 1 \) MHz linewidth in a 5 ms integration time, corresponding to \( \sim 1 \times 10^{-3} \) relative uncertainty in the detuning. By error propagation using equation (1), this corresponds to \( \sim 8 \times 10^{-4} \) relative uncertainty in the detected number of atoms; note that \( 8 \times 10^{-4} \cdot \sqrt{N} \approx 1.9 \) for \( N = 5 \times 10^6 \), in agreement with figure 6. The integration time for the line-width measurement is similar to the time between consecutive images, and we therefore expect this technical noise contribution to be independent of the pulse duration. The technical fluctuations in the imaging pulse amplitude at the \( 4 \times 10^{-3} \) level (see section 3) can be regarded as a next order effect, in that these fluctuations are present in both the signal and reference ROIs and hence accounted for in equation (3). The fluctuations in the imaging pulse amplitude manifest themselves instead as small variations in the mean atom decay rate for each imaging pulse.

In summary, we have investigated the noise properties in the individual Faraday images and characterized this noise with an imaging model that includes photon shot noise, atom loss noise and technical noise associated primarily with the imaging laser linewidth. The noise arising from stochastic atom loss is negligible for our imaging parameters, such that only the mean atom loss noise is manifest in the two-sample variance. The optimal imaging pulse duration of approximately 0.45 ms corresponds to the point where the photon shot noise and mean atom loss terms are equal, as one would expect when adding two quantities in quadrature. For this optimal setting, we achieve a single image noise level that is approximately five times the atom shot noise level. If we instead work with the normalized atom number \( N \), such that the mean atom loss noise contribution is absent, the imaging noise is reduced by 20% and approaches the technical noise floor for longer imaging pulses.

5.2. Multi-image detection

While the single image noise decreases as the pulse duration is increased when the data is corrected for the mean atom loss, the pulse duration is in practice limited by the dynamic range of the camera. The problem of finite dynamic range is exacerbated in our experimental realization by the fact that significant EM-gain is necessary to render the readout noise negligible when imaging small rotation angles. To artificially increase the dynamic range of the camera, we can exploit the ‘non-destructive’ nature of our detection, and take multiple images and average over the result. For uncorrelated noise, the standard error of the mean decreases with the number of images \( n \) as \( 1/\sqrt{n} \). This is the case for the shot noise and a fair assumption for the technical noise \( \sigma_{\text{loss}} \). It is not the case, however, for stochastic atom loss noise: this noise is correlated because the width of the atom number distribution reflects the random walk in atom number during all previous images.

To obtain experimental measures of the detection precision when averaging over several images, we break the data used for the single image characterization into two halves: the first set contains images 1:50 and the second set contains images 51:100. The number of images over which we average can be varied for \( 1 \leq n \leq 50 \). The normalized atom number averaged over images \( [50 - (n + 1), 50] \) is denoted by \( N_{150}^2 \), and over images \( [51, 51 + (n - 1)] \) by \( N_{51}^2 \). We focus on an intermediate imaging pulse duration, \( t = 0.88 \text{ ms} \).

Figure 7(a) shows \( N_{50}^2 \) as a function of \( N_{2}^1 \), i.e., we compare the results of the first and second image set where the average is performed over all 50 images. The 10% standard deviation in the initial atom number for this data set (determined from the first Faraday image in each run) is due to fluctuations in the experimental apparatus. It is evident that the normalized atom numbers in the first and second image sets are highly correlated. To extract the imaging noise, we fit a quadratic function to this data and evaluate the deviations from the fit, which we denote \( \Delta N_{50}^2 \). A quadratic fit is used because the atom number extracted from the signal retains vestiges of the temperature dependence of \( \Sigma_0 \) (see section 4). However, this systematic calibration effect is extremely small, as is clear from the data, and a linear fit yields comparable results.

The need for the two-sample variance in order to characterize the imaging noise while rejecting drifts in the experimental apparatus is illustrated in figure 7(b), which
shows $\Delta N^{50}$ as a function of run number. These deviations exhibit the noise we wish to determine, on top of a slight upward trend. This trend is caused by environmental drift that is slow compared to the $\sim 1$ minute duration of an experimental run. The primary cause of drift in the experiment is currently $\sim 0.1 ^\circ \text{C}$ variations in the water used to cool the magnetic trap: as the temperature of the cooling water drifts, the magnetic trap coils expand and contract slightly, leading to small drifts in the trapping parameters. In particular, the magnetic trap bottom drifts with respect to the RF knife that determines the height of the trapping potential, leading to slight changes in the atom loss rate due to free evaporation.

Figure 7(c) shows the measured imaging noise as a function of the number of images included in the average. The blue line shows the measured imaging noise, evaluated by calculating the two-sample variance on $\Delta N^{50}$, ordered by run number, for $1 \leq n \leq 50$. The shaded region denotes the uncertainty in the measured noise calculated by bootstrapping [26]. As expected, the noise is seen to decrease with the number of images $n$ used in the average.

The measured imaging noise is well described by a noise model comprised of contributions from the photon shot noise, stochastic atom loss noise and technical noise, evaluated to include the effect of averaging. For the photon shot noise and technical noise, we use the results from the single image precision, scaled by $1/n$. A noise contribution $\sigma_{\text{stoch}}$ to describe the effect of stochastic atom loss when averaging over several images is derived in appendix B. The noise model with these contributions is given by $\sigma^2 = \frac{\sigma_{\text{ph}}^2}{n} + \sigma_{\text{tech}}^2/n + \sigma_{\text{stoch}}^2$. The agreement between the noise model and the data is excellent, given that the model contains no free parameters: it is calculated directly from the statistics of the experimental data and an independent estimate of the imaging light frequency stability. The experimental noise and model plotted in figure 7(c) are again shown as the square root of the two-sample variance, scaled by the mean atom number in images 51 and 52. With this presentation, it is evident that we achieve sub-atom shot noise detection for averages over 10 or more images. Additionally, by reducing the imaging laser linewidth to a realistic value of $\sim 100$ kHz, the technical contribution to the imaging noise could be rendered negligible and we could attain an imaging noise level of $\sigma/\sqrt{N} \approx 0.4$.

The same imaging performance analysis can be performed using $\Sigma D$, whereby the atom numbers are extracted by directly fitting the signal distributions in the Faraday images. The result is shown as the red (fine dashed) line and shaded region in figure 7(c). We also obtain excellent precision with this method, but as expected the fitting routine introduces extra uncertainty, and we do not achieve sub-shot noise limited detection. Interestingly, it is beneficial to average over more images when using this approach. Finally, we note that this approach, up to a scaling factor, allows the total atom number to be evaluated with high precision even in cases where cross-calibration with absorption images is not possible.

6. Conclusion

In this work, we have demonstrated that a simple dispersive imaging technique can measure the atom number in a large, ultracold cloud with a precision more than a factor of two below the atom shot noise level. The minimally destructive character of the technique allows us to take multiple images of the same cloud. The benefit of this is two-fold. First, by averaging over several images, we improve the precision of the atom number measurement. Second, by analyzing the statistics of multiple measurements, we obtain an in situ estimate of the measurement precision. To understand the achievable precision of the method, we have developed an imaging noise model. This model has two fundamental contributions: the photon shot noise on the detected light and the noise associated with the single atom loss that is induced by the imaging. Through our analysis, we obtain the size of these two contributions directly from the images, in contrast to previous work in which they were estimated from model fitting [18]. We find that an additional technical noise contribution is necessary to quantitatively describe the imaging noise. We ascribe the origin of this noise to the finite linewidth of the imaging laser, and will in future work render this contribution negligible with an improved laser system.

By comparing the atom number from Faraday imaging with that obtained from independently calibrated absorption imaging, we achieve accurate as well as precise atom counting. The accuracy of this approach is supported by the quantitative agreement of the noise model with the observed imaging noise: the atom loss contribution to the noise model reflects the Poissonian statistics of the atom number distribution, thereby providing a link between the mean and the variance of of atom number.

Acknowledgments

We acknowledge support from the Danish Council for Independent Research, the European Research Council, and the Lundbeck Foundation. CK acknowledges support from the
Appendix A. Atom loss during imaging

The observed loss of atoms due to imaging is caused by spontaneous photon scattering. There are two main processes that arise from this scattering to consider: recoil heating inducing loss due to finite trap depth, and decay into untrapped states. To determine the size of these contributions, we first estimate the number of scattered photons per atom. To simplify the analysis, we focus on an intermediate imaging pulse duration $t = 0.66$ ms. This pulse duration leads to an 11% loss of atoms over 50 images (see figure 5(a)).

Since the bias magnetic field is oriented along the probing direction, the linearly polarized imaging light is most easily described as an equal superposition of $\sigma_+$ and $\sigma_-$ circular polarizations. The scattering of photons on the $\sigma_+$ transition does not lead to a change in the atom’s internal state, since by selection rules the atom must decay back to $|F = 2, m_F = 2 \rangle$ as seen in figure 8. The scattering rate on this cycling transition is given by

$$ R_{\text{scat},+} = \frac{\Gamma}{2 + \frac{1}{1 + I_s/I_{\text{sat}}} + 4(\Delta_{2,3}/\Gamma)^2}, $$

where the saturation intensity of the transition is $I_{\text{sat}} = 1.669$ mW cm$^{-2}$, and the other quantities were introduced in sections 2 and 3. The light intensity of the $\sigma_+$ polarized light is $I_s = I_0/2$, where the total light intensity typically used in our experiment is $I_0 = 0.5$ mW cm$^{-2}$. For 50 images, the number of scattered $\sigma_+$ photons per atom is

$$ N_{\text{ph},+} = 50 \cdot t \cdot R_{\text{scat},+} \approx 0.59. $$

To estimate the number of photons scattered by the $\sigma_-$ light, and with a view to calculating the probability for an atom to be optically pumped to an untrapped state, we must calculate the scattering rate into the three possible excited states: $|F'=j, m_F=1 \rangle$, $j=1, 2, 3$. Depending on the excited state, spontaneous decay can occur into the $F = 1$ and $F = 2$ ground state manifolds. The decay probability $P_{jk}$ for the $k = 1, \ldots, 5$ final states is proportional to the square of the matrix element for the transition

$$ P_{jk} \propto \left| \begin{array}{l} k \\ 1|F = 1, m_F = 0 \rangle \\ 2|F = 1, m_F = 1 \rangle \\ 3|F = 2, m_F = 0 \rangle \\ 4|F = 2, m_F = 1 \rangle \\ 5|F = 2, m_F = 2 \rangle \\ \end{array} \right| \frac{\Gamma'}{j} $$

$$ \text{with}\quad P_{j'} = j $$. (A3)

In this notation, the strength of the $\sigma_-$ transition is $P_+ \propto 1/2$. To estimate the number of photons scattered on the individual $\sigma_-$ transitions, $N_{\text{ph},j}$, we also need to take into account the detuning. In the off-resonant limit $(\Delta/\Gamma)^2 \gg 1$, we can approximate the number of scattered photons as

$$ N_{\text{ph},j} \approx N_{\text{ph},+} \frac{(\Delta_{2,3}/\Delta_{2,j})^2 P_{j'.5}/P_{j'.1}}{P_{j'.5}/P_{j'.1}}. $$

The total number of photons scattered on the $\sigma_-$ transition is then

$$ N_{\text{ph},-} = \sum_{j=1}^{3} N_{\text{ph},j} \approx 0.032 + 0.066 + 0.039 = 0.137, $$

which is about a factor of four lower than $N_{\text{ph},+}$.

Thus, the total number of photons scattered during an imaging series comprised of 50 pulses with $t = 0.66$ ms is given by $N_{\text{ph},+} + N_{\text{ph},-} \approx 0.7$. Assuming that each (absorption and emission) scattering event heats an atom by the recoil temperature $T_R \approx 0.36 \mu K$, the increase in temperature is small compared to the temperature of the cloud: with the initial temperature of the cloud at $18 \mu K$, this corresponds to a temperature increase of about 1.5%. The loss of atoms through recoil heating and evaporation is therefore much smaller than the observed 11% atom loss in the experiment.

We will now estimate the loss due to state transfer. As shown in (A3), the $\sigma_-$ light can optically pump the atoms into three possible excited states. Decay into any state but $|F = 2, m_F = 1 \rangle$ and $|F = 2, m_F = 2 \rangle$ leads to the loss of an atom from the magnetic trap, whereby the number of scattering events leading to loss can be estimated as

$$ N_{\text{ph},\text{loss}} \equiv \sum_{j=1}^{3} \sum_{k=1}^{5} P_{j,k} \approx 0.092. $$

This result is consistent with the observed loss of $\sim 11\%$ when one takes into account the additional residual background loss of approximately 2% in the imaging time.

Appendix B. Stochastic noise from atom loss when averaging over several images

To see how the stochastic noise grows as the number of images increases, we assume for the moment that the images are independent random variables with variance $\sigma^2 = N_0/\langle N \rangle$.

The mean atom number $\langle N \rangle$ in the $k$th image is

$$ \langle N \rangle = \langle N_{k-1} \rangle p = N_0 p^k. $$

Due to our assumption that stochastic atom loss is the only noise source, $N_0$ denotes the initial ‘true’ atom number. The loss of atoms leads to a stochastic step $\Delta N_k$ during every image, which results in a random walk away from the mean behavior. The actual atom number detected in the $k$th image is therefore

$$ N_k = N_{k-1}p + \Delta N_k = N_0 p^k + \sum_{j=1}^{k} \Delta N_j. $$

$\Delta N_j$ are independent random variables with variance $\sigma^2 = \langle \Delta N \rangle = N(1 - p)$ and mean $\langle \Delta N \rangle = 0$ (see equation (6)).
This realizes a modified Poisson distribution displaced by the mean value to be centred around 0. Estimating the atom number at the beginning of the imaging series from the \( k \)th image, \( \bar{N}_{0,k} \), we obtain
\[
\bar{N}_{0,k} = N_k p^{-k} = N_0 - p^{-k} \sum_{j=1}^{k} \Delta N_j. \tag{B3}
\]
The mean over all \( n \) images is thus
\[
\bar{N}_0 = \frac{1}{n} \sum_{k=1}^{n} \bar{N}_{0,k} = N_0 - \frac{1}{n} \sum_{k=1}^{n} p^{-k} \sum_{j=1}^{k} \Delta N_j. \tag{B4}
\]
The order of summation can be swapped since \( j \leq n \) and \( k \geq j \). Defining \( \Delta \bar{N}_0 = \bar{N}_0 - N_0 \), we obtain
\[
\Delta \bar{N}_0 = \frac{1}{n} \sum_{k=1}^{n} \Delta \bar{N}_0 \sum_{j=1}^{k} p^{-k} = \sum_{j=1}^{n} w_{n,j} \Delta N_j, \tag{B5}
\]
where the ‘weighting factor’ is defined as \( w_{n,j} = \sum_{k=j}^{n} p^{-k} / n \).

The summation corresponds to a geometric series, and is thus given by
\[
w_{n,j} = \sum_{k=j}^{n} p^{-k} = \frac{1}{n} \frac{p^{-j} - p^{-(n+1)}}{1 - p^{-1}}. \tag{B6}
\]
The variance of the mean due to the stochastic atom loss is then
\[
\sigma_{\text{stoch}}^2 = \text{Var}(\Delta \bar{N}_0) = \sum_{j=1}^{n} w_{n,j}^2 \text{Var}(\Delta N_j), \tag{B7}
\]
and using the fact that \( \text{Var}(\Delta N_j) = \bar{N}_j(1 - p) \),
\[
\sigma_{\text{stoch}}^2 = \sum_{j=1}^{n} w_{n,j}^2 \bar{N}_j(1 - p) = N_0(1 - p) \sum_{j=1}^{n} w_{n,j}^2 p^j. \tag{B8}
\]

This shows that the stochastic atom loss noise for an average over \( n \) images is a factor \( \sum_{j=1}^{n} w_{n,j}^2 p^j \) larger than the single image case. For a typical imaging series containing 50 images and an imaging pulse duration \( t = 0.66 \) ms [18], the stochastic atom loss noise is a factor of 108 larger than that of a single image.

References

[1] Andrews M R, Mewes M-O, van Druten N J, Durfee D S, Kurn D M and Ketterle W 1996 Science 273 84
[2] Bradley C C, Sackett C A and Hulet R G 1997 Phys. Rev. Lett. 78 985
[3] Wigley P B et al 2016 Opt. Lett. 41 4795–8
[4] Andrews M R, Kurn D M, Miesner H-J, Durfee D S, Townsend C G, Inouye S and Ketterle W 1997 Phys. Rev. Lett. 79553
[5] Matthews M R, Anderson B P, Haljan P C, Hall D S, Wieman C E and Cornell E A 1999 Phys. Rev. Lett. 832498
[6] Bons P C, de Haas R, de Jong D, Groot A and van der Straten P 2016 Phys. Rev. Lett. 116 173602
[7] Gajdacz M, Pedersen P L, Merch T, Hilliard A J, Arlt J and Sherson J F 2013 Rev. Sci. Instrum. 84 083105
[8] Bradley C C, Sackett C A and Hulet R G 1997 Phys. Rev. A 55 3951
[9] Kaminski F, Kampel N, Steenstrup M, Griesmaier A, Polzik E and Müller J 2012 Eur. Phys. J. D 66 1
[10] Hope J J and Close J D 2004 Phys. Rev. Lett. 93 180402
[11] Lewandowski H J, Harber D M, Whitaker D L and Cornell E A 2003 J. Low Temp. Phys. 132 309
[12] Bertelsen J F, Andersen H K, Mai S and Budde M 2007 Phys. Rev. A 75 013404
[13] Esslinger T, Bloch I and Hänsch T W 1998 Phys. Rev. A 58 R2664
[14] Schneemann U, Engler H, Grimm R, Weidemüller M and Zielonkowski M 1999 Rev. Sci. Instrum. 70 242
[15] Reinaudi G, Lahaye T, Wang Z and Guéry-Odelin D 2007 Opt. Lett. 32 3143
[16] Muesel W, Strobelt H, Joos M, Nicklas E, Stroescu I, Tomkovij J, Hume D B and Oberthaler M K 2013 Appl. Phys. B 113 69
[17] van Kampen N G 2007 Stochastic Processes in Physics and Chemistry 3rd edn (Amsterdam: Elsevier) (doi:10.1016/b978-0-444-52965-7.x5000-4)
[18] Gajdacz M, Hilliard A J, Kristensen M A, Pedersen P L, Klempt C, Arlt J J and Sherson J F 2016 Phys. Rev. Lett. 117 073604
[19] Hilliard A J, Fung Y H, Sompet P, Carpentier A V and Andersen M F 2015 Phys. Rev. A 91 053414
[20] Pethick C J and Smith H 2008 Bose–Einstein Condensation in Dilute Gases 2nd edn (Cambridge: Cambridge University Press)
[21] Bason M G, Heck R, Napolitano M, Elafsson O, Müller R, Thorsen A, Zhang W-Z, Arlt J and Sherson J F 2016 arXiv:1607.02934
[22] Ketterle W and Druten N V 1996 Evaporative cooling of trapped atoms Advances In Atomic, Molecular, and Optical Physics (vol 37) (New York: Academic) pp 181–236
[23] Robbins M S and Hadwen B J 2003 IEEE Trans. Electron Devices 50 1227
[24] Hynecek J, Member S and Nishiwaki T 2003 IEEE Trans. Electron Devices 50 239
[25] Hume D B, Stroescu I, Joos M, Muesel W, Strobel H and Oberthaler M K 2013 Phys. Rev. Lett. 111 253001
[26] Efron B 1979 Ann. Statist. 7 1