# Measuring Image Resolution in Optical Nanoscopy
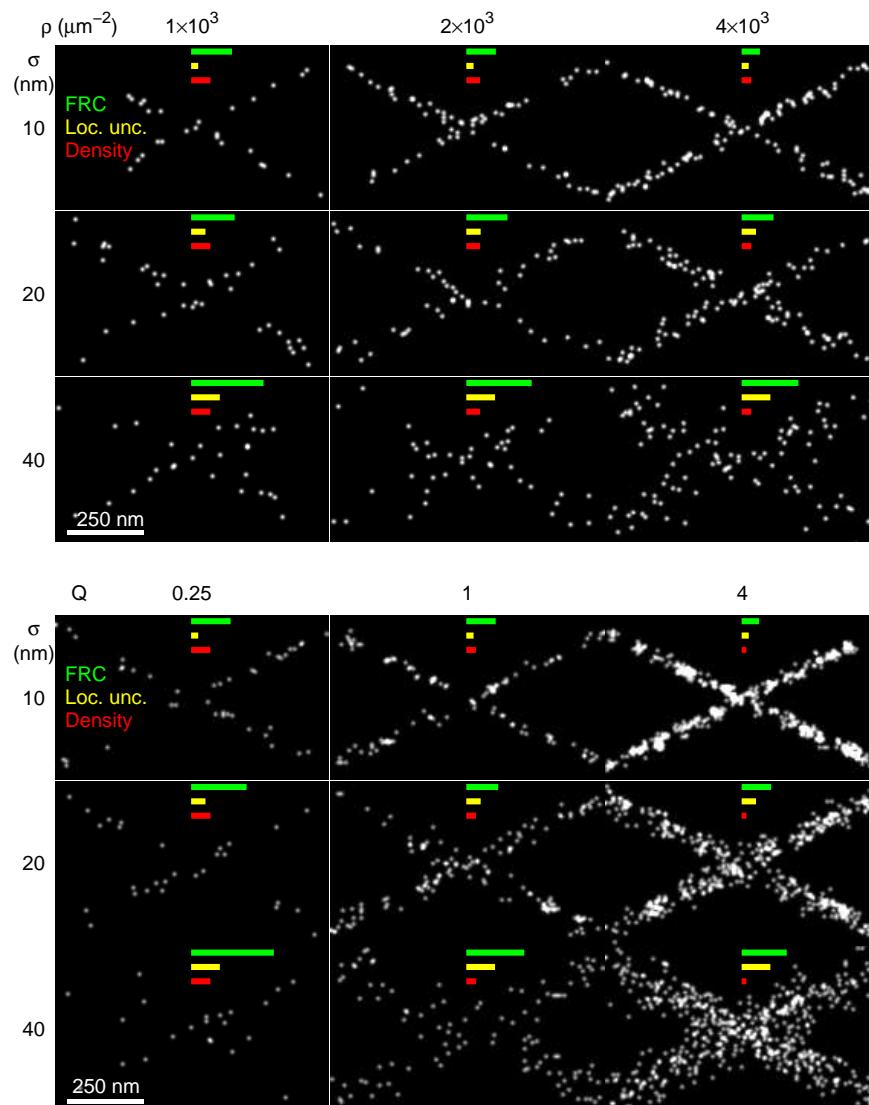
Robert P.J. Nieuwenhuizen, Keith A. Lidke, Mark Bates, Daniela Leyton Puig,
David Grünwald, Sjoerd Stallinga and Bernd Rieger

Supplementary Fig. 1
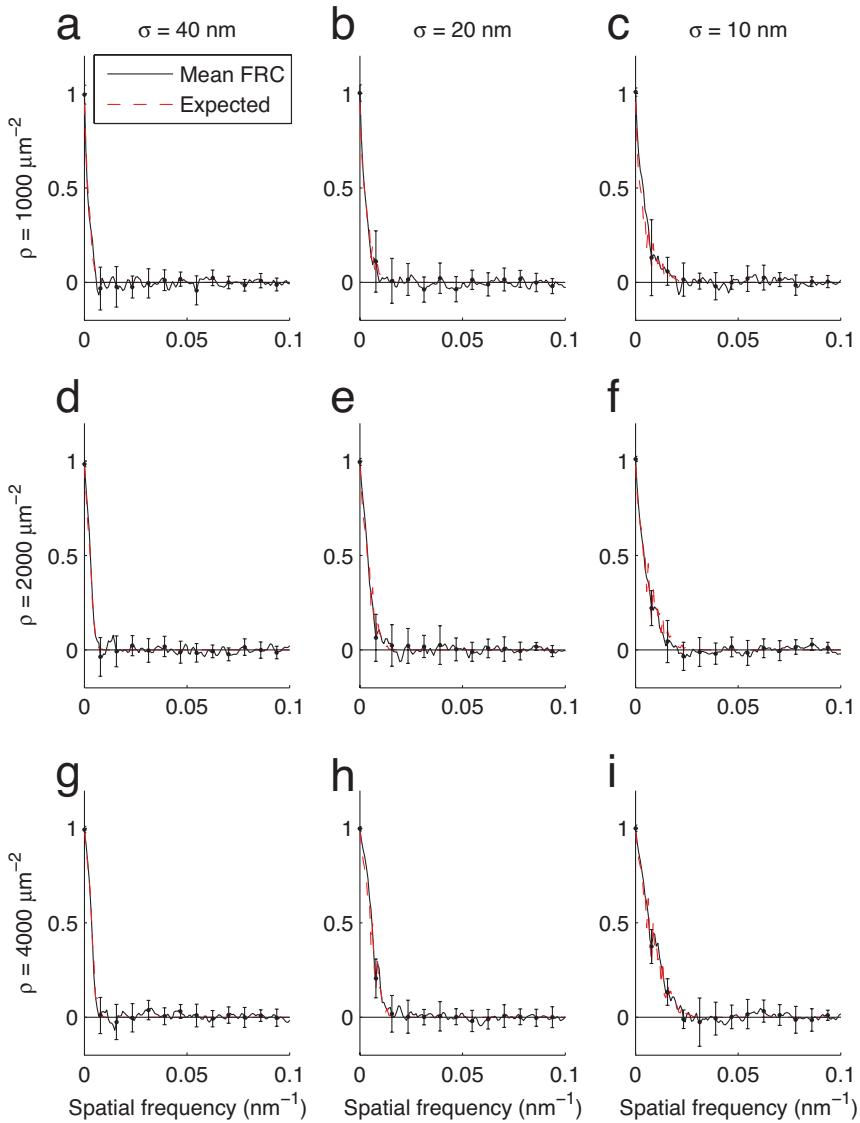
Supplementary Figure 1: **Evaluation of resolution threshold criteria.** Simulations without multiple localizations per emitter (i.e. $Q \ll 1$) (**b**) for a Siemens star shaped object (**a**), for a density of localizations $\rho = 2 \times 10^3 \ \mu m^{-2}$ and a localization uncertainty $\sigma = 20$ nm. A circle denotes where the arms of the star can just be distinguished according to the resolution computed with the most common threshold criteria (**Supplementary Note 5**): $1/7$ threshold (83 ± 3 nm, green), half-bit threshold (100 ± 5 nm, magenta), $1/2$ threshold (130 ± 7 nm, yellow), and 3σ threshold (186 ± 9 nm, cyan). White scalebar: 100 nm. For the correct threshold, the arms of the star should be distinguishable outside the corresponding circle and not distinguishable inside the circle. (**c-f**) show the regions of **b** within the different circles of the top right image. They are shown to illustrate the separability of the star arms in the absence of the visual aid from the regions outside the circles. Clearly, the $1/7$ threshold is most appropriate in these images.

2

Supplementary Fig. 2
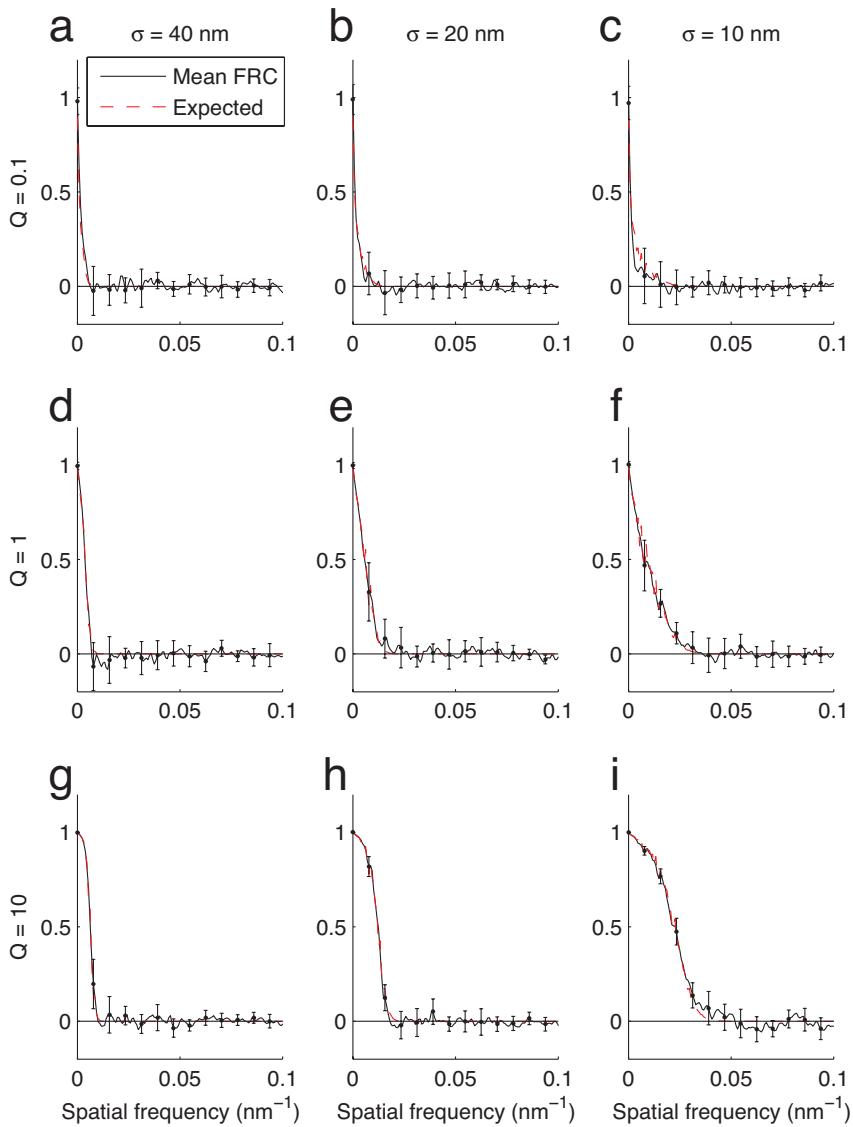
Supplementary Figure 2: **Qualitative validity of FRC resolution for crossing lines.** Simulations with a single localization per emitter of two crossing lines for different densities of localized labels ρ and localization uncertainties σ (top) and with multiple localizations per emitter for different numbers of localizations per emitter Q and localization uncertainties σ (bottom). Columns show constant ρ or Q and rows constant σ. The resolution according to the Nyquist density is depicted as a red bar, the full width at half maximum (FWHM) of the localization error distribution as a yellow bar, and the FRC resolution as a green bar. As expected, the FRC resolution indicates the critical distance at which two crossing lines can just be separated. Scale bar: 250 nm.

Supplementary Fig. 3
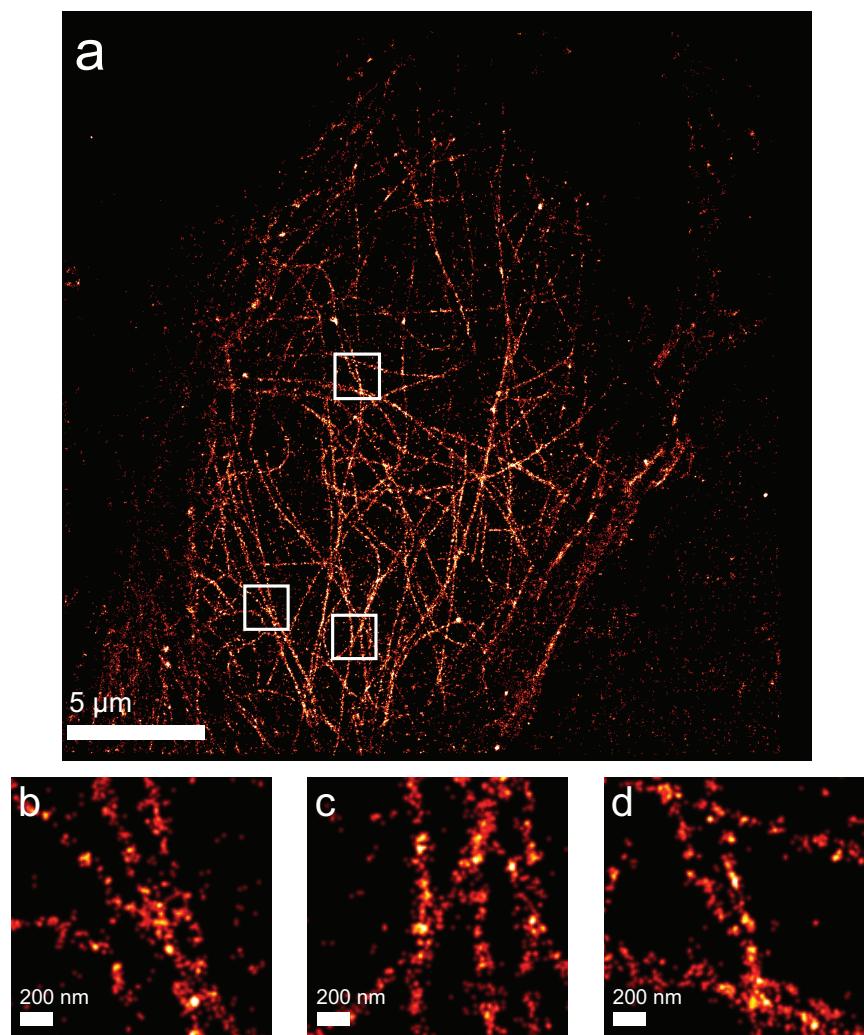
Supplementary Figure 3: **Validation of Eq. (2) for the expected FRC curve** $Q = 0$**.** The graphs show FRC curves for simulations with a single localization per emitter (i.e. with $Q \ll 1$) with the crossing lines from **Supplementary Fig. 16a**. The figure shows the theoretically expected FRC from Eq. (2) (red) and the mean FRC in 10 simulations for each of the 3 values of the localization precision σ and the density of localized emitters on the lines ρ. The FRC curves fall between the errorbars (s.d., $n = 10$) in all graphs, thus confirming the validity of Eq. (2).

Supplementary Fig. 4
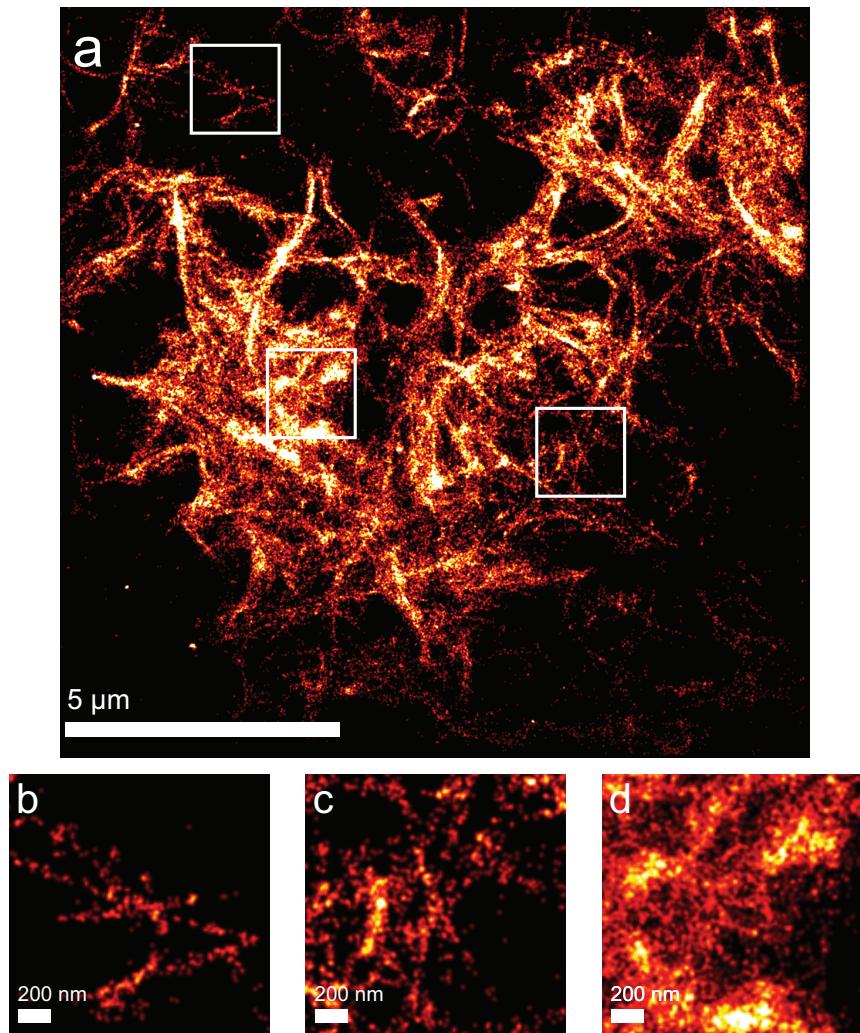
Supplementary Figure 4: **Validation of Eq. (2) for the expected FRC curve, $Q \gg 1$.** The graphs show FRC curves for simulations with multiple localizations per emitter with the crossing lines from **Supplementary Fig. 16a**. The figure shows the theoretically expected FRC from Eq. (S.26) (red) and the mean FRC in 10 simulations (black) for each of the 3 values of the localization precision σ and the number of times each emitter is localized Q. The density of localized emitters on the lines $\rho = 4 \times 10^4 \ \mu m^{-2}$. The FRC curves fall between the errorbars (s.d., $n = 10$) in all graphs, thus confirming the validity of Eq. (2).

Supplementary Fig. 5
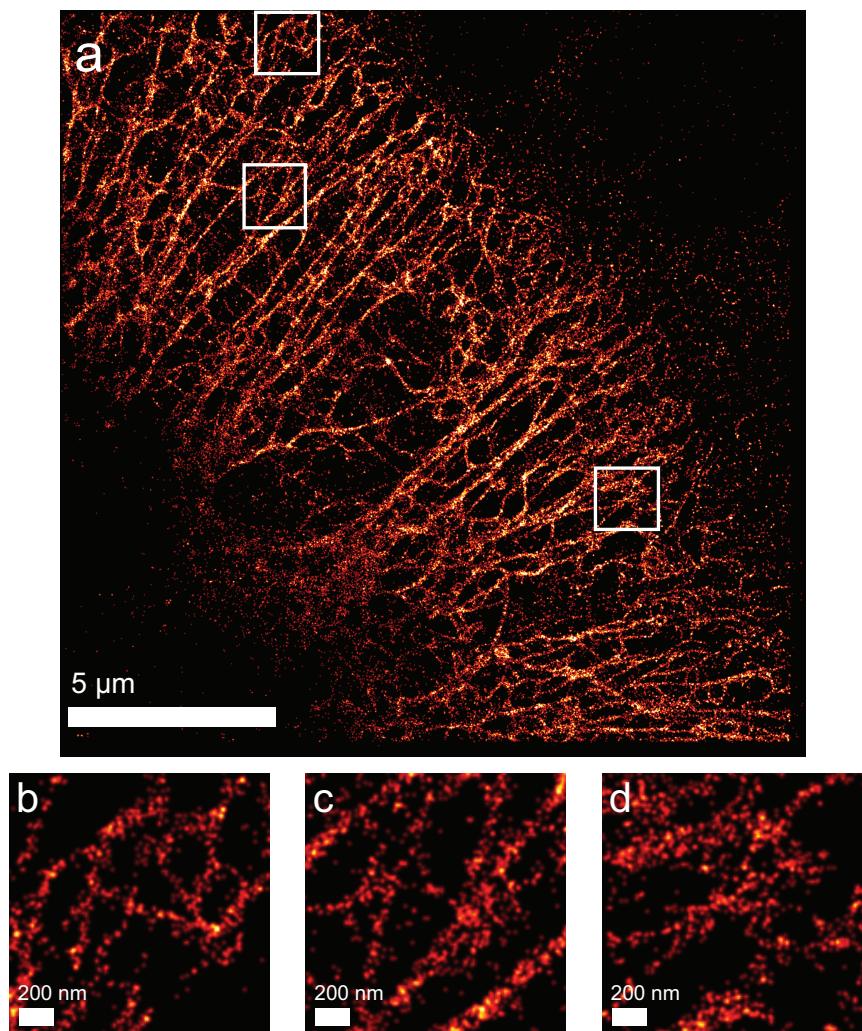
Supplementary Figure 5: **Localization microscopy image of tubulin in a HeLa cell labeled with Alexa Fluor 647.** The image is obtained from 19,000 time frames with localization uncertainty (after grouping nearby localizations in subsequent image frames) $\sigma = 7.4$ nm and a density of localizations $\rho = 1.3 \times 10^2 / \mu\text{m}^2$ (**a**). Higher magnification insets (**b-d**) show several regions with crossing lines. The calculated resolution for this image is 65 ± 2 nm, which seems reasonable for this image.

Supplementary Fig. 6
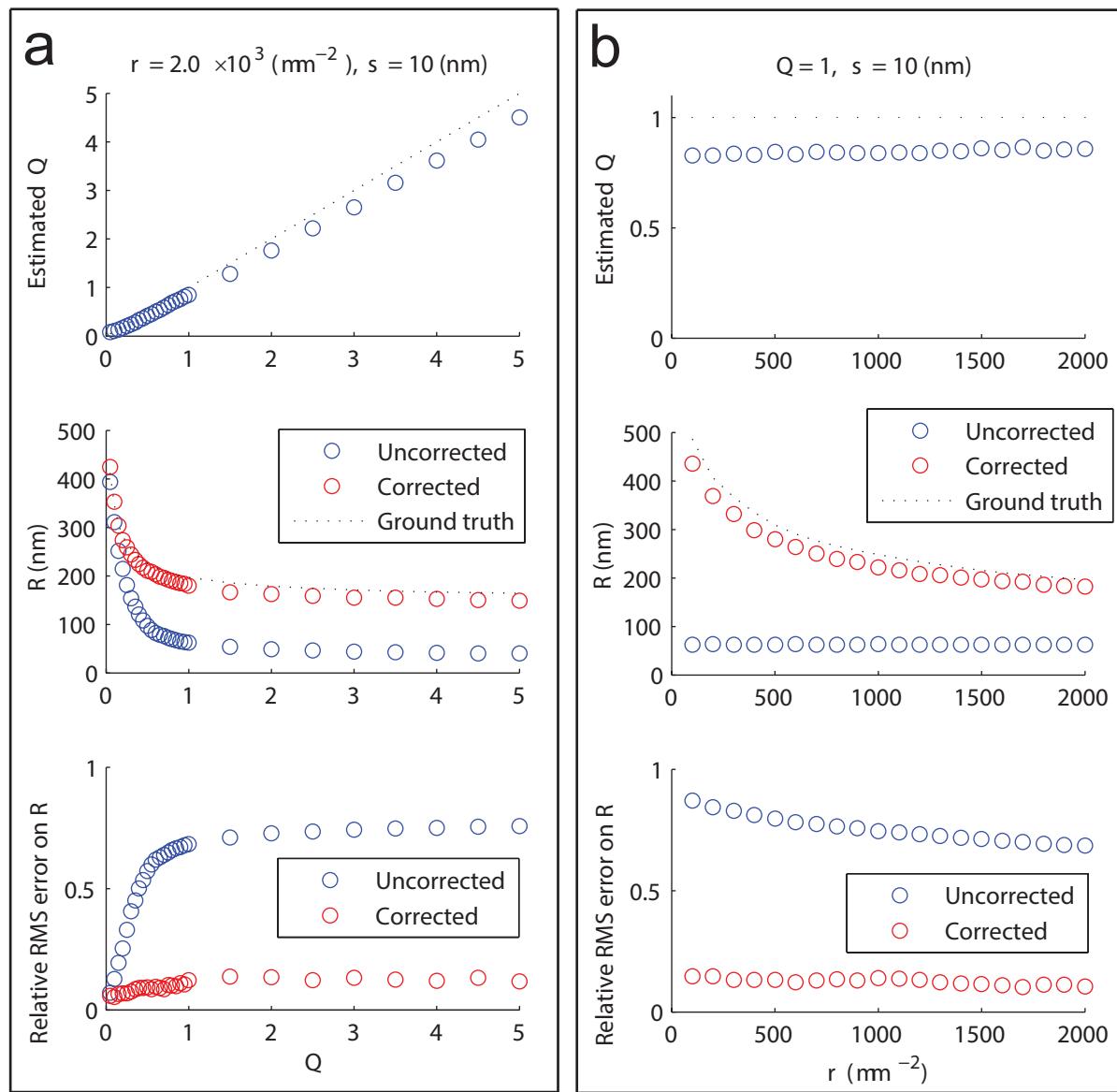
Supplementary Figure 6: **Localization microscopy image of actin filaments in a RB cell labeled with Alexa Fluor 647.** The image is obtained from 100,000 time frames with localization uncertainty (after grouping nearby localizations in subsequent image frames) $\sigma = 10.4$ nm and a density of localizations $\rho = 1.1 \times 10^3$ μm^{-2} (**a**). Higher magnification insets (**b-d**) show several regions with crossing lines. the colormaps were scaled differently in these images to guarantee sufficient contrast. The calculated resolution for this image is $R = 88 \pm 2$ nm. This seems acceptable for the sparser regions of the image (e.g. insets **b-c**) but too optimistic for the parts with more dense labeling, such as in **d**. This may be caused by false positive localizations due to simultaneous emissions of nearby emitters in those densely labeled regions, or because sparse three dimensional structures appear as dense unstructured regions when imaged in 2D.

Supplementary Fig. 7
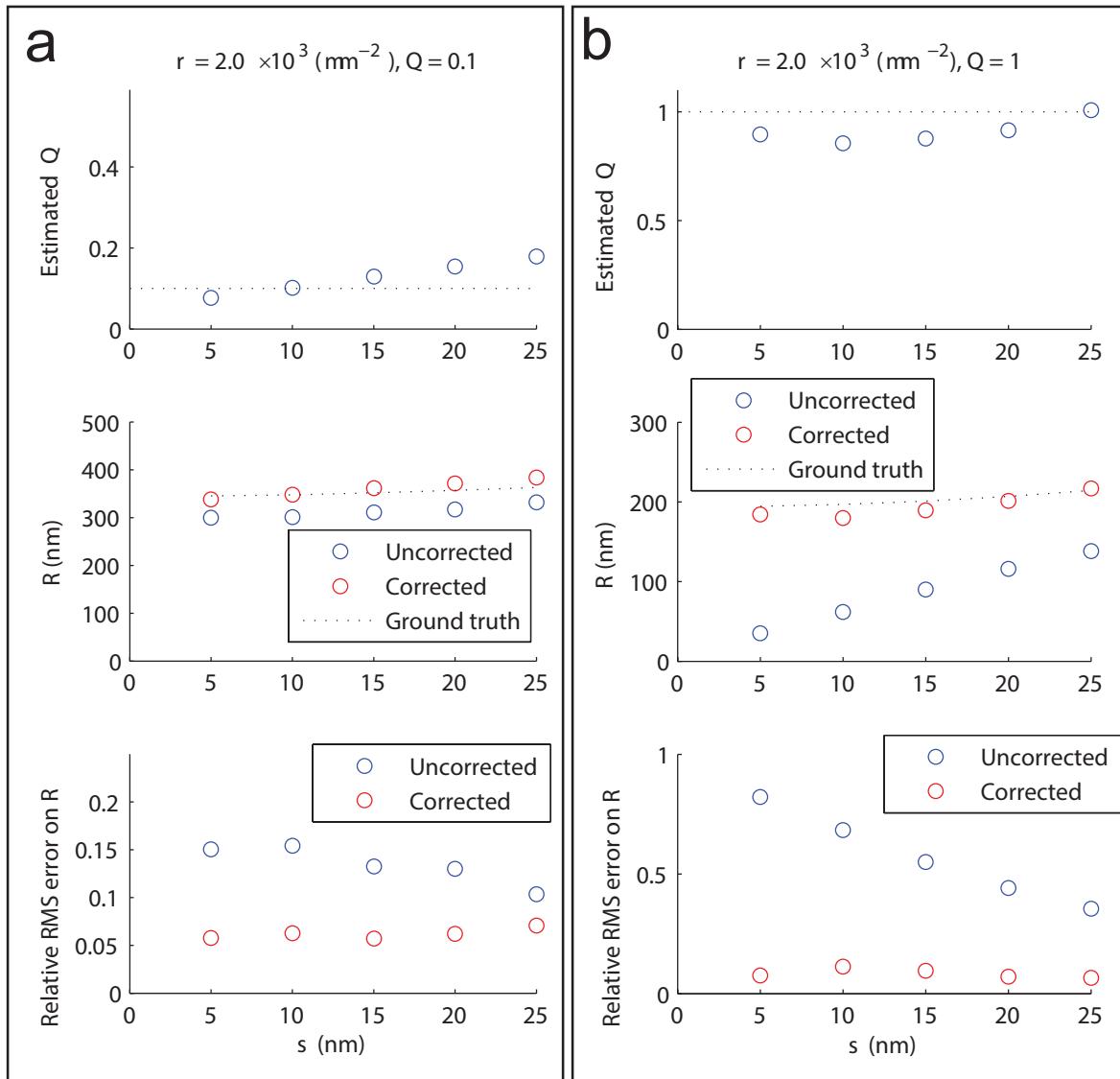
Supplementary Figure 7: **Localization microscopy image of a PA-JEB cell in TIRF microscopy showing keratine labeled with Alexa Fluor 532.** The image is obtained from 8,871 time frames with localization uncertainty (after grouping nearby localizations in subsequent image frames) $\sigma = 9.4$ nm and a density of localizations $\rho = 3.4 \times 10^2 \ \mu\text{m}^{-2}$ (**a**). Higher magnification insets (**b-d**) show several regions with crossing lines. The calculated resolution for this image is $R = 102 \pm 3$ nm, which appears to be reasonable for this image.

Supplementary Fig. 8
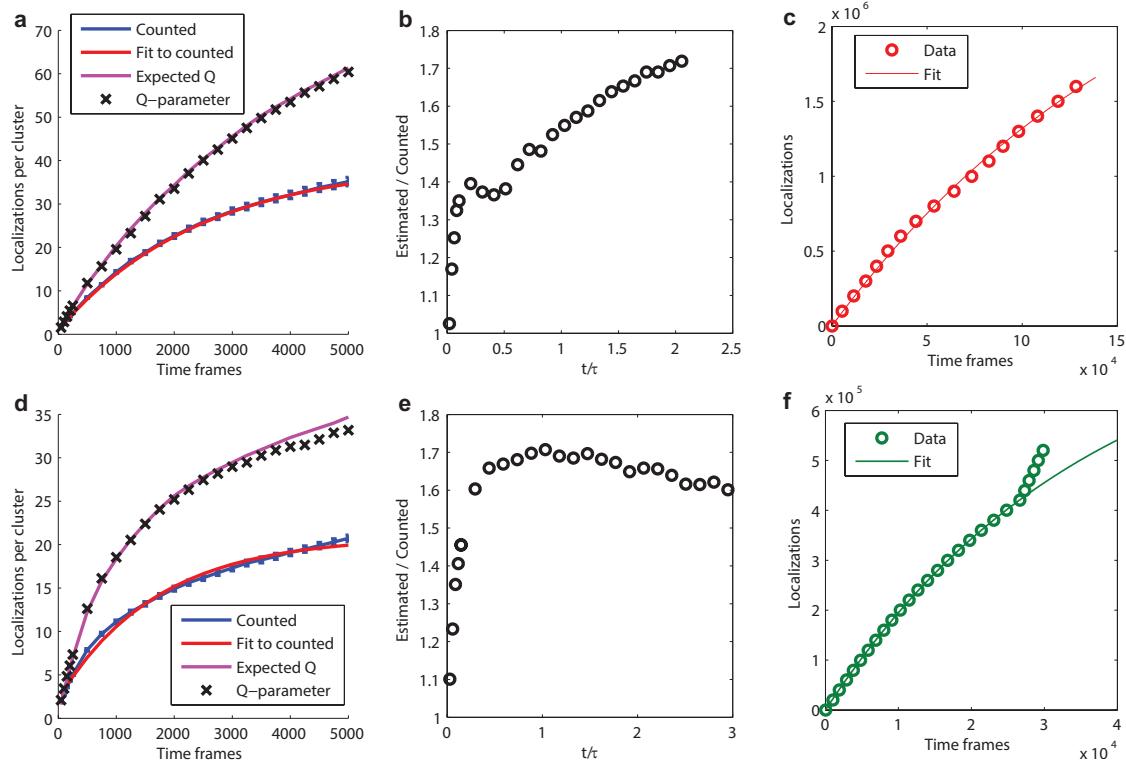
Supplementary Figure 8: **Simulation of the effect of Q and labeling density on spurious correlation correction.** The top graphs show the estimate for Q, the middle graphs show the theoretical resolution R according to equation S.28 as well as the estimates with and without correction for spurious correlations. The bottom graphs show the mean squared error in these estimates relative to the theoretical resolution. In these simulations, (**a**) $\rho = 2.0 \times 10^3 \ \mu m^{-2}$, $\sigma = 10$ nm and Q is varied, or (**b**) $Q = 1$, $\sigma = 10$ nm and ρ is varied. The error in resolution estimation levels to a plateau for increasing $Q > 1$.

Supplementary Fig. 9
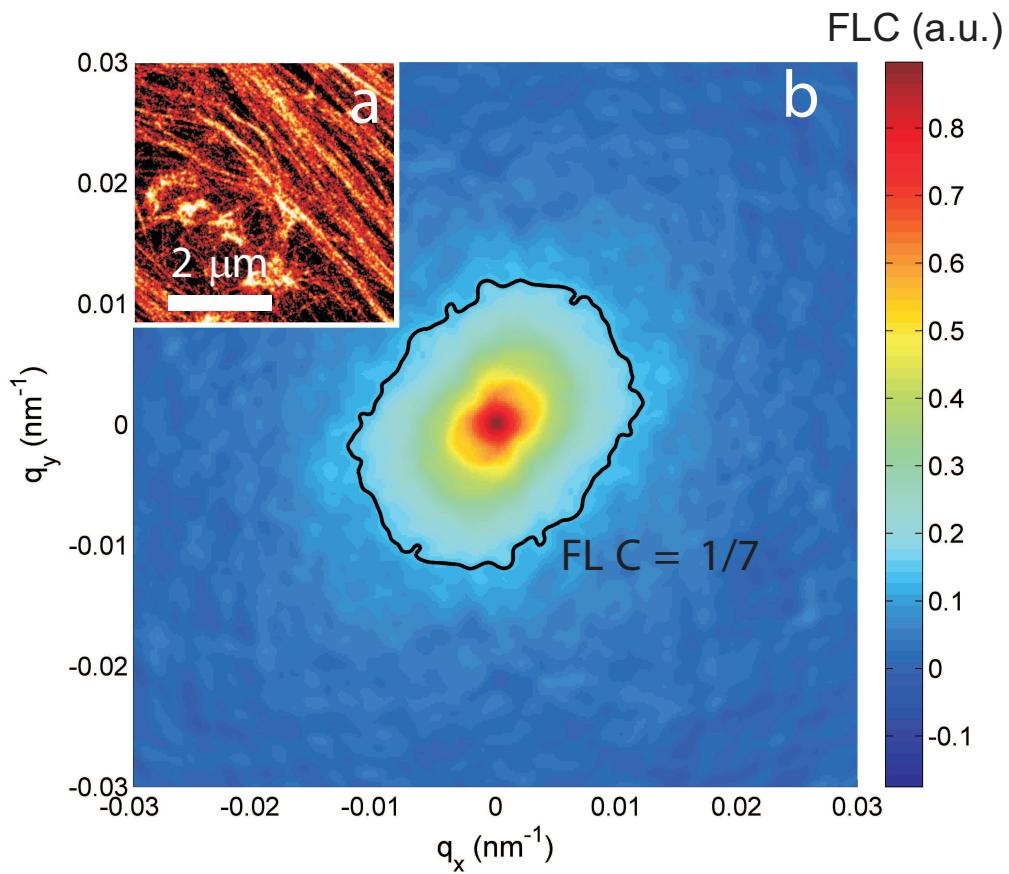
Supplementary Figure 9: **Simulation of the effect of localization uncertainty on spurious correlation correction.** The top graphs show the estimate for Q, the middle graphs show the theoretical resolution R according to Eq. (S.28) as well as the estimates with and without correction for spurious correlations. The bottom graphs show the mean squared error in these estimates relative to the theoretical resolution. In these simulations, σ is varied, $\rho = 2.0 \times 10^3 \ \mu m^{-2}$ and (**a**) $Q = 0.1$ or (**b**) $Q = 1$. The spurious correlation parameter is underestimated for small σ, and overestimated for large σ, the latter at least for $Q < 1$. This is caused by the noise floor on the numerator of the FRC which for large σ is reached before Q can become dominant in the numerator.

Supplementary Fig. 10
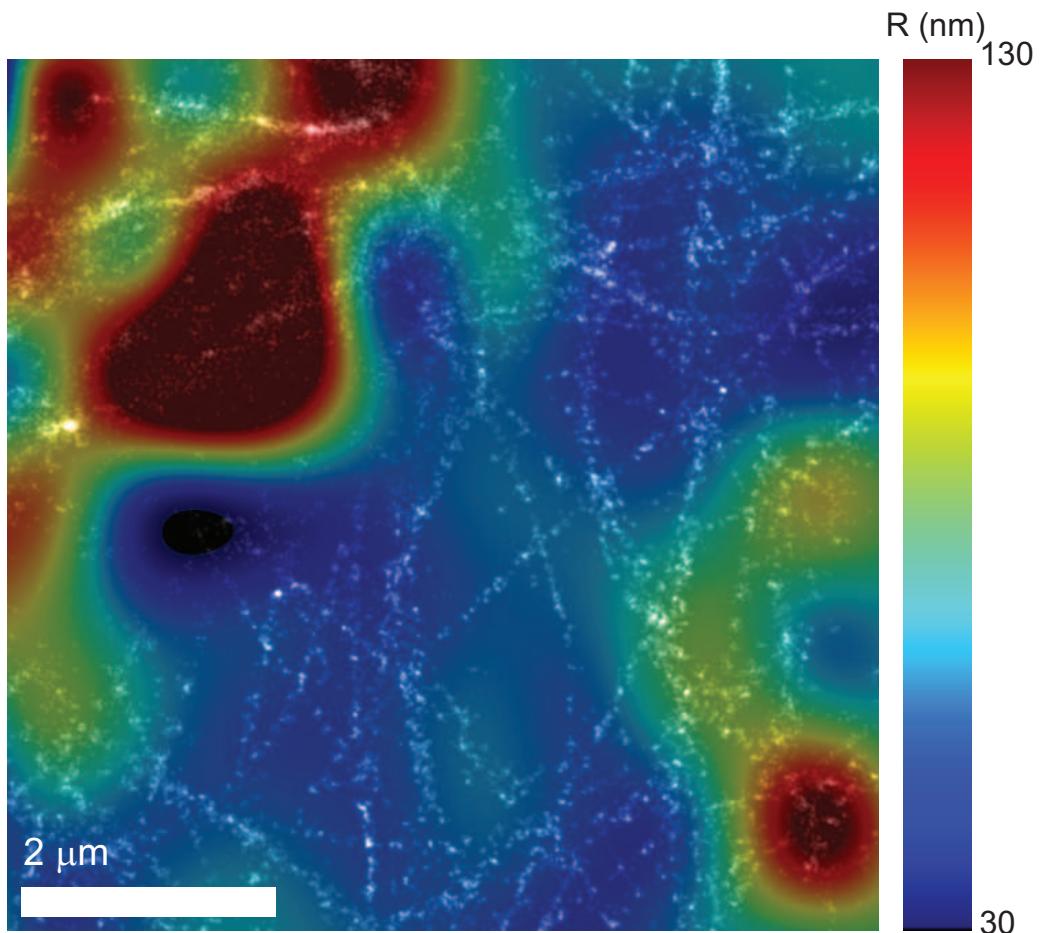
Supplementary Figure 10: **Control experiments for estimating the number of localizations per emitter.** Localization microscopy acquisitions were carried out on a glass surfaces of labeled antibodies that were distributed sparsely enough to enable counting of localizations from different antibodies. Results for Alexa Fluor 647 are presented in **a-c** and for Alexa Fluor 750 in **d-f**. (**a**,**d**) The counted number of localizations and the number estimated with the Q-parameter in clusters with at least one localization, as well as the expected Q-parameter based on Eq. S.20 and a fit of the counted localizations per molecule with a model of the form $A(1 - \exp(-t/\tau))$. The bleaching time constant τ in these control experiments was found to be 3.4×10^3 frames for Alexa Fluor 647 and 1.7×10^3 frames for Alexa Fluor 750. (**b**, **e**) The ratio of the number of localizations per cluster that is estimated with the Q-parameter and the number of counted localizations per cluster in **a** and **d**, respectively, showing a value close to one for small t (Poisson statistics regime) increasing to a value between one and two for large t due to photobleaching, in agreement with expectations (Supplementary Note 1). (**c**,**f**) The cumulative number of total localizations for the Alexa Fluor 647 and Alexa Fluor 750 data of Fig. 3, respectively, as well as a fit with a model of the form $A(1 - \exp(-t/\tau))$. The bleaching time constant τ was estimated to be 1.8×10^5 frames for Alexa Fluor 647 and 3.6×10^4 frames for Alexa Fluor 750. This leads to values $t/\tau = 0.78$ (Alexa Fluor 647) and $t/\tau = 0.83$ (Alexa Fluor 750) at the end of the data acquisition, the ratio of Q and the mean number of localizations per emitter from the control experiment follows as 1.5 and 1.7 respectively. The discontinuity in the number of Alexa Fluor 750 localizations at the end of the acquisition is due to an increase in the activation laser power. Since this part of the acquisition was relatively brief, it was ignored when τ was estimated.

11
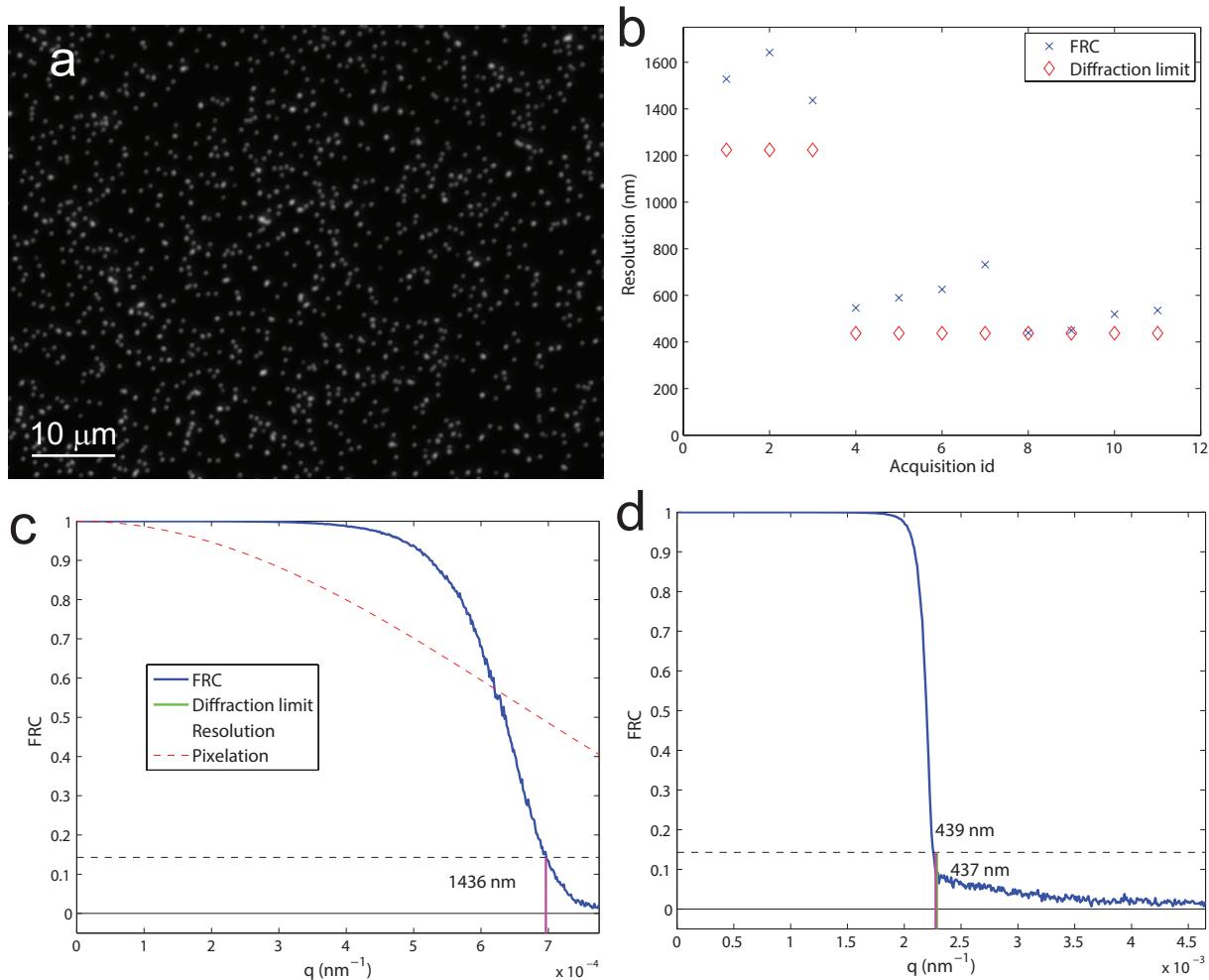
Supplementary Figure 11: **FLC for the data of Figure 2j.** (**a**) Part of the dataset of Fig. 2j. (**b**) FLC of this dataset, as well as the contour $FLC = 1/7 \approx 0.143$. The FLC has been smoothed with a Gaussian filter with a kernel of two pixels wide. As expected, the correlation and therefore also the resolution are lower in the direction of the filaments than in the orthogonal direction.

Supplementary Fig. 12
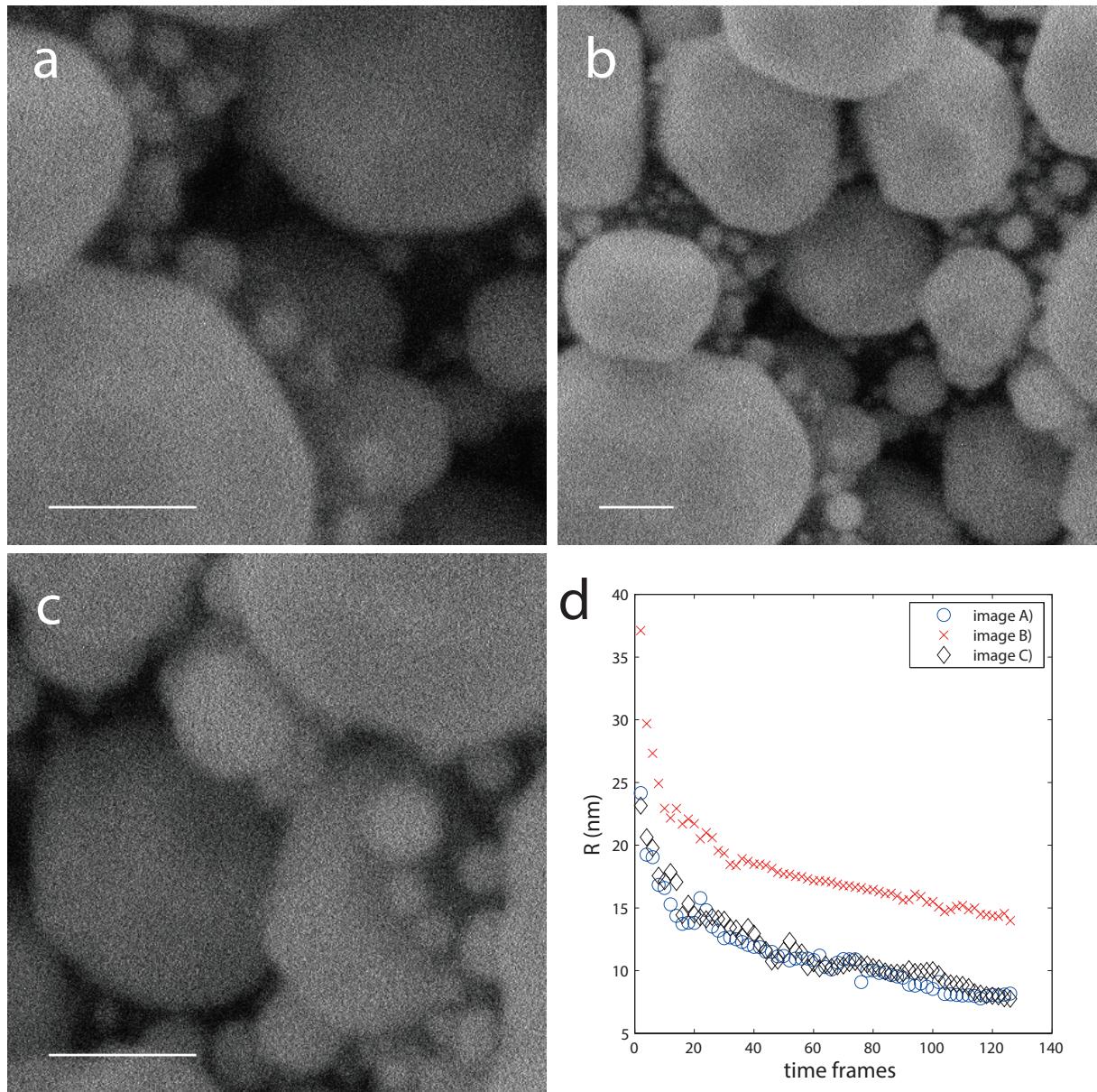
Supplementary Figure 12: **Local FRC resolution for the data of Figure 2a.** The local FRC resolution R was computed for square subregions of 128×128 superresolution pixels which were displaced by multiples of 84 pixels horizontally or vertically with respect to each other, i.e. 25% of the pixels in each region overlap with 25% of the pixels in each adjacent region. The resulting image was rescaled using spline interpolation and is shown here in false color overlaying the image of the filaments.

Supplementary Fig. 13
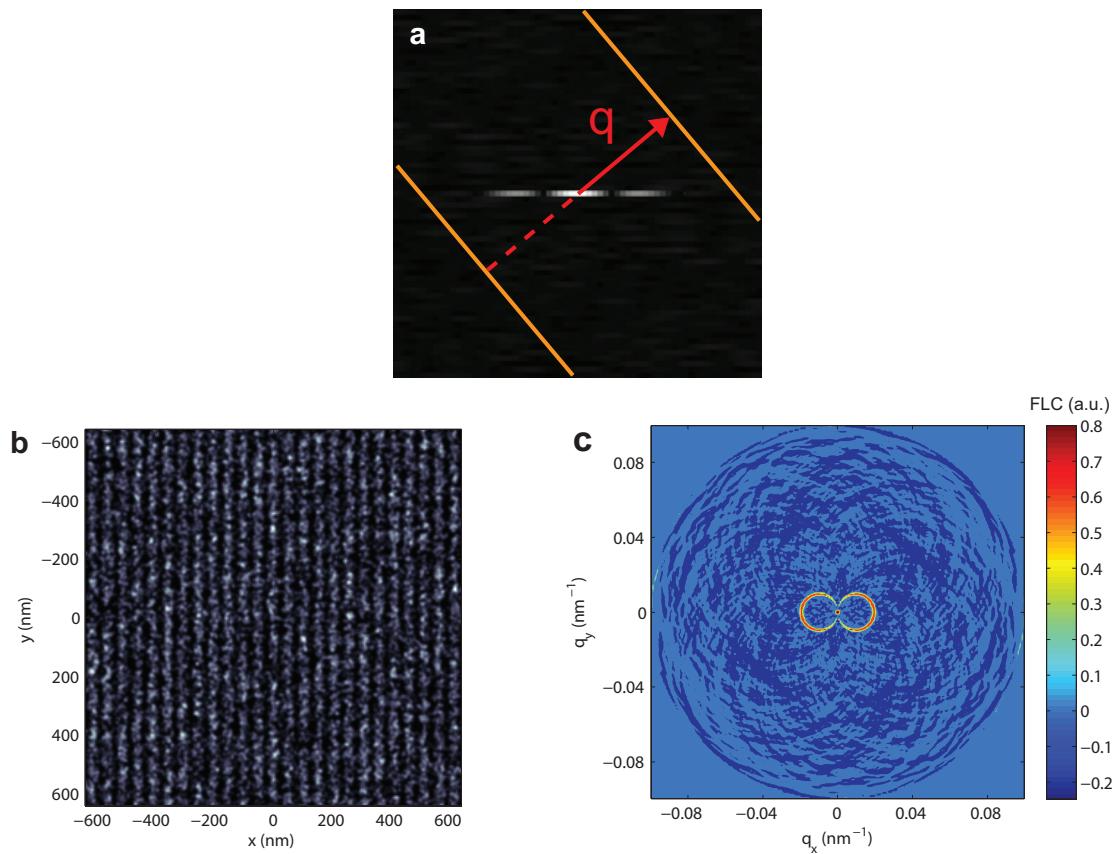
Supplementary Figure 13: **FRC curves and resolution values for widefield acquisitions.** (**a**) Part of an image of 1024×1024 pixels of fluorescent beads of 200 nm diameter. (**b**) Diffraction limit and computed FRC resolution for different acquisition settings and samples. For acquisition series ids 1-3 we used a 10x 0.25 NA lens resulting in slight undersampling. This sampling damps the transfer at higher spatial frequencies as depicted by the red dashed line. The found FRC resolution values are therefore ∼20% larger than the diffraction limit. For acquisition ids 4-11 we used a 60x 0.7 NA lens resulting in two times oversampling avoiding any pixelation related damping effects. The acquisition ids 4-8 show an increase in FRC resolution whereas the diffraction limit remains constant. This can be attributed to bleaching giving a decrease of SNR over time. During the experiment we kept the ADU on the camera roughly constant while increasing the gain of the camera. From Eq. (S.76) we expect a decrease in FRC resolution for increased influence of noise. For acquisition id 8 (shown in **a**) and id 9 we find resolution values at the diffraction limit as apparently the sample has sufficient transfer up to the OTF cut-off and noise is not the limiting factor. For acquisition ids 10 and 11 only less than 20 beads are visible in the whole field-of-view, which may have led to reduced high frequency contents of the sample but also directly influences the resolution according to Eq. (S.76). (**c**) FRC curve for acquisition id 3 and backprojected pixel size of 645 nm (612 nm Nyquist sampling). The red dashed line shows the transfer damping due to pixelation as Eq. (S.79). (**d**) FRC curve for image **a**, acquisition id 8. Backprojected pixel size is 107.5 nm (219 nm Nyquist sampling), so pixelation is not relevant here.

Supplementary Fig. 14
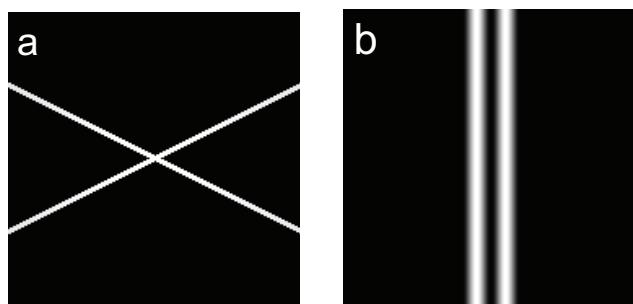
Supplementary Figure 14: **FRC resolution values in SEM acquisitions.** (**a-c**) cumulative image of 128 SEM acquisitions of single electron counting images. Scale bar 100 nm. (**d**) FRC resolution as a function of used time frames. As expected, the resolution improves over time due to the increase in SNR with acquisition time. The resolution for **c** is slightly worse than for **a** and **b**, which seems to be qualitatively correct.

Supplementary Fig. 15

Supplementary Figure 15: **Illustration of FLC principle.** (**a**) Illustration of the summation in Eq. (10) for FLC or FPC in Fourier space. All pixels on the orange lines perpendicular to $\vec{q}$ are used to compute the correlation averages $G_{jl}(\vec{q})$. For discrete implementation the width of the line is chosen to be one pixel. (**b-c**) Simulated localization microscopy dataset of a grating object without multiple localizations per fluorophore and the corresponding FLC as a function of the 2D-spatial frequency.

Supplementary Fig. 16

Supplementary Figure 16: **Test objects in simulations.** Objects used for the simulation results in Figs. 1a and 1b-c respectively. The dimensions of the images are (**a**) 180×180 and (**b**) 1024×1024 pixels, and the pixel sizes used were (**a**) 5 nm and (**b**) 1 nm. The pixel value of the white lines in image **a** is equal to ρ, and the width of the lines is 17.5 nm (i.e. 3.5 pixels). The maxima of the sinusoidal lines in **a** are 100 pixels (i.e. 200 nm) apart.

17

Supplementary Note 1: Average and variance of FRC

1.1 Imaging model

The object to be imaged, the 'ground truth' $o(\vec{r})$, is labeled with fluorescent probes. The distribution of fluorescent labels is described by the labeling density function:

$$\psi(\vec{r}) = \sum_{j=1}^{K} \delta\left(\vec{r} - \vec{r}_j^{\,em}\right), \tag{S.1}$$

and depends on the set of positions $\left\{\vec{r}_j^{\,\mathrm{em}} | j = 1, \ldots, K\right\}$ of the K labels.

Label j is activated and localized M_j times giving a total number of:

$$N = \sum_{j=1}^{K} M_j, \tag{S.2}$$

localizations, at the set of positions $\{\vec{r}_j \mid j = 1, \ldots, N\}$. It is assumed that localizations of the same emitter in subsequent image frames have been grouped into a single localization event such that there is only a single position estimate of an emitter each time it is activated. The probability density for localization of an emitter at position $\vec{r}$ is given by:

$$P_{\mathrm{loc}}(\vec{r}) = \frac{1}{K} \int d^n r'\, h(\vec{r} - \vec{r}')\, \psi(\vec{r}'), \tag{S.3}$$

where $h(\vec{r})$ is the localization Probability Distribution Function (PDF), which is taken to be a Gaussian:

$$h(\vec{r}) = \frac{1}{2\pi\sigma^2} \exp\left(-|\vec{r}|^2/2\sigma^2\right), \tag{S.4}$$

where the width satisfies $\sigma^2 = \sigma_0^2/n_{\mathrm{ph}}$ in the absence of background. Here σ_0 is a measure for the width of the PSF of the optical system and n_{ph} is the number of photons per emitter.

The usual way to argue that the localization PDF must be a Gaussian with variance decreasing as $1/n_{\mathrm{ph}}$ is that a measurement with only one photon gives the PSF as localization PDF, with variance σ_0^2, so by repeating the measurement n_{ph} times the localization PDF must be a Gaussian with variance given by $\sigma_0^2/n_{\mathrm{ph}}$. However, this argument does not apply because the actual PSF has infinite variance as the integral $\int dx dy\, PSF(x, y)\left(x^2 + y^2\right)$ diverges (the PSF decays with the second power of the coordinates times an oscillating function). A different argument is related to the asymptotic normality of the MLE-estimation of location. In case the number of signal photons is large the statistical error in the position estimation is small. Then we may approximate the log-likelihood with a parabolic function centered on the optimum. This means that the likelihood function (which is equal to the localization PDF) may be approximated with a Gaussian. In case of a non-zero background the same conclusion may be drawn, albeit with a different dependence of the localization uncertainty than the simple $1/\sqrt{n_{\mathrm{ph}}}$ relation, and provided that the number of signal photons is sufficiently large. We have used numerical analysis and the Kolmogorov-Smirnov (KS) test for finding out the similarity between the actual localization PDF and the Gaussian distribution. We have performed numerical tests using fully vectorial modeling of the PSF of a freely rotating dipole emitter with zero aberrations and zero background, and 500 signal photons [49], and the MLE localization routine implemented on GPU for speed [24]. The KS-statistic (maximum difference between the CDF and the Gaussian CDF) is typically $5 - 10 \times 10^{-3}$ for 5,000 localizations. The statistical significance of this residual deviation is characterized by the p-value, which takes values between 40% and 100% for the 5,000 sample runs. The validity of the scaling of the localization uncertainty with the inverse square root of the photon count follows from the fact that the variance of the localization error for MLE-estimation follows the CRLB for a Gaussian ground truth PSF over a wide range of photon counts [24]. It is noted that the Gaussian nature of the localization PDF would imply an infinite spatial frequency content (no cut-off, as opposed to the PSF). However, extrapolating to spatial frequencies corresponding to the sub-nm scale is not physically meaningful, but down to that length scale the Gaussian provides an excellent description of the localization PDF, provided of course the emitter is not too dim.

18
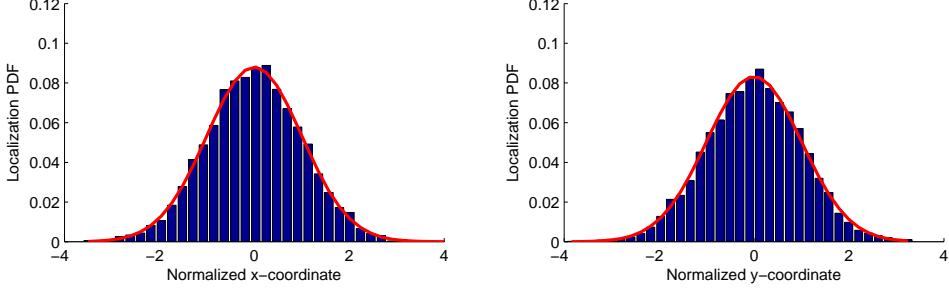
Figure S.17: Localization PDF and standard normal distribution, for 5,000 MLE-localizations for a freely rotating dipole, the simulated spots have on average 500 signal photons.

1.2 Average FRC

Each localization event is assigned to one of two groups with probability 1/2, leading to a split $N = N_1 + N_2$ that is governed by the binomial distribution. The M_j localizations of emitter j are also split into two parts $M_j = M_{1j} + M_{2j}$ according to the binomial distribution. The two sub-images are described by the image functions:

$$f_m(\vec{r}) = \sum_{j=1}^{N_m} \delta(\vec{r} - \vec{r}_j), \tag{S.5}$$

for $m = 1, 2$. The statistical averages of the two sub-images are simply:

$$\langle f_m(\vec{r}) \rangle = \frac{N}{2} P_{loc}(\vec{r}), \tag{S.6}$$

In the Fourier domain these relations are:

$$\left\langle \hat{f}_m(\vec{q}) \right\rangle = \frac{N}{2K} \hat{h}(\vec{q}) \hat{\psi}(\vec{q}), \tag{S.7}$$

with:

$$\hat{h}(\vec{q}) = \exp\left(-4\pi^2 \sigma^2 |\vec{q}|^2\right). \tag{S.8}$$

The definition for the FRC-curve in Eq. (1) was given by:

$$FRC(q) = \frac{\sum\limits_{\vec{q} \in \text{ring}} \hat{f}_1(\vec{q}) \hat{f}_2(\vec{q})^*}{\sqrt{\sum\limits_{\vec{q} \in \text{ring}} \hat{f}_1(\vec{q})^2} \sqrt{\sum\limits_{\vec{q} \in \text{ring}} \hat{f}_2(\vec{q})^2}}.$$

Consequently, to obtain the average FRC-curve we need to calculate the 2×2-correlation matrix elements:

$$\left\langle \hat{f}_m(\vec{q}) \hat{f}_n^*(\vec{q}) \right\rangle = \left\langle \sum_{j=1}^{N_m} \sum_{l=1}^{N_n} \exp\left(-2\pi i \vec{q} \cdot (\vec{r}_j - \vec{r}_l)\right) \right\rangle, \tag{S.9}$$

for $m, n = 1, 2$. In order to evaluate these averages we must consider pairs of localization events r_j and r_l. The N_m^2 pairs appearing in the diagonal correlation matrix elements can be divided in three distinct groups. These are the N_m pairs of identical localization events, the $\sum_{j=1}^{K} M_{m,j}(M_{m,j} - 1)$ pairs of different localization events of the same emitter, and the $N_m(N_m - 1) - \sum_{j=1}^{K} M_{m,j}(M_{m,j} - 1)$ pairs of different localization events of different emitters. A similar division can be made for the $N_m N_n$ pairs in the off-diagonal correlation matrix elements. The $N_m N_n$ pairs are necessarily all from different localization events, namely $\sum_{j=1}^{K} M_{m,j} M_{n,j}$ from the same emitter

19

and $N_m N_n - \sum_{j=1}^{K} M_{m,j} M_{n,j}$ from different emitters. The probability density for two localizations of the same emitter is:

$$P_{\text{loc,S}}(\vec{r}_1, \vec{r}_2) = \frac{1}{K} \int d^2r\, h(\vec{r}_1 - \vec{r})\, h(\vec{r}_2 - \vec{r})\, \psi(\vec{r}), \tag{S.10}$$

and the probability density for localizations from different emitters is:

$$P_{\text{loc,D}}(\vec{r}_1, \vec{r}_2) = \frac{1}{K^2} \int d^2r d^2r'\, h(\vec{r}_1 - \vec{r})\, h(\vec{r}_2 - \vec{r})\, \psi(\vec{r})\, \psi(\vec{r}'), \tag{S.11}$$

The Fourier transforms of these probability distributions are needed further on, and are given by:

$$\begin{aligned}
\hat{P}_{\text{loc,S}}(\vec{q}_1, \vec{q}_2) &= \int d^2r_1 d^2r_2\, P_{loc,S}(\vec{r}_1, \vec{r}_2) \exp\left(-2\pi i \vec{q} \cdot (\vec{r}_1 - \vec{r}_2)\right) \\
&= \frac{1}{K} \hat{h}(\vec{q}_1)\, \hat{h}(\vec{q}_2)^*\, \hat{\psi}(\vec{q}_1 - \vec{q}_2), \tag{S.12} \\
\hat{P}_{\text{loc,D}}(\vec{q}_1, \vec{q}_2) &= \int d^2r_1 d^2r_2\, P_{loc,D}(\vec{r}_1, \vec{r}_2) \exp\left(-2\pi i \vec{q} \cdot (\vec{r}_1 - \vec{r}_2)\right) \\
&= \frac{1}{K^2} \hat{h}(\vec{q}_1)\, \hat{h}(\vec{q}_2)^*\, \hat{\psi}(\vec{q}_1)\, \hat{\psi}(\vec{q}_2)^*. \tag{S.13}
\end{aligned}$$

Combining all these ingredients gives:

$$\begin{aligned}
\left\langle \left| \hat{f}_m(\vec{q}) \right|^2 \right\rangle &= \langle N_m \rangle + \left\langle \sum_{j=1}^{K} M_{m,j}(M_{m,j} - 1) \right\rangle \hat{P}_{\text{loc,S}}(\vec{q}, \vec{q}) \\
&\quad + \left\langle N_m(N_m - 1) - \sum_{j=1}^{K} M_{m,j}(M_{m,j} - 1) \right\rangle \hat{P}_{\text{loc,D}}(\vec{q}, \vec{q}), \tag{S.14} \\
\left\langle \hat{f}_m(\vec{q})\, \hat{f}_n^*(\vec{q}) \right\rangle &= \left\langle \sum_{j=1}^{K} M_{m,j} M_{n,j} \right\rangle \hat{P}_{\text{loc,S}}(\vec{q}, \vec{q}) \\
&\quad + \left\langle N_m N_n - \sum_{j=1}^{K} M_{m,j} M_{n,j} \right\rangle \hat{P}_{\text{loc,D}}(\vec{q}, \vec{q}). \tag{S.15}
\end{aligned}$$

Evaluating the averages over the binomial distribution of the localizations over the two sub-images gives:

$$\begin{aligned}
\left\langle \hat{f}_m(\vec{q})\, \hat{f}_n^*(\vec{q}) \right\rangle &= \frac{N}{2} \delta_{mn} + \frac{1}{4} \left\langle \sum_{j=1}^{K} M_j(M_j - 1) \right\rangle \hat{P}_{\text{loc,S}}(\vec{q}, \vec{q}) \\
&\quad + \frac{1}{4} \left[N(N-1) - \left\langle \sum_{j=1}^{K} M_j(M_j - 1) \right\rangle \right] \hat{P}_{\text{loc,D}}(\vec{q}, \vec{q}). \tag{S.16}
\end{aligned}$$

In case each emitter is localized only once it holds that:

$$\sum_{j=1}^{K} M_j(M_j - 1) = 0, \tag{S.17}$$

and as a result this term drops from the correlation averages. When emitters can be localized more than once this term will be non-zero. This term represents the *spurious correlations* between the two image halves, since having localizations from the same emitters in both halves violates the assumption that they are independent.

In order to find an expression for the ensemble average of $M_j(M_j - 1)$ a statistical model for on-off switching is needed. It appears that the number of activation cycles M_j for each emitter j is a Poisson distributed variable in case

20

the on-off switching statistics is described by a first order process and provided the ratio of the transition rates between the on and off-states is very different from one. This can be proven in the framework of the so-called asymmetric Random Telegraph Signal (RTS) model [50]. Also in a STORM acquisition where activator-reporter dye pairs are switched on with the same probability over many switching cycles, M_j would be Poisson distributed. In practice, the switching kinetics are much more subtle and the approximation as a Poisson process is merely a working assumption. Suppose now that there is a reservoir of K_r emitters and the average number of activation cycles is given by Q. Then a total of $K = (1 - \exp(-Q)) K_r$ emitters is localized at least once and a total of $N = QK_r$ localizations is generated. The average number of localizations per emitter for the group of emitters that has been localized at least once is:

$$Q_{\text{loc}} = \frac{N}{K} = \frac{Q}{1 - \exp(-Q)}. \tag{S.18}$$

Clearly, if the Poisson-rate Q goes to zero then the average number of localizations per emitter (restricted to the group of emitters that has been localized at least once) goes to one. It now follows that:

$$\left\langle \sum_{j=1}^{K} M_j (M_j - 1) \right\rangle = K_r Q^2 = NQ. \tag{S.19}$$

For a more general distribution of M_j, this is modified to:

$$\left\langle \sum_{j=1}^{K} M_j (M_j - 1) \right\rangle = N \left(\langle M_j \rangle + \frac{Var(M_j)}{\langle M_j \rangle} - 1 \right) \tag{S.20}$$

Consequently, M_j may be overestimated or underestimated if $Var(M_j) \neq \langle M_j \rangle$. For example, if almost all emitters have been localized and the dominant cause of variation of M_j is photobleaching, then M_j becomes geometrically distributed which implies that $Q \approx 2 \langle M_j \rangle - 2$. This means that $\langle M_j \rangle$ would be underestimated for $\langle M_j \rangle < 2$. For $\langle M_j \rangle > 2$, $\langle M_j \rangle$ is overestimated up to 100%. In general, photobleaching will lead to a distribution of M_j that is in between the Poisson distribution and the geometrical distribution and will therefore also lead to an overestimation of up to 100%.

We may now further develop the expressions for the correlation averages using that $N \gg 1$ and $K \gg 1$ and filling in the expressions for the localization pair distribution functions:

$$\left\langle \hat{f}_m (\vec{q}) \hat{f}_n^* (\vec{q}) \right\rangle = \frac{N}{2} \delta_{mn} + \frac{N}{4} \left[Q + \frac{N}{K^2} \left| \hat{\psi}(\vec{q}) \right|^2 \right] \left| \hat{h}(\vec{q}) \right|^2. \tag{S.21}$$

These statistical averages must next be integrated over Fourier space with the Fourier ring weight function:

$$\int d^2q' \, D(|\vec{q}'| - q) \left\langle \hat{f}_m(\vec{q}') \hat{f}_n^*(\vec{q}') \right\rangle = \frac{N}{2} \delta_{mn} + \frac{N}{4} [Q + NS(q)] \exp\left(-4\pi^2 \sigma^2 q^2\right), \tag{S.22}$$

with:

$$S(q) = \frac{1}{K^2} \int d^2q' \, D(|\vec{q}'| - q) \left| \hat{\psi}(\vec{q}') \right|^2, \tag{S.23}$$

and where the isotropy and explicit form of the Gaussian localization PDF is used. The Fourier ring weight function may be expressed as:

$$D(|\vec{q}'| - q) = \frac{\delta(|\vec{q}'| - q)}{2\pi q}, \tag{S.24}$$

Recalling the definition of the FRC in this notation as:

$$\langle FRC \rangle = \frac{\int d^2q' \, D(|\vec{q}'| - q) \left\langle \hat{f}_1(\vec{q}') \hat{f}_2^*(\vec{q}') \right\rangle}{\left(\int d^2q' \, D(|\vec{q}'| - q) \left\langle \hat{f}_1(\vec{q}') \hat{f}_1^*(\vec{q}') \right\rangle \right)^{1/2} \left(\int d^2q' \, D(|\vec{q}'| - q) \left\langle \hat{f}_2(\vec{q}') \hat{f}_2^*(\vec{q}') \right\rangle \right)^{1/2}} \tag{S.25}$$

21

This results in the final expression for the statistical average of the FRC:

$$\langle FRC \rangle = \frac{(Q + NS(q)) \exp\left(-4\pi^2\sigma^2 q^2\right)}{2 + (Q + NS(q)) \exp\left(-4\pi^2\sigma^2 q^2\right)}, \tag{S.26}$$

The ratio of the terms in the FRC nominator representing the genuine intrinsic image correlations and the spurious correlations at the resolution threshold is $NS(q_{\mathrm{res}})/Q = KS(q_{\mathrm{res}})$. It follows that the spurious correlations may be neglected provided the number of emitters K is sufficiently high and the sample has spectral signal content at q_{res}. Solving the resolution threshold $\langle FRC \rangle = 1/7$ then gives an apparent resolution:

$$R = \frac{2\pi\sigma}{\sqrt{\log(3Q)}}. \tag{S.27}$$

Clearly, the apparent resolution is simply a linear factor times the localization uncertainty.

A possible way to correct for the spurious correlations and come to a more realistic image resolution follows from a hypothetical division of the entire group of localization events such that all localizations of an emitter appear in either the first or the second data halve, but not mixed. Such a procedure results in a modified average FRC:

$$\langle FRC' \rangle = \frac{NS(q) \exp\left(-4\pi^2\sigma^2 q^2\right)}{2 + (2Q + NS(q)) \exp\left(-4\pi^2\sigma^2 q^2\right)}. \tag{S.28}$$

This modified FRC can also be produced if the average residual correlation $\sim Q \exp\left(-4\pi^2\sigma^2 q^2\right)$ is estimated from the data and then subtracted from the FRC nominator and added to the FRC denominator. This procedure for estimating the spurious correlation parameter Q is outlined in the Online Methods. Note that an effect of the parameter Q is still present in Eq. (S.28), but now represents the trade-off between the number of localizations per emitter and the total number of localized emitters at constant total number of localization events. In this trade-off it is more favorable to localize more emitters a fewer number of times.

The current model may be expanded by taking a distribution of localization uncertainties into account instead of a unimodal value. The correlation averages $\langle \hat{f}_m(\vec{q}) \hat{f}_n^*(\vec{q}) \rangle$ for a single value of the localization uncertainty must then be convolved with the distribution function of the localization uncertainty. Taking that distribution to be Gaussian with mean σ_{m} and width $\Delta\sigma$ we must replace the localization PDF factor $H(q) = \exp\left(-4\pi^2\sigma^2 q^2\right)$ by:

$$
\begin{aligned}
H(q) &= \int d\sigma \frac{1}{\sqrt{2\pi}\Delta\sigma} \exp\left(-\frac{(\sigma - \sigma_{\mathrm{m}})^2}{2\Delta\sigma^2}\right) \exp\left(-4\pi^2\sigma^2 q^2\right) \\
&= \frac{1}{\sqrt{1 + 8\pi^2\Delta\sigma^2 q^2}} \exp\left(-\frac{4\pi^2\sigma_{\mathrm{m}}^2 q^2}{1 + 8\pi^2\Delta\sigma^2 q^2}\right),
\end{aligned} \tag{S.29}
$$

in all expressions for the expectation value of the FRC.

1.3 Variance of the FRC

The variance of the FRC curve can be computed by using the following formula for the variance of the correlation coefficient $C = \vec{v_1} \cdot \vec{v_2} / \sqrt{|\vec{v_1}|^2 |\vec{v_2}|^2}$ between two random vectors $\vec{v_1}$ and $\vec{v_2}$ with the same mean [51]:

$$Var(C) = \frac{\left(2\alpha^2 + 1\right)\left(\alpha^2 + 1\right)^2 - \alpha^4\left(2\alpha^2 + 3\right)}{n\left(\alpha^2 + 1\right)^4}. \tag{S.30}$$

Here n is the dimension of the vectors and α^2 is defined as:

$$\alpha^2 \equiv \frac{\sum_{i=1}^{n} \langle v_{1,i} \rangle^2}{\sum_{i=1}^{n} Var(v_{1,i})} = \frac{\sum_{i=1}^{n} \langle v_{2,i} \rangle^2}{\sum_{i=1}^{n} Var(v_{2,i})}. \tag{S.31}$$

22

This equation holds under the conditions that n is large and that the components of $\vec{v_1}$ and $\vec{v_2}$ are independent. If $\vec{v_1}$ and $\vec{v_2}$ are associated with the real and imaginary parts of the frequency components $\hat{f}_1(\vec{q})$ and $\hat{f}_2(\vec{q})$ in the Fourier rings, then it appears that:

$$\langle FRC \rangle = \left\langle \frac{\sum\limits_{\vec{q} \in \text{ring}} \hat{f}_1(\vec{q}) \hat{f}_2(\vec{q})^*}{\sqrt{\sum\limits_{\vec{q} \in \text{ring}} \hat{f}_1(\vec{q})^2} \sqrt{\sum\limits_{\vec{q} \in \text{ring}} \hat{f}_2(\vec{q})^2}} \right\rangle \approx \frac{\sum\limits_{\vec{q} \in \text{ring}} \left| \left\langle \hat{f}_1(\vec{q}) \right\rangle \right|^2}{\sum\limits_{\vec{q} \in \text{ring}} \left| \hat{f}_1(\vec{q}) \right|^2}$$

$$= \frac{\alpha^2}{1 + \alpha^2}, \tag{S.32}$$

so that:

$$\alpha^2 = \frac{\langle FRC \rangle}{1 - \langle FRC \rangle}. \tag{S.33}$$

The quantity α^2 can be interpreted here as the signal-to-noise ratio of a half data image in the Fourier ring, (i.e. half the SNR of the total dataset). Substituting this into Eq. (S.30) then gives:

$$Var(FRC) = \frac{1}{n} (1 - \langle FRC \rangle)^2 \left(1 + 2\langle FRC \rangle - \langle FRC \rangle^2 \right), \tag{S.34}$$

where n is the number of pixels in a Fourier ring. Since the width of a Fourier pixel is $1/L$, where L is the linear size of the field of view, $n = 2\pi qL$ and therefore one gets finally:

$$Var(FRC) = \frac{1}{2\pi qL} (1 - \langle FRC \rangle)^2 \left(1 + 2\langle FRC \rangle - \langle FRC \rangle^2 \right). \tag{S.35}$$

The careful reader may notice that this equation suggests that increasing L through zero padding of the superresolution images would decrease the variance. However, because neighboring Fourier pixels are actually slightly correlated, increasing L through zero padding will increase the sampling in Fourier space but also leads to stronger correlations among neighboring pixels. Therefore, the effective number of independent pixels in a Fourier ring remains the same and the variance is not affected by zero padding.

Supplementary Note 2: FRC resolution for periodic object models

In this section we will work out explicit expressions for the FRC resolution for periodic structures (gratings) with different cross-sections. In all cases we assume that the Q-parameter is small, so $Q \ll 1$. Under this condition, setting the right-hand side of Eq. (S.26) equal to the $1/7$ threshold criterion leads to the following equality:

$$NS\left(q_{\text{res}}\right)\exp\left(-4\pi^2\sigma^2 q_{\text{res}}^2\right) = \frac{1}{3}. \tag{S.36}$$

This expression now needs to be solved to obtain an explicit expression for the FRC resolution.

2.1 Grating with a cosine cross-section

Consider a periodic structure consisting of $M \geq 1$ periods with period d and length L, and with an average density of localized labels ρ. The total number of localized labels is then $K = M\rho dL$. The labeling density function is given by:

$$\psi\left(\vec{r}\right) = \begin{cases} \rho\left[1 - \cos\left(2\pi x/d\right)\right] & \text{if } |x| < Md/2 \text{ and } |y| < L/2 \\ 0 & \text{otherwise} \end{cases} \tag{S.37}$$

Fourier transforming results in:

$$\begin{aligned}
\hat{\psi}\left(\vec{q}\right) &= K\left[\operatorname{sinc}\left(\pi Mq_x d\right) + \frac{1}{2}\operatorname{sinc}\left(\pi Md\left(q_x - 1/d\right)\right)\right.\\
&\quad \left. + \frac{1}{2}\operatorname{sinc}\left(\pi Md\left(q_x + 1/d\right)\right)\right]\operatorname{sinc}\left(\pi q_y L\right).
\end{aligned} \tag{S.38}$$

If L is sufficiently large, this gives the following expression for $S(q)$ at $q = 1/d$:

$$S\left(q\right) = \frac{1}{2\pi K^2}\int_0^{2\pi} d\varphi_k \left|\hat{\psi}_{em}\left(\vec{q}\right)\right|^2 \approx \frac{d}{\pi K^2}\int_{-\infty}^{+\infty} dq_y \left|\hat{\psi}_{em}\left(\vec{q}\right)\right|^2, \tag{S.39}$$

where the last integral is evaluated at $q_x = q = 1/d$, and so:

$$S\left(q\right) \approx \frac{d}{4\pi}\int_{-\infty}^{+\infty} dq_y \operatorname{sinc}^2\left(\pi q_y L\right) = \frac{d}{4\pi L}. \tag{S.40}$$

The resolution $R = 1/q_{\text{res}}$ follows from solving the threshold criterion Eq. (S.36). If the grating can just be resolved, then $R = d$. It follows that an expression may be derived for the required density of localized labels ρ for achieving a resolution $R = d$, given the localization uncertainty σ:

$$\rho = \frac{4\pi}{3MR^2}\exp\left(4\pi^2\sigma^2/R^2\right). \tag{S.41}$$

The dependence on the ratio σ/R is rather steep. For example, when $\sigma/R = 1$ the exponential factor is already on the order 10^{17}! Realistic minimum labeling densities arise when σ/R is less than approximately 1/4. So, even though the resolution is not determined solely by the localization uncertainty, the necessity of having practically achievable labeling densities does imply that the minimum resolution is of the order of the localization uncertainty. Eq. (S.41) can also be solved for the resolution R as a function of the density of localized labels and the localization uncertainty:

$$R = \frac{2\pi\sigma}{\sqrt{W\left(3\pi M\rho\sigma^2\right)}}, \tag{S.42}$$

where $W\left(\cdot\right)$ is the Lambert W-function [52], which is the inverse of $y = x\exp(x)$. A useful analytical approximation for the Lambert W-function is $W(y) \approx \ln\left(2y\right) - \ln\left(\ln\left(1 + 2y\right)\right)$ [21].

It appears that an increase in M leads to a better resolution. This implies that adding grating lines to the global image improves the ability to distinguish neighboring lines. The underlying reason for this is that a grating is a perfectly periodic structure. Therefore, if M increases then the contributions of the different lines add up in Fourier space and the SNR increases.

24

2.2 Grating with a square cross-section

In a similar way as above, it is also possible to derive the resolution for a grating object consisting of $M \geq 1$ lines of width a, length L and separation $d > a$. The density of localised labels is ρ and the total number of labels is $K = M\rho aL$. For such an object, the Fourier transform of the labeling density function $\psi(\vec{r})$ reads:

$$\hat{\psi}(\vec{q}) = \rho aL \frac{\sin(\pi M q_x d)}{\sin(\pi q_x d)} \operatorname{sinc}(\pi q_x a) \operatorname{sinc}(\pi q_y L). \tag{S.43}$$

The ring average of the spectral density is found for sufficiently large line length L as:

$$S(q) = \frac{1}{2\pi K^2} \int_0^{2\pi} d\psi \left|\hat{\psi}(\vec{q})\right|^2 \approx \frac{1}{\pi q K^2} \int_{-\infty}^{+\infty} dq_y \left|\hat{\psi}(\vec{q})\right|^2, \tag{S.44}$$

where the last integral is evaluated at $q_x = q$, and so:

$$S(q) \approx \frac{\operatorname{sinc}^2(\pi q a)}{\pi q} \left[\frac{\sin(\pi M q d)}{M \sin(\pi q d)}\right]^2 \int_{-\infty}^{+\infty} dq_y \operatorname{sinc}^2(\pi q_y L) = \frac{\operatorname{sinc}^2(\pi q a)}{\pi q L} \left[\frac{\sin(\pi M q d)}{M \sin(\pi q d)}\right]^2. \tag{S.45}$$

The equation for solving the resolution is:

$$\frac{3M\rho a}{\pi q_{\text{res}}} \operatorname{sinc}^2(\pi q_{\text{res}} a) \left[\frac{\sin(\pi M q_{\text{res}} d)}{M \sin(\pi q_{\text{res}} d)}\right]^2 \exp\left(-4\pi^2 \sigma^2 q_{\text{res}}^2\right) = 1. \tag{S.46}$$

The grating can just be resolved ($q_{\text{res}} = 1/d$) if the density of localized labels satisfies:

$$\rho = \frac{\pi}{3MaR} \frac{\exp\left(4\pi^2 \sigma^2 / R^2\right)}{\operatorname{sinc}^2(\pi a / R)}. \tag{S.47}$$

In the limit of thin lines ($a \ll d$) this may be inverted to give a resolution:

$$R = 2\pi\sigma \sqrt{\frac{2}{W(72M^2 \rho_{\text{lin}}^2 \sigma^2)}}, \tag{S.48}$$

where $\rho_{lin} = \rho a$. For a single line ($M = 1$) the equation for solving the resolution simplifies to:

$$\frac{3\rho a}{\pi q_{\text{res}}} \operatorname{sinc}^2(\pi q_{\text{res}} a) \exp\left(-4\pi^2 \sigma^2 q_{\text{res}}^2\right) = 1. \tag{S.49}$$

In the limit of large density of localized labels ($\rho a^2 \gg 1$) and small localization uncertainty ($\sigma \ll a$) this equation may be solved to give $R = a$, i.e. the resolution is equal to the line width. This may seem counter-intuitive, but it reflects the fact that no details smaller than the line width are present in the object itself. The predicted resolution from M grating lines might be tested if super-resolution data is obtained for a sample with a very sophisticatedly engineered ground truth [53].

2.3 Effect of spurious correlations

The preceding analysis provides some quantitative insights into the relevance of the spurious correlations between the two sub-images for techniques with multiple localizations per emitter. The equation for the resolution for a grating with the cosine cross-section for a non-zero Q-parameter is:

$$\left[Q + \frac{M\rho_{\text{loc}}R^2}{4\pi}\right] = \exp\left(4\pi^2 \sigma^2 / R^2\right), \tag{S.50}$$

where $\rho_{\text{loc}} = \rho Q / (1 - \exp(-Q))$ is the density of localizations, proportional to the labeling density and the average number of localizations per emitter (Supplementary Note 1). The effects of spurious correlations on resolution may be

neglected provided the ratio of the second and first term on the l.h.s. of Eq. (S.50) is sufficiently large. This happens when ρ_{loc} satisfies:

$$\rho_{\mathrm{loc}} \gg \frac{\pi Q}{M}\left(\frac{2}{R}\right)^2. \tag{S.51}$$

This regime is typically found when the resolution is much worse than the Nyquist resolution following from the density of localized labels ρ, i.e. $R \gg 2/\sqrt{\rho}$. Eq. (S.51) can be used as a self-consistency test for the need to correct for spurious correlations. The resolution found without Q-correction and the estimated value of Q can be confronted given the experimentally found density of localizations. If the inequality is satisfied for, say $M = 2$, then there is probably no need to recompute the resolution with Q-correction, as the magnitude of the correction does not outweigh the added uncertainties of the additional processing steps.

2.4 Relation to Nyquist sampling

Consider Eq. (S.42) in the limit where $\sigma \to 0$. The resolution becomes:

$$R = \sqrt{\frac{4\pi}{3M\rho}}, \tag{S.52}$$

which is nearly equal to the Nyquist resolution $R_{\mathrm{Nyquist}} = 2/\sqrt{\rho}$:

$$R = \sqrt{\frac{\pi}{3M}} R_{\mathrm{Nyquist}} \approx R_{\mathrm{Nyquist}} \quad \text{for } M = 1, \tag{S.53}$$

The FRC resolution is not exactly equal to the Nyquist resolution in this limit, because it is conceptually different: it describes for which spatial frequency q there is a sufficiently high signal-to-noise ratio rather than an absence of aliasing due to undersampling. However, the Nyquist sampling theorem does not strictly apply since localizations do not constitute samples of a bandwidth limited function. Therefore it is not surprising that these two concepts give slightly different values for the resolution.

Supplementary Note 3: Labeling-localization trade-off

3.1 Trade-off for a periodic object model

The expression in Eq. (S.42) can be used to analyze the impact of improvement in the density of localized labels ρ and the localization uncertainty σ on the resolution. Clearly obtaining infinitely many localizations yields $R = 0$, whereas perfect localization at $\sigma = 0$ was shown in the preceding section to result in a finite resolution value. This is the consequence of the imaging model, in which structures can have infinitely many labels attached to them. Localizing all these labels yields a blurry image without counting noise due to finite localization densities, which implies that the signal-to-noise ratio is infinite for all spatial frequencies.

Alternatively one could ask whether marginal improvements in ρ or σ yield greater improvements in resolution. The trade-off point in which both improvements are equivalent is found by requiring that the relative change in resolution for a given relative change in Nyquist area $\sim 1/\rho$ is equal to a relative change in resolution for a given relative change in localization uncertainty area σ^2, i.e. if we change either quantity with a given percentage, the resulting percentile change in resolution must be the same:

$$\frac{\sigma^2}{R} \frac{\partial R}{\partial \sigma^2} = -\frac{\rho}{R} \frac{\partial R}{\partial \rho}. \tag{S.54}$$

Evaluating the derivatives results in:

$$\frac{\sigma^2}{R} \frac{\partial R}{\partial \sigma^2} = \frac{W\left(3\pi M \rho \sigma^2\right)}{2\left(1 + W\left(3\pi M \rho \sigma^2\right)\right)}, \tag{S.55}$$

$$\frac{\rho}{R} \frac{\partial R}{\partial \rho} = -\frac{1}{2\left(1 + W\left(3\pi M \rho \sigma^2\right)\right)}. \tag{S.56}$$

It follows that:

$$W\left(3\pi M \rho \sigma^2\right) = 1, \tag{S.57}$$

implying that the resolution must be:

$$R = 2\pi\sigma, \tag{S.58}$$

and that the optimum trade-off occurs for:

$$\rho\sigma^2 = \frac{e}{3\pi M}. \tag{S.59}$$

For a two-line object ($M = 2$) this corresponds to $\rho\sigma^2 = e/6\pi \approx 0.144$. For M parallel lines we obtain a lower value, from which it may be inferred that for any intricate but irregular object structure the trade-off occurs for a value smaller than 0.144. For that reason the value s$\rho\sigma^2 = e/6\pi \approx 0.144$ should be considered in practice only as a rule-of-thumb.

3.2 Trade-off in general

The preceding analysis of marginal improvements in ρ and σ can be generalized by not assuming a specific object model. For this analysis it is useful to revisit Eq. (S.28). At $q = q_{\text{res}}$, the expected FRC curve drops below the threshold, which means that at this point $q_{\text{res}} = 1/R$ is an increasing function of:

$$A\left(q_{\text{res}}\right) = N \exp\left(-4\pi^2 \sigma^2 q_{\text{res}}^2\right). \tag{S.60}$$

Therefore R has its minimum whenever $A\left(q\right)$ has its maximum. Consider now the relative changes in A due to changes in N and n_{ph}:

$$\frac{N}{A} \frac{dA}{dN} = 1, \tag{S.61}$$

$$\frac{n_{\text{ph}}}{A} \frac{dA}{dn_{\text{ph}}} = -4\pi^2 n_{\text{ph}} q_{\text{res}}^2 \frac{d\sigma^2}{dn_{\text{ph}}}. \tag{S.62}$$

27

If $\sigma^2 \propto 1/n_{\mathrm{ph}}$ then these expressions are equal if $2\pi\sigma q_{\mathrm{res}} = 1$ so that $R = 2\pi\sigma$. This is therefore the point where obtaining 1% more localizations has about the same effect on A (and R) as obtaining 1% higher photon counts per localization. Hence $R = 2\pi\sigma$ marks the boundary between the regime $R > 2\pi\sigma$ where the resolution is limited by the number of localizations (labeling density) and the regime $R < 2\pi\sigma$ where the resolution is limited by localization uncertainty.

If a limited amount of time is available, then $R = 2\pi\sigma$ also marks the value of σ for which R is optimal. The localization uncertainty can be improved by increasing n_{ph}. However, when this is accomplished by increasing the on-time of the fluorophores, this also reduces the total number of labels that can be localized in a given acquisition time. This is implied by the requirement of having a sufficiently large distance between individual activated emitters at any point in time [22]. The decrease of σ thus has a positive effect on resolution, whereas the decrease of N has a negative effect on resolution, implying that an optimum can be found by balancing the two effects. This trade-off was already identified in the first publication on localization microscopy [3], where it was noted that 'Including fewer, but brighter, molecules results in higher localization and crisper images, but at a reduced molecular density giving less complete information about the spatial distribution of the target protein'.

This argument can be made quantitative by setting $N \propto \tau_{\mathrm{on}}^{-1}$ and $n_{\mathrm{ph}} \propto \tau_{\mathrm{on}}$ and subsequently considering the following equation:

$$
\begin{aligned}
\frac{\tau_{\mathrm{on}}}{A}\frac{dA}{d\tau_{\mathrm{on}}} &= \frac{\tau_{\mathrm{on}}}{A}\frac{dA}{dn_{\mathrm{ph}}}\frac{dn_{\mathrm{ph}}}{d\tau_{\mathrm{on}}} + \frac{\tau_{\mathrm{on}}}{A}\frac{dA}{dN}\frac{dN}{d\tau_{\mathrm{on}}} \\
&= \frac{n_{\mathrm{ph}}}{A}\frac{dA}{dn_{\mathrm{ph}}} - \frac{N}{A}\frac{dA}{dN}.
\end{aligned}
\tag{S.63}
$$

This derivative is equal to zero when $R = 2\pi\sigma$, which means that this expression marks the point where the best resolution has been obtained for the given amount of time available. Since the resolution is assumed to always increase as more time becomes available, this also implies that $R = 2\pi\sigma$ marks when the resolution R was obtained in the shortest possible amount of time.

The tuning of the switching kinetics of emitters outlined above is typically not a very important issue yet for localization microscopy experiments. Often the choice of fluorescent labels is constrained by the biological context, and under these constraints it is possible that some labels provide more photons in a shorter on-time than other labels. Moreover, imaging does not always take place at the optimal density of simultaneously active emitters, which means that increasing the on-time does not require that fewer emitters are simultaneously active. However, with the rapid developments of new fluorescent dyes and proteins, this trade-off described above will become more important in the future. Tuning may then be done by the choice of the fluorescent label or buffer composition. Tuning of emitter switching kinetics independent of brightness was demonstrated for example for oxazine dyes using the concentrations of the reducing and oxidizing agents in the imaging buffer [54]. Alternatively, if PALM imaging is combined with a triplet state relaxation scheme [55], the bleaching rate could be decoupled from the excitation intensity, giving the bleaching rate as tuning parameter. Then, in a fixed total time fewer but brighter single emitter events yield more accurate localizations but at the expense of a lower recorded emitter density [3].

3.3 Minimal time to resolution

The rule $R = 2\pi\sigma$ indicates for which σ the resolution R is obtained as quickly as possible. This is the conjugate of the result of Fig. 1c where it is shown that $R = 2\pi\sigma$ indicates when the highest possible resolution is obtained for a fixed total measurement time. To support this insight, the simulations results from 1c were taken, but this time lines of constant resolution R were calculated by taking the contour lines in the $\sigma T_{\mathrm{total}}$-plane where $FRC\,(q = 1/R) = 1/7$. The resulting graph is shown in Fig. S.18. As in Fig. 1, c, the lines from the simulated data show good agreement with the theoretically predicted lines. Moreover, Fig. S.18 shows that the red curve corresponding to $R = 2\pi\sigma$ does indeed seem to go through the points where the measurement time T_{total} is minimal for each resolution in the simulations.

28
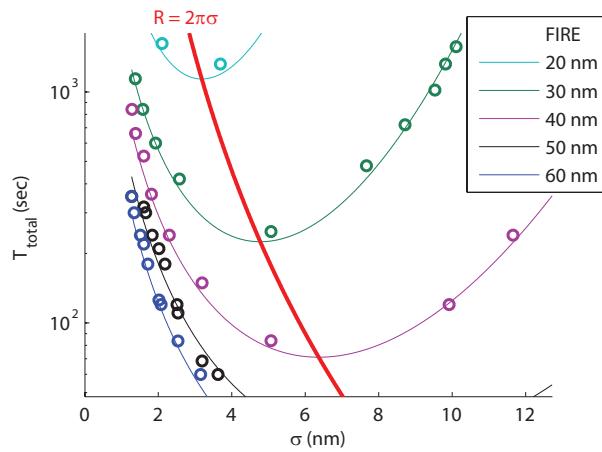
Figure S.18: Solid lines indicate constant resolution as a function of the total measurement time T_{total} and localization precision σ as predicted by Eq. (3) for two lines. The circles represent simulations averaged over 100 realizations. The red separation line $R = 2\pi\sigma$ goes through the shortest measurement time for each resolution in theoretical and simulated data. Left of the separation not enough emitters are collected, whereas to the right the emitters have not been localized precisely enough.

Supplementary Note 4: Application of FRC resolution to other imaging modalities

The FRC resolution concept can be applied to imaging modalities for which the image formation theory may be centered around the conventional concept of the Point Spread Function (PSF). This applies to both diffraction-limited modalities such as confocal or widefield fluorescence imaging, as well as to diffraction-unlimited modalities such as STED. There the resolving power of the imaging system is given by a modification of Abbe's formula[36]

$$d = \frac{\lambda}{2NA\sqrt{1 + I_{\text{STED}}/I_0}}, \tag{S.64}$$

where the crucial parameter is now the ratio of the intensity of the STED beam I_{STED} to the saturation intensity I_0. This measure characterizes optics, properties of the fluorophore and imaging conditions, but does not take the sample structure into account. We demonstrate here that the FRC resolution depends on the Optical Transfer Function (OTF), the spatial frequency content of the sample and the noise level. The resolution is always inferior to the spatial cut-off frequency for which the optical transfer goes to zero. This maximum resolution is only reached in the limit of high signal-to-noise ratio (SNR). So, for diffraction-limited modalities the FRC resolution coincides with the Abbe-resolution in the limited of high SNR. For the STED case it coincides with the formula of Eq. (S.64) in the limit of dense spatial frequency content of the sample.

Starting point of the proof is the labeling density function for an object consisting of fluorescent labels located at positions $\vec{r}_j^{em}$ for $j = 1, 2, \ldots K$:

$$\psi(\vec{r}) = \sum_{j=1}^{K} \delta\left(\vec{r} - \vec{r}_j^{em}\right). \tag{S.65}$$

We assume that the on average n_{ph} photons are detected per emitter. Then the signal part of the image is given by:

$$s(\vec{r}) = n_{\text{ph}} \int d^2 r' \, h(\vec{r} - \vec{r}') \psi(\vec{r}), \tag{S.66}$$

with $h(\vec{r})$ the PSF, which is normalized such that:

$$\int d^2 r \, h(\vec{r}) = 1. \tag{S.67}$$

In the Fourier domain the relation between the signal and the labeling density function is:

$$\hat{s}(\vec{q}) = n_{\text{ph}} \hat{h}(\vec{q}) \hat{\psi}(\vec{q}). \tag{S.68}$$

The FRC requires two measured images:

$$f_j(\vec{r}) = s(\vec{r}) + n_j(\vec{r}), \tag{S.69}$$

for $j = 1, 2$ where $n_j(\vec{r})$ is the noise. In order to simplify the analysis we assume that the only noise source is shot noise from photon statistics. In that case the correlation function for the two images is:

$$\langle f_j(\vec{r}) f_l(\vec{r}') \rangle = s(\vec{r}) s(\vec{r}') + s(\vec{r}) \delta(\vec{r} - \vec{r}') \delta_{jl}, \tag{S.70}$$

which gives after Fourier transformation:

$$\langle \hat{f}_j(\vec{q}) \hat{f}_l(\vec{q}')^* \rangle = \hat{s}(\vec{q}) \hat{s}(\vec{q}')^* + \hat{s}(\vec{q} - \vec{q}') \delta_{jl}. \tag{S.71}$$

If $\vec{q} = \vec{q}'$ then:

$$\begin{aligned}
\langle \hat{f}_j(\vec{q}) \hat{f}_l(\vec{q})^* \rangle &= |\hat{s}(\vec{q})|^2 + \hat{s}(0) \delta_{jl} \\
&= n_{\text{ph}}^2 \left|\hat{h}(\vec{q})\right|^2 \left|\hat{\psi}(\vec{q})\right|^2 + K n_{\text{ph}} \delta_{jl}. \tag{S.72}
\end{aligned}$$

Taking the average over rings in Fourier space, using the assumption that the OTF is rotationally symmetric, and defining (identical to the case of localization microscopy, Eq. (S.23)):

$$S\left(q\right) = \frac{1}{K^2} \int d^2 q' \, D\left(|\vec{q}'| - q\right) \left|\hat{\psi}\left(\vec{q}'\right)\right|^2,$$ (S.73)

with

$$D\left(|\vec{q}'| - q\right) = \frac{\delta\left(|\vec{q}'| - q\right)}{2\pi q},$$ (S.74)

the Fourier ring weight function, we obtain an expression for the expected value of the FRC:

$$\langle FRC \rangle = \frac{K^2 n_{\mathrm{ph}}^2 \left|\hat{h}\left(\vec{q}\right)\right|^2 S\left(q\right)}{K^2 n_{\mathrm{ph}}^2 \left|\hat{h}\left(\vec{q}\right)\right|^2 S\left(q\right) + K n_{\mathrm{ph}}}.$$ (S.75)

Clearly, the FRC decays to zero for large $\vec{q}$ as the OTF $\hat{h}\left(q\right)$ also goes to zero for large $\vec{q}$ and is zero for $|\vec{q}| \geq q_{cut}$. The FRC resolution for the sum image is obtained by setting the FRC equal to the threshold value of 1/7. This results in the implicit equation for q_{res}:

$$6 K n_{\mathrm{ph}} \left|\hat{h}\left(q_{\mathrm{res}}\right)\right|^2 S\left(q_{\mathrm{res}}\right) = 1.$$ (S.76)

So, in general, the FRC resolution depends on the OTF $\hat{H}(q)$, the spatial frequency content of the object $S(q)$, and the level of noise (photon count). In the limit of infinite SNR the resolution resulting from this analysis is given by:

$$\hat{h}\left(q_{\mathrm{res}}\right) = 0,$$ (S.77)

which corresponds to the Abbe-resolution for diffraction-limited imaging systems. It is stressed that the conclusions described here are not altered by obvious generalizations of the model, e.g. effects of finite pixel size and additional noise sources such as readout noise.

The analysis for diffraction-limited systems can be readily extended to diffraction-unlimited methods such as STED, taking the effective transfer at a given I_{STED} for the OTF-function in Eq. (S.76). Interestingly, previous theoretical analyses of STED[36] have shown before that the sample structure ultimately influences the resolution in the case of finite intensities. For STED as well as for fluorescence nanoscopy techniques it should be noted that this conclusion is contingent on the assumption made here that the labels on the sample are the relevant signal source. If the labeled structure rather than the labels themselves is considered to be the signal source, then the labeling process becomes a source of noise as well. Repeating the above analysis with the labeling process included as a noise source then results in an extra term $2 K n_{\mathrm{ph}}^2 \left|\hat{h}\left(\vec{q}\right)\right|^2$ in the denominator of Eq. (S.75). From this it can be concluded that the labeling density, rather than the OTF, may become the limiting factor to resolution for high photon counts in e.g. STED.

31

Supplementary Note 5: Computation of FRC and FRC resolution

5.1 Data splitting

For this work, half data sets were obtained by splitting the timeseries into blocks of 500 frames and assigning an equal number of blocks randomly to each half set. Alternatively, localizations could also have been assigned randomly to the half sets or the timeseries could have been split into two parts to obtain two sets with half the localizations. Both these methods have some issues though. Random assignment may cause spurious correlations due to localizations of a single emitter activation event being assigned to both half sets. This will also happen if the block size for splitting the data is too short compared with the on-time τ_{on} of the emitters.

Splitting the timeseries in two parts may cause the half sets to be overly dissimilar due to systematic differences between the halves of the timeseries. Next to drift, for example, photobleaching may occur and then the localizations in the second half set arise from fewer emitters or from more photostable emitters. On the other hand, photobleaching may also reduce spurious correlations due to reactivation of previously active emitters in the second half set for this splitting method. We have nevertheless chosen for splitting the experimental timeseries in smaller blocks because it only has the problem of spurious correlations and because it is more accurate in the absence thereof.

5.2 Discretization

Two issues regarding the discrete computation of the FRC curve require some further discussion here. The first issue is the masking of the binned images. This masking is needed to suppress high frequency components in the Fourier transforms that may result from the edges of the binned images. It is a common technique in signal processing [56] to suppress edge artifacts resulting from the digital Fourier transformation which cannot be avoided otherwise. If an object which is imaged extends up to the edge of an image and no mask is applied, then the finite extent of the image acts as a rectangular window $M(\vec{r})$ which is 1 inside the field of view and 0 outside. In the frequency domain this has the effect on the images $\hat{f}(\vec{q})$ that:

$$\hat{f}(\vec{q}) \rightarrow \hat{f}(\vec{q}) * \hat{M}(\vec{q}) = \hat{f}(\vec{q}) * L^2 \text{sinc}(\pi q_x L) \text{sinc}(\pi q_y L) . \tag{S.78}$$

Here $\text{sinc}(x) \equiv \sin(x)/x$ and $*$ is the convolution operation. The window extends the highly correlating low spatial frequency components into the higher spatial frequencies due to the finite width and oscillating tails of the sinc functions. Therefore it increases the FRC at those frequencies and leads to a higher threshold frequency q_{res}. In order to reduce this effect, a mask $M(\vec{r})$ with a smooth drop-off to 0 at the edges has to be applied which is narrower in the frequency domain. Note that q_{res} will then still be slightly overestimated, proportional to the finite width of $\hat{M}(\vec{q})$.

A second issue worth mentioning here is the effect of the pixel size in the binned images (i.e. the superresolution pixel size) on the FRC. As a first consideration, pixel sizes larger than $R/2$ will result in aliasing at spatial frequencies around q_{res}. This will appear as additional correlation in the FRC and may therefore artificially increase the estimated resolution. Furthermore, the pixel binning also acts as a low pass filter. If l is the pixel size then the effect is approximately as if the exponential in Eq. (2) is replaced by:

$$\exp\left(-4\pi^2\sigma^2 q^2\right) \rightarrow \text{sinc}(\pi q l)^2 \exp\left(-4\pi^2\sigma^2 q^2\right) . \tag{S.79}$$

At $q = q_{\text{res}}$ this is the same as replacing σ in Eq. (2) by:

$$\sigma^2 \rightarrow \sigma_{\text{eff}}^2 = \sigma^2 - \frac{\log\left(\text{sinc}(\pi q l)^2\right)}{4\pi^2 q^2} . \tag{S.80}$$

If we assume that $R \approx 2\pi\sigma$ and we require that the loss in the resolution from having a finite pixel size should not be greater than 10%, the this leads to the requirement that

$$\Delta R = 2\pi(\sigma_{\text{eff}} - \sigma) < 0.1R = 0.1 \cdot 2\pi\sigma, \tag{S.81}$$

32

which implies that

$$4\pi^2\sigma_{\mathrm{eff}}^2 q_{\mathrm{res}}^2 - 4\pi^2\sigma^2 q_{\mathrm{res}}^2 = -\log\left(\mathrm{sinc}\left(\pi q_{\mathrm{res}}l\right)^2\right) < 0.21, \tag{S.82}$$

from which it can be deduced that

$$l < \frac{0.25}{q_{\mathrm{res}}} = \frac{R}{4}. \tag{S.83}$$

The recommendation is therefore to keep the superresolution pixel size smaller than $R/4$.

5.3 Data visualization

The images used for resolution estimation are computed by binning localizations. For the experimental results presented in this article, however, the binned images were Gaussian blurred to make them visually more appealing for the various figures. One might wonder whether using these blurred images would lead to a different FRC resolution estimate. It turns out that this is not the case. The blurring does suppress high spatial frequencies in the images. The FRC though is only sensitive to the correlation between these components in the half images and not to the magnitudes of these components. Therefore the blurring does not influence the resolution. This was tested for the data in Fig. 2a by blurring the binned half images and calculating the FRC with these, but the typical difference in the FRC values was on the order of 10^{-4}; this is negligible compared to the variance in the FRC values. Note that the above implies that other isotropic linear filtering approaches than Gaussian blurring would also not have influenced the resolution, provided that the frequency response of the filter is nonzero for all spatial frequencies.

The influence on resolution estimation of non-linear visualization methods, such as adaptive bin size histograms or methods based on Delauney triangulation [57], is more difficult to analyze. Two general points can be made. First, the density of localizations is proportional to the number of labeled sites on a structure. Therefore, unbiased visualization methods should be linear in the density of localizations. Second, visualization methods cannot distinguish between variations in local localization density due to the underlying sample structure and variations due to noise. Consequently they cannot *both* be unbiased and improve the FRC between two half data image and can therefore not provide a more accurate resolution estimate than the binning method used here. It may be concluded that the non-linear visualization methods that have been proposed in the literature [57, 58] do not increase the useful information content of the image.

5.4 Thresholds

In the field of single-particle electron microscopy, there is no general consensus on what threshold $\theta\left(q\right)$ should be used for the FRC or FSC. Three main kinds of thresholds are used: The first is fixed thresholds (e.g. $\theta\left(q\right) = 1/7 \approx 0.143$ [20] or $\theta\left(q\right) = 0.5$ [18, 19]). The second kind is sigma factor curves ($\theta\left(q\right) \propto 1/\sqrt{qL}$ [15, 59]), which require the FRC to be larger than some multiple of the standard deviation of the FRC for white noise (for which $\langle FRC\rangle = 0$), with L the field-of-view. The third kind is information level curves[60]. The curves were derived based on the RMS value of the numerator and denominator of the FRC curve which leads to the following approximation:

$$FRC\left(q\right) \approx \frac{SNR\left(q\right) + \left(2/SNR\left(q\right)+1\right)/\sqrt{qL}}{SNR\left(q\right) + \left(2/SNR\left(q\right)\right)/\sqrt{qL}+1}. \tag{S.84}$$

The requirement that the information content per pixel $\log_2\left(1+SNR\right)$ is larger than a certain number of bits (e.g. 0.5 or 1) results in the desired curves. Note here that for 0.5 bits of information per pixel and large L, the threshold rapidly converges to $\theta\left(q\right) = 0.1716$, which is close to the $1/7$ threshold. The sigma factor and information level curves are conceptually conservative: they are chosen such that even if due to noise $FRC\left(q\right)$ is larger than its expected value, the image should still be resolved if it is above the threshold. However, since the FRC gives a single resolution figure for an entire image, it should be seen as giving the length scale at which details are resolved *on average* rather than *with certainty*. Moreover, these curves do not take into account that the noise on $FRC\left(q\right)$ is heavily suppressed by means of smoothing. Therefore these conservative thresholds are inappropriate. The most commonly used fixed threshold is 0.5 [61]. However, recent work [37] suggests that this threshold appears to give a realistic resolution estimate because the FSC in single-particle electron microscopy is often overoptimistic due to spurious correlations. In the absence of spurious correlations the threshold of $1/7 \approx 0.143$ was found to be more adequate. Ultimately though, the correct

threshold is determined empirically: it is the one that corresponds best with what intuitively appears to be resolved in actual images. From the results in Supplementary Fig. 1 we conclude that the 1/7 threshold is the most appropriate choice of threshold.

Supplementary Note 6: Supplementary simulations

6.1 Qualitative validity in simulations

The qualitative validity of the FRC resolution method in simulations was verified for the object shown in Supplementary Fig. 16a. Simulations with a single localizations per emitter (i.e. with $Q \ll 1$) were done for various values of the localization uncertainty σ and the density of localized labels ρ on the lines (i.e. the pixel values of the white parts of the 'ground truth' image were equal to ρ). In the simulations, the superresolution pixel size of 5 nm and the expected photon count per localization $n_{\mathrm{ph}} = 500$. Note that the empty areas in the images of the simulation results were removed for display. The Nyquist resolution was taken to be $R_{\mathrm{Nyquist}} = 2/\sqrt{\rho}$. The FRC resolution value could not be calculated from the simulation data since the field of view and number of localizations are too small for accurate resolution determination. Instead, the resolution was calculated from the expected FRC curve in Eq. (2). The results are shown in Supplementary Fig. 2. From this figure it becomes clear that the FRC resolution value does indeed show for which separation the two crossing lines can just be separated.

Simulations with multiple localizations per emitter were set up in the same way as above but with a few changes: the density of labels was constant at $\rho = 4 \cdot 10^4 \ \mu m^{-2}$ (N.B. this is the density of all labels, both localized and not localized). However, the average number of times Q that emitters (including unlocalized ones) were localized in the simulations was equal to 1/4, 1 and 4. The FRC curve that was used to compute the resolution was the theoretical FRC from Eq. (S.28) instead of the one calculated from the simulated data, since the number of localizations was too small to accurately determine the resolution from the data. Note that Eq. (S.28) gives the expected resolution when the Q-correction for spurious correlations is perfect. The Nyquist resolution was calculated as $R_{\mathrm{Nyquist}} = 2/\sqrt{Q\rho}$. The results are shown in Supplementary Fig. 2. From this figure it is clear that also in the case where emitters are localized multiple times, the FRC method with perfect Q-correction provides a resolution that seems to correspond qualitatively well with the distance for which the crossing lines can just be distinguished. This supports the qualitative validity of the FRC number.

Another test of the validity of the simulation method is in the sample dependence of the image resolution. Eq. (2) suggests that the image resolution depends on the sample itself. This insight was not originally considered for classical resolution limits such as the Abbe resolution. Simulations were carried out for pairs of lines where the spacing between the lines is kept constant but the width of the lines is varied. The localization precision σ was equal to 20 nm. The results of these simulations are shown in Fig. S.19. In Fig. S.19a, the two line pairs have the same localization density $\rho = 10^4 \ \mu m^{-2}$ and spacing $d = 30$ nm, but the width of the lines $w = 10$ nm to the left and $w = 30$ nm to right. In Fig. S.19b, the two line pairs have $N = 100$ localizations per line and spacing $d = 20$ nm, but the width of the lines $w = 20$ nm to the left and $w = 100$ nm to right. In both cases, it is evident that the two line pairs are not equally easy to distinguish, even though (a) σ and ρ or (b) σ and N are kept constant. This implies that the resolution must depend on the width of these lines, and therefore on the spatial structure of the object.

An extreme example of the sample dependence is shown in figure S.20, where the result of a simulation of an object without structure is shown (i.e. an object with constant expected labeling density). The resulting FRC curve also indicates that no structure can be interpreted in the image. This means that no interpretable details can be distinguished in the image, which is reasonable since no details are present in the structure of the object itself.

6.2 Supplementary Figs. 8 and 9: Correction for spurious correlations

In these simulations, the algorithm for correcting spurious correlations was applied 100 times for various values of the density of labels ρ, the localization uncertainty σ (which was the same for all localizations in a single simulated experiment) and Q for the Siemens star object in Fig. 1a. The object image was 4096×4096 pixels with a pixel size of 2.5 nm, whereas the localizations were binned into pixels of 10 nm in size. This was done in order to make sure that effects of pixelation in the object did not influence the numerator of the FRC curve.

A number of relevant experimental effects is not taken into account in these simulations: imperfect correction for drift and the heterogeneity of the switching kinetics leading to e.g. variations in the distribution of localization uncertainties and in the number of switching cycles all influence the drop-off of the spurious correlation term in
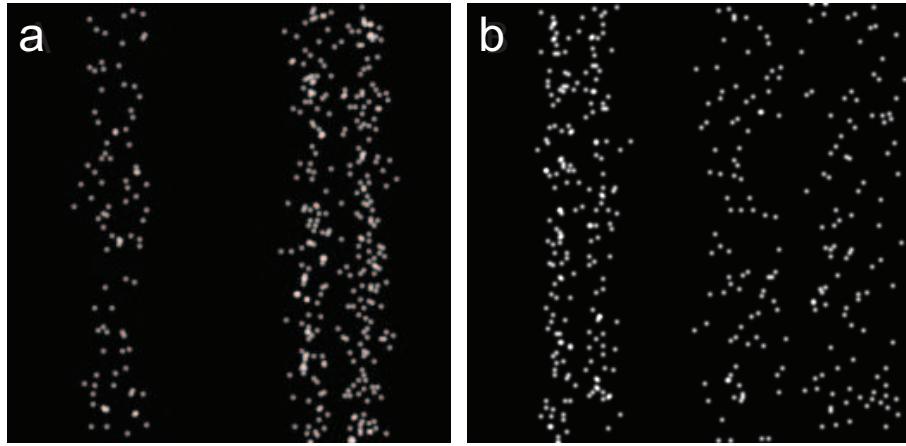
Figure S.19: Simulations without multiple localizations per emitter of line pairs with $\sigma = 20$ nm. In **a**, the localization density $\rho = 10^4 \ \mu m^{-2}$ and the spacing $d = 30$ nm for both line pairs, but the width $w = 10$ nm to the left and $w = 30$ nm to right. In **b**, all lines contain 100 localizations and the spacing $d = 20$ nm for both line pairs, but the width $w = 20$ nm to the left and $w = 100$ nm to right.
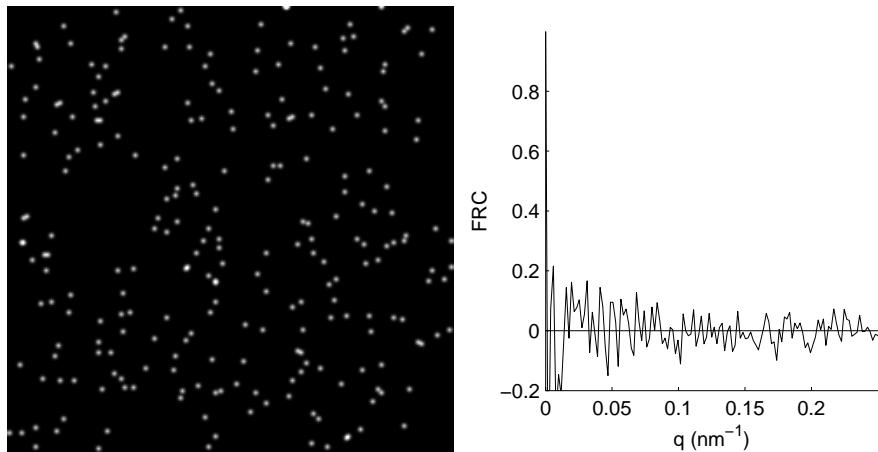

Figure S.20: Simulation without multiple localizations per emitter with constant density of localized labels $\rho = 10^3 \ \mu m^{-2}$ everywhere and $\sigma = 20$ nm. The FRC curve drops immediately below 0. The resolution is therefore larger than half the size of the field of view which is 512 nm (256 pixels of 2 nm). Therefore it more or less indicates that no structural information can be interpreted in the image.

36

Eq. (S.26) and may therefore have an influence on the performance of the algorithm. Also, the noise floor on the numerator of the FRC is inversely proportional to the size of the field of view. Therefore the algorithm will perform better for larger fields of view, all other things being equal. Finally it should be noted that the Siemens star object in these simulations has a rather typical shape: about 97% of the emitters reside in parts of the object where the arms are wider than 100 nm. Therefore the high frequency content of the object itself is relatively low. Consequently, the resolution values in these figures are not representative of the values that will generally be obtained in experiments with similar values of ρ, σ, and Q.

6.3 Supplementary Fig. 15: Illustration of FLC

The FLC-concept is illustrated in Supplementary Fig. 15. The figure shows a simulation of a grating structure proportional to $\sin^2(\pi x/p)$ in a 1.285 μm field of view with 25 grating periods (i.e. $p = 51.4$ nm). In the simulations, the average density of localizations $\rho = 10^4\ \mu\text{m}^{-2}$ and the localization uncertainty $\sigma = 5$ nm. The FLC is calculated by splitting the localizations randomly into two half data images. The FLC shows a distinctive ∞-symbol shape which can be described by $q = |\cos\phi|/p$, in correspondence with the apparent pitch $p/|\cos\phi|$ in azimuthal direction ϕ. Note that the FLC-plot always has point symmetry due to the real valuedness of the image data. It appears that the noise level at high spatial frequencies takes values up to 0.1 (root-mean-square value) at the edges of the Fourier domain. The threshold level $1/7 = 0.14$ used for the FRC or FSC may therefore not be directly appropriate for the FLC or FPC measures. The cause for this may be the noise averaging which is less complete for lines or planes normal to spatial frequency vectors with large magnitude compared to circles or spheres with that particular radius. A sensible remedy may be to smooth the evaluated FLC or FPC (e.g. through Gaussian filtering) before thresholding, since smoothing was also applied to the FRC curves. Another remedy may be to introduce a weighted average over the lines or planes in Fourier space such that the contribution of noise in the regions of high spatial frequency is suppressed. As weight we would envision a Gaussian type weight with center at q that should account for the differences in SNR of the Fourier pixels over the line or plane. Another reason why the 1/7 threshold level may be inappropriate is that the signal-to-noise-ratio in the Fourier pixels varies along the lines or planes in Fourier space, which also suggests the use of weighted averages.

Supplementary Fig. 15 shows that the FLC gives an indication of image resolution that corresponds to the anisotropy of the sample. However, other ways to extend the average over rings can be envisioned, e.g. using segments of rings/cones or Gabor wavelets. These other approaches can localize strongly in Fourier space, which would yield a more or less two point response to the simulated grating of Fig. 15. We argue that the resolution in a observed direction of e.g. $\phi = 45°$ to the grating normal is not zero (as would be assessed by the wavelets or ring segments) but just less than perpendicular to it by $\cos(45°) = 1/\sqrt{2}$ as the effective pitch distance increases by $1/\cos(\phi) = \sqrt{2}$. The FLC or FPC gives this desired behavior, as opposed to the approach relying on wavelets or ring segments.

37

Supplementary Note 7: Supplementary experimental setup and methods

7.1 Supplementary Fig. 5: Tubulin filaments in a HeLa cell

Supplementary Fig. 5 shows another example of an acquisition of tubulin filaments in a human epithelial cervical cancer (HeLa) cell. In this case, both anti-alpha tubulin and anti-beta tubulin antibodies were used. The instrumental setup and the localization and image rendering steps were the same as before.

In the sample preparation protocol for this second acquisition, the fixing step was the same as before. However, the fixing was quenched using 50 mM Tris(hydroxymethyl)amino methane (Tris-Cl). Afterwards the cells were washed thrice in PBS and permeabilised (0.5%v/v Triton X-100 with 2% Bovine Serum Albumin (BSA) in PBS) for 5 minutes, washed twice with PBS and put in 2% BSA for 20 minutes to reduce non-specific binding. Then the cells was were washed three times in PBS again and subsequently they were labeled with Alexa Fluor 647 molecules conjugated to anti-beta tubulin antibodies (9F3 rabbit monoclonal, Cell Signaling Technology Inc.) in PBS at a concentration of approximately 1.0 μg/ml, after which they were washed thrice in PBS and put in 2% PFA to fix the antibodies in place. Next the cells were washed once with Tris-Cl and twice with PBS and put for 1 hour in a 1:350 dilution of anti-alpha tubulin acetylated antibodies (Sigma-Aldrich Corp.) in PBS. Then they were washed thrice in PBS and labeled in a 1:100 dilution of 2 mg/ml donkey anti-mouse IgG antibodies conjugated to Alexa Fluor 647 dye molecules (Invitrogen Corp.) in PBS. Finally the cells were washed thrice in PBS, put in 2% PFA again, washed once with Tris-Cl again and twice with PBS.

7.2 Supplementary Fig. 6: Actin filaments in an RBL cell

Supplementary Fig. 6 shows another additional experimental result of a localization microscopy data acquisition of actin filaments in Rat Basophil Leukemia (RBL) cells.

RBL-2H3 cells were plated on 18 mm cover slips overnight in standard DMEM phenol free media. Cells were rinsed once with warm PBS and then fixed in 0.3% PFA, 0.25% triton 5 mM glucose in cytoskeleton buffer (10 mM MES, 138 mM KCl, 3 mM MgCl, 2 mM EGTA, 0.32 M sucrose) for 1 minute and then with 4% PFA, 5 mM glucose in cytoskeleton buffer for 10 minutes. Cells were rinsed 4 times for 10 minutes with PBS and then labeled for 2 hours with 250 mM phalloidin-Alexa Fluor 647 (Invitrogen Inc.) in PBS. Immediately before imaging, cells were rinsed 2 times with PBS for two minutes and then mounted on a slide with imaging buffer (10% (w/v) glucose, 50 mM Tris, 10 mM NaCl, pH 8.5, glucose oxidase, catalyze, and 50 mM MEA) and sealed using nail polish. Single molecule imaging experiments were performed using an inverted microscope (IX71, Olympus) with a 1.45 NA TIRF objective (U-APO 150x, Olympus America Inc.). A 635 nm diode laser (Radius 635, Coherent) provided excitation and the emission path included a 685/40 emission filter (Semrock) and an electron multiplying CCD camera (iXon 897, Andor Technologies) Data was collected at 46 frames per second.

The localization and image rendering algorithms that were used were the same as above, except for the values used for filtering the localizations because of the different sample and dye used. The minimum value for n_{ph} was set to 50, the maximum value for σ to 54 nm, all values of σ_{PSF} were accepted and a maximum number of 50 estimated background photons per pixel was used for filtering localizations.

7.3 Supplementary Fig. 7: Keratin filaments in a PA-JEB cell

Supplementary Fig. 7 shows a localization microscopy image of keratin filaments in immortalized β4-deficient keratinocytes derived from a patient with pyloric atresia associated with junctional epidermolysis bullosa (PA-JEB) [62, 63]. PA-JEB cells stably expressing β4 integrin were cultured on #1.5 coverslips. After 48 hours cells were washed briefly with PBS and fixed with 4%PFA in PBS for 10 minutes at room temperature. Samples were extensively washed with PBS and blocked with 5% BSA for 30 minutes at room temperature. Immunolabelling was performed at room temperature using first a polyclonal antibody against Keratin 14 (Covance, Catalog Number PRB-155P) diluted in 5%BSA in PBS to a final concentration of 1.25 μg/ml and later a Goat Anti-Rabbit IgG antibody coupled to Alexa

Fluor 532 (Invitrogen) diluted in 5%BSA in PBS to a final concentration of 0.2 mg/ml. Cells were washed using first 0.1% Tween-20 in PBS and then PBS

The same instrumental setup as described in sec. was used for this acquisition, except for the EM gain of 16.8 that was used for this acquisition instead of 50.6. The localization and image rendering algorithms that were used were also the same, except for the values used for filtering the localizations because of the different dye used. The minimum value for n_{ph} was set to 100, and the accepted values of σ_{PSF} ranged from 89 nm to 133 nm.

7.4 Supplementary Fig. 13: Widefield imaging

Red fluorescent beads (Duke Scientific Corp.,) with a diameter of 200 nm (excitation maxima 542 nm, emission maxima 612 nm) were diluted and put onto a coverslip. Images were recorded on a ORCA-ER CCD camera (Hamamatsu Photonics) mounted to an inverted microscope (IX71 Olympus). The sample was illuminated with a Hg lamp. The filter set consisted of a 560/20 excitation filter, 585 nm DRLP dichroic mirror and 630/40 emission filter (all Chroma Technology). The pixel size of the camera is 6.45 μm and for imaging we used 10x 0.25 NA (acquisition series ids 1-3) and 60x 0.7 NA (acquisition ids 4-11) air objectives. The exposure time was between 0.2 and 0.5 s. For each field-of-view a series of 20 consecutive images was recorded. A second series of images with identical acquisition parameters was taken with no light incident on the camera. This background estimation was averaged over time, then subtracted from each image of the data series on a pixel by pixel basis. The displayed FRC curves and resolution values are computed from the average of 19 FRC curves from the 20 subsequent images of the same scene. The standard deviation of the values is below 1% within one image series.

7.5 Supplementary Fig. 14: SEM imaging

The resolution in SEM depends on many different aspects, such as the electron beam size, its aberrations [64, 65], and very importantly, the beam-sample interaction. The cross-correlation between images in a time series of the same scene, in particular its decay as a function of spatial frequency has been proposed as a resolution measure for SEM [66]. We have applied the FRC resolution concept to SEM time series in which only single electrons are recorded due to the very short pixel dwell time and the low probe current. The FRC method should also be applicable to images acquired with the newest generation Transmission Electron Microscopy (TEM) cameras that can operate in single electron counting mode.

Images of gold on carbon were acquired on a FEI Quanta FEG SEM (FEI Company) at 5 kV acceleration voltage in high vacuum mode. We recorded time series of 128 images each of size 512×442 pixels with an effective pixel size of 1.7 nm or 0.83 nm. The pixel dwell time was set to 50 ns and single secondary electrons were recorded with an Everhart-Thornley detector. Due to the very small pixel dwell time and the comparably slow detector response, single events show up as 1-3 pixel wide streaks along the scan line. These events were replaced with one single peak of unit height at the maximum pixel response position.

7.6 Additional analyses of the data in Fig. 3

7.6.1 Checks of the spurious correlation correction in Fig. 3

An estimate of the localization uncertainty for the data in Fig. 3 was obtained by analyzing the distribution of localizations of several bright, isolated clusters of localizations, which may be interpreted as originating from the same binding site. The histograms of the found localizations for Alexa Fluor 647 and Alexa Fluor 750 are compiled from the data of 56 and 61 hand-picked clusters, respectively, and fitted with a 2D-Gaussian in order to determine the localization uncertainty. The fit points to localization uncertainties equal to 7.7 nm for Alexa Fluor 647 and 11.9 nm for Alexa Fluor 750 (see histograms in Fig. S.21), in good agreement with the spurious correlation analysis (localization uncertainties 9.2 nm for Alexa Fluor 647 and 12 nm for Alexa Fluor 750). The histograms of the number of localizations per cluster is shown in Fig. S.21, giving rise to a mean number of localizations per cluster equal to 22±9 for Alexa Fluor 647 and 31±22 for Alexa Fluor 750, which is a factor of two higher than the values for the Q-parameter estimated from the spurious correlations (10 for Alexa Fluor 647 and 18 for Alexa Fluor 750). The reason for this discrepancy is that there is a substantial bias towards higher numbers of localizations per cluster, as only relatively

39
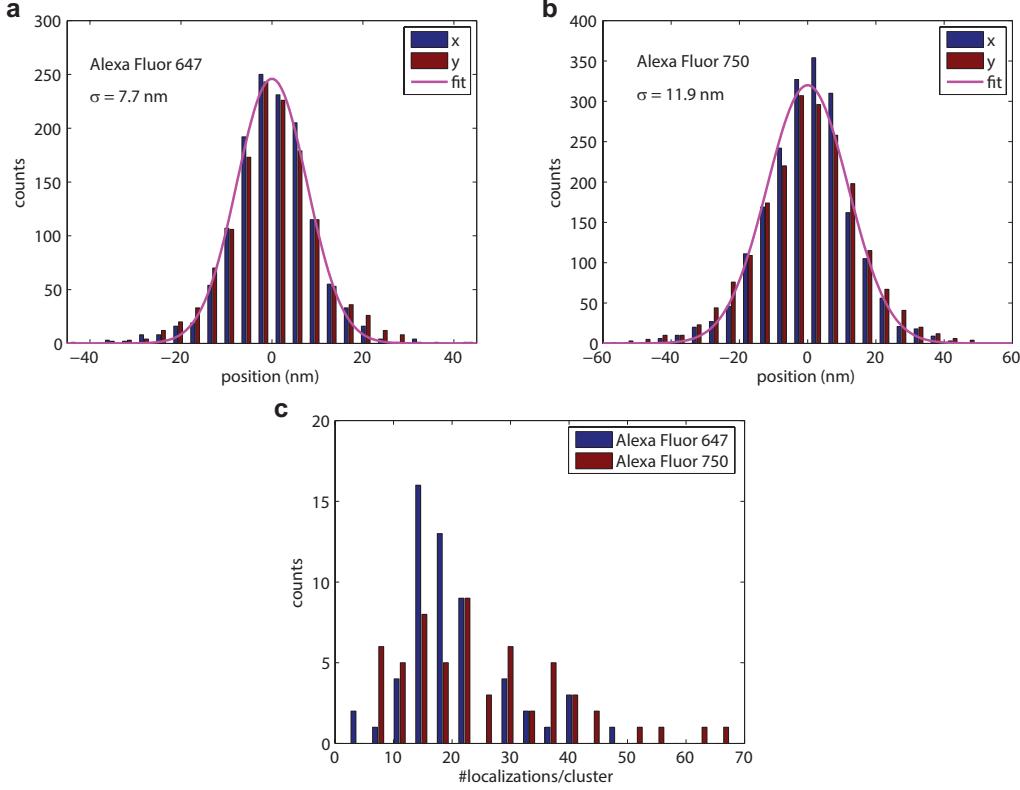
Figure S.21: **Cluster analysis for the data in Figure 3**. Histograms of found x and y positions of localizations in several bright, isolated clusters of localizations in Fig 3 for Alexa Fluor 647 (**a**) and Alexa Fluor 750 (**b**) and Gaussian fits to the 2D-histograms, as well as the histogram of the number of localizations per cluster for both species (**c**). The estimated localization uncertainties, 7.7 nm for Alexa Fluor 647 and 11.9 nm for Alexa Fluor 750, are in good agreement with those used in the spurious correlation analysis for Fig. 3 (respectively 9.2 nm and 12 nm). Also the ratio between the number of Alexa Fluor 647 localizations and Alexa Fluor 750 localizations per cluster of 1.4 agrees well with the ratio of 1.8 between the Q-values in Fig. 3.

bright clusters are picked for the analysis. However, the ratio of the mean number of localizations per cluster for Alexa Fluor 647 and Alexa Fluor 750 (equal to 1.4) does fit reasonably with the ratio of the Q-parameters estimated for both species (equal to 1.8).

The values for Q are also consistent with the uncorrected resolution values. It follows from Eq. (2) that the FRC resolution without spurious correlation correction is equal to $2\pi\sigma/\sqrt{\log(3Q)}$ in the regime where spurious correlations dominate the FRC curve (Supplementary Note 1). Using this formula and the found values for σ and Q we obtain 31 nm for Alexa Fluor 647 and 37 nm for Alexa Fluor 750, close to the uncorrected resolution values of 25 nm and 34 nm respectively.

An additional check that was made to confirm the validity of the correction for spurious correlations involved a comparison between the corrected FRC curves and cross-channel FRC curves. This comparison involves a theoretical insight obtained in the context of colocalization analysis. There it was shown that if two independent noisy images A_1 and A_2 are available for color channel A and two independent noisy images B_1 and B_2 for color channel B and all images show the same structure, then:

$$\frac{1}{4}\left(r(A_1, B_1) + r(A_1, B_2) + r(A_2, B_1) + r(A_2, B_2)\right) \approx \sqrt{r(A_1, A_2)r(B_1, B_2)}, \qquad \text{(S.85)}$$

where $r(.,.)$ denotes the cross-correlation between two zero-mean images [67]. Since the sum of all the pixel

40
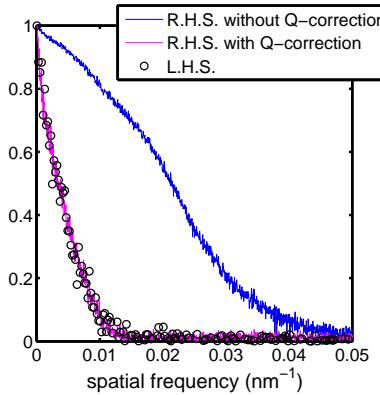
Figure S.22: **Validation of Q-correction**. Plots of the left-hand side and right-hand side of Eq. (S.86) for the data of Fig. 3. The right-hand side is computed with the uncorrected FRC-curves from Fig. 3c and the corrected FRC-curves from Fig. 3d; the images used to compute those FRC curves were also used to compute the left-hand side of Eq. (S.86). This figure suggests that Eq. (S.86) is correct after correction for spurious correlation, thus indicating that the correction was successful.

values over a ring in Fourier space of an image is typically very small, the FRC can be considered to be the same kind of correlation as $r(.,.)$. Therefore it follows that if two images $f_{A647,1}$ and $f_{A647,2}$ are obtained from the Alexa Fluor 647 localizations and two images $f_{A750,1}$ and $f_{A750,2}$ from the Alexa Fluor 750 localizations then:

$$\frac{1}{4}\sum_{i,j=1,2} FRC\left(\hat{f}_{A647,i}\left(\vec{q}\right),\hat{f}_{A750,j}\left(\vec{q}\right)\right) \approx \sqrt{FRC\left(\hat{f}_{A647,1}\left(\vec{q}\right),\hat{f}_{A647,2}\left(\vec{q}\right)\right)FRC\left(\hat{f}_{A750,1}\left(\vec{q}\right),\hat{f}_{A750,2}\left(\vec{q}\right)\right)}.$$
(S.86)

Fig. S.22 shows that if the Alexa Fluor 647 and Alexa Fluor 750 datasets are split into half and $FRC(\hat{f}_{A647,1},\hat{f}_{A647,2})$ and $FRC(\hat{f}_{A750,1},\hat{f}_{A750,2})$ are corrected for spurious correlations, then the left-hand side is approximately equivalent to the right-hand side of Eq. (S.86). This implies that the corrected FRC curves give rise to the same equivalence in Eq. (S.86) as when FRC curves were obtained with independent images in both color channels, and therefore that the correction for spurious correlations was successful.

7.6.2 Calibration sample for emitter estimation

In order to test how well the spurious correlation correction method performs in estimating the number of localizations per emitter, calibrations experiment were carried out with fluorescent dyes on a glass surface.

For the first experiment, streptavidin molecules were labeled with a combination of two amine-reative fluorescent dyes: Cy3 and Alexa Fluor 647. The average degree of labeling for each streptavidin molecule was 3.1 Cy3 and 0.6 Alexa Fluor 647 fluorophores per streptavidin. The labeled streptavidin was bound to a glass surface coated with biotinylated BSA, at a concentration low enough such that single molecules of fluorescent streptavidin were distinguishable on the surface when viewed through the microscope. By repetitive cycling of a weak activation light source (a 532 nm laser) and an excitation light source (a 642 nm laser), the Alexa Fluor 647 fluorophores were switched between the fluorescent (on) state and the dark (off) state. The cycling continued until all of the molecules in the field of view had photobleached. This dataset was analyzed by a conventional localization microscopy analysis, resulting in an image where each molecule appears as a small cluster of localizations. The number of localizations per cluster corresponds to the number of times each molecule switched on before bleaching.

In the second experiment, streptavidin was labeled with Cy3 and Alexa Fluor 750, with an average degree of labeling of 2.0 Cy3 and 0.07 Alexa Fluor 750 per streptavidin. The rest of this second experiment proceeded analogous to the first, but with a 752 nm laser instead of a 642 nm laser as an excitation light source.

41

To count the number of localizations per cluster, these clusters first had to be identified. This was done by binning localizations to make superresolution images of the localizations from the first 5,000 frames and searching for connected regions in those images with at least 4 nonzero superresolution pixels. Pixels were considered to be connected if at least 1 of their 8 neighbors had a nonzero value. The reason for taking only the first 5,000 frames was to limit the number of false clusters from false positive localizations. Subsequently regions arising from multiple emitters were filtered out based on their ellipticity by computing the largest eigenvalue of the inertia tensor of the distribution of localizations in each region. The average number of localizations in the previously identified clusters was then computed for clusters containing at least one localization for various fractions of the total acquisition time. Finally, the Q-parameter was estimated on the segmented localizations for the same fractions of the total acquisition time, using a mean and width of the distribution of localization uncertainties equal to 5.0 and 4.0 nm.

Supplementary Fig. 10 shows a comparison between the counted number of localizations and the estimated value of the Q-parameter for the first 5,000 frames. From this it becomes clear that the Q-parameter provides an accurate estimate of the number of localizations per emitter when few emitters have bleached. When more emitters have bleached, the Q-parameter increasingly overestimates the number of localizations per emitter, even though it does correspond well to its expected value based on Eq. (S.20). The amount of overestimation was calibrated by considering the extent of photobleaching of the emitter population. The number of localizations per emitter over time is proportional to $1 - \exp(-t/\tau)$ for some decay time τ which is estimated from the data (3.4×10^3 frames in Supplementary Fig. 10). Since the laser power was different in the control experiments and the experiments for Fig. 3, the ratio t/τ was used to quantify the extent of photobleaching in a way that is invariant to these differences in the illumination. Fig. 10 shows the overestimation of the number of localizations per emitter by the Q-parameter as a function of t/τ for the Alexa Fluor 647 and the Alexa Fluor 750 dataset separately.

When the data from Fig. 3 are considered, it turns out that the cumulative number of localizations over time can also be fitted with a model of the form $A(1 - \exp(-t/\tau))$ for $\tau = 1.8 \times 10^5$ frames for the Alexa Fluor 647 and $\tau = 3.6 \times 10^4$ frames for the Alexa Fluor 750. Based on the acquisition times, it then follows that the Q-parameter overestimates the number of localizations per emitter by a factor 1.5 for the Alexa Fluor 647 and 1.7 for the Alexa Fluor 750.

[1] Hell, S. W. & Wichmann, J. Breaking the diffraction limit resolution by stimulated emission: stimulated-emission-depletion microscopy. *Opt. Lett.* **19**, 780–783 (1994).

[2] Hofmann, M., Eggeling, C., Jakobs, S. & Hell, S. W. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17565–17569 (2005).

[3] Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1643–1645 (2006).

[4] Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Meth.* **3**, 793–795 (2006).

[5] Fölling, J. *et al.* Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nat. Meth.* **5**, 943–945 (2008).

[6] Heilemann, M. *et al.* Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Ange. Chemie* **47**, 6172–6176 (2008).

[7] Lidke, K. A., Rieger, B., Jovin, T. M. & Heintzmann, R. Superresolution by localization of quantum dots using blinking statistics. *Opt. Express* **13**, 7052–7062 (2005).

[8] Dertinger, T., Colyer, R., Iyer, G., Weiss, S. & Enderlein, J. Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI). *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22287–22292 (2009).

[9] Ram, S., Ward, E. S. & Ober, R. J. Beyond Rayleighs criterion: A resolution measure with application to single-molecule microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4457–4462 (2006).

[10] Löschberger, A. *et al.* Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution. *J. Cell Sci.* **125**, 571–575 (2012).

[11] Kanchanawong, P. *et al.* Nanoscale architecture of integrin-based cell adhesions. *Nature* **468**, 580–584 (2010).

[12] van de Linde, S., Wolter, S., Heilemann, M. & Sauer, M. The effect of photoswitching kinetics and labeling densities on super-resolution fluorescence imaging. *J. Biotech.* **149**, 260–266 (2010).

[13] Cordes, T. *et al.* Resolving single-molecule assembled patterns with superresolution blink-microscopy. *Nano Lett.* **10**, 645–651 (2010).

[14] Fitzgerald, J. E., Lu, J. & Schnitzer, M. J. Estimation theoretic measure of resolution for stochastic localization microscopy. *Phys. Rev. Lett.* **109**, 048102 (2012).

[15] Saxton, W. O. & Baumeister, W. The correlation averaging of a regularly arranged bacterial cell envelope protein. *J. Micros.* **127**, 127–138 (1982).

[16] van Heel, M. Similarity measures between images. *Ultramicros.* **21**, 95–100 (1987).

[17] Unser, M., Trus, B. L. & Steven, A. C. A new resolution criterion based on spectral signal-to-noise ratio. *Ultramicros.* **23**, 39–52 (1987).

[18] Beckmann, R. *et al.* Alignment of conduits for the nascent polypeptide chain in the ribosome-sec61 complex. *Science* **278**, 213–2126 (1997).

[19] Böttcher, B., Wynne, S. A. & Crowther, R. A. Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature* **386**, 88–91 (1997).

[20] Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Bio.* **333**, 721–745 (2003).

[21] Barry, D. A. *et al.* Analytical approximations for real values of the Lambert W-function. *Math. Comput. Simul.* **53**, 95–103 (2000).

[22] Small, A. R. Theoretical limits on errors and acquisition rates in localizing switchable fluorophores. *Biophys. J.* **92**, L16–L18 (2009).

[23] Wolter, S. *et al.* rapidSTORM: accurate, fast open-source software for localization microscopy. *Nat. Meth.* **9**, 1040–1041 (2012).

[24] Smith, C. S., Joseph, N., Rieger, B. & Lidke, K. A. Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nat. Meth.* **7**, 373–375 (2010).

[25] Mlodzianoski, M. J. *et al.* Sample drift correction in 3D fluorescence photoactivation localization microscopy. *Opt. Express* **19**, 15009–15019 (2011).

[26] Bates, M., Dempsey, G. T., Chen, K. H. & Zhuang, X. Multicolor super-resolution fluorescence imaging via multi-parameter fluorophore detection. *ChemPhysChem* **13**, 99–107 (2012).

[27] Lando, D. *et al.* Quantitative single-molecule microscopy reveals that CENP-ACnp1 deposition occurs during G2 in fission yeast. *Open Bio.* **2**, 120078 (2012).

[28] Sengupta, P. *et al.* Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis. *Nat. Meth.* **8**, 969–975 (2011).

[29] Veatch, S. *et al.* Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *Plos ONE* **7**, e31457 (2012).

[30] Annibale, P., Vanni, S., Scarelli, M., Rothlisberger, U. & Radenovic, A. Identification of clustering artifacts in photoactivated localization microscopy. *Nat. Meth.* **8**, 527–528 (2011).

[31] Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Meth.* **8**, 1027–1031 (2011).

[32] von Middendorff, C., Egner, A., Geisler, C., Hell, S. W. & Schönle, A. Isotropic 3D nanoscopy based on single emitter switching. *Opt. Express* **16**, 20774–20788 (2008).

[33] Toprak, E. *et al.* Defocused orientation and position imaging (DOPI) of myosin V. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6495–6499 (2006).

[34] Gustafsson, M. G. L. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J. Micros.* **198**, 82–87 (2000).

[35] Mukamel, E. & Schnitzer, M. Unifed resolution bounds for conventional and stochastic localization fluorescence microscopy. *Phys. Rev. Lett.* **109**, 168102 (2012).

[36] Hell, S. W. Towards fluorescence nanoscopy. *Nat. Biotech.* **21**, 1347–1355 (2003).

[37] Scheres, S. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Meth.* **9**, 853–854 (2012).

[38] Huang, F., Schwartz, S. L., Byars, J. M. & Lidke, K. A. Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. *Biomed. Opt. Express* **2**, 1377–1393 (2011).

[39] Holden, S. J., Uphoff, S. & Kapanidis, A. N. DAOSTORM: an algorithm for high-density super-resolution microscopy. *Nat. Meth.* **8**, 279–280 (2011).

[40] Zhu, L., Zhang, W., Elnatan, D. & Huang, B. Faster STORM using compressed sensing. *Nat. Meth.* **9**, 721–726 (2012).

[41] Tukey, J. W. *Analysis of time series*, chap. An introduction to the calculations of numerical spectrum analysis, 25–46 (Wiley, New York, 1967).

[42] Cleveland, W. Robust locally weighted regression and smoothing scatterplots. *J. Amer. Stat. Ass.* **74**, 829–836 (1979).

[43] Wolter, S., Endesfelder, U., Linde, S. v. d., Heilemann, M. & Sauer, M. Measuring localization performance of super-resolution algorithms on very active samples. *Opt. Express* **19**, 7020–7033 (2011).

[44] Xu, K., Babcock, H. & Zhuang, X. Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nat. Meth.* **9**, 185–190 (2012).

[45] Ghosh, R. & Webb, W. Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules. *Biophys. J.* **66**, 1301–1318 (1994).

[46] Lee, G., Ishihara, A. & Jacobson, K. Direct observation of brownian motion of lipids in a membrane. *Proc. Natl. Acad. Sci. USA* **88**, 6274–6278 (1991).

[47] Bates, M., Jones, S. & Zhuang, X. *Imaging: A Laboratory Manual*, chap. Stochastic Optical Reconstruction Microscopy (STORM): a method for super-resolution fluorescence imaging (Cold Spring Harbor Laboratory Press, 2011).

[48] Hanser, B., Gustafsson, M., Agard, D. & Sedat, J. Phase-retrieved pupil functions in wide-field fluorescence microscopy. *J. Micros.* **216**, 32–48 (2004).

[49] Stallinga, S. & Rieger, B. Accuracy of the Gaussian point spread function model in 2D localization microscopy. *Opt. Express* **18**, 24461–24476 (2010).

[50] Fitzhugh, R. Statistical properties of the asymmetric random telegraph signal, with applications to single-channel analysis. *Math. Biosci.* **64**, 75–89 (1983).

[51] Saxton, W. *Computer Techniques for Image Processing of Electron Microscopy* (Academic Press, New York, 1978).

[52] Corless, R., Gonnet, G., Hare, D., Jeffrey, D. & Knuth, D. On the lambert-W function. *Adv. Computat. Mathe.* **5**, 329–359 (1996).

[53] Berro, A., Berglund, A., Carmichael, P., Kim, J. & Liddle, J. Super-resolution optical measurement of nanoscale photoacid distribution in lithographic materials. *ACS Nano* **11**, 9496–9502 (2012).

[54] Vogelsang, J., Cordes, T., Forthmann, C., Steinhauer, C. & Tinnefeld, P. Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. *Proc. Natl. Acad. Sci. USA* **106**, 8107–8112 (2009).

[55] Donnert, G. *et al.* Macromolecular-scale resolution in biological fluorescence microscopy. *Proc. Natl. Acad. Sci. USA* **103**, 11440–11445 (2006).

[56] Oppenheim, A., Willsky, A. & Young, I. *Signals and Systems* (Prentice-Hall International, London, 1983).

[57] Baddeley, D., Cannell, M. & Soeller, C. Visualization of localization microscopy data. *Micros. Microana.* **16**, 64–72 (2010).

[58] Krizek, P., Raska, I. & Hagen, G. Minimizing detection errors in single molecule localization microscopy. *Opt. Express* **19**, 3226–3235 (2011).

[59] Orlova, E. *et al.* Structure of keyhole limpet hemocyanin type 1 (klh1) at 15 a resolution by electron cryomicroscopy and angular reconstitution. *J. Mole. Bio.* **271**, 417–437 (1997).

[60] Heel, M. v. & Schatz, M. Fourier shell correlation threshold criteria. *J. Struct. Bio.* **151**, 250–262 (2005).

[61] Liao, H. & Frank, J. Definition and estimation of resolution in single-particle reconstructions. *Structure* **18**, 768–775 (2010).

[62] Niessen, C. *et al.* Deficiency of the $\beta 4$ subunit in junctional epidermolysis bullosa with pyloric atresia: consequences for hemidesmosome formation and adhesion properties. *J. Cell Sci.* **109**, 1695–1706 (1996).

[63] Schaapveld, R. *et al.* Hemidesmosome formation is initiated by the $\beta 4$ integrin subunit, requires complex formation of $\beta 4$ and h.d1/plectin, and involves a direct interaction between $\beta 4$ and the bullous pemphigoid antigen. *J. Cell Bio.* **142**, 271–284 (1998).

[64] Barth, J. & Kruit, P. Addition of different contributions to the charged particle probe size. *Optik* **101**, 101–109 (1996).

[65] Lorusso, G. & Joy, D. Experimental resolution measurement in critical dimension scanning electron microscope metrology. *Scanning* **25**, 175–180 (2003).

[66] Frank, J. *Computer Processing of Electron Microscope Images*, chap. The Role of Correlation Techniques in Computer Image Processing, 197–233 (Springer Verlag, 1980).

[67] Bergholm, F., Adler, J. & Parmryd, I. Analysis of bias in the apparent correlation coefficient between image pairs corrupted by severe noise. *J. Mathe. Imag. Vis.* **37**, 204–219 (2010).