

Single-atom counting in a two-color magneto-optical trapMartin Schleder¹,* Alexandra Mozdzen, Thomas Lompe, and Henning Moritz²*Institut für Laserphysik, Universität Hamburg, 22761 Hamburg, Germany**and The Hamburg Centre for Ultrafast Imaging, Universität Hamburg, 22761 Hamburg, Germany*

(Received 19 November 2020; accepted 5 February 2021; published 12 March 2021)

Recording the fluorescence of a magneto-optical trap (MOT) is a standard tool for measuring atom numbers in experiments with ultracold atoms. When trapping few atoms in a small MOT, the emitted fluorescence increases with the atom number in discrete steps, which allows one to measure the atom number with single-particle resolution. Achieving such single-particle resolution requires stringent minimization of stray light from the MOT beams, which is very difficult to achieve in experimental setups that require in-vacuum components close to the atoms. Here, we present a modified scheme that addresses this issue: Instead of collecting the fluorescence on the MOT (D2) transition, we scatter light on an additional probing (D1) transition and collect this fluorescence with a high-resolution microscope while filtering out the intense MOT light. Using this scheme, we are able to reliably distinguish up to 17 ^{40}K atoms with classification fidelities of $\sim 98\%$ for up to 5 atom numbers and fidelities of more than 85% for up to 17 atoms.

DOI: [10.1103/PhysRevA.103.033308](https://doi.org/10.1103/PhysRevA.103.033308)**I. INTRODUCTION**

In recent years, ultracold atoms have emerged as a unique and powerful platform to experimentally study few-body physics. They provide synthetic model systems where key properties such as the particle number, the strength of the interparticle interactions, and the confining potential can be tuned nearly at will [1]. This has enabled pioneering works that study, e.g., entanglement generation [2] or the emergence of correlations and quantum phases with increasing particle number [1,3].

An essential requirement for such experiments is the ability to count the number of atoms in the trap with high fidelity [4]. Collecting fluorescence light while laser cooling and trapping the atoms in a magneto-optical trap (MOT) has proven to be a reliable and well-established method to measure atom numbers with single-particle resolution [5–15] and is especially suited for experiments studying few atoms confined in optical tweezers, where *in situ* imaging would lead to photoassociation losses. As the number of photons emitted by a single atom is quite limited, these experiments require considerable effort to maximize the signal-to-noise ratio of the fluorescence signal. Common techniques for achieving this include the use of large magnetic field gradients to compress the MOT and thereby spatially concentrate the fluorescence signal, using optics with high numerical aperture to maximize the number of collected photons, as well as

employing highly sensitive and low-noise cameras. At the same time, detrimental stray-light collection is suppressed as far as possible by employing laser-cooling beams with small diameters, avoiding the presence of unwanted scattering surfaces close to atoms, and spatial filtering of the fluorescence signal.

Using these techniques, single-atom resolution for atom numbers in excess of 300 has been achieved [13,14]. However, the stringent limits on the amount of stray light present in the system render this method unsuitable for many experimental setups that were not specifically designed for this purpose. Suppressing stray-light detection is particularly challenging in experimental setups that feature in-vacuum components such as electrodes, high-resolution optics, or atom chips, as spatial filtering of stray light that originates close to the atoms is inefficient.

In this work, we describe a technique that addresses this issue by using two different transitions. In brief, we cool and trap fermionic ^{40}K atoms using light on the D2 transition, but collect fluorescence on the D1 transition. The D2 MOT light propagates along all three spatial directions and hence generates a large amount of stray light, which can be filtered out before the camera with dichroic mirrors. Meanwhile, D1 fluorescence photons are generated by a D1 pumping beam whose size and beam path have been chosen specifically to avoid scattering surfaces and thereby minimize the amount of stray light that is generated. This scheme allows us to perform single-atom counting of up to 17 atoms with an average classification fidelity of 95% in an experimental setup featuring a high-resolution microscope objective placed inside the vacuum chamber very close to the atoms.

This work is structured as follows: the experimental setup is described in Sec. II, followed by a discussion of the image postprocessing and atom-number determination in Sec. III. The classification fidelity and the limiting factors of the

*Corresponding author: mschlede@physnet.uni-hamburg.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

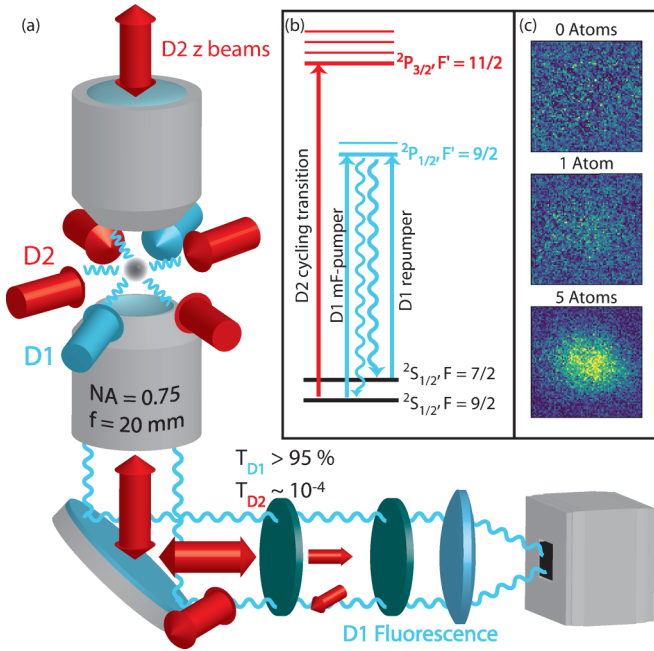


FIG. 1. (a) Sketch of the experimental setup. The detection MOT is formed by three pairs of D2 trapping beams (dark red), while a pair of D1 beams (light blue) generates D1 fluorescence for detection. The D1 fluorescence light is collected by an in-vacuum microscope objective (NA = 0.75, $f = 20$ mm), separated from the D2 trapping beams using dielectric filters, and finally imaged on an sCMOS camera. (b) Simplified level scheme showing the atomic transitions utilized in the detection MOT: The D2 line is used for cooling and trapping the atoms, while the D1 repumper addresses atoms that decay into the $F = 7/2$ state. The D1 mF-pumper excites atoms from the ground state to the $2P_{1/2}$ manifold and thereby enhances the D1 fluorescence. (c) Typical fluorescence images containing zero, one, and five atoms.

presented scheme are discussed in Sec. IV, and Sec. V summarizes the presented work.

II. SETUP

The central element of our setup to study ultracold few-body systems is a pair of microscope objectives with high numerical aperture [16] placed inside the vacuum chamber; see Fig. 1(a). These objectives are ideally suited to generate arbitrary potentials and image single atoms with high spatial resolution. However, at the same time, they represent a significant challenge for performing atom counting with a MOT.

The first constraint placed by the objectives is that they cover such a large fraction of the solid angle that the vertical MOT beams have to pass through the objectives. This precludes using the objectives to detect the fluorescence created by the MOT light, as the fluorescence signal will be completely drowned out by the MOT beams. At the same time, the geometry of the vacuum chamber, which has to accommodate the in-vacuum components, severely limits the optical access for fluorescence light collection along other directions. Combined with the large background of MOT light scattered by the objectives, this makes it exceedingly difficult to measure

the fluorescence on the MOT transition with a signal-to-noise ratio sufficient for single-atom resolution.

Fortunately, both of these issues can be addressed by using our two-color approach. By adding a retroreflected D1 pumping beam that propagates orthogonal to the objective axis, we create fluorescence light on the D1 transition while causing very little D1 stray light to enter the objective. As the wavelengths of the D1 and D2 transition differ by ~ 3 nm, this allows us to collect the D1 fluorescence photons with one of the *in vacuo* objectives, while filtering out the D2 MOT light with two dichroic long-pass filters [17] with very steep edges transmitting, in total, more than 90% of the D1 light at 770.1 nm, but less than 10^{-8} of the D2 light at 766.7 nm [18,19].

In the following, the D2 MOT and the D1 excitation scheme are described in more detail. The D2 MOT is formed by three beam pairs addressing the cycling transition between the $F = 9/2$ and $F = 11/2$ hyperfine states. The horizontal D2 MOT beams have $1/e^2$ beam waists of 1.5 mm and are retroreflected, while the vertical beams are independent of each other and have waists of only 340 and 480 μm , as the short focal length of the microscope strongly constrains the maximal achievable beam size. The magnetic quadrupole field for the MOT is provided by a set of magnetic field coils in an anti-Helmholtz configuration. To decrease the influence of background light on the atom counting, we use a strong quadrupole field with a gradient of $dB/dz = 132$ G/cm in the vertical axis to make the MOT as small as possible and thereby concentrate the fluorescence signal in the smallest possible area. For our setup, we find an optimum at a combined saturation parameter for the D2 light of 12 and 31 for the horizontal and vertical beams, respectively, and a detuning $\Delta\omega_{D2}$ of -1.2Γ , where $\Gamma/2\pi = 6$ MHz is the natural linewidth of the excited state. The detuning of the D2 MOT beams is a compromise between minimizing the MOT size, which is minimal for small detunings, and a stable scattering rate that is not too sensitive to detuning changes, which is the case at larger detunings. We note that our D2 MOT differs from a standard MOT in one aspect: There are no D2 beams that repump atoms that have decayed into the $F = 7/2$ hyperfine ground state and thereby became dark to the cycling transition. Instead, this repumping is performed by one of the D1 excitation beams.

The D1 beams propagate in the horizontal plane along an axis having a 45° angle to the D2 beams and have waists of 1.2 mm. The first beam addresses the $F = 7/2 \leftrightarrow F' = 9/2$ transition. This repumper brings atoms that have decayed into the $F = 7/2$ hyperfine state back to the $F = 9/2$ ground state. It operates with a saturation parameter of 14 and a detuning $\Delta\omega_{D1}$ of $+0.6\Gamma$.

However, using only this repumper does not produce enough D1 fluorescence for imaging as the $F = 7/2$ hyperfine state is populated only slowly while cooling on the D2 cycling transition [20]. Hence, we use a second D1 beam to drive the $F = 9/2 \leftrightarrow F' = 9/2$ transition. This so-called mF-pumper addresses the prevalent $F = 9/2$ ground-state atoms and excites them to the $F' = 9/2$ manifold. From there, the atoms decay into the $F = 9/2$ in about 30% and the $F = 7/2$ manifold in about 70% of the cases. If the atom is pumped into the $F = 7/2$ manifold, the subsequent repumping results in an average of three more photons being emitted on the D1

line before the atom is pumped back to the $F = 9/2$ manifold. By varying the strength of this mF-pumper, we can therefore control the amount of D1 fluorescence emitted by the atoms. For our imaging, the mF-pumper transition is driven resonantly with a saturation parameter of 7. The resulting significant amount of light pressure on the atoms is canceled by retroreflecting the beam.

The D1 fluorescence light emitted by the atoms is collected by one of the microscope objectives and imaged onto an sCMOS camera; examples of typical fluorescence images obtained with the setup are shown in Fig. 1(c). The total collection efficiency of the imaging system is estimated to be $12.7 \pm 1.6\%$, with the dominant effect being the relative solid angle covered by the microscope (17%). Together with the quantum efficiency of 38% [21], this leads to $4.8 \pm 0.6\%$ of the photons actually being detected. On average, the camera detects 2.3×10^4 photoelectrons per atom for an exposure time of 1 s, corresponding to a scattering rate of 470 kHz on the D1 transition [22]. Adding a retroreflected D1 beam pair along the second horizontal axis did not result in a significant improvement in the signal-to-noise ratio. While the fluorescence increased by about 20%, adding the additional beams also increased the amount of background light, resulting in a similar signal-to-noise ratio.

The number of background photons due to scattered light is ~ 175 /pixel, which, for typical regions of interest of 120×120 pixels corresponding to $150 \times 150 \mu\text{m}$, amounts to about 2.5×10^6 photons. This background level is dominated by stray light from the D1 beams; the background originating from the D2 beam directly incident onto the objective accounts for less than 10% of the background light. To achieve this low background, it is essential to not only perform the obvious filtering of D2 light with filters in front of the camera, but also to suppress the broad pedestal of incoherent light emitted by the semiconductor lasers and amplifiers generating the D2 beams. This is achieved by placing two ultranarrow 780 nm bandpass filters [23] that are angle tuned to 766.7 nm in the beam path between the tapered amplifier and the fibers.

III. IMAGE POSTPROCESSING AND ATOM-NUMBER EXTRACTION

To characterize our imaging scheme and establish a protocol for determining the atom number from the fluorescence signal, we generate a large number of sample images of MOTs containing a small number of atoms. We achieve this by releasing atoms from a far-detuned optical trap and then recapturing them with the MOT. This leads to a random number of atoms in the imaging MOT, where the mean value can be controlled via the atom number in the optical trap and the time delay between releasing the atoms and switching on the MOT. While the two-color scheme described in Sec. II yields fluorescence signals of single atoms that are significantly above the background noise [see Fig. 1(c)], some steps of image postprocessing are required to extract reliable particle numbers from this data. The following detailed description can be summarized into three main points: In the first step, the background of scattered light is subtracted from the fluorescence images and the raw fluorescence signal is determined by summing up the camera counts within a defined

region around the current center of the MOT. Slow drifts in the fluorescence per atom and the zero atom offset are then determined and corrected by time-resolved fitting of these two parameters, thus providing a normalized fluorescence signal. A histogram of the normalized fluorescence signals [see Fig. 2(b)] shows clearly discernible peaks corresponding to different atom numbers. To assign atom numbers to individual images, we first quantify the positions and widths of the peaks by fitting a sum of Gaussians to the histogram. We then define classification intervals around each peak based on their widths and assign atom numbers to all fluorescence measurements that fall within these intervals.

We now go through each step of the analysis in more detail, starting with the subtraction of the background signal. To achieve this, we make use of a principal component analysis, as this approach reduces noise compared to subtracting a reference image taken shortly after each atom image. To this end, a set of reference images without atoms, yet otherwise the same experimental conditions, is taken and the main components are determined using singular-value decomposition. For the data shown in Fig. 2, 100 reference images were taken at random times during the $\sim 10\,000$ runs and used for the singular-value decomposition. The components with the ten highest singular values were then projected out of the actual data images [24] by subtracting these basis vectors, weighted with the scalar product between them and the atom image. Using more reference images or projecting out more components does not result in a significant improvement in the signal-to-noise ratio, as the components are not necessarily orthogonal to the fluorescence signal.

Next, the current MOT position on each of these background-free images is determined by fitting a Gaussian to an averaged image obtained from the current and the 10 preceding and following images. This allows us to compensate for slow drifts of the MOT position over time and use a tight region of interest of only $150 \times 150 \mu\text{m}$ for summing up camera counts. The resulting camera counts recorded by the experiment are shown as a function of time in Fig. 2(a) (red dots). While discrete steps in the fluorescence for different atom numbers can be clearly made out by eye, there are slow drifts in the signal levels over time and simple thresholds are not sufficient to reliably distinguish between different atom numbers with high fidelity. To provide the time-resolved values of the zero atom offset and of the fluorescence per atom, we determine the current average count levels for 0, 1, 2, and 3 atoms from the 20 realizations closest in time containing these atom numbers, respectively, and fit a linear relation to these levels. The camera counts are then corrected for these drifts, resulting in the black data points shown in Fig. 2(a).

Figure 2(b) shows a histogram of these corrected fluorescence levels, which we then fit with a sum of Gaussians with centers x_N and standard deviations σ_N . We classify events within $x_N \pm 2\sigma_N$ to have N atoms present; events outside these limits are discarded. For the data shown in Fig. 2, about 6% of the events are rejected. Tighter classification intervals lead to higher identification fidelity, but also to a larger fraction of events being rejected. We find that the mean fluorescence level x_N grows linearly with the atom number N , as shown in Fig. 3(a), indicating no sign of density-dependent effects.

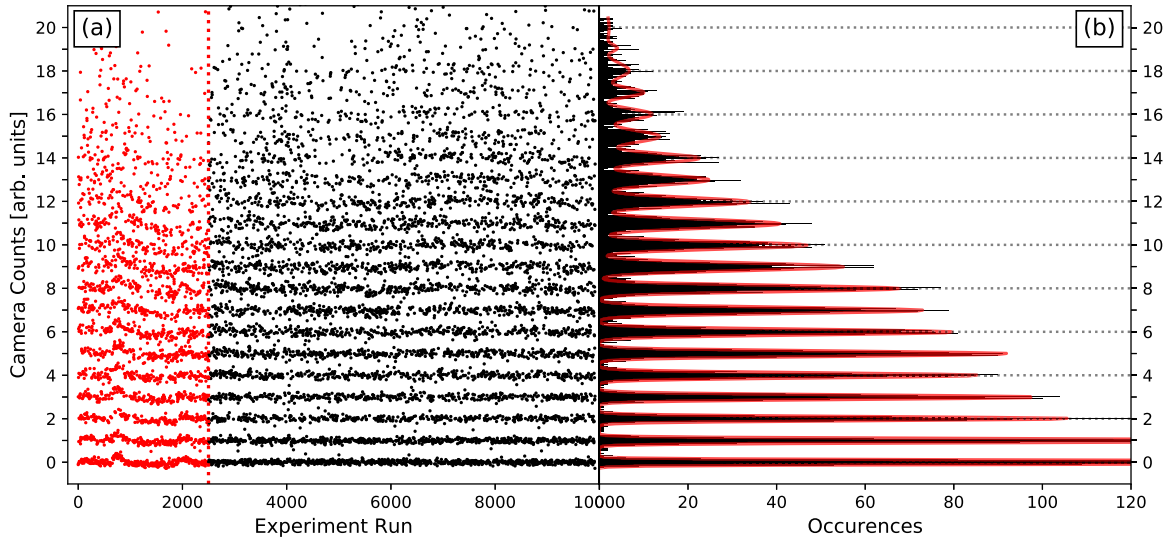


FIG. 2. (a) Timeline of the fluorescence level over $\sim 10\,000$ experiment runs, where each dot shows the result of a single experiment run. The red dots to the left of the vertical dotted line show the raw fluorescence levels without correcting for time-dependent changes in offset and scattering rate; the black dots are corrected for these effects. (b) Histogram of the corrected fluorescence levels. The red line shows a fit to the data, consisting of a sum of 20 Gaussians with independent centers x_N , widths σ_N , and amplitudes. We can reliably determine the atom number for systems with up to about 17 atoms. Above this number, both the limited statistics of the data set and the width of the distributions prevent a reliable classification.

IV. RESOLUTION LIMIT AND CLASSIFICATION FIDELITY

After establishing our postprocessing procedure, we now analyze the data to determine the fidelity of our atom-number determination and find the limiting factors of the detection scheme. For our evaluation, we assume that the peaks in the histogram are clearly distinguishable and we can assign reliable atom numbers as long as the separation of the peaks is larger than 4σ . Hence the maximal atom number we can resolve is limited by the observed growth of the peak width σ_N with the atom number N ; see Fig. 3(b).

The peak width is determined by three noise sources, which can be assumed to be independent and therefore are added up quadratically [14]. These are a constant background noise stemming from the shot noise of the background light (including D1 and D2 light) and the read-out noise of the camera, the shot noise of the fluorescence photons collected by the camera, which scales with the square root of the atom number, and shot-to-shot fluctuations in the average fluorescence rate per atom, which scale linearly with the atom number.

When analyzing the data, we find that for atom numbers < 20 , the shot noise of the fluorescence photons contributes $\sigma_{\text{ph}}^2 \lesssim 0.01^2$ to the variance and the atom-number independent background noise is about $\sigma_{\text{bg}}^2 \approx 0.07^2$. These noise sources therefore only play a small role in limiting our atom-number determination.

Instead, the dominant effect is a linear growth of the peak width σ_N with the atom number N [see Fig. 3(b)], which strongly suggests that our atom-number determination is limited by fluctuations in the atom fluorescence. While long-term drifts are removed by the rescaling method described in Sec. III, shot-to-shot fluctuations or drifts on timescales shorter than a few tens of shots are not significantly

suppressed with this method. The observed level of fluctuations is compatible with either laser frequency noise with a rms level of ~ 150 kHz or rms beam power fluctuations of $\sim 3\%$, or a combination of both. We note that one-color schemes [5,8,9,12–15] are less sensitive to frequency or power fluctuations due to the high saturation of the single cooling transition. In contrast, in our two-color scheme, the $F = 9/2$ ground state is addressed by two transitions, i.e., the D2 cooler and the D1 mF-pumper, which thus compete with each other. Hence, relative power fluctuations directly translate into fluctuations of the fluorescence per atom and lead to a faster growth of σ_N than in one-color schemes.

For the data set shown in Fig. 3, the threshold value $\sigma_N = 0.25$ where the $\pm 2\sigma$ intervals of neighboring histogram peaks start to overlap is reached for $N \approx 17$. Beyond this threshold, we define the boundary between N and $N + 1$ atoms as $(x_N + 2\sigma_N + x_{N+1} - 2\sigma_{N+1})/2$, and our ability to separate the fluorescence signals belonging to different atom numbers substantially decreases. We therefore consider $\sigma_N < 0.25$ to roughly indicate the threshold for fluorescence measurements with single atom-number resolution. We note that this condition is significantly more conservative than the requirement $\sigma_N < 1$ used by other authors [13–15]. Extrapolating our observed peak width growth indicates that we reach $\sigma_N = 1$ for $N \approx 80$.

In addition to the noise sources discussed above, which broaden the fluorescence peaks and therefore limit our ability to reliably distinguish between different atom numbers for larger values of N , there is another effect that can limit the fidelity of our atom-number determination: loss of atoms from the MOT during the imaging period. As the lifetime of atoms in our imaging MOT is of the order of 80 s, much longer than the imaging time of 1 s, loss events are very rare when measuring small atom numbers, but become more

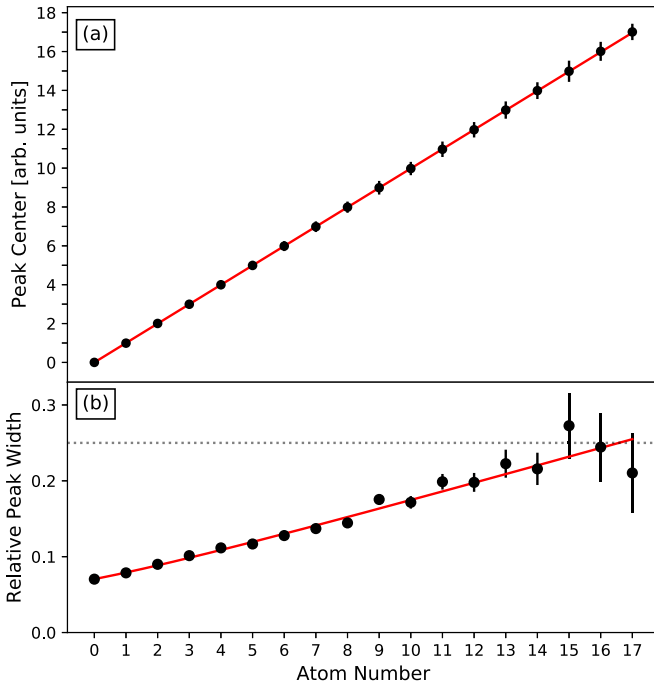


FIG. 3. (a) Mean value x_N and (b) width σ_N of the fluorescence signal as a function of atom number, as determined from the fit to the histogram shown in Fig. 2(b). The mean fluorescence increases linearly with atom number (red solid line), indicating that density-dependent effects are negligible. The error bars represent the $\pm 2\sigma_N$ confidence intervals. The widths σ_N of the peaks in the histogram are normalized to the average fluorescence per atom and shown in (b). We assume the peak widths to be a quadratic sum of a constant background noise, photon shot noise scaling with \sqrt{N} , and a linear noise contribution due to short-term variations in the intensity and detuning of the imaging light. Fitting the data with this model (red curve) shows that the noise is dominated by the linear term. The dotted line shows the point where the $\pm 2\sigma_N$ classification regions of two neighboring peaks start to overlap, which roughly indicates the point where we can no longer reliably assign an atom number to an event.

relevant for measuring larger systems. Also, depending on the number of atoms in the system and at which time during the measurement it occurs, not every loss event necessarily leads to an incorrect determination of the atom number. To obtain a quantitative estimate of the influence of the various noise sources on our classification fidelity, we therefore perform a phenomenological Monte Carlo simulation of our system using the following model for the counts c_i observed in run i :

$$c_i = \alpha_i + \beta_i \sum_{n=1}^N \gamma_{i,n} \min\left(1, \frac{\tau_{i,n}}{t_{Im}}\right), \quad (1)$$

where α_i , β_i , and $\gamma_{i,n}$ are Gaussian distributed random numbers with mean 0, 1, and 1, respectively. The standard deviations σ_α , σ_β , and σ_γ of these three random numbers are *a priori* free parameters that model the noise sources of the fluorescence signal described above, where σ_α corresponds to the atom-number independent noise, σ_β to the global and hence correlated fluctuations in fluorescence per atom, and σ_γ to the fluorescence shot noise from each uncorrelated atom.

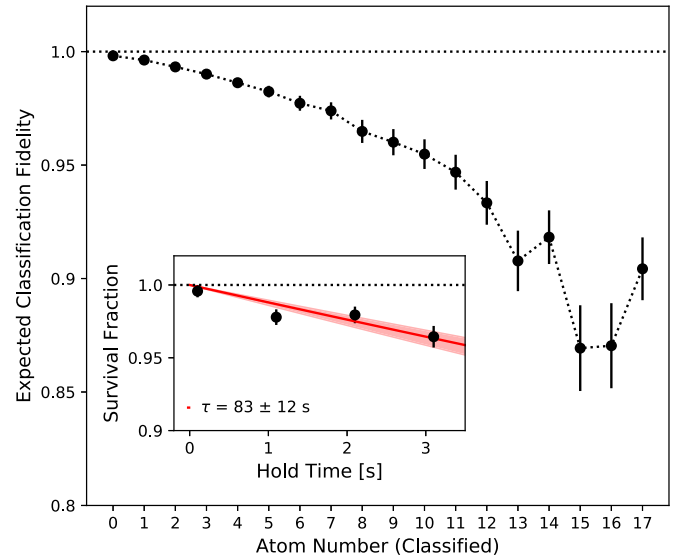


FIG. 4. Simulated classification fidelities for systems with different atom numbers using the noise characteristics determined in Fig. 2(b) and an independent measurement of the MOT lifetime (see inset; assuming an exponential decay of the atom number). The classification fidelity is limited by the ratio of imaging time to lifetime: loss events during the imaging time reduce the number of collected fluorescence photons and thus can cause the fluorescence signal to fall into the classification bin of a lower atom number.

The values of these parameters are fitted to best reproduce the observed peak widths in the fluorescence count histogram in Fig. 3(b) [25]. The possibility of atom loss is taken into account by comparing the imaging time t_{Im} to a time $\tau_{i,n}$ after which atom n is lost from the MOT in run i that is drawn from an exponential distribution using the experimental lifetime of 83 ± 12 s, and taking the shorter of the two times to calculate the fluorescence of each atom [26]. The counts from the simulation are processed and classified based on the experimentally determined intervals as described in Sec. III. The classification fidelity is estimated by comparing the initial atom number with the classification result for each simulated run.

Figure 4 shows the resulting classification fidelities, which are close to 100% for low atom numbers, decrease to $\sim 95\%$ for 10 atoms and to $\sim 87\%$ for the highest atom numbers we can reliably distinguish. The vast majority of the incorrectly classified events stem from atom loss, especially having one more atom initially than classified: when working with, at most, 10 (17) atoms, this is the case for $>99\%$ ($>95\%$) of the incorrectly identified atom numbers.

One interesting insight that can be gained from the simulation is that while the classification fidelity strongly depends on the MOT lifetime, this is not directly visible in the peak widths: a limited lifetime leads to a broad background in simulated histograms, emulating the one shown in Fig. 2(b). The peaks are still quite narrow and clearly identifiable, but loss of atoms during imaging can move the fluorescence count into the classification region of a lower atom number. This can, in principle, be improved by increasing the lifetime of our MOT as the fraction of incorrectly identified events scales linearly with the inverse lifetime, but in our case the MOT

lifetime is already close to the vacuum-limited lifetime of the system. It seems more realistic to reduce the width of the fluorescence peaks by improving the power and frequency stability. This would increase the classification fidelity as smaller peak widths decrease the probability of misclassifying events involving atom loss. Alternatively, such an improved stability would allow one to reduce the imaging time, while keeping the classification fidelity the same.

V. CONCLUSION

In this work, we have presented a two-color MOT for counting small numbers of atoms. With this scheme, we are able to reliably perform atom-number measurements with single-particle resolution for systems of up to about 17 atoms with classification fidelities of $\sim 98\%$ when working with up to 5 atoms, and fidelities of $>85\%$ when working with up to 17 atoms. The number of atoms we can currently distinguish

is determined by the stability of the fluorescence rate on the probing transition, and thus by the technical performance of the frequency locks and power stabilizations used, which can be improved if a need to count larger atom numbers arises. The two-color scheme significantly reduces the influence of light scattered from the MOT beams and thereby enables such single-atom resolved measurements in systems that have not been purpose-built for this application and, for example, contain in-vacuum components.

ACKNOWLEDGMENTS

We thank P. Wieburg for his work in building and characterizing the quantum gas apparatus. This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the framework of SFB-925, Project No. 170620586, and the excellence cluster “Advanced Imaging of Matter,” EXC 2056, Project ID No. 390715994.

-
- [1] L. Bayha, M. Holten, R. Klemt, K. Subramanian, J. Bjerlin, S. M. Reimann, G. M. Bruun, P. M. Preiss, and S. Jochim, *Nature (London)* **587**, 583 (2020).
 - [2] L. Pezze, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, *Rev. Mod. Phys.* **90**, 035005 (2018).
 - [3] A. Wenz, G. Zürn, S. Murmann, I. Brouzos, T. Lompe, and S. Jochim, *Science* **342**, 457 (2013).
 - [4] H. Ott, *Rep. Prog. Phys.* **79**, 054401 (2016).
 - [5] Z. Hu and H. J. Kimble, *Opt. Lett.* **19**, 1888 (1994).
 - [6] F. Ruschewitz, D. Bettermann, J. Peng, and W. Ertmer, *Europhys. Lett.* **34**, 651 (1996).
 - [7] C. Chen, Y. Li, K. Bailey, T. O’Conner, L. Young, and Z.-T. Lu, *Science* **286**, 1139 (1999).
 - [8] D. Haubrich, H. Schadwinkel, F. Strauch, B. Ueberholz, R. Wynands, and D. Meschede, *Europhys. Lett.* **34**, 663 (1996).
 - [9] N. Schlosser, G. Reymond, I. Protsenko, and P. Grangier, *Nature (London)* **411**, 1024 (2001).
 - [10] W. Alt, D. Schrader, S. Kuhr, M. Müller, V. Gomer, and D. Meschede, *Phys. Rev. A* **67**, 033403 (2003).
 - [11] S. Yoon, Y. Choi, S. Park, J. Kim, J.-H. Lee, and K. An, *J. Phys.: Conf. Ser.* **80**, 012046 (2007).
 - [12] F. Serwane, G. Zürn, T. Lompe, T. Ottenstein, A. Wenz, and S. Jochim, *Science* **332**, 336 (2011).
 - [13] D. B. Hume, I. Stroescu, M. Joos, W. Muessel, H. Strobel, and M. K. Oberthaler, *Phys. Rev. Lett.* **111**, 253001 (2013).
 - [14] I. Stroescu, D. B. Hume, and M. K. Oberthaler, *Phys. Rev. A* **91**, 013412 (2015).
 - [15] A. Hüper, C. Pür, M. Hetzel, J. Geng, J. Peise, I. Kruse, M. Kristensen, W. Ertmer, J. Arlt, and C. Klempt, [arXiv:1912.05689](https://arxiv.org/abs/1912.05689).
 - [16] Special Optics Type 54-31-20, NA = 0.75, working distance = 2.5 mm, effective focal length $f = 20$ mm, achromatic for 770, 870, and 590 nm.
 - [17] Alluxa LC-HBP770.7/3, nominal center wavelength 770.7 nm, FWHM 3 nm, transmission $>95\%$ at 770.1 nm, blocking $>OD4$ at 766.7 nm.
 - [18] A similar two-color MOT has been used, but not described, in Ref. [19] to measure small atom numbers.
 - [19] C. Shkedrov, Y. Florshaim, G. Ness, A. Gandman, and Y. Sagi, *Phys. Rev. Lett.* **121**, 093402 (2018).
 - [20] This is due to the fact that decay from the excited state $F' = 11/2$ of the D2 MOT cycling transition to the $F = 7/2$ state is forbidden and one has to rely on off-resonant excitations to the $F' = 9/2$ state in the D2 manifold, which is 8.5Γ detuned from the D2 laser frequency.
 - [21] Andor Zyla 5.5.
 - [22] The imaging time of 1 s is needed to obtain a large enough signal-to-noise ratio to distinguish 17 atoms. For smaller atom numbers, shorter imaging times would be sufficient.
 - [23] Laseroptics Garbsen L-08729 IF780nm/6, FWHM ≈ 0.3 nm, transmission $\approx 90\%$.
 - [24] C. F. Ockeloen, A. F. Tauschinsky, R. J. C. Spreeuw, and S. Whitlock, *Phys. Rev. A* **82**, 061606(R) (2010).
 - [25] $\sigma_\alpha = \sigma_{bg} = 0.07 \pm 0.01$, $\sigma_\beta = 0.034 \pm 0.012$, $\sigma_\gamma = \sigma_{ph} = 0.012 \pm 0.005$.
 - [26] The lifetime has been measured by loading 0, 1, or 2 atoms into the MOT and comparing the fluorescence level of two measurements with some variable time delay between them; see inset of Fig. 4. In the ~ 4000 runs used for this, not a single event of loading an atom from the background gas has been observed, thus limiting the background loading rate to be below $3 \times 10^{-4} \text{ s}^{-1}$ at 68% CL, thus not affecting the classification fidelity.