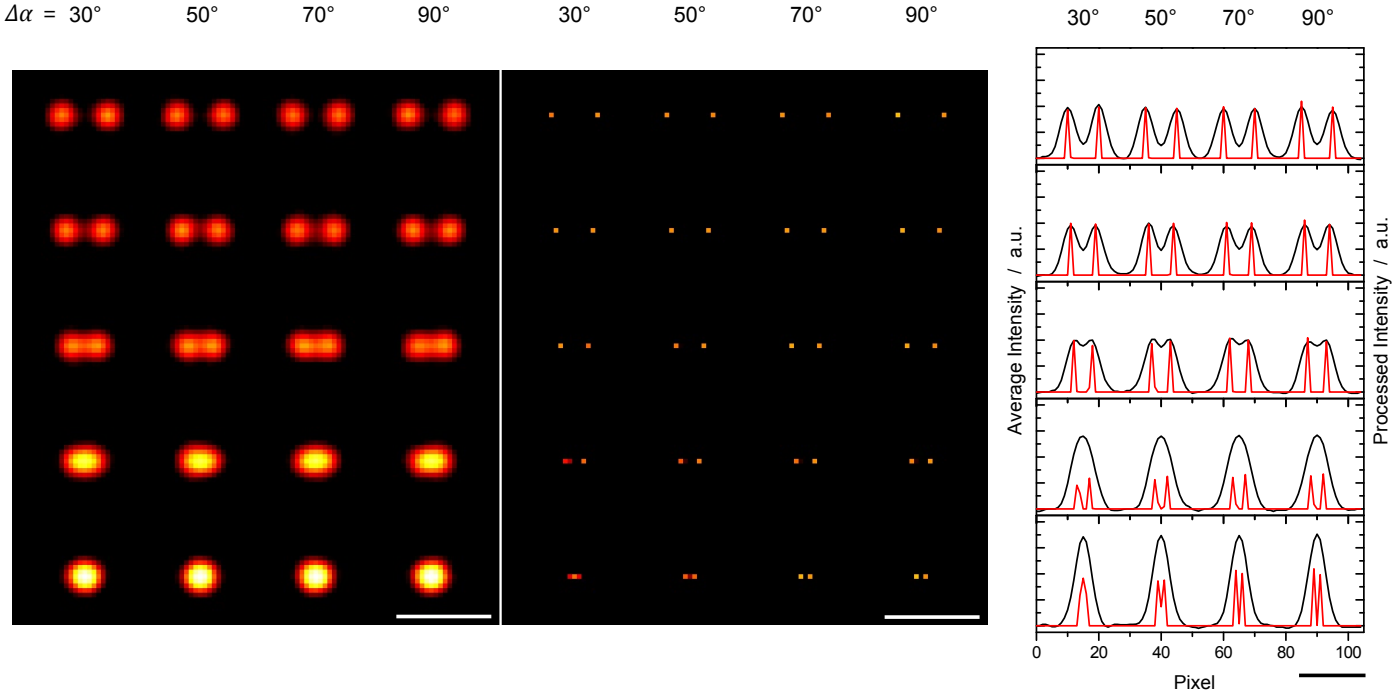
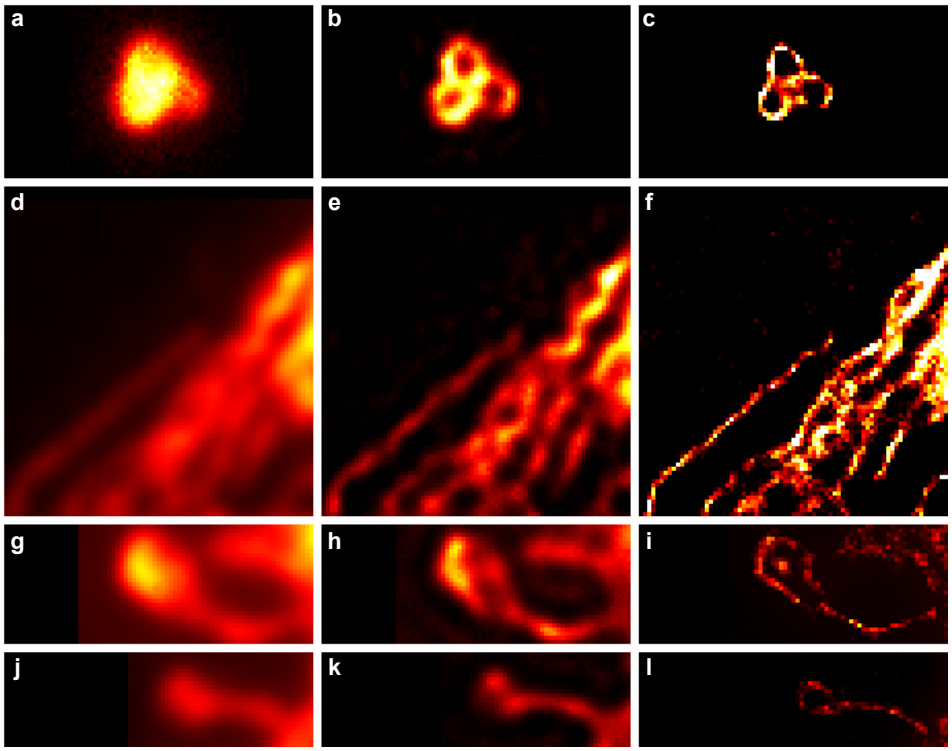
Supplementary Figure 1 | Simulations and high-resolution analysis of single-molecule pairs with different distances and orientations.



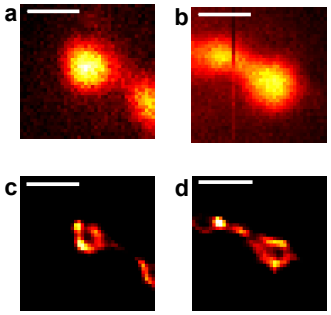
Average intensity images, high resolution images and intensity profiles of simulated signals of various molecule pairs with decreasing distance and increasing phase differences, $\Delta\alpha$, including modulation, shot noise and background noise. Scale bar corresponds to ~1 µm when assuming an emission wavelength of 620 nm and a numerical aperture of 1.3.

Supplementary Figure 2 | Comparison of diffraction limited fluorescence images with Richardson-Lucy deconvolution and SPoD-images from various samples.



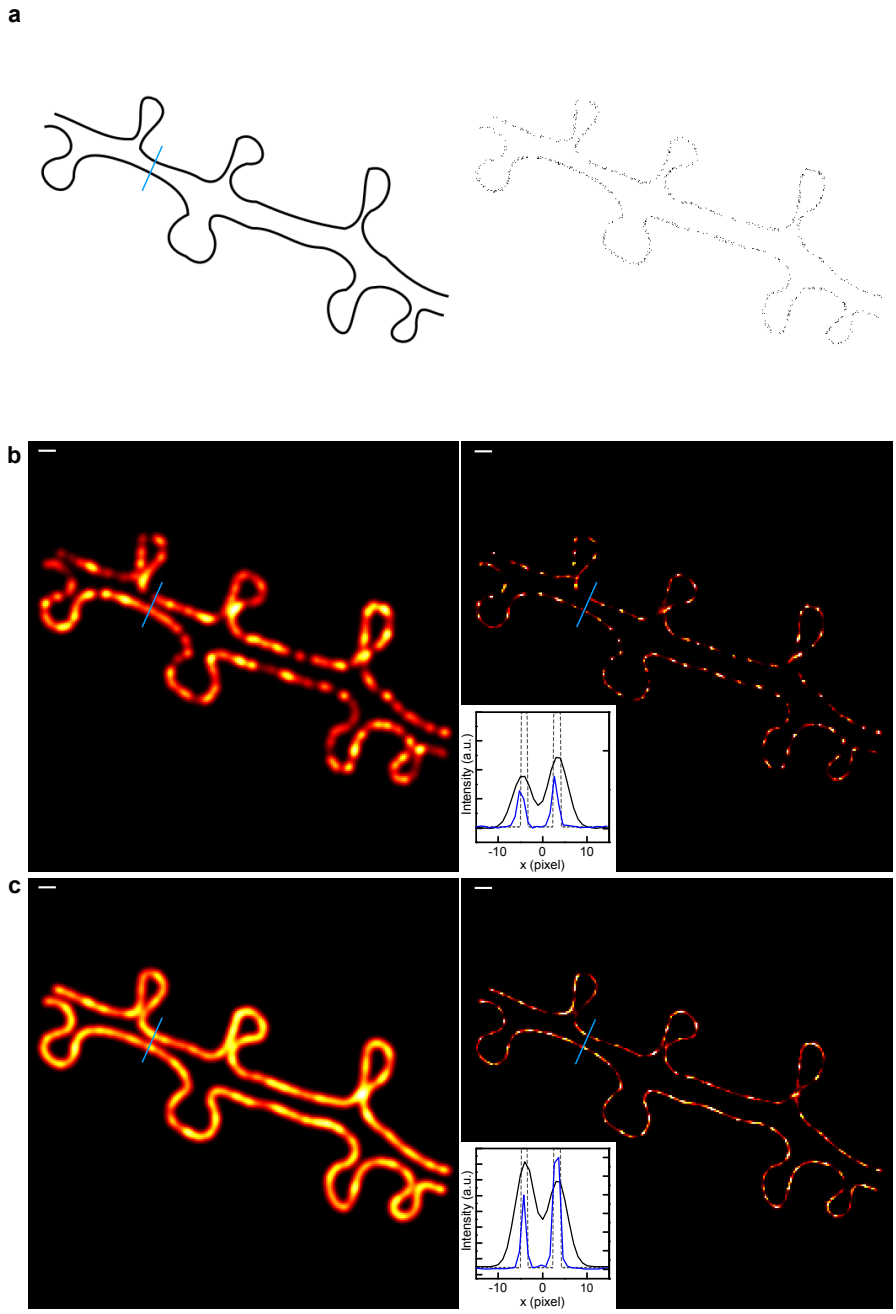
Comparison of diffraction limited fluorescence images (**a**, **d**, **g**, **j**) with Richardson-Lucy deconvolution (**b**, **e**, **h**, **k**) and SPoD-images (**c**, **f**, **i**, **l**) from various samples shown in **Figs. 2-4**. (**a**), (**b**), (**c**) Comparison of the experimental diffraction limited fluorescence image with Richardson-Lucy deconvolution and the SPoD-image shown in **Fig. 2e** and **g**. (**d**), (**e**), (**f**) Corresponding comparison for the experimental images shown in the insets of **Fig. 3a** and **3b**. (**g**), (**h**), (**i**) Corresponding comparison for the spine marked in **Fig. 4b** by the arrow on the left. (**j**), (**k**), (**l**) Corresponding comparison for the spine shown in **Fig. 4a** and **4b** on the top left.

Supplementary Figure 3 | Two-photon imaging in thick brain slices



(**a**), (**b**) two-photon imaging of spine heads from transgenic mice brain slices labeled with pfEGFP. Images were recorded at tissue depths of **a** 0.4 mm and **b** 0.6 mm using two-photon excitation at 940 nm and a special deep tissue multi-photon microscope objective (25 x, NA =1.05). (**c**), (**d**) SPoD images of data in **a** and **b**. Scalebars represent 1 µm.

Supplementary Figure 4 | Simulations and high-resolution analysis of membrane-like sectional views.



(**a**) Structural mask and single molecule positions for the signal simulation in **b**. (**b**) Average intensity image, high resolution image and intensity profiles of a simulation of 800 single molecules at positions as marked in **a** including modulation, shot noise, background noise and single-molecule blinking as well as phase (orientation) fluctuations. (**c**) As in **b** but with ten times more molecules. Insets represent intensity profile plots along the blue lines. For the profile plots only a pixel oversampling factor of two in comparison to the original image pixels was used, i.e. the actual resolution is higher. Scale bar corresponds to ~ 1 µm when assuming an emission wavelength of 620 nm and a numerical aperture of 1.3.

Supplementary Note 1

Demodulation Algorithm

The emission from a single molecule or nanoarea which modulates with the amplitude A and phase Δt (or $\Delta\alpha$) around an average intensity I, is blurred over several pixels by the diffraction limited PSF of width ~250 nm. Thus, in the raw data each nanoarea (each containing a certain number n of molecules) with different phases contributes dominantly to about 25 pixels of size 50 x 50 nm^2 with different contributions according to the PSF. The intensity, I, in the final image with 50 x 50 nm^2 nanoareas is proportional to the number of molecules present in the real structure in these nanoareas. The task of the analysis is to find the best set of parameters A, I, Δt for each of these ~25 nanoareas that fits the actual overall modulation $I(pixel,t)$ observed in all ~25 pixels simultaneously. All other parameters that are needed to describe the observed picture (frequency, pixel size, PSF) are fixed. In contrast to steady state excitation in which only one intensity is observable per pixel, the algorithm here has notably more than one data point available per pixel and each data point carries additional information due to the modulation.

Our statistical data analysis method is therefore based on an explicit model of the observation process. The i-th molecule, located at position r_i in the focal plane, emits photons with a rate $k(\alpha(t)-\alpha_i)$ which depends on time t due to the temporal variation of the polarization angle and its difference to the transition dipole moment orientation α_i of the molecule ($\alpha(t)-\alpha_i=\Delta\alpha$). These photons are blurred by the point spread function $U(r)$ of the microscope, resulting in a Poisson distributed number of photons registered by the camera at location r. The average value of photons registered is $U(r-r_i)\cdot k(\alpha(t)-\alpha_i)$. Summing over all molecules the resulting image is

$$I(r,t)\sim\text{Poisson}(I_0 \sum U(r-r_i)\cdot k(\alpha(t)-\alpha_i)) \quad (4)$$

The constant I_0 is a prefactor containing such influences as the excitation intensity, the exposure time of the camera, the pixel size and so on.

Since the polarization angle is a linear function of time (the polarizer is rotated with constant velocity), we may express the angle dependent function $k(\alpha(t)-\alpha_i)$ by the time dependent function $f(t-t_i)$ with a suitably chosen t_i, the time shift of molecule i. The function $f(t)$ is periodic because after the time T corresponding to half a rotation of the polarizer the emission rate repeats itself. In the simplest case (no ExPAN) the function $f(t)$ is

$$f(t) = cos^2\left(\frac{\pi t}{T}\right) \quad (5)$$

It is more convenient to rewrite the above expression in terms of a density $g(r, t)$ of molecules at r with time shift t,

$$(r,t) = \sum_i \delta(r-r_i)\delta(t-t_i) \quad (6)$$

where $\delta(r)$ is a two-dimensional Dirac δ-function and $\delta(t)$ a one-dimensional one. Then the registered number of photons in Eq. 4 is distributed according to

$$I(r,t)\sim Poisson\left(I_0 \int_0^T \int U(r-r')f(t-t')\,dr'dt'\right) \quad (7)$$

Now the additional information due to the polarization modulation becomes apparent: without polarization modulation, only the marginal density $\hat{g}(r) = \int_0^T g(r,t)dt$ can be used for the reconstruction whereas SPoD allows for additional time resolution resulting in a much sparser representation. This can be seen most easily by considering a discretized version of $g(r, t)$ as it is done for the algorithmic implementation. For N molecules and S phase or time dependent images with P pixels the fraction of nonzero entries (measuring the sparsity) of the

Nature Methods: doi:10.1038/nmeth.2919

marginal density is N/P (which may be of order 1) whereas it is N/PS for the full density, i.e. reduced by a factor of S. The density $\hat{g}(r)$ is what a Richardson-Lucy algorithm could aim to reconstruct directly in an experimental setup without the rotating polarizer, hence we see that the gain of the method compared to conventional imaging consists of a sparsification of the data by allowing for a representation in terms of $g(r, t)$.

The analysis becomes even more robust when the following experimental effects are also considered. Due to the properties of the dichroitic mirrors used in the experimental setup the overall intensity of the excitation laser beam varies somewhat with the polarization angle, hence the intensity of the emitted light also varies accordingly. This results in a time dependent but periodic correction factor $I_0(t)$ instead of the constant I_0. Moreover, light from molecules outside of the focal plane is also registered to some degree but it is heavily blurred. Due to the strong blurring its time shift information is lost since at any particular point of the image the contributions from many such molecules are added. This unmodulated background can be modeled by a time independent function $b(r)$. Since it represents the blurred background, it is expected to be slowly varying in space. This property can be enforced by demanding sparsity of its cosine transform $\tilde{b}(k)$ (we use the cosine transform instead of the Fourier transform in order to avoid unnecessary complications due to the complex values appearing in the Fourier transform). Our complete model hence is

$$I(r,t) \sim Poisson\,(\mu(r,t)) \tag{8}$$

where

$$\mu(r,t) = I_0(t)\left(\int_0^T \int U(r-r')f(t-t')g(r',t')dr'dt' + b(r)\right) \tag{9}$$

In order to estimate the density $g(r, t)$ and the background $b(r)$, we apply penalized maximum likelihood estimation for Poisson statistics enhanced by sparsity enforcing terms for both $g(r, t)$ and $\tilde{b}(k)$. The negative log likelihood of Poisson statistics with intensity μ for an observation of n photons is $l(\mu; n) = \mu - n \log \mu$. In our model we have an observation $I(r, t)$ at each r and t with parameter $\mu(r, t)$. In order to estimate $\mu(r, t)$ from the data we then suggest to minimize the functional $L[g, b; I]$ given by

$$L[g,b,I] = \int_0^T \int l\big(\mu(r,t); I(r,t)\big)\,dr\,dt + \lambda_1|g|_1 + \lambda_2\big|\tilde{b}\big|_1 \tag{10}$$

with respect to $g(r, t)$ and $\tilde{b}(k)$. The first term is the negative log likelihood of Poisson statistics and the sparsity enforcing terms are the ℓ_1 -norms $|g|_1$ and $|\tilde{b}|_1$. The parameters λ_1 and λ_2 are empirical parameters that depend, for example, on the above mentioned back-ground modulation caused by the overall excitation beam variation in the specific experimental set-up used. In general, we term the method that is used here to analyse the modulation data “Sparsity penalty enhanced estimation by demodulation" (SPEED).

The guaranteed sparsity of the representation is a great help in reconstructing the true molecule distribution due to the compressive sensing paradigm. Although it has been shown that this is more difficult for Poisson statistics than for, e.g., Gaussian noise[1], the shrinkage property of the ℓ_1-norm remains and the reconstruction error as measured by a weaker measure than the ℓ_2-norm will still be small under suitable assumptions.[2] However, a full detailed theoretical analysis of the situation of this experimental setup in the compressive sensing framework is not yet available and is the subject of future research.

The functional $L[g, b; I]$ consists of a convex smooth part (the Poisson likelihood) and a convex non-smooth part (the ℓ_1-norms). Fast minimization of such a functional can be accomplished with the Beck-Teboulle algorithm[3], which we used for this publication.

There are several factors which are not included in this simplified model, for instance possible jumps in the dipole moment orientation α_i or photobleaching. The inclusion of these additions is left for future work.

1. Raginsky, M., Willett, R. M., Harmany, Z. T. & Marcia, R. F. Compressed sensing performance bounds under poisson noise. *Signal Processing, IEEE Transactions on* **58**, 3990-4002, (2010).
2. Aspelmeier, T., Ebel, G. & Engeland, U. Transmission image reconstruction and imaging using Poissonian detector data. Germany. PCT/EP2011/003507 (2013).
3. Beck, A. & Teboulle, M. A Fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences* **2**, 183-202, (2009).

Supplementary Note 2

Estimate of resolution

An exact mathematical description for the resolution that can be achieved under various different experimental conditions is subject of ongoing studies. However a first order estimate for an upper limit of the resolution of SPoD images (without considering the modulation enhancement available with ExPAN) can be obtained as follows. We first assume two spots at which N_{mol} molecules with random orientations are positioned. The number of photons observed from a single spot corresponds on average to $N_{ph} = N_{mol} \cdot N_{ph/mol} \cdot t$ photons where $N_{ph/mol}$ is the average number emitted from a single molecule per unit time and t the data acquisition time. During polarization rotation, the signal observed from such a spot modulates around this average value (Compare also to the black curves in **Fig. 1b**). Due to averaging of the modulation from individual emitters, the relative amplitude of this modulation becomes the smaller the more molecules, N_{mol}, contribute to one spot. On average the number of photons in the modulated signal is $N_{mod} = N_{ph} \cdot (1/N_{mol})^{1/2}$. During polarization rotation the total signal observed from one spot is thus represented by a constant photon fraction $N_{const} = N_{ph} - N_{\mathrm{mod}}$ plus a modulated photon fraction N_{mod} (see e.g. **Fig. 1g**).

The analysis described in the **Supplementary Note 1** derives the final images by a simultaneous analysis of the time-dependent modulation and constant signals and their different space-dependent contribution to different pixels. However, to provide a analysis -independent estimate for the upper limit of the resolution enhancement available from the modulation information it is conceptually easier to divide the analysis into two steps. In the first step, the differences in the modulation phase of the spots are used to distribute the intensity from both spots to a series of phase-dependent images in which the signals emitted from one and the other spot are still spatially blurred but occur only in that image which belongs to its phase of the modulation.

When the difference in the modulation phases of the two spots is maximal ($\Delta\alpha = |\alpha_1 - \alpha_2| = 90°$), they appear maximally separated in the phase-dependent image series as one spot in one image and another spot in another image with total intensities corresponding to N_{mod} photons. For smaller phase differences $\Delta\alpha$ and when the distance of both spots, r, is small compared to the standard deviation of the point-spread function, s, the amount of information available for distinguishing the spots is proportional to $\sin\Delta\alpha$, i.e. for close-by spots of similar phase only a certain fraction of the signal can be assigned reliably to the different phase dependent images. Maximum information is available for $\Delta\alpha = 90°$ or when the distance r is significantly larger then s. Thus, the information that can be used to determine the position of one spot is contained in the position of at least $\sin\Delta\alpha \cdot N_{mod}$ detected photons or on average of $(2/\pi)\cdot N_{\mathrm{mod}}$ photons.

In addition, each spot is also affected by the shot noise from the constant (non-modulated) background corresponding maximally to $2 \cdot N_{Const.}$ from both spots. For our upper limit estimate we conservatively assume that this total background noise is entirely contributing as background to the single phase-dependent image that belongs to the phase of the modulated fraction of one spot even though in reality the background for each phase image is reduced since it is distributed over several phase-dependent images. This conservative assumption also accounts for the possibility that the intensity of the modulated signal, N_{mod}, of one spot might be blurred to some extent over several phase-images. In any case, the number of photons that can be used to determine the position of one spot corresponds on average at least to $(2/\pi)\cdot N_{mod}$. The constant photon number per one pixel in this phase-dependent image is approximately equal to $N_{const}^{per\ pixel} \approx N_{\mathrm{const}}/(2s/q)^2 \cdot N_{spots,}$ where q is the pixel size in the image space and $N_{\mathrm{spots}} = 2$ the number of spots that overlap in a diffraction limited zone.[4]

Assuming photon shot noise statistics for the background noise per pixel, $\boldsymbol{b}$, yields: $b = (N_{const}^{per\ pixel})^{1/2}$. Then the positioning accuracy of one spot in its phase dependent image relative to other spots is given by:

$$\sigma_{x,y} = \sqrt{\frac{s^2 + q^2/12}{N_{mod} sin\Delta\alpha} + \frac{8\pi s^4 b^2}{q^2 {N_{mod}}^2 sin^2\Delta\alpha}} = \sqrt{\frac{s^2 + q^2/12}{\sqrt{N_{Mol}} \cdot N_{Ph/mol} \cdot sin\Delta\alpha \cdot t} + \frac{2\pi s^2\left(1 - \frac{sin\Delta\alpha}{\sqrt{N_{Mol}}}\right) \cdot N_{spots}}{N_{mol}^{Ph} \cdot t \cdot sin^2\Delta\alpha}} \qquad (11)$$

Please note that N_{mod} is the total number of photons in the modulated signal per spot, not per pixel whereas b is the background noise per pixel, (not the background intensity, N_{const}). Please note that here a modified definition for $N_{const} = N_{ph} - sin\Delta\alpha \cdot N_{mod}$ was used that considers also the phase difference $\Delta\alpha$. In case of only one molecule per spot (N_{mol} =1) and $\Delta\alpha = 90°$ this formula becomes identical to the localization precision in STORM and PALM without background noise.

This estimate is still valid for more than two spots per diffraction limited area as long as the spots can be separated by their phase information. Then still each spot shows up in a single phase-dependent image and can be localized according to eq. 11. There is also no big conceptual difference between considering spots containing a certain number N_{mol} molecules or small nanoareas of a pixel size of, e.g., 50 x 50 nm^2 in which N_{mol} molecules are located. For example, for the data shown in **Fig. 2g**, **Fig. 3b** and large parts in **Fig. 4b** the structure of the modulating signals in the focal plane usually do not exceed much more than N_{Spots}=5 nanoareas or pixels in a diffraction limited area of ~5 x 5 Pixels = 250 x 250 nm^2. This is a number of nanoareas per diffraction limited region that is usually easily assigned to different phases by the analysis at least if the modulation is averaged over several periods or if ExPAN is applied. Therefore eq. 11 gives a reasonable upper limit for the resolution also in these cases. Indeed, when assuming for the data shown in **Fig. 2e-g** N_{mol} ~20 per pixel (as inferred from the total intensity), N_{Spots} ~5 nanoareas per diffraction limited area, a minimum phase difference of $\Delta\alpha$~10°, $N_{ph/mol} \approx$ 25000 photons per second, for s ~250 nm, a data acquisition time of t = 1 s and a pixel size of q = 48 nm then eq. 11 yields an upper resolution limit of $\sigma_{x,y}$ ~40 nm that corresponds very well to our experimental observations. The equation predicts also very well the resolution obtained with the two single molecules shown in **Fig. 2d**. With N_{mol}=1, N_{Spots}=2, a phase difference of ~20°, $N_{ph/mol} \approx$ 50000 photons per second, for s ~250 nm, a data acquisition time of t = 0,3 s and a pixel size of q = 48 nm eq. 11 predicts a resolution limit of ~18 nm that corresponds also very well to our experimental observations.

Eq. 11 also predicts that the resolution increases by a factor of $\sim (2)^{1/2}$ when doubling the data acquisition time or the photons emitted per molecule, $N_{ph/mol}$, as well as approximately a factor of ~2 when increasing the phase difference, $\Delta\alpha$, by a factor of two. The influence of the number of molecules per pixel on the resolution decreases quickly for $N_{mol} > 5$. If single molecules can be separated by their phase with or without ExPAN the equation for this estimate becomes identical to the localization precision achievable by PALM and STORM.

However, for higher densities of spots or nanoareas the estimate given by eq. 11 breaks down because the actual resolution enhancement is based on the simultaneous analysis of the entire time-dependent modulation and time-independent signals and their different space-dependent contribution to different pixels. Also, data generated with ExPAN (see red lines in **Fig. 2b and in Fig. 1 g**) contains a multitude of additional information that can be used to robustly reconstruct the original structure and which is not considered in the first order estimate provided by eq. 11. For these experimental conditions a much more elaborated estimate than eq. 11 has to be employed which is the subject of ongoing studies.

4. Thompson, R. E., Larson, D. R. & Webb, W. W. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J.* **82**, 2775-2783, (2002).