New Journal of Physics

The open access journal at the forefront of physics

CrossMark

PAPER

Super-resolution microscopy of single atoms in optical lattices

Andrea Alberti[1,3], Carsten Robens[1], Wolfgang Alt[1], Stefan Brakhane[1], Michał Karski[1], René Reimann[1], Artur Widera[2] and Dieter Meschede[1]

1   Institut für Angewandte Physik der Universität Bonn, Wegelerstr. 8, D-53115 Bonn, Germany
2   Fachbereich Physik und Landesforschungszentrum OPTIMAS, Gottlieb-Daimler-Str., D-67663 Kaiserslautern, Germany
3   Author to whom any correspondence should be addressed.

E-mail: alberti@iap.uni-bonn.de

Abstract

We report on image processing techniques and experimental procedures to determine the lattice-site positions of single atoms in an optical lattice with high reliability, even for limited acquisition time or optical resolution. Determining the positions of atoms beyond the diffraction limit relies on parametric deconvolution in close analogy to methods employed in super-resolution microscopy. We develop a deconvolution method that makes effective use of the prior knowledge of the optical transfer function, noise properties, and discreteness of the optical lattice. We show that accurate knowledge of the image formation process enables a dramatic improvement on the localization reliability. This allows us to demonstrate super-resolution of the atoms' position in closely packed ensembles where the separation between particles cannot be directly optically resolved. Furthermore, we demonstrate experimental methods to precisely reconstruct the point spread function with sub-pixel resolution from fluorescence images of single atoms, and we give a mathematical foundation thereof. We also discuss discretized image sampling in pixel detectors and provide a quantitative model of noise sources in electron multiplying CCD cameras. The techniques developed here are not only beneficial to neutral atom experiments, but could also be employed to improve the localization precision of trapped ions for ultra precise force sensing.

1. Introduction

Detection and manipulation of individual atoms on neighboring sites of an optical lattice have attracted great interest in recent years for applications in quantum information processing [1–8], quantum simulations [9–13], and very recently for studying strongly correlated Fermi systems at the single particle level [14–17]. Resolving atom positions with single-site resolution represents a technological challenge, since in optical lattices the distance between two lattice sites is on the order of the optical lattice wavelength. In fact, experiments relying on atoms tunneling between lattice sites require short lattice constants since the tunneling rate decreases exponentially with larger intersite separation.

Previously, we demonstrated that the number of lattice sites between well-isolated atoms in a one-dimensional (1D) optical lattice can be resolved with high reliability even with objective lenses of moderate numerical aperture (NA) [4]. Single-atom localization methods are employed in our laboratories ever since to measure the spatial probability distribution of atoms performing discrete-time quantum walks [12, 18, 19]. Recent experiments in our laboratory beyond single particle physics require resolving the position of each individual atom in small clusters at high filling factors, even when each lattice site is occupied [20]. By exploiting the discreteness of the atoms' positions in the lattice, we demonstrate in this manuscript new methods that enable resolving clusters of atoms with high reliability. Super-resolving a cluster of atoms constitutes a much bigger challenge than resolving the distance between exactly two atoms, as we originally demonstrated [4].

Besides presenting a conceptually simple introduction to super-resolved fluorescence imaging of atoms and to the related deconvolution problem, this work develops several new methods with respect to our original work

[4], which include: Wiener deconvolution of fluorescence images combined with a robust spectral-density-estimation algorithm for a first estimation of the atoms' positions (section 6.3), a weighted nonlinear least squares estimation of positions accounting for the experimentally characterized noise (section 5.2), and inclusion of constraints of the atoms' positions on the periodic lattice (section 6.4), as well as an optimal algorithm for the iterative reconstruction of the line spread function of the imaging system (the analogue of the point spread function (PSF) for 1D imaging) (section 4.1) with mathematical treatment of the convergence limit (appendix C), and a measurement of optical aberrations from single atom images (section 4.2).

In general, our methods for the analysis of fluorescence images are closely related to those employed in superresolution microscopy of biological structures [21–23], or in astronomy, where stars appear as point-like radiation sources [24]: knowing the exact number of emitters in an observed region of interest allows us to determine the center position of each emitter with an uncertainty much smaller than the width of the PSF of the imaging system. Super-resolution imaging relies on the precise knowledge of the properties of point-like atomic emitters trapped in an optical lattice as well as the detailed properties of background noise.

Light sources separated by less then an Abbe radius, $r_A = \lambda_f/(2\mathrm{NA})$ (λ_f is the fluorescence radiation wavelength and NA is the NA of the imaging system) form blurred, overlapping images, which cannot be optically distinguished and require super-resolution techniques to be resolved. In the past few years, the Abbe diffraction limit has prompted ultracold atom experiments to develop objective lenses with larger NAs in the range of $0.6 < \mathrm{NA} < 0.8$ to significantly enhance the optical resolution, thus allowing single lattice sites to be optically resolved [10, 11]. These experiments have developed different solutions to the deconvolution problem: for example, linear least-squares fit of fluorescence intensities by a sum of reconstructed PSFs on a fixed lattice combined with threshold criterion [10, 13, 15–17], deconvolution through a kernel function containing the PSF information (though not noise information) and threshold criterion [25], deconvolution through a Gaussian kernel combined with a threshold criterion [26], deconvolution by fitting different configurations of occupancies on a fixed lattice (though without specifically mentioning methods to account for noise) [11], or simply using a threshold criterion on the integrated fluorescence in the pixels corresponding to each lattice site [14]. A different approach based on electron microscopes has demonstrated even higher spatial resolution to resolve atoms in an optical lattice, though without reaching the sensitivity level needed for detecting single atoms yet [27].

However, experiments working under conditions of a low signal-to-noise ratio or employing a moderately large NA (as is our case NA = 0.23) require methods that can extract the maximum physical information on the positions of atoms, especially when bunched in closely packed clusters. The Fisher information matrix provides the mathematical instrument to identify the fundamental limit on the information that an estimator of positions can extract from a fluorescence image (Cramér–Rao information bound) [28]: if such an estimator exists, than this estimator maximizes the likelihood function associated with the estimated quantity (i.e. positions). In this sense, the maximum-likelihood estimator defines the gold standard for any image analysis technique. As argued in section 6.3, we can approach this limit relying on a accurate knowledge of the line spread function and noise characteristics [29].

Furthermore, we would like to remark that the techniques and results demonstrated in this work could have an impact even beyond neutral atom experiments. For example, the majority of techniques for photoactivated localization microscopy have been optimized for the situation of fully separated fluorophore, while we here deal specifically with the opposite situation of bunched emitters (atoms) [30]. Our methods can also find application to improve the localization precision of trapped ions, where displacements of the equilibrium position are recorded to sense extremely tiny forces on the yoctonewton scale [31–34]. Recently, the same techniques presented in section 4.2 to quantitatively reconstruct optical aberrations have been employed, concurrently with this work, to characterize the imaging system of single trapped ions [34].

2. The deconvolution problem

The steps involved in the image formation—from the point-like atomic radiation sources to the final image on the digital camera—are schematically depicted in figure 1 and summarized below:

(i) Optical diffraction by the imaging microscope transforms all radiation sources into blurred spatial distributions. This process can be mathematically expressed as the convolution of the original distribution $O(x, y)$ with the PSF $P(x, y)$, which represents the characteristic intensity distribution of an ideal point source recorded by the imaging system. In case of a system whose optical response is translationally invariant (isoplanatic behavior), the fluorescence image distribution reads

$$I_{\mathrm{fluo}}(x, y) = (P * O)(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x - u, y - v)\, O(u, v)\, \mathrm{d}u\, \mathrm{d}v. \tag{1}$$

In case of a hard circular aperture, for example, the PSF of diffraction-limited aberration-free optical
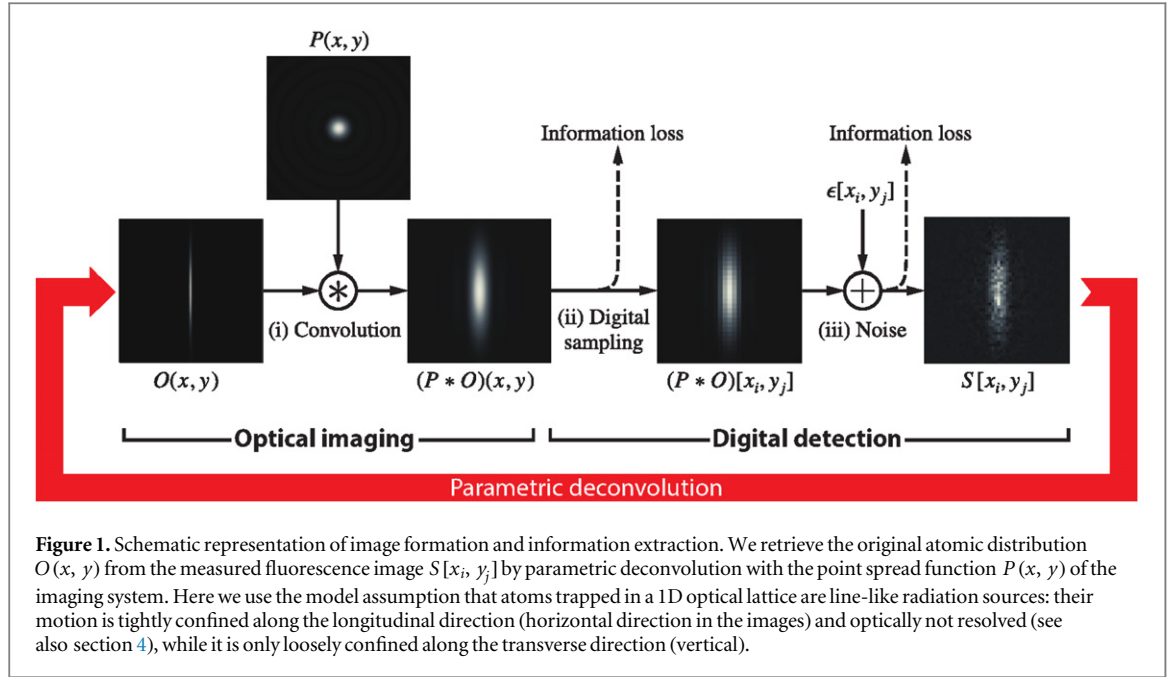
Figure 1. Schematic representation of image formation and information extraction. We retrieve the original atomic distribution $O(x, y)$ from the measured fluorescence image $S[x_i, y_j]$ by parametric deconvolution with the point spread function $P(x, y)$ of the imaging system. Here we use the model assumption that atoms trapped in a 1D optical lattice are line-like radiation sources: their motion is tightly confined along the longitudinal direction (horizontal direction in the images) and optically not resolved (see also section 4), while it is only loosely confined along the transverse direction (vertical).

system is the well-known Airy disc pattern [35], whose first minimum corresponds to $1.22 r_A$, with r_A being the Abbe radius defined above.

(ii) A CCD detector samples and digitizes the image distribution $I(x, y) \mapsto I[x_i, y_j]$, where x_i and y_j denote the integer-valued horizontal and vertical positions of a sampling bin and the squared brackets in our notation distinguish discrete from continuous distributions because in general $I(x, y) \neq I[x_i, y_j]$. In fact, the discrete and continuous distribution are mathematically related through

$$I[x_i, y_j] = \sum_{n,m=-\infty}^{\infty} \mathcal{R}_s(x_i - n\, \Delta_s, y_j - m\, \Delta_s) \int_{-\infty}^{\infty} \mathcal{R}_p(n\, \Delta_s - u, m\, \Delta_s - v) I(u, v)\, du\, dv, \qquad (2)$$

where Δ_s represents the sampling spacing in the object plane, $\mathcal{R}_p(u, v)$ is the CCD pixel rectangular function equal to $1/\Delta_p^2$ for u and v in the interval $[-\Delta_p/2, \Delta_p/2]$ and zero outside, with $\Delta_p < \Delta_s$ being the size of the pixel projected onto the object plane (today's CCD detector have $\Delta_p \sim \Delta_s$). Likewise, $\mathcal{R}_s(u, v)$ is the sampling rectangular function equal to one for both u and v in the interval $[-\Delta_s/2, \Delta_s/2]$ and zero elsewhere. Equation (2) represents the convolution of the continuous intensity distribution with the pixel response function $\mathcal{R}_p(u, v)$ (digitization), which is multiplied by the 2D comb function with spacing Δ_s (discrete sampling) [36–38]. In order to prevent information loss by discrete sampling, the Nyquist-Shannon sampling theorem shows that the PSF—or, more precisely, the Abbe radius—must be imaged onto more than two CCD pixels, i.e., $r_A > 2\Delta_s$ [39, 40].

(iii) Physical information contained in the recorded signal $S[x_i, y_j]$ is partially lost due to diverse noise sources affecting the image formation process. The effect of noise sources (see table 2 for a complete list) is taken into accounted through an additive stochastic noise term $\epsilon[x_i, y_j]$, which is added to the fluorescence intensity distribution:

$$S[x_i, y_j] = I_{\text{fluo}}[x_i, y_j] + \epsilon[x_i, y_j] = (P * O)[x_i, y_j] + \epsilon[x_i, y_j]. \qquad (3)$$

Here we assumed that the homogeneous background (by digitization offset, stray light, and dark currents discussed in section 5.1) is subtracted from the signal so that the average value of the noise vanishes, $\langle \epsilon[x_i, y_j] \rangle = 0$.

In order to retrieve the original information $O(x, y)$ from $S[x_i, y_j]$, we need to invert equation (3) through deconvolution. However, deconvolution problems are in general ill-conditioned, especially in the presence of noise. A physical model assumption—known as regularization—must be employed to constrain the solutions. For example, atoms which are strongly confined in an optical lattice are modeled as identical, isolated emitters characterized by only two parameters: positions and fluorescence intensity. Numerous deconvolution strategies exist in the literature, differing in their effectiveness in constraining solutions and in the computational resources required [41]. Our specific parametric deconvolution approach mainly relies on a maximum-
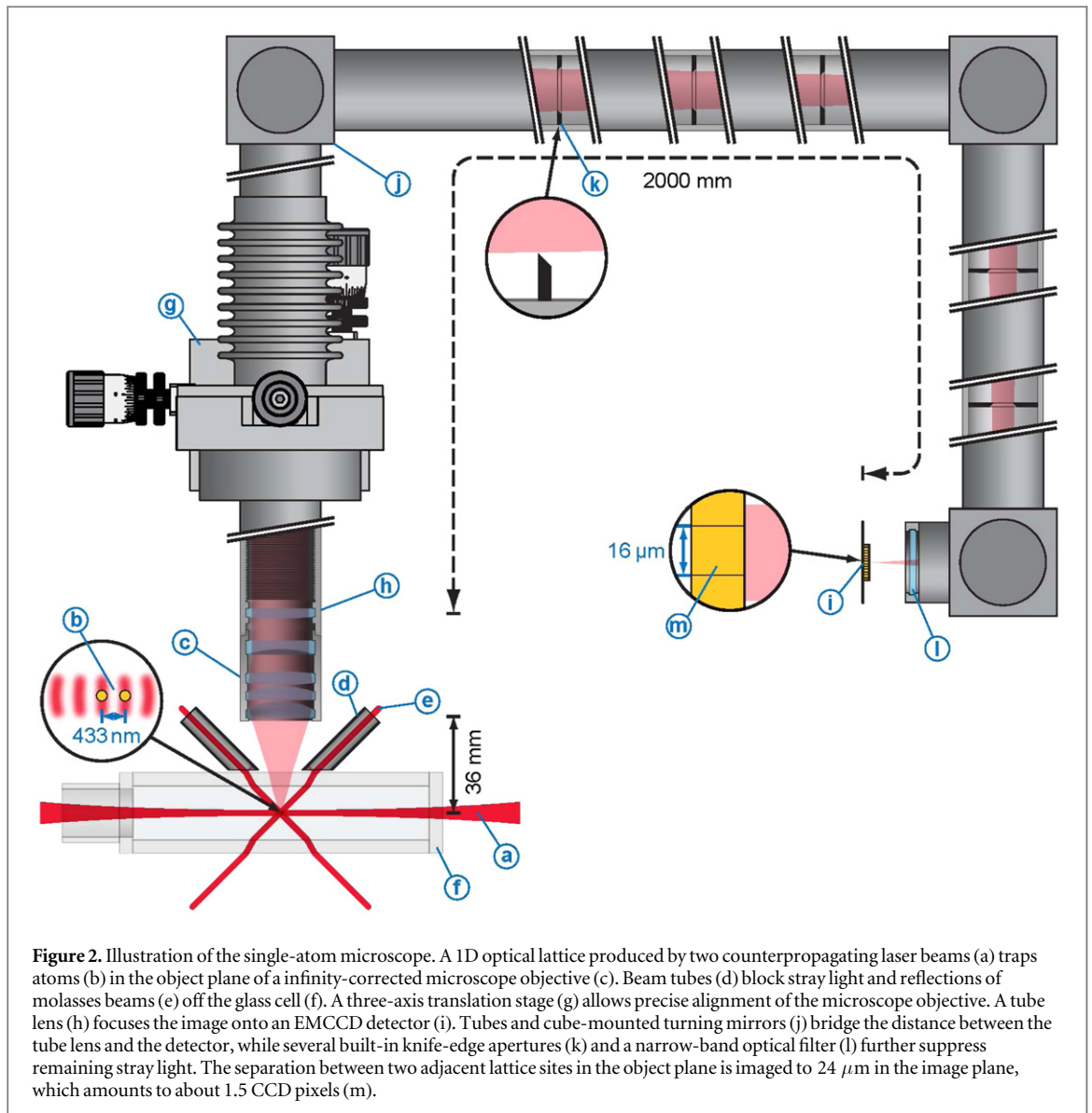
Figure 2. Illustration of the single-atom microscope. A 1D optical lattice produced by two counterpropagating laser beams (a) traps atoms (b) in the object plane of a infinity-corrected microscope objective (c). Beam tubes (d) block stray light and reflections of molasses beams (e) off the glass cell (f). A three-axis translation stage (g) allows precise alignment of the microscope objective. A tube lens (h) focuses the image onto an EMCCD detector (i). Tubes and cube-mounted turning mirrors (j) bridge the distance between the tube lens and the detector, while several built-in knife-edge apertures (k) and a narrow-band optical filter (l) further suppress remaining stray light. The separation between two adjacent lattice sites in the object plane is imaged to 24 μm in the image plane, which amounts to about 1.5 CCD pixels (m).

likelihood estimation constrained on a discrete lattice, for which a first estimation of the atoms' positions is provided by the so-called MUSIC (multiple signal classification) algorithm (see section 6).

3. The detection apparatus

3.1. The optical microscope

The imaging system depicted in figure 2 realizes an infinity-corrected microscope: the fluorescence light emitted by the atoms at $\lambda_f = 852$ nm is collimated by a diffraction-limited objective lens (effective focal length $f_1 = 36$ mm) [42] and imaged onto an EMCCD detector by a plano-convex tube lens (focal length $f_2 = 2$ m). The magnification of the microscope is $f_2/f_1 \sim 55$, so that the Abbe radius ($r_A = 1.9$ μm) of the PSF is imaged over >6 pixels of the CCD camera, thus fulfilling the requirement by the Nyquist–Shannon sampling theorem to avoid information loss [39, 40].

We assembled the microscope objective from four off-the-shelf spherical lenses, which are stacked into a one-inch aluminum holder. By design the objective lens operates at the diffraction limit with a NA of $NA_{obj. lens} = 0.29$. The objective lens was characterized with a shear-plate interferometer resulting in a peak-valley wavefront distortion of less than $\lambda_f/4$ over 90% of the clear aperture fulfilling Rayleigh's quarter wavelength rule [43]. The long working distance (36.5 mm) allows the objective lens to be mounted outside of the vacuum sufficiently far away from the vacuum cell to prevent reflected light from molasses laser beams from reaching the camera. The microscope objective is mounted on a three-axis translation stage and connected through blackened tubes to the EMCCD detector. Inside the tubes five sooted knife-edge apertures are placed

with gradually decreasing inner diameters to block stray light. To further suppress the remaining stray light a narrow-band (3 nm FWHM) optical filter with a transmission of 98% at the wavelength λ_f is placed in front of the EMCCD detector. Over the past few years, single aspheric lenses with long working distance (a few cm) have become widely available, and their utilization represents a good alternative to build an imaging system with a moderate NA (0.2 < NA < 0.5).

3.2. Localization of trapped atoms with small number of photons

According to Abbe's diffraction limit, the optical resolution of our imaging system is $r_A = 1.9~\mu$m. However, to achieve single-site resolution we need to extract the position of single trapped atoms with an uncertainty smaller than the lattice spacing a (433 nm). In analogy to super-resolution imaging in biological systems, we can determine the position of our atoms beyond the optical resolution by precisely knowing its PSF and the underlying noise. Following [44], in one-dimension the localization precision of the fluorescence peak produced by a single atom can be estimated by

$$(\Delta x)^2 = \frac{\mathrm{rms}_{\mathrm{PSF}}^2 + \Delta_{\mathrm{p}}^2/12}{N} + \frac{4\sqrt{\pi} \cdot \mathrm{rms}_{\mathrm{PSF}}^3\,\sigma_{\mathrm{b}}^2\,n_\perp}{\Delta_{\mathrm{p}} N^2}, \tag{4}$$

where it is assumed that the fluorescence signal is integrated over $n_\perp$ pixels in the direction transverse to the lattice, and that $\mathrm{rms}_{\mathrm{PSF}}$ is the rms width of a Gaussian PSF, Δ_{p} is the size of a camera pixel in the object plane, N is the average number of recorded photons per atom, and σ_{b} is the rms background noise (see equation (9) in section 5.1). In the literature, extensions of the result in equation (4) can be found for two-dimensions [45] and, using the statistical theory based on the Fisher information matrix, for a generic disc PSF (e.g., Airy disc) [28]. The Fisher information approach, which for a Gaussian PSF yields exactly equation (4), produces the fundamental theoretical localization limit that can be attained (Cramér–Rao information bound). Note also that the localization precision in equation (4) concerns only a single localized emitter, which is the case, for example, of an isolated fluorophore in photoactivated localization microscopy or of a very sparsely filled optical lattice. Section 6 is in particular concerned to super-resolve the position of emitters (atoms) clustered in small ensembles, which constitutes a significantly more demanding task. In addition, it should be noticed that, when employing an electron multiplying CCD (EMCCD) camera (as is the case of the present work), a factor 2 must to be added in front of $\mathrm{rms}_{\mathrm{PSF}}^2$ in equation (4) to account for the effectively halved quantum efficiency due to the electron multiplying excess noise (see section 5.1) [46].

In the following, we intend to give an estimate of the localization precision of our imaging system based on equation (4): the $\mathrm{rms}_{\mathrm{PSF}}$ of our imaging system is $\sim 1.5~\mu$m (see section 4) and the parameter Δ_{p} can be calculated by dividing the pixel size (16 μm) by the magnification (~ 55, see section 3.1). The number of photoelectrons (ph. e$^-$) recorded on the EMCCD sensor per single atom can be estimated by knowing the photon scattering rate, the solid angle of the microscope objective into which photons are emitted, and the exposure time. Atoms illuminated with nearly resonant light at λ_f emit photons at the maximal rate of $\Gamma/2$ for strong saturation, with $\Gamma \sim 2\pi \times 5$ MHz being the radiative decay rate for cesium. However, to prevent atoms from hopping along the lattice during imaging, the saturation parameter is typically chosen much smaller [47, 48] ($s \sim 0.01$), which reduces the scattering rate by a factor of 10 or more [49]. The solid angle directly depends on the NA of the imaging system according to the formula $\Omega/4\pi = \left(1 - \sqrt{1 - \mathrm{NA}^2}\right)/2 \sim 1\%$. By additionally taking into account the finite quantum efficiency of the CCD camera QE$(\lambda_f) \sim 30\%$ (see section 3.3) as well as photon losses ($\sim 6\%$) due to both reflections from optical surfaces (e.g. the vacuum glass cell) and the transmission of the narrow-band optical filter (see section 3.1), we expect to detect about 1000 ph. e$^-$ per atom for a single fluorescence image with an exposure time of $T = 1$ s. For comparison, in our experiments we record about 1300 ph. e$^-$ s^{-1} per atom as discussed in section 6.1. The measured background-noise distribution, which is analyzed in section 5.2, has a rms width σ_{b} of about 0.6 ph. e$^-$ per camera pixel. Since we integrate the fluorescence images along the direction transverse to the 1D optical lattice (see section 4.1), the variance of the background noise σ_{b}^2 is multiplied by the number of transverse pixels $n_\perp$ (typically $n_\perp \sim 40$). Hence, based on equation (4) we expect a localization precision of $\Delta x \sim 60$ nm, which is sufficiently smaller than the separation between two lattice sites. By using longer exposure times it is possible to improve the resolution even further, however, at the cost of decreasing the duty cycle and increasing the probability for atoms to either hop to adjacent lattice site or to be lost because of heating and background gas collisions. Moreover, a slow drift of the entire lattice with respect to the imaging system ($\leqslant 20$ nm s^{-1} [5]) is responsible for the existence of an optimal exposure time (estimated around 2 s for our system) beyond which the localization precision deteriorates instead of improving, if the lattice drift is not suitably tracked and accounted for. Such drifts are especially notable in case of imaging systems with very high optical resolution, as recently demonstrated through the measurement of the Allan variance associated with the position uncertainty of trapped ions [34].
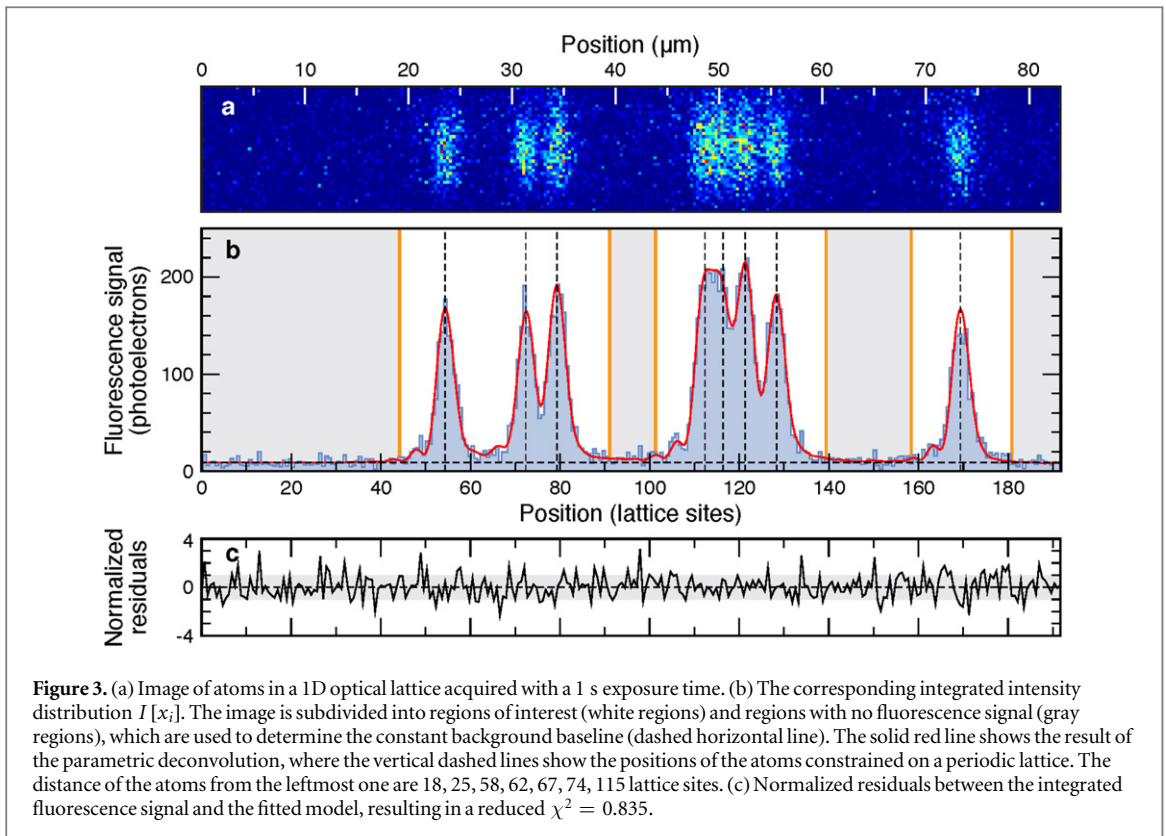
Figure 3. (a) Image of atoms in a 1D optical lattice acquired with a 1 s exposure time. (b) The corresponding integrated intensity distribution $I[x_i]$. The image is subdivided into regions of interest (white regions) and regions with no fluorescence signal (gray regions), which are used to determine the constant background baseline (dashed horizontal line). The solid red line shows the result of the parametric deconvolution, where the vertical dashed lines show the positions of the atoms constrained on a periodic lattice. The distance of the atoms from the leftmost one are 18, 25, 58, 62, 67, 74, 115 lattice sites. (c) Normalized residuals between the integrated fluorescence signal and the fitted model, resulting in a reduced $\chi^2 = 0.835$.

3.3. The EMCCD detector

In our experiment, the fluorescence signal of the atoms is detected using an EMCCD camera (Andor iXon DV897DCS-FI), whose read-out noise is more than one order of magnitude smaller compared to that of scientific-grade CCD sensors. In fact, scientific-grade CCD sensors are at present limited by a background noise $\sigma_b > 6$ ph. e$^-$ per pixel. The increased noise would deteriorate the localization precision estimated for our system by a factor of 6 (see equation (4)), therefore preventing reliable single-site resolution. To detect signals of few photons, such as fluorescence of single atoms, alternative types of imaging sensors exist, which include intensified CCD (ICCD) sensors and CMOS sensors. In appendix A we provide a review of sensors suited for few-photon-signal detection, and discuss technical noise sources inherent the different technologies are discussed. Our EMCCD camera employs a front-illuminated, frame-transfer, buried channel CCD sensor (L3Vision CCD97) containing 512×512 active pixels with a pixel size of 16×16 μm^2. A measurement comparing front- with back-illuminated sensors is given in appendix B. The quantum efficiency at the imaging wavelength λ_f with the EMCCD sensor cooled to -70 °C is QE$(\lambda_f) \sim 30\%$. It should be noted that the efficiency decreases at lower temperatures for wavelengths >800 nm due to a temperature dependence of the silicon band gap [50].

In EMCCD sensors, suppression of read-out noise is achieved through the serial electron multiplying (EM) register, which amplifies charges by impact ionization at each shift step, similar to a staircase avalanche photodiode (see appendix A). The EM register of our camera comprises $N = 536$ shift steps. Even though the probability of impact ionization at each individual step is only $p_{imp} \sim 1.5\%$, due to the large number of steps, the mean gain of the cascaded multiplication process $g = (1 + p_{imp})^N$ can reach values well above 1000. The effect of its stochastic nature on the detection noise is discussed in section 5.1.

4. The atomic sources

To acquire fluorescence images with the detection apparatus described in section 3 we trap atoms in a deep lattice potential with lattice constant $a = 433$ nm and illuminate them with a three-dimensional optical molasses. The saturation parameter of the optical molasses and the lattice trap depth ($U/k_B = 0.4$ mK) are chosen as such to prevent atoms from hopping along the lattice direction. Figure 3(a) exemplarily shows a fluorescence image of eight trapped atoms, which are loaded into a 1D optical lattice in stochastic positions and subsequently imaged with an illumination time of 1 s. The intensity distribution for each atom exhibits a characteristic elliptical shape elongated along the radial direction of the optical lattice with an aspect ratio of

about 6:1 (FWHM along the axial direction of 1 μm). The elongated shape originates from the thermal motion of trapped atoms ($\sim$30 μK by sub-Doppler cooling in the optical molasses) in the radial direction, along which the confinement is weaker. Along the lattice direction, instead, trapped atoms can be regarded as localized point sources with a Dirac-delta longitudinal distribution

$$O(x, y) = O(y)\delta(x) \tag{5}$$

with a certain radial intensity distribution $O(y)$, because the extent of the axial thermal motion (FWHM $\sim$60 nm) as well as the drift of the optical lattice ($\leqslant$20 nm s^{-1} [5]) is one order of magnitude smaller than the optical resolution. Being primarily interested in extracting the precise position of atoms along the optical lattice, we integrate the acquired images along the radial direction ($I[x_i] = \sum_j I[x_i, y_j]$) as depicted in figure 3(b), which reduces the deconvolution problem to a 1D one. The continuous curve overlapped with the integrated fluorescence signal shows the end result of the parametric deconvolution problem presented in section 6, which yields the atoms' positions with single lattice-site precision. Figure 3(c) shows the residuals between the reconstructed distribution and the measured signal, normalized to the expected noise strength. The uniform distribution of residuals with rms spread around one attests the quality of the parametric deconvolution, which is ensured by an accurate knowledge of the LSF of the imaging system as well as of the noise model. The methods to reconstruct the LSF are presented in following section 4.1, while the physical noise model is presented in section 5 and the parametric deconvolution process is illustrated in the last section 6.

4.1. Reconstructing the line spread function with sub-pixel resolution

One key element to achieve a resolution beyond the diffraction limit is the accurate knowledge of the response of our imaging apparatus. More precisely, it is important for the parametric deconvolution problem to know exactly the imaged fluorescence intensity distribution of a single illuminated atom. The importance of an accurate knowledge of this distribution is quantitatively demonstrated in section 6.4. In a 1D optical lattice, the problem of reconstructing the positions of atoms can be reduced to one-dimension by exploiting the factorized form of the single-atom distribution in equation (5). In fact, a single atom positioned at $x = 0$ yields (see equation (1)) a fluorescence distribution that integrated along the radial direction reads

$$\int_{-\infty}^{\infty} I_{\text{fluo}}(x, y)\,\mathrm{d}y = L(x) \int_{-\infty}^{\infty} O(y)\,\mathrm{d}y, \tag{6}$$

where $L(x) = \int_{-\infty}^{\infty} P(x, y)\,\mathrm{d}y$ represents the so-called line spread function. As argued in section 2, the response function required in the deconvolution problem is the convolution of the optical line spread function $L(x)$ with the 1D CCD pixel function $\mathcal{R}_{\text{p}}(x)$ [51],

$$L_{\text{CCD}}(x) = (\mathcal{R}_{\text{p}} * L)(x). \tag{7}$$

In the following we present our method to reconstruct the L_{CCD} function with increased signal-to-noise ratio and sub-pixel resolution, which is based on superimposing multiple intensity distributions of sufficiently isolated atoms (for example, the rightmost atom in figure 3(b)). The superimposing process is generally referred to as image registration in digital signal processing.

We make use of a recursive algorithm to process single-atom images, whose end result should ideally converge to L_{CCD} in equation (7). The algorithm is composed of a preparatory procedure followed by an iterative one. The first step of the preparatory procedure consists in identifying those regions of interest (ROI) containing exactly one atom well separated from other atoms by several Abbe radii (typically 10) in order to allow us not only to reconstruct the central peak of the LSF but also the wings containing the diffraction fringes. In the next step, we apply a Fourier filter to each single-atom image to remove high-spatial-frequency noise. The filter utilizes the fact that every optical system with a hard aperture has a cutoff in the optical transfer function (OTF), defined as the Fourier transform of L_{CCD}, exactly at the Abbe frequency $1/r_{\text{A}} = 2\text{NA}/\lambda_{\text{f}}$. After discrete Fourier transformation (DFT) of the integrated intensity distributions, the filter sets the amplitude of all frequencies beyond the Abbe cut-off (typically $>1.2/r_{\text{A}}$ to reduce Fourier artifacts) to zero because these frequencies components do not carry physical information (OTF $= 0$ in this region). The effect of Fourier filtering is significant for our imaging system because the Abbe frequency is three times smaller than the Nyquist frequency of 0.5 pixel^{-1}—the frequency up to which noise appears if not filtered out. The last step of the preparatory procedure to reconstruct the LSF consists in interpolating the noise-filtered single-atom distributions with sub-pixel resolution, which allows us to reposition them in the subsequent iterative procedure with high precision. Because of the finite bandwidth of the OTF, the integrated fluorescence signal can be interpolated with an arbitrary spatial resolution using the Whittaker–Shannon interpolation formula: we extend the DFT fluorescence distribution in Fourier space beyond the Abbe cut-off with zero values (zero padding), so that the number of points in the Fourier space is increased by an integer factor s with respect to the original number. The inverse DFT of the zero-padded signal results in an upsampled distribution, where the width of a sub-pixel is equal $1/s$ of the original pixel's width. The size of the sub-pixel is chosen smaller than the estimated localization
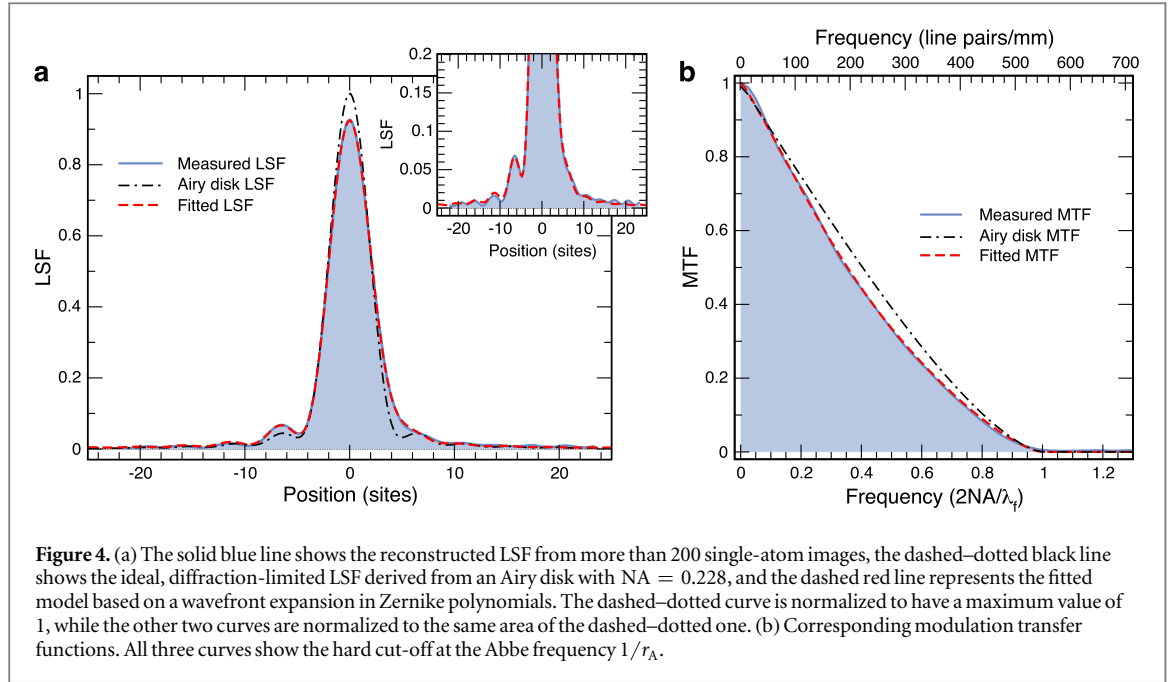
Figure 4. (a) The solid blue line shows the reconstructed LSF from more than 200 single-atom images, the dashed–dotted black line shows the ideal, diffraction-limited LSF derived from an Airy disk with NA = 0.228, and the dashed red line represents the fitted model based on a wavefront expansion in Zernike polynomials. The dashed–dotted curve is normalized to have a maximum value of 1, while the other two curves are normalized to the same area of the dashed–dotted one. (b) Corresponding modulation transfer functions. All three curves show the hard cut-off at the Abbe frequency $1/r_A$.

precision (typically $s = 8$ so that 1/8 pixel = 37 nm $< \Delta x$, see section 3.2). An alternative yet equivalent application of the Whittaker–Shannon interpolation formula operates directly in position space by convolving the spatial distribution with a sinc function.

The iterative part of the reconstruction algorithm consists chiefly of two steps. In the first one, we obtain the position of each atom by a nonlinear least squares fit of the model distribution L_{CCD} to the recorded fluorescence signal (see section 6 for more details). The precise (unrounded) value of the atom position is used to shift and align all noise-filtered sub-pixel-interpolated intensity distributions. Hence, superimposing all images gives a reconstruction of the fluorescence distribution of a single atom with a signal-to-noise ratio enhanced by a factor $\sqrt{N_{at}}$, where N_{at} is the number of superimposed single atoms (typically a few hundreds). The reconstructed distribution L_{guess} provides us with a new estimate of L_{CCD}. The iterative algorithm stops when no change is observed (typically after 5–10 iterations). For the first iteration, we use a Gaussian function to determine the position of single atoms in case no LSF function is *a priori* known.

A mathematical derivation (see appendix C) shows that this algorithm converges to

$$L_{guess}(x) = (\mathcal{R}_x * \mathcal{R}_{sp} * \mathcal{R}_p * L)(x), \qquad (8)$$

instead of the desired expression in equation (7), where $\mathcal{R}_x$ is the probability distribution of the nonlinear least squares estimator of the single-atom position for an atom ideally positioned in the origin $x = 0$ (with a rms width $\Delta x \sim 60$ nm, see section 3.2), and $\mathcal{R}_{sp}$ the sub-pixel function equivalent to the pixel function $\mathcal{R}_p$ but s times narrower. Because the 'blurring' effect by both additional convolutions in equation (8) is on the order of a few tens of nanometers, we conclude that $L_{guess}(x) \sim (\mathcal{R}_p * L)(x)$ to a good approximation. For precise reconstruction of the L_{CCD} function, equation (8) shows that it is advantageous to increase the illumination time in order to decrease Δx. The reconstructed LSF and the related modulation transfer function (MTF = |OTF|) are displayed in figure 4 and analyzed in detail in the following section.

4.2. Analysis of the reconstructed line spread function

Besides its importance to retrieve the atoms' positions with the maximum localization precision (see section 6.4), the line spread function contains valuable information about the performance of the optical system. Since the influence of $\mathcal{R}_p$ in equation (7) is small (ensured by the Nyquist-Shannon condition $r_A > 2\Delta_s$), the optical line spread function L is well approximated by L_{CCD}. Figure 4(a) shows the reconstructed LSF obtained with the algorithm outlined in the foregoing section. In case of an aberration-free imaging system, the PSF is described by the well-known Airy disk, whose corresponding LSF is displayed for comparison in the same figure. Besides an overall agreement, the reconstructed LSF exhibits a lower maximum and a distinct asymmetry such that the higher-order diffraction peaks are only visible on the left-hand side. These differences arise from wavefront distortion caused by optical aberration. Mathematically, the PSF is defined by computing the modulus square of the Fourier transform of the electric field (wavefront) at the pupil (Fraunhofer diffraction). The wavefront contains all information about optical aberrations and can be expressed in the basis of Zernike polynomials [43]. To gain insight into the nature and amount of the optical aberrations affecting our optical

Table 1. Result of the wavefront fitting to the measured LSF expressed in terms of low-order Zernike polynomials. The overall wavefront distortion is obtained by adding the different contributions in quadrature. 1D fitting of our model to the LSF cannot prevent a certain ambiguity on the identification of wavefront distortion angles (not displayed).

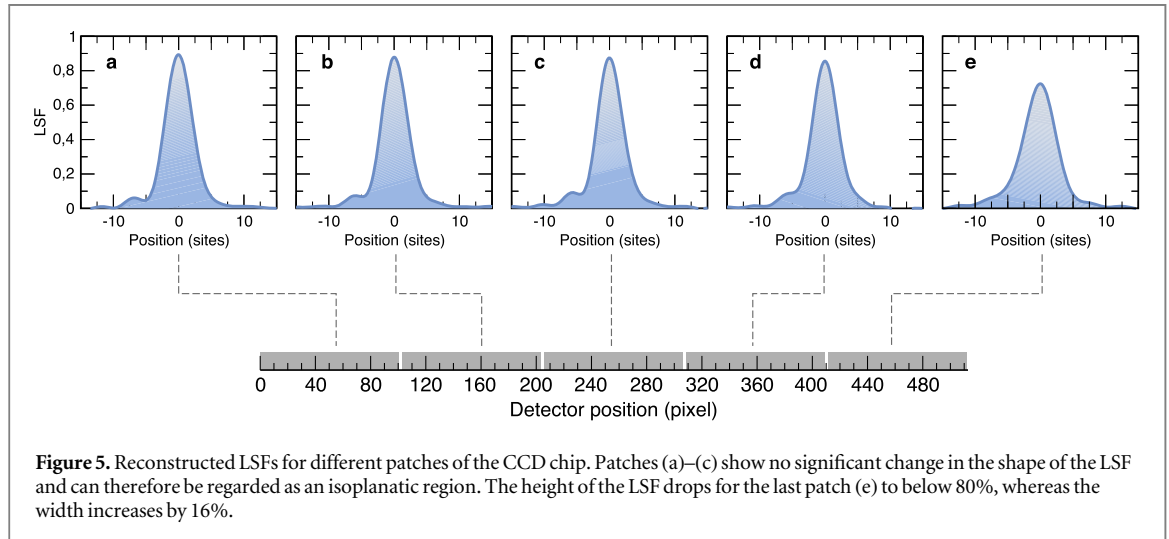|	Defocus	Astigmatism	Coma	Trefoil	Spherical
Orders (radial, azimuthal)	(2,0)	(2,2)	(3,1)	(3,3)	(4,0)
Rms wavefront distortion (λ units)	0.016(2)	0.048(2)	$-0.007(1)$	$-0.025(1)$	0.013(1)

Figure 5. Reconstructed LSFs for different patches of the CCD chip. Patches (a)–(c) show no significant change in the shape of the LSF and can therefore be regarded as an isoplanatic region. The height of the LSF drops for the last patch (e) to below 80%, whereas the width increases by 16%.

system, we fitted to the reconstructed LSF the one obtained from a wavefront expansion in low-order Zernike polynomials up to spherical aberration. The fitted LSF is displayed in the same figure, demonstrating a remarkable agreement with the experimental curve. The fit analysis reveals that the leading aberration contribution arises from astigmatism. The detailed list of Zernike coefficients is given in table 1. Combining all contributions in the table yields an overall rms wavefront error of $\sim\lambda/17$ (whereas the peak-valley deviation is $\lambda/3$), which corresponds to a Strehl ratio of 0.87 defined as the ratio between the maxima of the measured PSF and the ideal one. Note that the Strehl ratio, in general, differs from the ratio obtained analogously for the 1D LSF (about 0.92, see figure 4). In addition, the wavefront analysis gives an estimate of the actual NA of the optical system, NA $= 0.228(3)$. The deviation between the estimated NA and the one of the objective lens design (NA $= 0.29$) is most likely caused by clipping at the knife-edge apertures along the imaging path, see figure 2. Concurrently with this work, a similar wavefront analysis based on Zernike polynomials has been carried out to characterize the aberrations affecting two-dimensional fluorescence images of single trapped ions [34].

Figure 4(b) shows the modulation transfer function of the reconstructed LSF compared to that of an aberration-free optical system and of the fitted wavefront model. The MTF of an optical system with a hard aperture has a simple, direct geometrical interpretation, which explains the shape as well as the hard cut-off. In general, it can be shown that the MTF is directly computed by convolving the pupil function with itself, where displacements of the electric field distribution in the convolution integral directly translate into spatial frequency units of the MTF [35]. Therefore, an optical system with a hard aperture, outside of which the pupil function is strictly zero, results in a cut-off of the MTF at the Abbe frequency. This cut-off also provides a direct method to extract the actual NA of the optical system without resorting to fitting wavefront distortions.

4.3. Isoplanatic regions
The deconvolution problem described in section 2 relies on a translationally invariant response of the optical system. However, in real systems the LSF varies over the field of view because of optical aberrations primarily due to coma. Due to spatial variations, the localization precision of the atoms' positions is reduced if a single LSF is employed over the full field of view. In the literature, regions over which the LSF remains effectively unchanged are known as isoplanatic regions. To characterize the homogeneity of the LSF of our imaging system, we divide the full CCD region into five patches, each spanning over 100 pixels, where we reconstructed the LSF individually for each patch using the algorithm presented in section 4.1, see figures 5(a)–(e). The first three patches exhibit minor changes in the optical response, while the rightmost one shows a clearly visible broadening and decreased peak hight. Fluorescence images of atoms analyzed in this manuscript are from the three leftmost regions only. To even account for minor spatial deviations, the first three regions are further
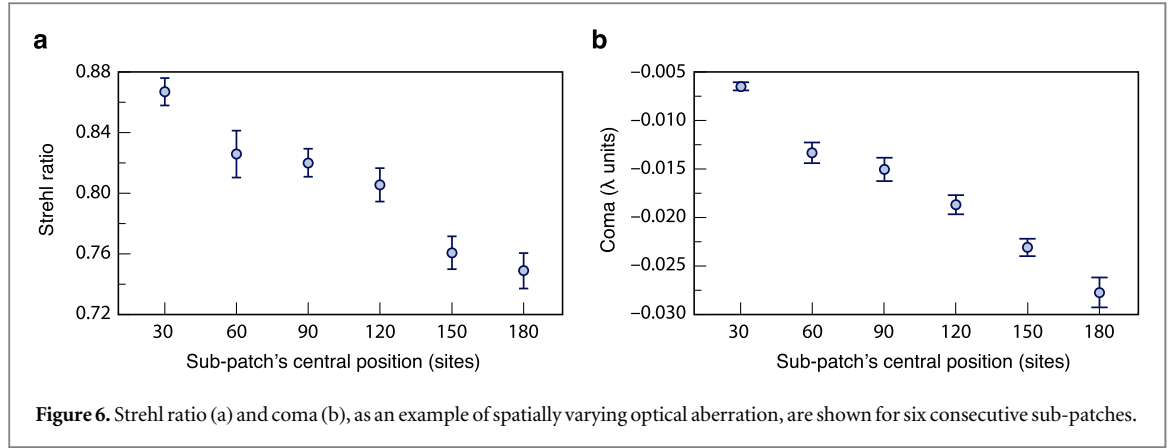
Figure 6. Strehl ratio (a) and coma (b), as an example of spatially varying optical aberration, are shown for six consecutive sub-patches.

Table 2. Noise contributions affecting single-atom imaging, a EMCCD detector cooled to $-70\,°C$, and 1 s exposure time. The overall noise σ is obtained from equation (9), which takes into account the EM amplification factor g and the excess noise factor F. Rms noise values extracted from: (a) figure 8, (b) figure 7. Rms noise values referring to technical specifications: (c) for inverted mode operation CCD with $T < -50\,°C$, (d) for EMCCD sensors, (e) 10 MHz read-out rate.

Noise type	Physical origin	Scaling	rms noise
Fluorescence photon shot noise	Uncorrelated scattered photons	$\langle S_{\text{fluo}} \rangle^{1/2}$	$\sigma_{\text{fluo}}^{(a)} \lesssim 4\ e^-\big/\sqrt{\text{pixel s}}$
Stray light shot noise	Spuriousreflections oflaser fields	$\langle S_{\text{stray}} \rangle^{1/2}$	$\sigma_{\text{stray}}^{(b)} \sim 0.5\ e^-\big/\sqrt{\text{pixel s}}$
Laser intensity noise	Technicalfluctuations	$\langle S_{\text{fluo}} + S_{\text{stray}} \rangle$	σ_{int}
Photo response non-uniformity	CCD substrateinhomogeneity	$\langle S_{\text{fluo}} + S_{\text{stray}} \rangle$	σ_{PRNU}
Dark current	Thermal chargegeneration	$\langle S_{\text{therm}} \rangle^{1/2}$	$\sigma_{\text{therm}}^{(c)} \sim 0.01 - 0.1\ e^-\big/\sqrt{\text{pixel s}}$
Clock induced charges	Charge transfer	$\langle S_{\text{CIC}} \rangle^{1/2}$	$\sigma_{\text{CIC}}^{(d)} \sim 0.05 - 0.1\ e^-\big/\sqrt{\text{pixel}}$
Read-out noise	Amplification and digitization processes	Read-out rate, temperature	$\sigma_{\text{readout}}^{(e)} \sim 50/M\ e^-\big/\sqrt{\text{pixel}}$
Excess noise	EM amplification	Constantfactor for $g \gg 1$	$F = \sqrt{2}$

divided into sub-patches, each of them with an individual local LSF used to reconstruct atoms' position. The result of the wavefront-deviation analysis of the sub-patches, according to the method presented in the previous section, is illustrated in figure 6.

5. Modeling the noise sources

In order to identify the exact lattice-site locations of individual atoms with high reliability, not only the optical response of the imaging system should be precisely known (see previous section 4) but also an accurate model of all relevant noise sources should be constructed. A well-designed imaging apparatus should strive to attain a rms noise limited by fluorescence photon shot noise, which represents the true fundamental noise contribution. In general, this can be reached (as is the case of this work) by understanding, and suppressing where required, the technical noise contributions affecting the image formation. Moreover, one should also be aware of the excess noise adding on top of shot noise, which is caused by the EM register in EMCCD cameras. This noise contribution effectively decreases the signal-to-noise ratio by a factor $\sqrt{2}$ and cannot be simply eschewed as it is intrinsic to the technology of EMCCD cameras. Alternative detectors such as CMOS cameras, which also feature small read-out noise, are discussed in appendix A. In this section, we discuss the relevant noise sources and show that the next noise contribution after fluorescence photon shot noise is photon shot noise caused by background stray light (≈ 0.5 ph. $e^-\big/\sqrt{\text{pixel}}$). A summary of individual noise components with their scaling and quantitive estimates is provided in table 2.

5.1. Noise sources in the detection process

Fluorescence photon shot noise originates from fluctuations in the number S_{fluo} of accumulated photoelectrons. This noise component originates from three independent stochastic processes: scattering of photons by atoms (Poissonian distributed), photoelectron generation with finite quantum efficiency $\text{QE}(\lambda_f)$

(binomially distributed), stochastic partitioning by imaging the PSF over several CCD pixels (binomially distributed). The resulting pixel's fluorescence distribution follows Poissonian statistics characterized by an average fluorescence signal $\langle S_{\text{fluo}}[x_i, y_j] \rangle = I_{\text{fluo}}[x_i, y_j] = \text{QE}(\lambda_f) F[x_i, y_j] \, T$, where $F[x_i, y_j]$ is the average photon flux directed onto a given pixel of the CCD sensor, T is the exposure time, and $I_{\text{fluo}}[x_i, y_j]$ the intensity distribution from equation (3). The rms shot noise is $\sigma_{\text{fluo}}[x_i, y_j] = \langle S_{\text{fluo}}[x_i, y_j] \rangle^{1/2}$.

Stray light contributes with Poissonian noise due to fluctuations in the number of photoelectrons $S_{\text{stray}}[x_i, y_j]$. Stray light can be minimized by shielding the objective and blocking reflections of the molasses laser beams as shown in figure 2. The remaining homogeneous stray-light background yields a rms noise $\sigma_{\text{stray}} = \langle S_{\text{stray}} \rangle^{1/2}$.

Illumination intensity noise is produced by temporal intensity fluctuations of the laser light that illuminates the atoms. The fluorescence emission rate is directly proportional to the illumination intensity for small saturation parameters (see section 3.2). Hence, fluctuations of the laser intensity result in a rms noise proportional to the detected signal, $\sigma_{\text{int}}[x_i, y_j] \propto \langle S_{\text{fluo}}[x_i, y_j] \rangle$.
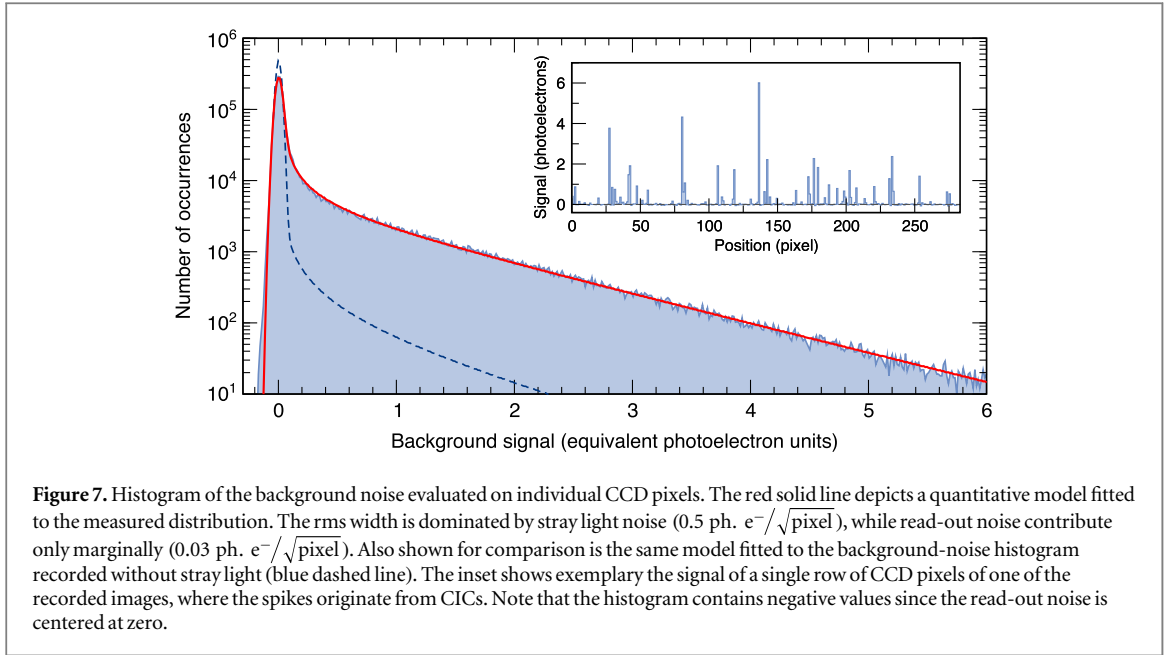
Photo-response non-uniformity (PRNU) is caused by variations in the pixel geometry and in the substrate material across the CCD chip. In back-illuminated EMCCD sensors, this also includes the so-called etaloning effect due to interference fringes in the a back-thinned silicon substrate (see appendix B). The rms noise is proportional to the overall incident photon flux, $\sigma_{\text{PRNU}}[x_i, y_j] \propto \langle S_{\text{fluo}}[x_i, y_j] \rangle + \langle S_{\text{stray}} \rangle$. Because of its static nature, this noise contribution can be reduced by calibrating the CCD sensor sensitivity with a uniform illumination source in order to remove pixel-to-pixel variations.

Read-out noise occurs in the charge-to-voltage amplification and analog-to-digital conversion process. Because this noise component σ_{readout} is not amplified by the EM register, it is effectively suppressed by setting the multiplication gain to large values.

Dark current noise arises from thermally generated charges. Buried-channel sensors are affected by two contributions—bulk and surface dark currents—depending whether electron–hole pairs are generated in the depletion region or at the silicon-silicon dioxide interface. Midgap edge states, either localized in the proximity of bulk impurities or at the front interface, strongly enhance the probability of electrons to thermally hop from the valence to the conduction band through a two-step trap-assisted process [52]. A third contribution, in general negligible at low temperatures, comes from diffusion of minority carriers (electrons) from the p-doped silicon substrate into the depletion region. Because of the large continuum of edge states at the interface with the silicon dioxide layer, the surface contribution dominates by about two orders of magnitude. Since all dark currents are amplified by the EM register, cooling of the CCD sensor is necessary to detect signals with few photons. For Peltier-cooled sensors ($T > -100\,^{\circ}\text{C}$), the contribution by surface dark currents is suppressed in inverted-mode EMCCD chips by applying a large negative bias voltage (multi-pinned phase mode [53]), which creates an inversion layer of holes at the surface preventing electrons from filling the midgap states, and thus suppressing the charge generation process. Fluctuations of the number of thermally generated electrons in the bulk, S_{therm}, adds a Poissonian noise component with rms noise $\sigma_{\text{therm}} = \langle S_{\text{therm}} \rangle^{1/2}$.

Clock-induced charges (CICs) are a spurious electronic signal, S_{CIC}, generated during the charge transfer process in the CCD sensor when the clock voltage switches the pixel from inversion to non-inversion mode. The process accelerates holes at the inversion layer back to the heavily doped p-type regions (channel stops), which produce charge carriers through impact ionization. Despite their ubiquitous presence in every CCD sensor, CICs can only be detected in EMCCD sensors due to the extremely low read-out noise. CICs are reduced by high parallel transfer rates, small slew rates, and small clock swing [54], while they cannot be suppressed by cooling the EMCCD chip (the probability of impact ionization even increases with lower temperatures). By advanced clock-waveform shaping, modern EMCCD cameras can strongly reduce CICs produced by the vertical transfer process [55]. CICs produced in the serial and multiplication register cannot be simply explained in terms of impact ionization by the accelerated holes. In modern EMCCD cameras, the generation probability of horizontal CICs per shift step results similar to that of vertical CICs [56] in spite of what was originally deemed [57]. The stochastic generation of CICs is Poissonian distributed with a rms noise denoted by $\sigma_{\text{CIC}} = \sqrt{\langle S_{\text{CIC}} \rangle}$ [58].

Excess noise arises from the stochastic nature of the gain process in the EM register of EMCCD cameras, which causes an asymmetric broadening of noise distributions. The resulting noise distribution after amplification by the EM register has an rms noise increased by the so-called excess noise factor (denoted by F), which tends to $\sqrt{2}$ for a large number of multiplication stages ($g \gg 1$), as can be shown [59].

Figure 7. Histogram of the background noise evaluated on individual CCD pixels. The red solid line depicts a quantitative model fitted to the measured distribution. The rms width is dominated by stray light noise (0.5 ph. $e^-/\sqrt{\text{pixel}}$), while read-out noise contribute only marginally (0.03 ph. $e^-/\sqrt{\text{pixel}}$). Also shown for comparison is the same model fitted to the background-noise histogram recorded without stray light (blue dashed line). The inset shows exemplary the signal of a single row of CCD pixels of one of the recorded images, where the spikes originate from CICs. Note that the histogram contains negative values since the read-out noise is centered at zero.

The overall signal measured by the EMCCD camera is the sum of all components, $S[x_i, y_j] = S_{\text{fluo}}[x_i, y_j] + S_{\text{stray}} + S_{\text{therm}} + S_{\text{CIC}}$, whose variance σ^2 is the quadrature sum of all individual noise components

$$\sigma^2[x_i, y_j] = F^2(\sigma^2_{\text{fluo}}[x_i, y_j] + \sigma^2_{\text{stray}} + \sigma^2_{\text{int}}[x_i, y_j] + \sigma^2_{\text{PRNU}}[x_i, y_j] + \sigma^2_{\text{therm}} + \sigma^2_{\text{CIC}}) + \frac{\sigma^2_{\text{readout}}}{g^2}. \qquad (9)$$

Note that $\sigma^2[x_i, y_j]$ is also the variance of $\epsilon[x_i, y_j]$, which is defined as $\epsilon[x_i, y_j] = S[x_i, y_j] - \langle S[x_i, y_j] \rangle$ in equation (3). Equation (9) shows that high EM gains g improve the signal-to-noise ratio for signals of few photons by effectively removing the read-out noise component σ_{readout}, which in CCD sensors dominates over σ_{fluo}, instead. In turn, EMCCD cameras are affected by excess noise through the factor F, which effectively halves the quantum efficiency.
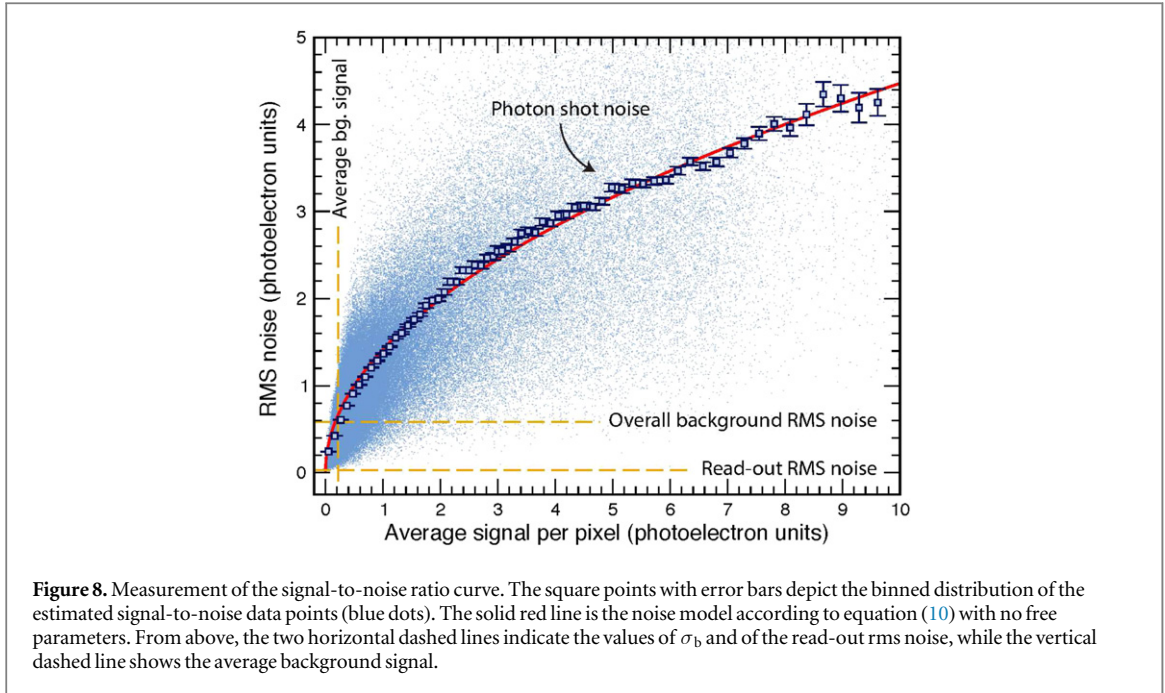
5.2. Noise model
The knowledge of each individual noise contribution and their physical scaling with respect to the spatially dependent fluorescence signal $S_{\text{fluo}}[x_i, y_j]$ (see table 2) permits to construct a noise model, which is used in the parametric deconvolution process in section 6.3. Based on the functional dependence of σ on S_{fluo}, we rewrite equation (9) in the form

$$\sigma(\langle S_{\text{fluo}} \rangle) = \sqrt{\sigma^2_{\text{b}} + c_1^2 \langle S_{\text{fluo}} \rangle + c_2^2 \langle S_{\text{fluo}} \rangle^2}, \qquad (10)$$

where the rms background noise σ_{b} as well as the coefficients c_1 and c_2 are parameters to be experimentally determined. When the signal S_{fluo} is given in photoelectron units, the coefficient c_1 is directly given by the excess noise factor $F = \sqrt{2}$. The coefficient c_2 is compatible with zero for our experimental apparatus as shown in the following.

We determine the coefficient σ_{b} by analyzing the background noise of the imaging system from a series of images recorded without atoms in the optical lattice (see inset of figure 7). The histogram in figure 7 shows the background signal per pixel binned by signal strength in units of photoelectrons. The characteristic shape of the histogram mainly arises from read-out noise (Gaussian peak) and stray light (exponential tail). Calculating the rms width of the recorded histogram yields the desired coefficient $\sigma_{\text{b}} = 0.6$ ph. $e^-/\sqrt{\text{pixel}}$ for 1 s illumination time. To obtain quantitative insight into the different noise components of the background signal, we model it in terms of Poisson-distributed CICs, which are stochastically amplified through the EM register, on top of which Gaussian-distributed read-out noise is added [60]. The red line in figure 7 shows the fitted model reproducing closely the recorded background-noise histogram, whereas the dashed blue line shows the same model fitted to a background-noise histogram for images recorded with the camera shutter closed. Due to the blocked stray light, the exponential tail is reduced by more than one order of magnitude. However, it does not fully vanish because of dark currents and primarily CICs (see section 5.1). Fitting the background-noise histogram allows not only a more accurate estimate of the rms background noise (σ_{b}), but also a precise calibration of the EM gain g, which is a free fitting parameter. In fact, the parameter g enters the model through the probability to record y electrons

Figure 8. Measurement of the signal-to-noise ratio curve. The square points with error bars depict the binned distribution of the estimated signal-to-noise data points (blue dots). The solid red line is the noise model according to equation (10) with no free parameters. From above, the two horizontal dashed lines indicate the values of σ_b and of the read-out rms noise, while the vertical dashed line shows the average background signal.

after the EM register for x initial electrons—comprising both spurious electrons and photoelectrons—which is given by [61]

$$\mathcal{P}(y|x) = \frac{y^{x-1}\exp(-y/g)}{g^x(x-1)!}. \tag{11}$$

To validate the noise model given in equation (10), we perform a measurement of the signal-to-noise relationship [62]. Based on >1000 sets of five consecutive fluorescence images (3 s exposure time each) containing a small ensemble of atoms, we estimate for every CCD pixel the average signal as well as the standard deviation (rms noise) in each set of images. Note that we consider only image sets where the distribution of atoms remains constant (neither atom hopping nor losses). The cloud of dots in figure 8 shows the correlation between the estimated average signal and the estimated rms noise, both expressed in photoelectron units. By binning the signal-to-noise data points by their signal strength, we obtain a precise reconstruction of the signal-to-noise relationship, as shown in the figure. The measurement is in good agreement with the noise model calculated using the coefficients σ_b, c_1 and c_2 given above, therefore confirming the square-root dependence of the rms noise on the fluorescence signal strength S_{fluo}. Furthermore, because no linear dependency is discernible in the recorded signal-to-noise relationship, we set c_2 to zero.

6. Localization of atoms by parametric deconvolution

We retrieve the position of atoms in the optical lattice using a parametric deconvolution process, which comprises several stages: (1) the 1D integrated fluorescence images are divided into ROI, each with a small number of atoms. (2) The number of atoms is determined for each ROI based on the total number of photoelectrons. (3) We create a model function of the fluorescence distribution for the given number of atoms and (4) use it to obtain a first estimate of the positions of atoms employing a spectral-density estimation algorithm. (5) The estimated positions provide the starting values for a nonlinear least squares estimate, which yields the location of atoms with improved precision. (6) We further enhance the localization accuracy by an additional stage that constrains the atoms' positions to the discreteness of the optical lattice and merges all ROIs together.

6.1. Counting atoms in ROI
The knowledge of the exact number of atoms is a necessary prerequisite in order to determine the positions. While the identification of the number of atoms is relatively straightforward when the atoms are well separated from each other, e.g. by counting the number of peaks in the intensity distribution, it is more difficult when the atoms cluster together with separations smaller then the optical resolution. In such a situation, the peaks of individual atoms are no longer discernible. Hence, instead of counting peaks, we estimate the number of atoms
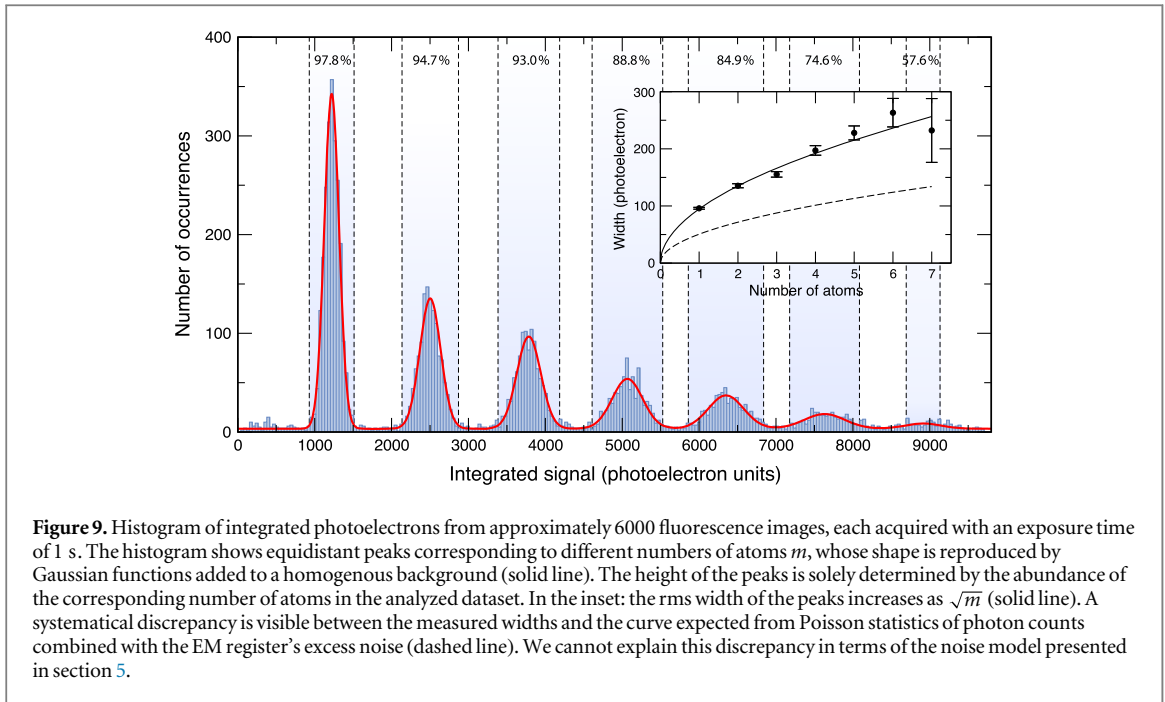
Figure 9. Histogram of integrated photoelectrons from approximately 6000 fluorescence images, each acquired with an exposure time of 1 s. The histogram shows equidistant peaks corresponding to different numbers of atoms m, whose shape is reproduced by Gaussian functions added to a homogenous background (solid line). The height of the peaks is solely determined by the abundance of the corresponding number of atoms in the analyzed dataset. In the inset: the rms width of the peaks increases as $\sqrt{m}$ (solid line). A systematical discrepancy is visible between the measured widths and the curve expected from Poisson statistics of photon counts combined with the EM register's excess noise (dashed line). We cannot explain this discrepancy in terms of the noise model presented in section 5.

from the total number of recorded photoelectrons normalized to that of a single atom. Figure 9 shows the histogram of the integrated photoelectrons of a large number of ROIs, exhibiting well-separated equidistant peaks. Each peak corresponds to a different number of atoms m, with the leftmost peak marking the number of photoelectrons per atom (about 1300 ph. e$^-$ s^{-1}). The continuous curve in the figure shows the best fit to the recorded histogram based on the sum of seven Gaussian distributions combined with a homogeneous background, which is added to account for atom losses during the exposure time corresponding to the flat background. The width of the peaks obtained from the fitting model (see inset in figure 9) increases with $\sqrt{m}$. This broadening implies that adjacent peaks with $m > 5$ overlap significantly, thus preventing an unambiguous identification of the exact number of atoms m. This is one reason why we divide every integrated fluorescence distribution into smaller, well-separated ROIs, each containing at least one atom, as shown in figure 3(b). The width of each ROI is determined by thresholding method known as image segmentation algorithm. Besides, the subdivision in ROIs with a small number of atoms is also beneficial to reduce the computation time of the nonlinear last squares estimation of positions, which scales quadratically with the number of atoms in a ROI.

The parametric deconvolution problem requires as a precondition that the number of atoms is correctly determined. Therefore, it is important to identify acceptance regions R_i of the integrated photoelectron signal where the statistical hypothesis H_i—the analyzed ROI contains precisely i atoms—is verified with a probability higher than certain desired confidence levels. Moreover, we should also take into account the additional statistical hypothesis H_0 that one or more atoms are lost during the exposure time. Referring to the fitting model in figure 9, the ith Gaussian function describes the probability distribution when H_i occurs, while the homogenous background models the probability distribution when H_0 occurs. The acceptance regions R_i (shaded regions in the figure) are obtained by maximizing their width (i.e., the probability $\mathcal{P}(R_i|H_i)$ defining the power of the statistical test) under the provision that the statistical confidence remains higher than a certain desired confidence level (percentage numbers in the figure). The confidence level for the region R_i is defined as the probability that the hypothesis H_i is verified when the integrated photoelectron signal is detected in R_i, namely $\mathcal{P}(H_i|R_i)$[4]. From the analysis of acceptance regions in figure 9 we find that, given a certain confidence level, the width of the acceptance region is strongly determined by atom losses (homogeneous background). This result shows that is beneficial to minimize the atom loss probability during the illumination process ($\sim$1% in our case). By post-selection analysis, we only retain ROIs for which the integrated photoelectron signal lies inside one of the acceptance regions. Higher confidence levels can be chosen, but this implies narrower regions R_i and, therefore, more ROIs post-selected out. In order to achieve for the one-atom peak confidence levels >99% with small rejection rates (<1%), we segment the fluorescence image in smaller ROIs than those considered in the example displayed in figure 9.

The described method to determine the number of atoms, which relies on the total number of photoelectrons in each ROI, is relatively robust and simple to implement. However, it does not make optimal

---

[4] The conditional probability can be computed as $\mathcal{P}(H_i|R_i) = \mathcal{P}(H_i)\,\mathcal{P}(R_i|H_i)\big/\sum_j \mathcal{P}(H_j)\,\mathcal{P}(R_i|H_j)$.

use of physical information contained in the image since spatial information is lost after the integrating photoelectrons for each ROI. It is possible to improve the accuracy by incorporating the spatial information of the recorded fluorescence images through a Bayesian updating algorithm [63].

6.2. The model distribution

When illuminated by nearly resonant light, atoms in a deep optical lattice behave like identical light sources positioned at certain sites of the lattice. Supposing that the number of atoms m is exactly known (see section 6.1), we model the integrated signal of equation (3) as

$$I_{\mathrm{M}}[x_i] = \sum_{l=1}^{m} A_l L_{\mathrm{CCD}}(x_i - \xi_l),\tag{12}$$

where ξ_l are the positions of the atoms, and the amplitude factors A_l account for inhomogeneities from the the illumination lasers as well as from atom losses during the exposure time. L_{CCD} is the response function of the imaging system defined in equation (7) representing the fluorescence distribution of a single atom with sub-pixel resolution (numerical interpolation between sub-pixels permits its evaluation for any real-valued argument). Because we perform background subtraction on all integrated intensity distributions, the model in equation (12) does not require an additional constant offset. In addition, relying on the discreteness of positions in the optical lattice, we can express ξ_l as

$$\xi_l = a\, p_l + \delta_{\mathrm{L}},\tag{13}$$

where p_l are the desired integer positions in lattice-site units, a is the separation between lattice sites in CCD pixel units, and δ_{L} is the position offset of the optical lattice. For small optical aberrations (see sections 4.2 and 4.3), it is sufficient to consider a single value of a (1.47 pixels per lattice site) for the entire field of view. Moreover, losses by light-induced collisions prevents the detection of two (or more) atoms in the same lattice site in deep optical lattices [64, 65]. Hence, $p_l \neq p_{l'}$ for any pair of atoms l and l'.

6.3. The parametric deconvolution process

To retrieve the atoms' positions, we employ a nonlinear least squares optimization algorithm, which fits the model distribution I_{M} in equation (12) to the recorded fluorescence distributions. This parametric deconvolution approach allows us to make optimal use of physical information contained in the response function of the imaging system and in the noise model. However, nonlinear least squares optimizations require well-chosen starting parameters in order to guarantee the convergence to the global optimum. The parameters of our model are the amplitudes A_l and positions ξ_l of atoms. While an initial estimate of A_l can be directly obtained from the average number of photoelectrons per atom (see section 6.1), an estimate of positions ξ_l demands a separate procedure. Hence, to obtain the first estimate of positions for the nonlinear least squares optimization, we make use of the Wiener deconvolution combined with a spectral density estimation algorithm (MUSIC algorithm).

Wiener deconvolution. The main idea underlying our approach to obtain an estimate of ξ_l is understood by considering the Fourier transform of the model distribution in equation (12),

$$\mathcal{F}(I_{\mathrm{M}})[k] = \mathrm{OTF}_{\mathrm{CCD}}[k] \sum_{l=1}^{m} A_l\, \mathrm{e}^{\mathrm{i}2\pi k \xi_l} = \mathrm{OTF}_{\mathrm{CCD}}[k]\, f[k],\tag{14}$$

where the convolution theorem enables the optical transfer function $\mathrm{OTF}_{\mathrm{CCD}}[k]$ (see section 4.1) to be factored out of the sum. The function $f[k]$ is an oscillatory signal containing exactly m Fourier components, whose frequencies are exactly the positions ξ_l. This allows us to recast the problem of estimating the positions ξ_l in that of estimating a discrete number m of frequencies (spectral density estimation). The presence of noise in the recorded signal $S[x_i]$, however, makes $f[k]$ difficult to be computed from the ratio $\mathcal{F}(S)[k]/\mathrm{OTF}_{\mathrm{CCD}}[k]$. In fact, because the noise $\epsilon[x_i]$ has a white spectrum, if we divided the Fourier-transformed recorded signal $\mathcal{F}(S)[k]$ by the reconstructed optical transfer function $\mathrm{OTF}_{\mathrm{CCD}}[k]$, noise spectral components in the proximity of the Abbe frequency would be strongly amplified owing to the small magnitude of OTF at higher frequencies (see section 4.2). In order to obtain $f[k]$ avoiding noise amplification, we employ the Wiener deconvolution algorithm [66], which computes

$$f[k] = \frac{\mathcal{F}(S)[k]}{\mathrm{OTF}_{\mathrm{CCD}}[k]}\, \frac{\mathrm{MTF}[k]^2}{\mathrm{MTF}[k]^2 + 1/\mathrm{SNR}[k]},\tag{15}$$

where $\mathrm{SNR}[k] = f[k]^2/\sigma^2$ is the signal-to-noise ratio defined as the ratio between the estimated deconvolved signal $f[k]$, which is obtained by applying the filter iteratively (typically 10 iterations), and the integrated noise power in the analyzed ROI, which is estimated as $\sigma^2 = n_\perp n_\parallel \sigma_{\mathrm{b}}^2 + F^2 N$ (see also section 5.2). We recall that $n_\perp$ and $n_\parallel$ represent the number of CCD pixels in the ROI in the direction transverse and parallel to the lattice, respectively, and N is the integrated number of photoelectrons. The Wiener filter factor in equation (14) is
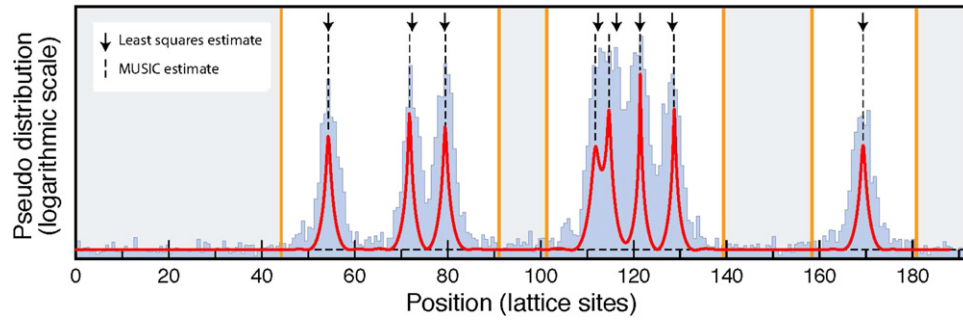
Figure 10. Pseudospectrum obtained by the MUSIC spectral density estimate algorithm (in logarithmic scale). The narrow peaks provide an estimate of the eight atoms' positions. The binned intensity distribution is the same as in figure 3(b) (in linear scale). The positions estimated by the MUSIC algorithm (vertical dashed lines) are very close within one lattice site to the positions determined by the least squares algorithm (arrows).

relevant only at higher frequencies, while it is approximately 1 at lower frequencies because $\text{SNR}[k]$ is very large (typically >100) and $\text{MTF}[k] \sim 1$ for $k\, r_A \ll 1$.

Spectral density estimate (MUSIC algorithm). Several methods exist in the literature to estimate the spectral density from a noisy signal $f[k]$. The simplest method known as periodogram method employs a DFT of $f[k]$ to determine the n dominant Fourier components, whose frequencies yield an estimate of the positions ξ_l. However, this method suffers from known deficiencies such as being a biased estimator and exhibiting spectral leakage. More refined methods known as subspace methods have been developed for the estimation of the spectral components when the signal contains exactly m dominant Fourier components (i.e. m atoms) with amplitudes well above the noise background. Among the subspace methods, the so-called MUSIC algorithm (multiple-signal classification) has been identified as the one exhibiting the highest spectral resolution [67]. MUSIC yields a pseudospectrum $f[k]$ exhibiting a negligible bias in case of sufficient signal-to-noise ratio [68] and not suffering from spectral leakage in contrast to non-parametric methods (e.g., periodogram). In particular, the strength of MUSIC algorithm for the first estimation of the atoms' position resides in its robustness against noise disturbances and in the fact that no prior knowledge of the parameters (i.e. the atoms' positions) is required. This differs from least-squares minimization procedures, which require good starting parameters to ensure a rapid converge to a global minimum. While our implementation of MUSIC algorithm requires the prior knowledge of the number of atoms, m, in the ROI, extensions of the algorithm exist in the literature that also estimate the number of sources [69], which could be helpful to handle very large ROIs with high filling factors. The solid line in figure 10 displays the estimated power spectral density obtained from the MUSIC algorithm applied to the fluorescence distribution shown in figure 3. The pseudospectrum exhibits eight sharp peaks much narrower than the diffraction-limit width, each approximately centered on the atoms' positions (note the logarithmic scale in the figure). The figure shows that the positions estimated by the MUSIC algorithm are very close to those determined by the more accurate nonlinear least squares estimator, which takes into account the dependence of noise on the signal.

Nonlinear least squares estimator. We use the position of the atoms estimated by the MUSIC algorithm as input parameters of the nonlinear least squares minimization

$$\min_{A_1\ldots A_m, \xi_1\ldots\xi_n} \left(\sum_{i=1}^{n_\parallel} \frac{(S_{\text{fluo}}[x_i] - I_M[x_i])^2}{\sigma^2(I_M[x_i])} \right), \tag{16}$$

where $S_{\text{fluo}}[x_i]$ is the background-subtracted fluorescence distribution, $I_M[x_i]$ the model distribution given in equation (12), σ the noise model presented in equation (10), and $n_\parallel$ is the number of pixels in the 1D ROI. In the minimization problem of equation (16), the positions ξ_l are treated as real-valued free parameters (compare section 6.4). Furthermore, we use the model distribution I_M instead of the measured signal S to estimate the noise variance σ^2 at the pixel x_i because the model function provides a better estimate of the fluorescence signal after a few iterations of the least squares minimization. Several algorithms exist to carry out the minimization in equation (16) such as the Levenberg–Marquardt method. In our case we employ a trust-region algorithm, which allows us to constrain the amplitude parameters A_l to physical boundaries (typically five times the width of the one-atom peak in figure 9). An example of the least squares parametric deconvolution is shown by the solid red line in figure 3(b). The accurate model in equation (12) constructed from the measured LSF and the weighting factors in equation (16) accounting for the correct variance of noise in each pixel ensure flat residuals with a variance around 1, as displayed in figure 3(c). For normal distributed residuals, the nonlinear least squares fit is equivalent to a maximum-likelihood estimator of positions [30], which defines the gold standard concerning the

extraction of physical information from fluorescence images, as argued in section 1. Because each 1D pixel of the integrated signal carries a large number of fluorescence photoelectrons (see figure 3(b)), the dominating Poisson-distributed shot noise is well approximated by a Gaussian distribution. However, excess noise in the EMCCD camera causes non-Gaussian deviations, which can be seen, for example, in the logarithmic graph of figure 7. Even neglecting this super-Poissonian noise characteristic, previous work using EMCCD cameras reports localization of single emitters with a precision attaining the Cramér–Rao information bound [70]. It is the purpose of future work to refine our estimation of the atoms' positions by maximizing the appropriate likelihood function in order to account for the EM excess noise [71] as well as for the Poissonian statistics in the limit of very small signals [72]. In addition, we find that the distribution of the sum of squared residuals obtained by analyzing the positions of atoms in > 5000 ROIs is well described by a χ^2 distribution with $n_{\parallel} - 2m$ degrees of freedom. This result suggests that the minimization procedure of equation (16) approaches the limit of the maximum-likelihood estimator of the atoms' positions.

6.4. Enhancing the localization precision at higher filling factors

The parametric deconvolution method outlined in section 6.3 works very reliably in case of ROIs containing only a few atoms separated by several lattice sites. This is the situation, for example, of single-particle experiments such as quantum walks [12, 18, 19]. However, the determination of the atoms' positions is less reliable for atoms clustered in small ensembles, where the lattice filling factor approaches unity [4]. In experiments investigating strongly interacting particles, it is particularly important to reconstruct the atoms' positions with a high reliability also when the spacing between particles is close to, or is even less than, the optical resolution of the imaging system [13]. By taking the discreteness of the lattice into account, we demonstrate that the previously presented parametric deconvolution method can be extended to achieve high success rates also for small ensembles of atoms that are closely packed. As argued in section 6.2, the fact that atoms are trapped in an optical lattice provides us with two pieces of information: (1) the distance between two atoms can only be a multiple of the intersite separation a (see equation (13)) and (2) two or more atoms cannot occupy the same lattice site. To exploit the discreteness of the optical lattice, the lattice constant a (433 nm) needs to be precisely known in units of CCD pixels. Its value can directly be computed from the magnification factor (see section 3.1) and the pixel size (see section 3.3) or more accurately measured by analyzing the distribution of distances between two atoms, which are obtained using the deconvolution method presented in the previous section.

To include the constraints (1) and (2), we adopt the following procedure: the positions of the atoms are initially determined by the parametric deconvolution process described in the previous section, and rounded to an integer multiple of lattice sites. We subsequently produce an array of all combinations of distances between atoms, where each distance is let vary by ±1 lattice site with respect to the initial estimate. Furthermore, we exclude all combinations where two atoms occupy the same lattice site. For each combination of distances, we perform a nonlinear least squares minimization of equation (16) with the amplitudes A_l and the lattice position offset δ_{L} as the only free parameters. We finally choose the combination of distances with the maximum likelihood, which provides us with the best guess of the positions of atoms. Moreover, the χ^2 distribution of the sum of squared residuals (see equation (16)) allows us to perform a likelihood-ratio test (e.g., Neyman–Pearson lemma) that rejects the reconstructed positions if the statistical confidence lies below a certain specified value. A related approach has also been reported in [11], however, without discussing how noise contributions are handled in the deconvolution problem.

In order to quantitatively benchmark the reliability of the discrete deconvolution method against the continuous one presented in the previous section, we employ Monte Carlo simulations of fluorescence images, which provide large statistics and the exact knowledge of the true positions. To simulate a pattern of atoms in the lattice, the model distribution in equation (12) is used to construct the fluorescence images, where the noise is randomly drawn to reproduce shot noise fluctuations of the fluorescence signal and the background noise distribution shown in figure 7. Our Monte Carlo simulations also incorporate the stochastic fluctuations produced by the EM register. In particular, we simulated four atoms equally spaced with the spacing varying from one to nine lattice sites. Figure 11 shows the success rate in determining the correct distance between the four atoms. The analysis of simulated images shows that the success rate of the continuous parametric deconvolution rapidly decreases for separations smaller than the Abbe radius r_{A} (diffraction limit). The drop in the success rate is even more evident when a Gaussian function is used instead of the precisely reconstructed LSF L_{CCD}, see section 4.1. In contrast, it is remarkable that the success rate of the discrete parametric deconvolution remains above 90% for almost all lattice filling factors. Moreover, the success rate even increases at unity filling since the number of possible configurations of the four atoms is strongly reduced.

We recently developed a new atom resorting technique that allows us to deterministically position a small ensemble of atoms in any arbitrary pattern on the 1D optical lattice. The experimental details of the resorting technique are the subject of a future publication [20]. We employed this technique to reposition, on demand, four atoms along the lattice, thus reproducing the same distributions studied in figure 11, with the atoms
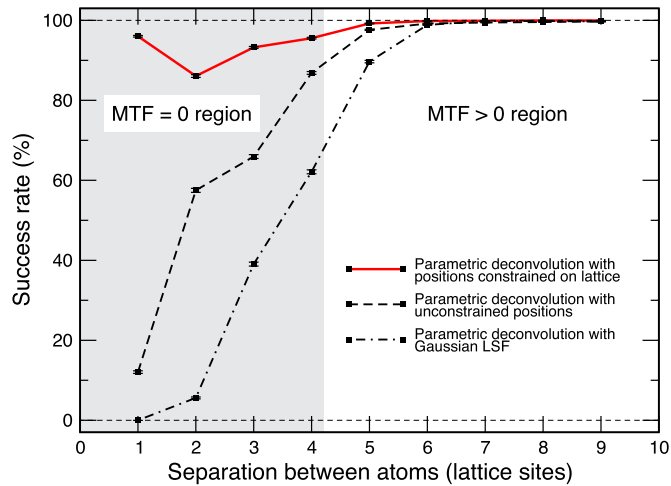
Figure 11. Success rates of finding the correct distances between four atoms equally spaced using different parametric deconvolution methods. Note that the success rates refer to the correct identification of all three distances separating the four atoms. For each point, we analyze 10 000 simulated images of four atoms with equidistant separations (1 s exposure time). The solid line includes the constraints on the atoms' positions by the optical lattice, the dashed line refers to the continuous deconvolution method with unconstrained positions, and the dotted–dashed line shows for comparison the success rate when Gaussian functions are used instead of the reconstructed LSF. All fits produce good success rates for separations larger than the Abbe limit (right-hand side), but only the discrete parametric deconvolution method achieves high success rates for all lattice filling factors.

separated by an equal number of sites (10, 5, 2 or even 1 lattice sites). Though based on a significantly smaller statistics, the experimental results are consistent with the theoretical findings obtained with a large number of simulated images, confirming an enhancement in the reliability of the parametric deconvolution if the positions are constrained onto the lattice.

7. Outlook

Optical diffraction imposes a stringent limit on the bandwidth of an optical system: physical information contained in the spatial frequency components above the Abbe frequency $1/r_A$ is not captured by the imaging system. However, this physical information is not irremediably lost as long as prior knowledge of the structure of the imaged object is available. Advances in image processing techniques—especially driven by the field of super-resolved fluorescence microscopy—have demonstrated that in this case the higher-frequency components of the imaged object's spectrum can be extrapolated from the imaged distribution. This principle is what enables information retrieval with spatial resolution beyond the diffraction limit. The prior knowledge is necessary to solve the deconvolution problem, which ideally reconstructs the original object's distribution (and its entire spectrum) by eliminating diffraction-induced blurring and noise effects. In this article, we presented state-of-the-art methods that solve the deconvolution problem for fluorescence images of neutral atoms in optical lattices making optimal use of the prior information on the physical system (sub-pixel-reconstructed line spread function, accurate noise model, discreteness of the optical lattice). The image processing methods we developed are applicable to any experimental apparatus for fluorescence imaging of trapped atoms and ions, and can be directly generalized to two-dimensions.

Our methods are particularly beneficial to improve the localization reliability in experiments with constraints on the number of scattered photons, or on the NA of the objective lens. For example, experiments imaging light fermions like lithium and potassium atoms suffer from low fluorescence scattering rates, which are generally more than one order of magnitude smaller than those achieved with heavier atoms like Cs and Rb. Recent experiments with light fermions have shown that even for large NAs NA > 0.8 the number of photoelectrons recorded per atom is around 1000 for an exposure time of 1 s [14–16]. Similar yields of photoelectrons are obtained in our apparatus where we employ an objective lens with much smaller NA (NA = 0.23). By taking advantage of our deconvolution methods, these experiments can minimize the exposure time while ensuring a high reliability to localize each individual atom in the lattice. Short exposure times reduce the probability of atoms to hop to neighboring sites, thus avoiding localization errors as well as losses of atoms colliding inelastically with a second neighboring atom.

In our laboratory, the construction of a new experimental apparatus for imaging single atoms with much higher NA (NA > 0.9) is underway, which ensures a twenty times higher collection efficiency and a four times narrower PSF. The analysis methods demonstrated in this article, when applied to the new imaging apparatus,

should enable single-site resolution with unprecedentedly short exposure times (<10 ms)[5], allowing us to directly discriminate between the two internal hyperfine states of Cs atoms (qubit states) [48, 63, 73]. Moreover, it is the goal of future work to investigate whether compressed sensing techniques, which rely on a completely different principle than our parametric deconvolution method, provide advantages for information retrieval beyond the diffraction limit [74, 75].

Acknowledgments

Appendix A. Single-photon cameras

With current technology, the sensitivity limit of conventional CCD sensors is determined by read-out noise, which is produced during the processes of charge-to-voltage and analog-to-digital conversion [76]. Low-noise CCD cameras cooled to low temperatures ($<-50\ ^{\circ}$C) and operating at high read-out rates (>10 Mpixel s^{-1})) have rms values of read-out noise equivalent to 6 ph. e$^-$. When operating at low read-out rates (<20 kHz), commercial state-of-the-art CCD sensors can attain lower read-out noise around 2 e$^-$, however, still above the value of one electron per pixel—the ultimate limit for single photon imaging. Recently, research prototypes showed that sub-electron read-out noise could be realized in the near future [77].

For weak radiation sources (e.g. fluorescing single atoms) the amount of shot noise can amount to very few electrons. One would ideally need CCD sensors with sub-electron sensitivity to avoid that the read-out noise dominates over shot noise. To overcome the technical limit imposed by read-out noise, three major technologies have been developed over the years and found widespread application: EMCCD cameras and ICCD cameras, and more recently CMOS sensors. All these technologies rely on the preamplification of the physical signal—the number of photoelectrons—prior to the read-out amplification stage. These three technologies are shortly reviewed in the following.

A.1. ICCD cameras
The basic idea underlying ICCD cameras dates back to the first half of the 20th century: it consists in employing a photocathode to convert the incoming photons into photoelectrons, whose number can then be multiplied by avalanche amplification. A gating electric field is used to precisely control the access of photoelectrons into a microchannel plate, where the avalanche multiplication process takes place. The electrons exiting the microchannel plate are accelerated towards a phosphor screen, upon which they recreate the same distribution of photons impinging at the photocathode. Secondary photons emitted by the phosphor screen are eventually imaged onto a low-noise CCD detector. Electron multiplication in the microchannel plate can readily reach amplification factors up to 10^4, which allow read-out noise to be effectively suppressed down to values as low as 10^{-3} e$^-$/pixel. For a more detailed account of ICCD cameras, please refer to [78] and references therein.

A.2. EMCCD cameras
The first practical demonstration of EMCCD cameras has been provided in 2001 [79–81]. In essence, EMCCD sensors are CCD sensors equipped with a low-noise electron multiplying (EM) register in addition to the conventional register used to transport electron charges. In comparison to conventional CCD registers, the EM register uses a higher clocking voltage to provide the electrons with sufficient kinetic energy to cause impact ionization. Through avalanche multiplication, the EM register is able to stochastically multiply the number of electrons by factors in the range of a few thousands. The read-out noise is thereby effectively reduced on the order of 10^{-2} e$^-$/pixel, which is smaller than noise produced by dark counts and CIC [82]. For a more detailed account, please refer to [83] and references therein. In addition, both CIC and read-out noise can be further suppressed by hardware binning the CCD pixels along the vertical direction [84]. It is understood that spatial

---

[5] Because the localization precision in equation (4) scales as $\Delta x \propto \lambda_{\mathrm{f}}/(\mathrm{NA}\sqrt{N})$ (assuming only shot noise), the twenty times larger collection solid angle and the four times narrower PSF should enable single-site resolution with exposure times around 1000 ms$/(16 \cdot 20) \approx 3$ ms.

resolution along the binned direction is reduced depending on how many pixels are vertically binned together. Note that hardware binning is not exploited in this work.

A.3. Comparison between ICCD and EMCCD cameras

In contrast to EMCCD sensors, ICCD sensors are virtually insensitive to spurious CIC and thermal dark electrons since these processes occur after the amplification process. They are, however, sensitive to dark electrons generated in the photocathode (so-called equivalent background illumination), whose rate is generally small $\ll 1$ e$^-$/pixel/s already at room temperature. Therefore, cooling of the sensor to low temperatures is not needed for short exposure times lasting only a few seconds. On the other hand, ICCD cameras have a lower quantum efficiency than EMCCD cameras (see also discussion in appendix B) because of the lower sensitivity of photocathode materials, especially at longer wavelengths. For instance, at our fluorescence wavelength $\lambda_f = 852$ nm, $QE(\lambda_f)$ does not exceed 20% at the present time. ICCDs allow much shorter gate times than EMCCDs, though this feature is not particularly relevant for imaging single atoms trapped in optical potentials. In addition, the finite radius of the microchannel plate in ICCD sensors limits the resolution to about 50 line pairs/mm; for the Abbe radius r_A of our microscope objective, for instance, a magnification factor of the order of 50 is required to avoid affecting the overall optical resolution. A detailed comparison of noise properties has been published elsewhere [85–87].
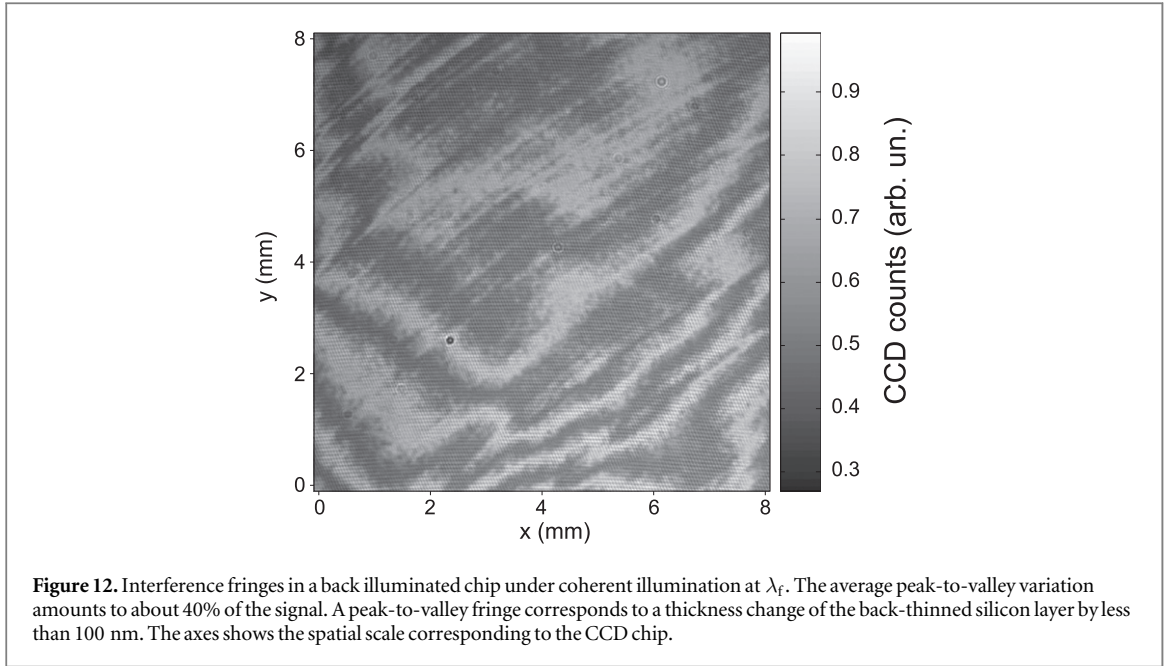
A.4. CMOS sensors

Due to significant advances over the past two decades in manufacturing microscale, ultralow-noise MOSFET devices, CMOS image sensors represent today an appealing alternative to conventional CCD detectors in low light imaging applications [88, 89]. The basic element consists here of an active pixel sensor, which provides the charge-to-voltage conversion electronics and the transistors needed for voltage buffering and pixel addressing [90]. The absence of the CCD shift register enables faster parallel read-out rates and excludes the noise contribution caused by CIC. Read-out noise in CCD sensors is dominated by Johnson noise at the charge amplifier, whose white spectrum is inevitably fed into the large video bandwidth. In contrast, parallel amplification in CMOS sensors makes it possible to directly amplify the signal at the active pixel location, where the signal is formed. In such a way, the bandwidth of the source-follower and column amplifier used to amplify the charge signal into a voltage signal can be limited through a low-pass filter. This prevents high-frequency noise components to be amplified and fed through the high-bandwidth analog-to-digital conversion circuitry [91, 92]. Commercially available CMOS cameras exhibit read-out noise as low as 1 e$^-$/pixel, which is nearly independent of the video frame rate. The presence of transistors in the pixel area significantly screen the silicon photosensitive area from the impinging photons, thus reducing the pixel's fill factor. To circumvent this problem, an array of microlenses is usually employed in scientific-grade sensors to efficiently collect photons in front of each pixel. Alternatively, back-illuminated CMOS sensors (see e.g. OmniVision Technologies, Inc.) can also be employed. A detailed comparison of CMOS cameras with CCD, ICCD, and EMCCD cameras has been carried out elsewhere [93, 94].

Appendix B. Front versus back illuminated CCD detectors

In silicon-based detectors, the quantum efficiency (QE) of CCD detectors is determined by the probability that an incoming photon is converted into an electron–hole pair inside the photosensitive region—a region that is completely depleted of mobile charge carriers—where electrons are efficiently collected into the pixel by means of a built-in electric field. For CCD detectors that are frontally illuminated, photons must first transverse the polycrystalline silicon structure of electrodes and a silicon-oxide insulating layer before they can reach the photosensitive region. Reflections and absorptions by the electrodes cause a reduction of QE, which can be avoided if the back side, i.e. the one opposite to the CCD electrodes, is turned towards the radiation source. This can be achieved by etching the chip to a thin layer around 10–20 μm thick. Back-illuminated back-thinned CCD detectors have thereby doubled the QE with peak values above 90%.

We have tested a commercial, state-of-the-art back-illuminated EMCCD sensor with read-out noise specified at <0.05 e$^-$/pixel and dark current noise at $<5 \times 10^{-3}$ e$^-$/pixel/s for a temperature below -65 °C. This sensor provides a QE of 60% at the fluorescence wavelengths of 852 nm, which is about the double of that achieved by our front-illuminated camera. As a downside, the sensor displays interference fringes caused by multiple reflections of monochromatic photons at the interface between the substrate and the silicon-oxide layer and at the interface between the substrate and the air. In fact, the silicon substrate behave like an etalon plate at longer wavelengths where the absorption depth of silicon increases. The etaloning effect is evidenced in figure 12, where variations of the substrate thickness results in visible fringes with a contrast $\sim 40\%$. We have verified that the interference pattern is stable and has only a weak dependence on temperature, as the

Figure 12. Interference fringes in a back illuminated chip under coherent illumination at λ_f. The average peak-to-valley variation amounts to about 40% of the signal. A peak-to-valley fringe corresponds to a thickness change of the back-thinned silicon layer by less than 100 nm. The axes shows the spatial scale corresponding to the CCD chip.

interference pattern shifted by about half a fringe for a temperature change of 20 °C. Despite the large intensity variations across the sensor, the stability of interference pattern allows us to filter it out by dividing the recorded signal by a calibrated spatial mask.

Larger reverse-bias voltages (~ 100 V) together with thicker substrates (~ 100 μm) of high-resistivity silicon results in much wider photosensitive regions (deep depletion) [95]. Thicker depletion regions permit to enhance the QE especially in the near-infrared, where the absorption depth of silicon is >10 μm. Because of the higher thickness, photons are converted into electron–hole pairs before reaching the silicon-oxide layer, with the result that the etaloning effect is strongly suppressed. However, the suppressed etaloning effect comes along with about a one hundred-fold increase of dark counts, which can be suppressed by cooling to low temperatures < -50 °C. Up to the present, there exists neither CMOS nor EMCCD cameras that are based on a deep-depletion back-thinned sensor, which would be ideal for detecting small signals in the near-infrared.

Appendix C. Asymptotic limit of the iterative PSF reconstruction

We study the LSF produced by the iterative reconstruction algorithm in the limit of infinitely many single-atom images that are overlapped with sub-pixel resolution according to the algorithm presented in section 4.1. Although the following discussion specifically focuses on the reconstruction of LSF in one-dimension, analogous results can be demonstrated for the PSF in two-dimensions.

The algorithm first constructs from the image of an atom positioned in x_0 a new real-valued distribution with sub-pixel resolution

$$I_{x_0}(x) = \sum_{n=-\infty}^{\infty} (L_{\text{CCD}}(n\,\Delta_\text{s} - x_0) + \epsilon[n\,\Delta_\text{s}])\;\mathcal{R}_\text{sp}(x - n\,\Delta_\text{s}), \qquad (17)$$

where L_{CCD} is the response of the imaging system defined in equation (7), ϵ is the additive noise, Δ_s is the spacing between CCD pixels, and $\mathcal{R}_\text{sp}(x)$ is the sub-pixel function (one for $-\Delta_\text{sp}/2 < x < \Delta_\text{sp}/2$ and zero elsewhere, where $\Delta_\text{sp} = \Delta_\text{s}/s$ for some integer s). In essence, equation (17) maps the signal $L_{\text{CCD}}(n\,\Delta_\text{s} - x_0)$ recorded at the nth CCD pixel to a s-fold narrower signal $\mathcal{R}_\text{sp}(x - n\,\Delta_\text{s})$, yielding a continuous distribution in the real-valued variable x. The function $L_{\text{guess}}^{(k-1)}$, which is produced by the algorithm at the iteration $k - 1$, is fitted to the recorded fluorescence distribution to provide a maximum-likelihood estimator $\tilde{x}_0$ of the atom position x_0. This estimator is a stochastic variable due to the noise term ϵ in equation (17). Because of symmetry reasons, its probability distribution is symmetrically centered on x_0 (unbiased estimator). Subsequently, the reconstruction algorithm translates with sub-pixel resolution the distribution in equation (17) by $-\tilde{x}_0$, producing a new distribution $I_{x_0}[x + \tilde{x}_0]$. The algorithm adds all repositioned single-atom intensity distributions to yield a new estimate of the LSF

$$L_{\text{guess}}^{(k)} = \int_{-\infty}^{\infty} dx_0 \; \mathcal{P}(x_0) \int_{-\infty}^{\infty} d\tilde{x}_0 \; \mathcal{P}(\tilde{x}_0 | x_0) I_{x_0}(x + \tilde{x}_0), \tag{18}$$

where $\mathcal{P}(x_0)$ is the probability of the single atom to be in x_0, and $\mathcal{P}(\tilde{x}_0 | x_0)$ is the conditional probability expressing the uncertainty distribution of the maximum-likelihood estimator $\tilde{x}_0$. In the asymptotic limit of infinitely many images, the noise contribution ϵ averages out so that the expression of equation (18) takes the form

$$L_{\text{guess}}^{(k)} = \int_{-\infty}^{\infty} dx_0 \; \mathcal{P}(x_0) \int_{-\infty}^{\infty} d\tilde{x}_0 \sum_{n=-\infty}^{\infty} \mathcal{P}(\tilde{x}_0 | x_0) L_{\text{CCD}}(n \, \Delta_{\text{s}} - x_0) \mathcal{R}_{\text{sp}}(x + \tilde{x}_0 - n \, \Delta_{\text{s}}). \tag{19}$$

This expression can be further simplified by assuming an isoplanatic response function of the imaging system such that $\mathcal{P}(\tilde{x}_0 | x_0) = \mathcal{R}_x(x_0 - \tilde{x}_0)$, where $\mathcal{R}_x$ is the uncertainty distribution of the maximum-likelihood estimator of the atom's position. In addition, we make the physical assumption that $\mathcal{P}(x_0)$ is uniformly distributed, which is ensured by the incommensurability of the optical lattice with respect to the CCD array and by small drifts in the time of the optical lattice. The latter condition is particularly important to guarantee that all sub-pixels of the reconstructed LSF are equally sampled. After some algebra, equation (19) can be recast in the form

$$L_{\text{guess}}^{(k)}(x) = (\mathcal{R}_x * \mathcal{R}_{\text{sp}} * L_{\text{CCD}})(x), \tag{20}$$

which proves the expression used in equation (8).

References

[1] Zhang C, Rolston S L and Das Sarma S 2006 Manipulation of single neutral atoms in optical lattices *Phys. Rev.* A **74** 042316

[2] Lee P J, Anderlini M, Brown B L, Sebby-Strabley J, Phillips W D and Porto J V 2007 Sublattice addressing and spin-dependent motion of atoms in a double-well lattice *Phys. Rev. Lett.* **99** 020402

[3] Bloch I 2008 Quantum coherence and entanglement with ultracold atoms in optical lattices *Nature* **453** 1016

[4] Karski M, Förster L, Choi J M, Alt W, Widera A and Meschede D 2009 Nearest-neighbor detection of atoms in a 1D optical lattice by fluorescence imaging *Phys. Rev. Lett.* **102** 053001

[5] Karski M, Förster L, Choi J M, Steffen A, Belmechri N, Alt W, Meschede D and Widera A 2010 Imprinting patterns of neutral atoms in an optical lattice using magnetic resonance techniques *New J. Phys.* **12** 065027

[6] Lee J H, Montano E, Deutsch I H and Jessen P S 2013 Robust site-resolvable quantum gates in an optical lattice via inhomogeneous control *Nat. Commun.* **4** 2027

[7] Wang Y, Zhang X, Corcovilos T A, Kumar A and Weiss D S 2015 Coherent addressing of individual neutral atoms in a 3D optical lattice *Phys. Rev. Lett.* **115** 043003

[8] Xia T, Lichtman M, Maller K, Carr A W, Piotrowicz M J, Isenhower L and Saffman M 2015 Randomized benchmarking of single-qubit gates in a 2D array of neutral-atom qubits *Phys. Rev. Lett.* **114** 100503

[9] Bloch I, Dalibard J and Nascimbène S 2012 Quantum simulations with ultracold quantum gases *Nat. Phys.* **8** 267

[10] Bakr W S, Gillen J I, Peng A, Fölling S and Greiner M 2009 A quantum gas microscope for detecting single atoms in a Hubbard-regime optical lattice *Nature* **426** 74

[11] Sherson J F, Weitenberg C, Endres M, Cheneau M, Bloch I and Kuhr S 2010 Single-atom-resolved fluorescence imaging of an atomic Mott insulator *Nature* **467** 68

[12] Karski M, Förster L, Choi J M, Steffen A, Alt W, Meschede D and Widera A 2009 Quantum walk in position space with single optically trapped atoms *Science* **325** 174

[13] Yamamoto R, Kobayashi J, Kuno T, Kato K and Takahashi Y 2016 An ytterbium quantum gas microscope with narrow-line laser cooling *New J. Phys.* **18** 023016

[14] Haller E, Hudson J, Kelly A, Cotta D A, Peaudecerf B, Bruce G D and Kuhr S 2015 Single-atom imaging of fermions in a quantum-gas microscope *Nat. Phys.* **11** 738

[15] Cheuk L W, Nichols M A, Okan M, Gersdorf T, Ramasesh V V, Bakr W S, Lompe T and Zwierlein M W 2015 Quantum-gas microscope for fermionic atoms *Phys. Rev. Lett.* **114** 193001

[16] Parsons M F, Huber F, Mazurenko A, Chiu C S, Setiawan W, Wooley-Brown K, Blatt S and Greiner M 2015 Site-resolved imaging of fermionic ^{6}Li in an optical lattice *Phys. Rev. Lett.* **114** 213002

[17] Omran A, Boll M, Hilker T A, Kleinlein K, Salomon G, Bloch I and Gross C 2015 Microscopic observation of Pauli blocking in degenerate fermionic lattice gases *Phys. Rev. Lett.* **115** 263001

[18] Genske M, Alt W, Steffen A, Werner A H, Werner R F, Meschede D and Alberti A 2013 Electric quantum walks with individual atoms *Phys. Rev. Lett.* **110** 1–5

[19] Robens C, Alt W, Meschede D, Emary C and Alberti A 2015 Ideal negative measurements in quantum walks disprove theories based on classical trajectories *Phys. Rev.* X **5** 011003

[20] Robens C, Zopes J, Alt W, Brakhane S, Meschede M and Alberti D A 2016 Low-entropy states of neutral atoms in polarization-synthesized optical lattices in preparation

[21] Huang B, Babcock H and Zhuang X 2010 Breaking the diffraction barrier: super-resolution imaging of cells *Cell* **143** 1047

[22] Thompson M A, Lew M D and Moerner W E 2012 Extending microscopic resolution with single-molecule imaging and active control *Annu. Rev. Biophys.* **41** 321

[23] Pertsinidis A, Zhang Y and Chu S 2010 Subnanometre single-molecule localization, registration and distance measurements *Nature* **466** 647

[24] Pantin E, Starck J L and Murtagh F 2007 Deconvolution and blind deconvolution in astronomy *Blind Image Deconvolution: Theory and Applications* ed P Campisi and K Egiazarian (Boca Raton, FL: CRC) p 277

[25] Greif D, Parsons M F, Mazurenko A, Chiu C S, Blatt S, Huber F, Ji G and Greiner M 2016 Site-resolved imaging of a fermionic Mott insulator *Science* **351** 953

[26] Miranda M, Inoue R, Okuyama Y, Nakamoto A and Kozuma M 2015 Site-resolved imaging of ytterbium atoms in a two-dimensional optical lattice *Phys. Rev.* A **91** 063414

[27] Gericke T, Würtz P, Reitz D, Langen T and Ott H 2008 High-resolution scanning electron microscopy of an ultracold quantum gas *Nat. Phys.* **4** 949

[28] Ober R J, Ram S and Ward E S 2004 Localization accuracy in single-molecule microscopy *Biophys. J.* **86** 1185

[29] Abraham A V, Ram S, Chao J, Ward E S and Ober R J 2009 Quantitative study of single molecule location estimation techniques *Opt. Express* **17** 23352

[30] Small A and Stahlheber S 2014 Fluorophore localization algorithms for super-resolution microscopy *Nat. Methods* **11** 267

[31] Biercuk M J, Uys H, Britton J W, VanDevender A P and Bollinger J J 2010 Ultrasensitive detection of force and displacement using trapped ions *Nat. Nanotechnol.* **5** 646

[32] Anetsberger G, Gavartin E, Arcizet O, Unterreithmeier Q P, Weig E M, Gorodetsky M L, Kotthaus J P and Kippenberg T J 2010 Measuring nanomechanical motion with an imprecision below the standard quantum limit *Phys. Rev.* A **82** 061804

[33] Gloger T F, Kaufmann P, Kaufmann D, Baig M T, Collath T, Johanning M and Wunderlich C 2015 Ion-trajectory analysis for micromotion minimization and the measurement of small forces *Phys. Rev.* A **92** 043421

[34] Wong-Campos J D, Johnson K G, Neyenhuis B, Mizrahi J and Monroe C 2015 High resolution adaptive imaging of a single atom arXiv:1512.06172

[35] Goodman J W 1996 *Introduction to Fourier Optics* (New York: McGraw-Hill)

[36] Wittenstein W, Fontanella J C, Newbery A R and Baars J 1982 The definition of the OTF and the measurement of aliasing for sampled imaging systems *Opt. Acta* **29** 41

[37] Park S K, Schowengerdt R A and Kaczynski M A 1984 Modulation-transfer-function analysis for sampled image systems *Appl. Opt.* **23** 2572

[38] Boreman G D 2001 *Modulation Transfer Function in Optical and Electro-Optical Systems* (Bellingham, WA: SPIE)

[39] Gaskill J D 1978 *Linear Systems, Fourier Transforms, and Optics* (Wiley: New York)

[40] Pawley J B 2006 Points, pixels, and gray levels: digitizing image data handbook of biological *Handbook of Biological Confocal Microscopy* 3rd edn ed J B Pawley (New York: Springer) p 59

[41] Starck J L, Pantin E and Murtagh F 2002 Deconvolution in astronomy: a review *Publ. Astron. Soc. Pac.* **114** 1051

[42] Alt W 2002 An objective lens for efficient fluorescence detection of single atoms *Optik* **113** 142

[43] Born M and Wolf E 2003 *Principle of Optics* 7th edn (Cambridge: Cambridge University Press)

[44] Bobroff N 1986 Position measurement with a resolution and noise-limited instrument *Rev. Sci. Instrum.* **57** 1152

[45] Thompson R E, Larson D R and Webb W W 2002 Precise nanometer localization analysis for individual fluorescent probes *Biophys. J.* **82** 2775

[46] Quan T, Zeng S and Huang Z L 2010 Localization capability and limitation of electron-multiplying charge-coupled, scientific complementary metal-oxide semiconductor, and charge-coupled devices for superresolution imaging *J. Biomed. Opt.* **15** 066005

[47] Miroshnychenko Y, Alt W, Dotsenko I, Förster L, Khudaverdyan M, Rauschenbeutel A and Meschede D 2006 Precision preparation of strings of trapped neutral atoms *New J. Phys.* **8** 191

[48] Fuhrmanek A, Bourgain R, Sortais Y R P and Browaeys A 2011 Free-space lossless state detection of a single trapped atom *Phys. Rev. Lett.* **106** 133003

[49] Hume D B, Stroescu I, Joos M, Muessel W, Strobel H and Oberthaler M K 2013 Accurate atom counting in mesoscopic ensembles *Phys. Rev. Lett.* **111** 253001

[50] Sperlich K and Stolz H 2013 Quantum efficiency measurements of (EM)CCD cameras: high spectral resolution and temperature dependence *Meas. Sci. Technol.* **25** 015502

[51] Anderson J and King I R 2000 Toward high- precision astrometry with WFPC2: I. Deriving an Accurate Point-Spread Function *Publ. Astron. Soc. Pac.* **112** 1360

[52] Balestra F and Ghibaudo G (ed) 2001 *Device and Circuit Cryogenic Operation for Low Temperature Electronics* (Boston, MA: Springer)

[53] Saks N S 1980 A technique for suppressing dark current generated by interface states in buried channel Ccd imagers *Electron Device Lett.* **1** 131

[54] Janesick J R 2001 *Scientific Charge-Coupled Devices* (Bellingham, USA: SPIE) ch 7

[55] Daigle O, Quirion P O and Lessard S 2010 The darkest EMCCD ever *Proc. SPIE* ed A D Holland and D A Dorn vol 7742 p 774203

[56] Djazovski O, Daigle O, Laurin D, Bedirian M, Ducharme M E, Artigau É and Doyon R 2013 Electronmultiplying CCDs for future space instruments *Proc. SPIE: Photonics North 2013* ed P Cheben and E A Schmid p 89150Q

[57] e2v Technologies 2014 *Dark Signal and Clock-Induced Charge* (Chelmsford, UK: e2v Technologies)

[58] Hirsch M, Wareham R J, Martin-Fernandez M L, Hobson M P and Rolfe D J 2013 A stochastic model for electron multiplication charge-coupled devices-from theory to practice *Plos One* **8** e53671

[59] Robbins M S and Hadwen B J 2003 The noise performance of electron multiplying charge-coupled devices *IEEE Trans. Electron Devices* **50** 1227

[60] Lantz E, Blanchet J L, Furfaro L and Devaux F 2008 Multi-imaging and Bayesian estimation for photon counting with EMCCDs *Mon. Not. R. Astron. Soc.* **386** 2262

[61] Basden A G, Haniff C A and Mackay C D 2003 Photon counting strategies with low-light-level CCDs *Mon. Not. R. Astron. Soc.* **345** 985

[62] Mullikin J C, van Vliet L J, Netten H, Boddeke F R, van der Feltz G and Young I T 1994 Methods for CCD camera characterization *Proc. SPIE: Image Acquisition and Scientific Imaging Systems* ed H C Titus and A Schmid vol 2173 p 73

[63] Martínez-Dorantes M 2016 Fast non-destructive internal state detection of neutral atoms in optical potentials, PhD Thesis, University of Bonn

[64] DePue M T, McCormick C, Winoto S L, Oliver S and Weiss D S 1999 Unity occupation of sites in a 3D optical lattice *Phys. Rev. Lett.* **82** 2262

[65] Schlosser N, Reymond G, Protsenko I and Grangier P 2001 Sub-poissonian loading of single atoms in a microscopic dipole trap *Nature* **411** 1024

[66] Rajagopalan A and Chellappa R 2014 *Motion Deblurring: Algorithms and Systems* (New York: Cambridge University Press)

[67] Stoica P and Moses R 2005 *Spectral Analysis of Signals* (Englewood Cliffs, NJ: Prentice-Hall)

[68] Xu X L and Buckley K M 1992 Bias analysis of the MUSIC location estimator *IEEE Trans. Signal Process.* **40** 2559

[69] Wax M, Shan T J and Kailath T 1984 Spatio-temporal spectral analysis by eigenstructure methods *IEEE Trans. Acoust. Speech Signal Process.* **32** 817

[70] Smith C S, Joseph N, Rieger B and Lidke K A 2010 Fast, single-molecule localization that achieves theoretically minimum uncertainty *Nat. Methods* **7** 373

[71] Chao J, Ram S, Ward E S and Ober R J 2013 Ultrahigh accuracy imaging modality for super-localization microscopy *Nat. Methods* **10** 335

[72] Ashida Y and Ueda M 2016 Precise multi-emitter localization method for fast super-resolution imaging *Opt. Lett.* **41** 72

[73] Gibbons M J, Hamley C D, Shih C Y and chman M S 2011 Nondestructive fluorescent state detection of single neutral atom qubits *Phys. Rev. Lett.* **106** 133002

[74] Szameit A *et al* 2012 Sparsity-based single-shot subwavelength coherent diffractive imaging *Nat. Mater.* **11** 455

[75] Tang G, Bhaskar B N, Shah P and Recht B 2013 Compressed sensing off the grid *IEEE Trans. Inf. Theory* **59** 7465

[76] Seitz P 2011 Fundamentals of noise in optoelectronics (*Springer Series in Optical Sciences* vol 160) ed P Seitz and A J Theuwissen (Berlin: Springer) p 1

[77] Fernández Moroni G, Estrada J, Cancelo G, Holland S E, Paolini E E and Diehl H T 2012 Sub-electron readout noise in a Skipper CCD fabricated on high resistivity silicon *Exp. Astron.* **34** 43

[78] Nützel G 2011 Single-photon imaging using electron multiplication in vacuum (*Springer Series in Optical Sciences* vol 160) ed P Seitz and A J Theuwissen (Berlin: Springer) p 73

[79] Jerram P, Pool P J, Bell R, Burt D J, Bowring S, Spencer S, Hazelwood M, Moody I, Catlett N and Heyes P S 2001 The LLCCD: low-light imaging without the need for an intensifier *Proc. SPIE: Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications II* ed M M Blouke *et al* vol 4306 p 178

[80] Hynecek J 2001 Impactron—a new solid state image intensifier *IEEE Trans. Electron Devices* **48** 2238

[81] Mackay C D, Tubbs R N, Bell R, Burt D J, Jerram P and Moody I 2001 Sub-electron read noise at MHz pixel rates *Proc. SPIE: Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications* vol 4306 p 289

[82] Ives D 2009 Electron multiplication CCDs for astronomical applications *Nucl. Instrum. Methods Phys. Res.* A **604** 38

[83] Robbins M 2011 Electron-multiplying charge coupled devices *Springer Series in Optical Sciences* ed P Seitz and A J Theuwissen vol 160 (Berlin: Springer) p 103

[84] Westra A H, Heemskerk J W T, Korevaar M A N, Theuwissen A J P, Kreuger R, Ligtvoet K M and Beekman F J 2009 On-chip pixel binning in photon-counting EMCCD-based gamma camera: a powerful tool for noise reduction *IEEE Trans. Nucl. Sci.* **56** 2559

[85] Denvir D J and Conroy E 2003 Electron-multiplying CCD: the new ICCD *Proc. SPIE: Low-Light-Level and Real-Time Imaging Systems, Components, and Applications* ed C B Johnson *et al* vol 4796 p 164

[86] Coates C G, Denvir D J, McHale N G, Thornbury K D and Hollywood M A 2003 Ultrasensitivity, speed, and resolution: optimizing low-light microscopy with the back-illuminated electron-multiplying CCD *Proc. SPIE: Confocal, Multiphoton, and Nonlinear Microscopic Imaging* ed T Wilson vol 5139 P 56

[87] Dussault D and Hoess P 2004 Noise performance comparison of ICCD with CCD and EMCCD cameras *Proc. SPIE: Infrared Systems and Photoelectronic Technology* ed E L Dereniak *et al* vol 5563 p 195

[88] Theuwissen A 2008 CMOS image sensors: state-of-the-art *Solid State Electron.* **52** 1401

[89] Zimmermann H 2009 *Integrated Silicon Optoelectronics* (*Springer Series in Optical Sciences* vol 148) (Berlin: Springer)

[90] Magnan P 2003 Detection of visible photons in CCD and CMOS: a comparative view *Nucl. Instrum. Methods Phys. Res.* **504** 199

[91] Kozlowski L, Luo J and Tomasini A 1999 Performance limits in visible and infrared imager sensors *IEDM '99 Technical Digest* p 867

[92] Janesick J 2002 Dueling detectors *SPIE Newsroom* **2** 30

[93] Hain R, Kähler C J and Tropea C 2007 Comparison of CCD, CMOS and intensified cameras *Exp. Fluids* **42** 403

[94] Krishnaswami V, Van Noorden C J F, Manders E M M and Hoebe R A 2014 Towards digital photon counting cameras for single-molecule optical nanoscopy *Opt. Nanoscopy* **3** 1

[95] Holland S E *et al* 2006 High-voltage-compatable fully depleted CCDs *High Energy* **6276** 10