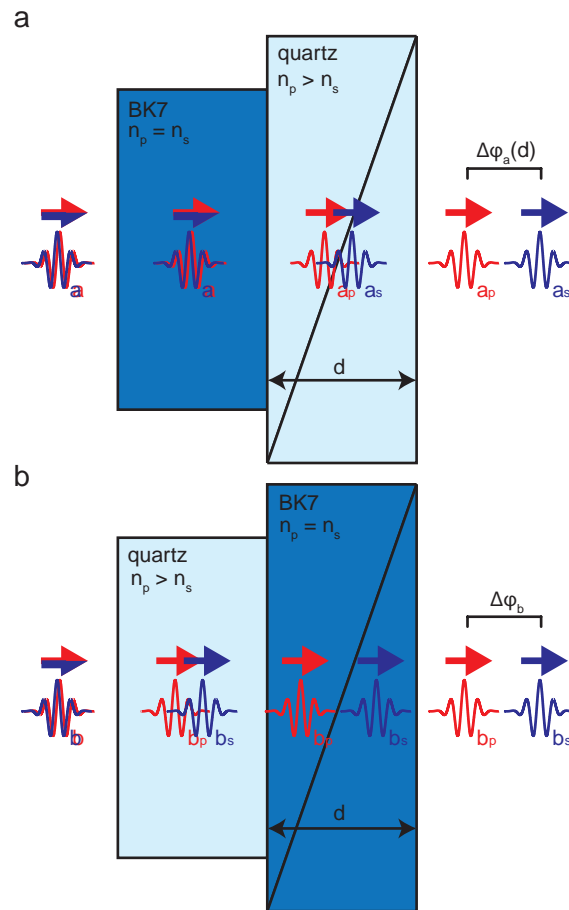


## **Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores**

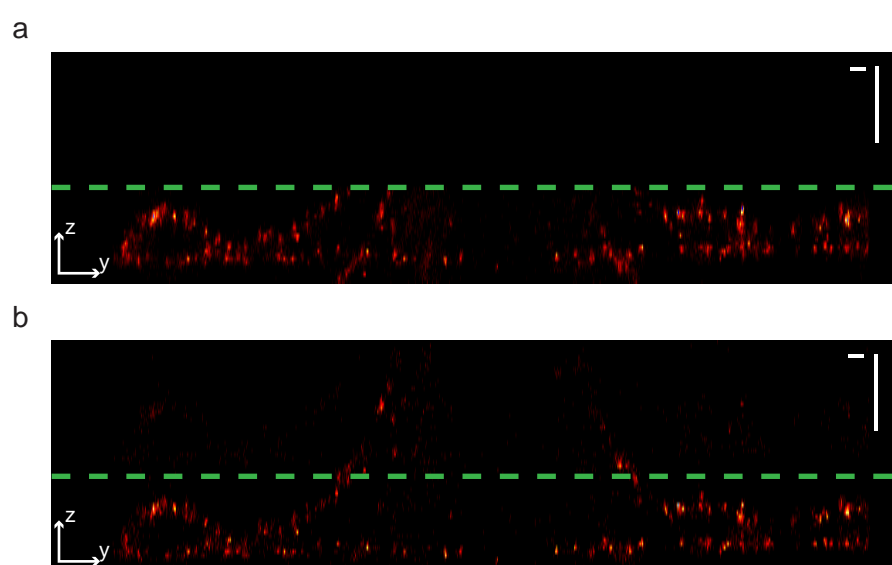
Daniel Aquino, Andreas Schönle, Claudia Geisler, Claas v Middendorff, Christian A Wurm, Yosuke Okamura, Thorsten Lang, Stefan W Hell & Alexander Egner

<b>Supplementary Figure 1</b>	Modified Babinet-Soleil compensators
<b>Supplementary Figure 2</b>	Gain in information obtained from higher order moment evaluation
<b>Supplementary Figure 3</b>	Modulation frequency as a function of moment order
<b>Supplementary Figure 4</b>	Phase pairs in $\varphi_0/\varphi_3$ -space
<b>Supplementary Figure 5</b>	Localization precision in the focal plane as a function of photon counts
<b>Supplementary Figure 6</b>	Additional rings of labels around tubulin
<b>Supplementary Figure 7</b>	Dye classification in s-p-space
<b>Supplementary Figure 8</b>	Validation of dye classification
<b>Supplementary Note 1</b>	Theory and Implementation of the Position Estimators
<b>Supplementary Note 2</b>	4Pi-SMS microscope system parts list

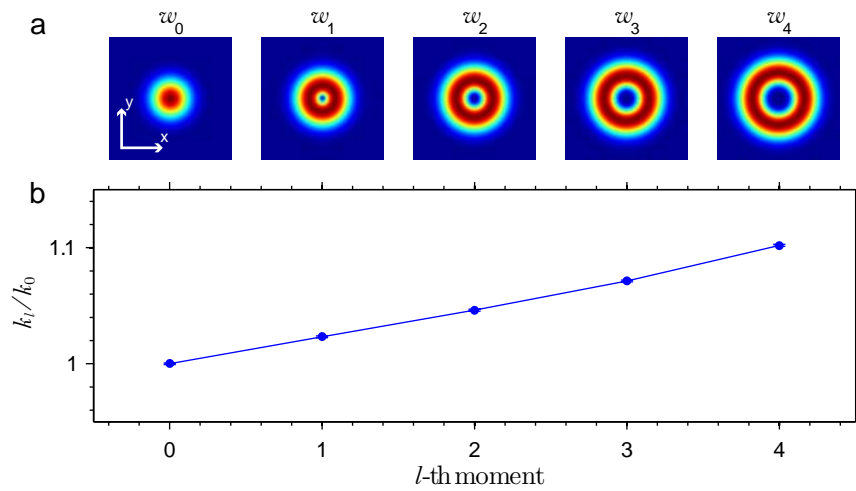
## Supplementary Figures



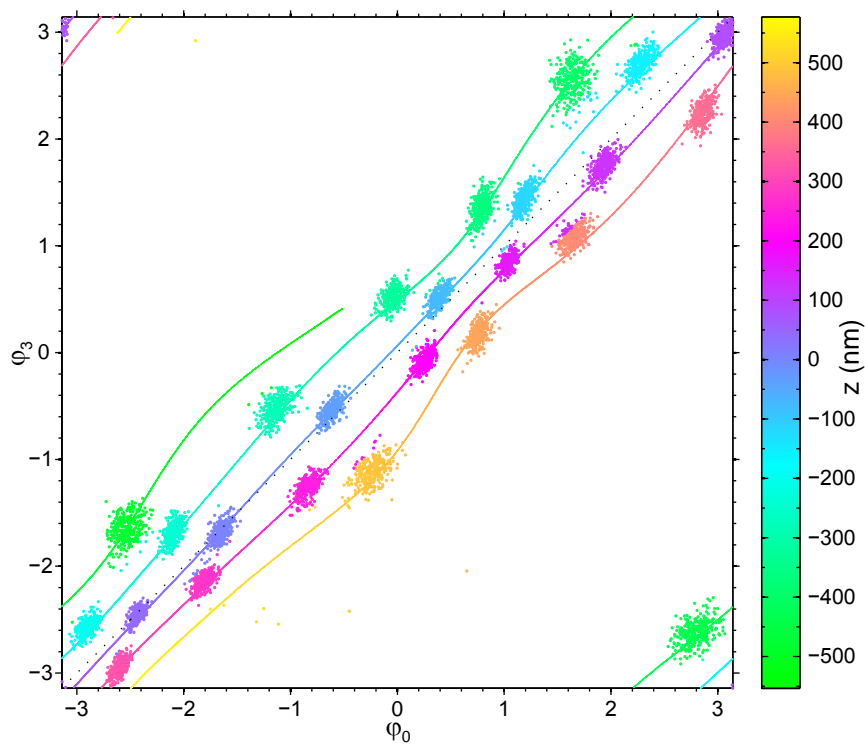
**Supplementary Figure 1** | Pair of modified Babinet-Soleil compensators as used behind the two objective lenses **a** and **b**. While compensator **a** consists of a parallel BK7 plate and two quartz wedges, compensator **b** consists of a parallel quartz plate and two BK7 wedges. By varying the thickness of the plate made up by the two quartz wedges, we can precisely tune the phase delay behind the lens **a** to achieve  $\Delta\varphi_s - \Delta\varphi_p = \pi/2$ . By varying the thickness of the variable BK7 plate made up of the two BK7 wedges behind lens **b**, we are able to correct the overall dispersion within the 4-Pi triangle.



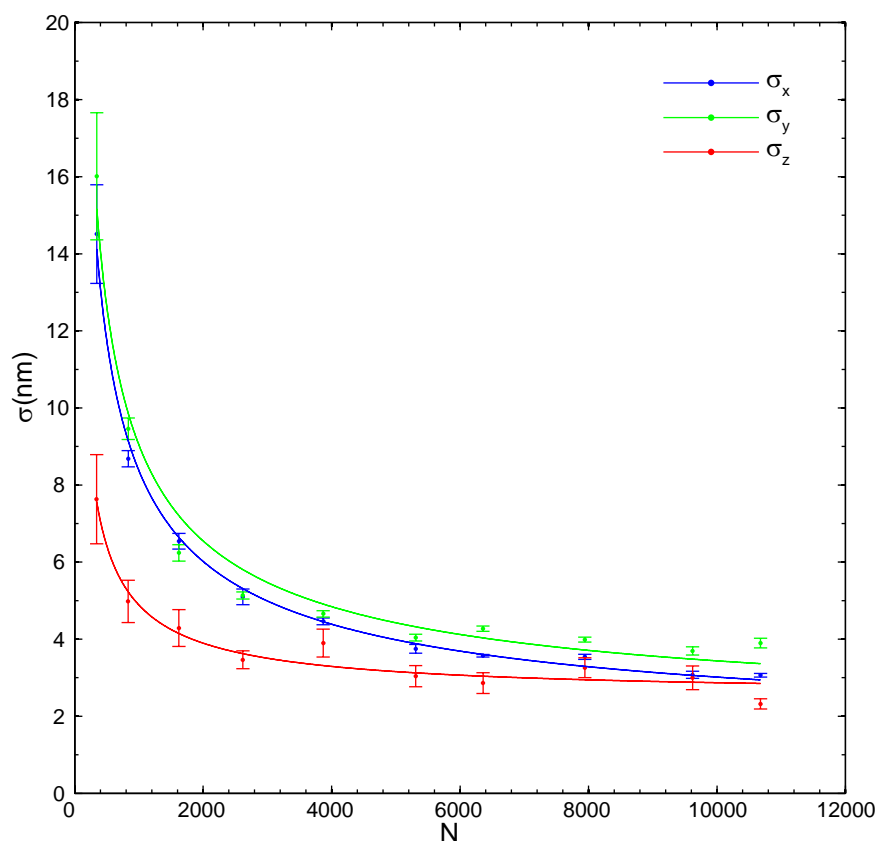
**Supplementary Figure 2** | Well-defined fluorophore localization by employing moments of higher orders illustrated by means of the data underlying Fig. 2g: **(a)** By considering the relative intensities of the detection channels only, molecular events from outside the area close to the coverslip (255 nm, green line) are falsely back projected into it. **(b)** By exploiting the high NA detection through including the 3rd central moment of the image into the analysis this ambiguity can be lifted, while preserving the high localization precision. Scale bars: (a,b) 200 nm in x-direction and z-direction.



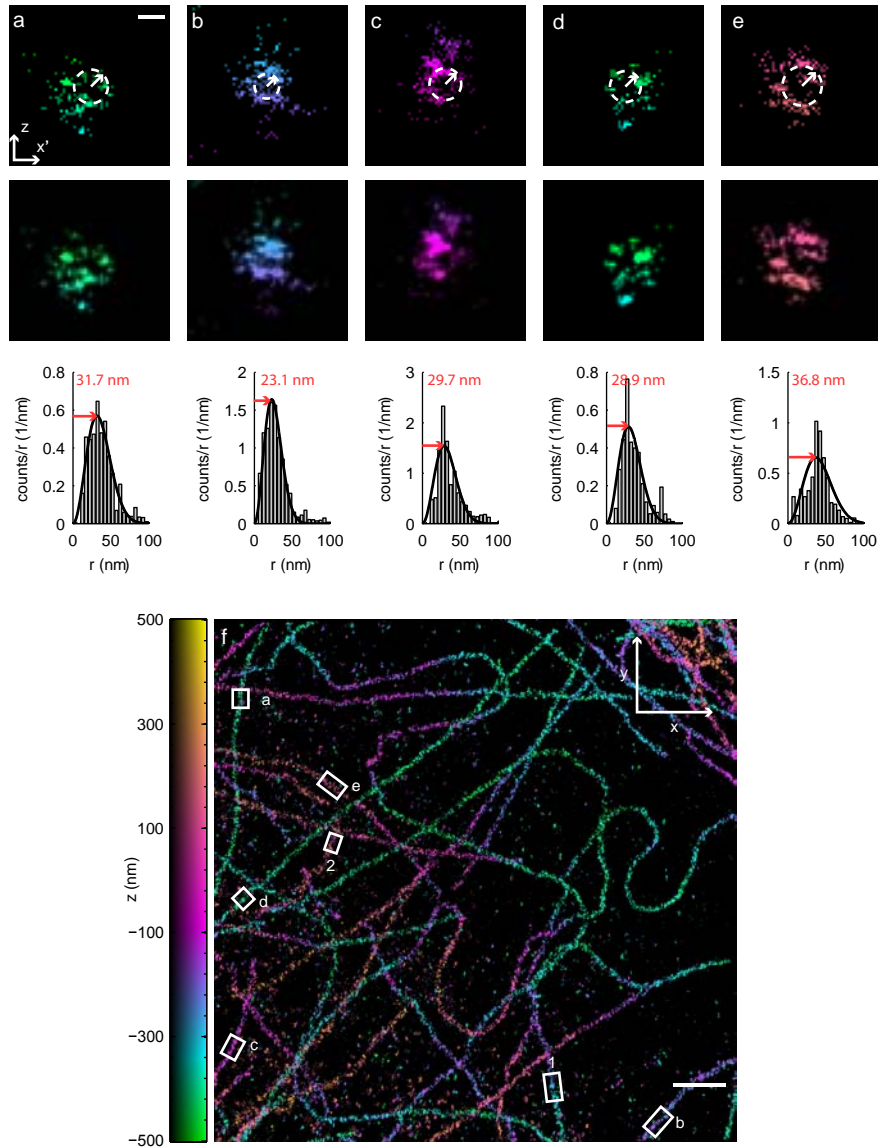
**Supplementary Figure 3** | The modulation speed of the moments  $M_l$  increases with their order. (a) Illustration of the weighting functions  $w_l$  used for the calculation of the moments  $M_l$ . (b) Spatial modulation frequency  $k_l$  along the optic axis of the different moments  $M_l(z)$  extracted from measuring a bead scanned along  $z$ .  $k_l$  was determined by fitting the reduced moments defined in equation (14) with  $b \cos(k_l z - \varphi_0)$ . The modulation is faster for higher moment orders  $l$  which emphasize the outer parts of the emitter's image. From a single phase  $k_l z$  the  $z$ -position can not be unambiguously inferred as a shift by  $2\pi/k_l$  reproduces the same value. However by measuring at least two phases this ambiguity is lifted allowing localization over a much larger range of  $z$ -values.



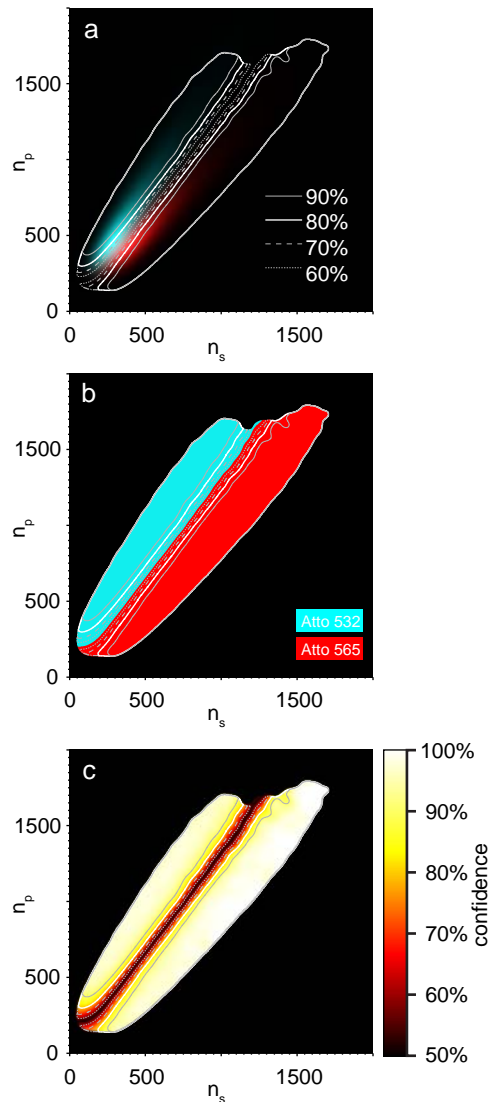
**Supplementary Figure 4** | Phase pairs  $\varphi_0/\varphi_3$  for the measurement shown in Fig. 1e. At each axial location 300 independent measurements were performed. The axial marker position is then estimated by determining the point on the gauge curve that minimizes the distance  $D(z)$  defined in equation (20). Note that in a properly aligned setup, the focal plane ( $z = 0$ ) is unambiguously defined by  $\varphi_0 = \varphi_3$ .



**Supplementary Figure 5** | Localization precision of an emitter in the focal plane as a function of its brightness. We performed 300 independent measurements for each emitter brightness and determined the average number of photons detected per measurement,  $N_{\text{ph}}$  and the standard deviation  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$  of the lateral and axial coordinates were determined by our algorithm. The data is well described by  $\sigma = [2\sigma_0^2/N + \sigma_{\text{vib}}^2]^{1/2}$ , where  $\sigma_0$  is the error due to the localization process and  $\sigma_{\text{vib}}$  is a standard deviation related to residual instabilities of the setup. The factor of 2 in the formula accounts for excess noise introduced by the electron amplification within the EMCCD. We determined  $\sigma_{\text{vib}}/\sigma_0$  to be 1.5/184, 2.0/197 and 2.6/93 nm in the x, y and z-directions respectively.

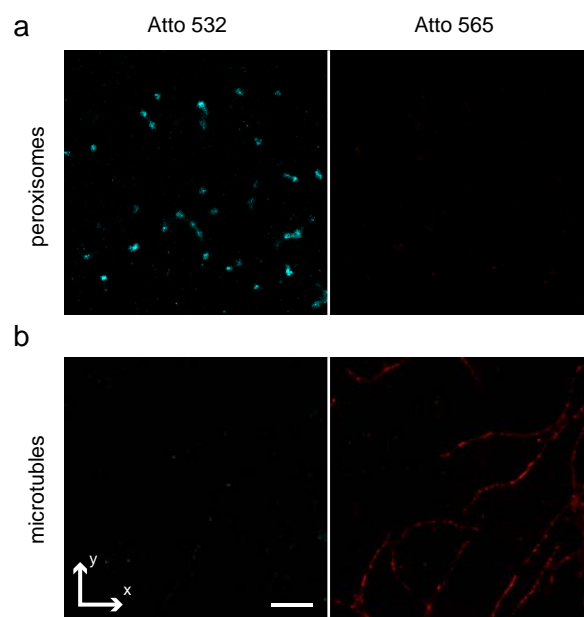


**Supplementary Figure 6** | Additional rings of labels around tubulin. (a-e) x'z histogram image of the distribution of labels located within the rectangles indicated in (f). The upper row shows the x'z histogram images, the second row exhibits the same images smoothed with a Gaussian ( $\sigma = 3.2$  nm) for better visibility. The bottom row shows the respective radial-position histograms centered at the ring center shown in the upper row. The ring radius (heuristically) calculated by fitting the function  $ax^2 \exp[-(x/b)^2]$  to the histogram (black line) and using the maximum (indicated by the red arrows). The ring-like structure is due to the fact that only the surface of the tubules was labeled. (f) Like Fig 3a, extended by boxes indicating the regions of (a-e). The rings shown in the main text are indicated by the boxes 1 and 2. The mean fit radius of all rings is  $29.7 \pm 4.1$  nm. Scale bars: (a-e) 50 nm, (f)  $1\mu\text{m}$ .



**Supplementary Figure 7 | Dye classification in s-p-space.** Data is from the gauge measurements used for Fig. 4 in the manuscript. **(a)** Smoothed and normalized histograms containing the events from the Atto532-labelled sample (cyan) and those from the Atto 565-labelled sample (red). These distributions estimate the probabilities  $p[n_s, n_p; \nu]$  of species  $\nu = \text{Atto532, Atto565}$  producing an event with  $n_p$  photons in the p-polarized channels and  $n_s$  photons in the s-polarized channels. **(b)** Classification map used in Fig. 4c of the manuscript. The cyan (red) region marks  $(n_s, n_p)$  values for which the estimator classifies an event as stemming from an Atto532 (Atto 565) molecule. **(c)** Confidence level, i.e. the probability of correctly assigning an event if both species have the same probability of occurring as a function of  $(n_s, n_p)$ . Regions where the histogram dropped below a value of 0.125 (black area in b) are set to 50 %. The confidence level is also indicated by contours at 60, 70, 80 and 90 percent in all panels.





**Supplementary Figure 8** | Superresolution histogram images generated from our gauge measurements. **(a)** Gauge measurement for peroxysomes labeled with Atto532. We applied the derived classification map to check for systematic cross-talk (e.g. correlated with position in the sample). The left panel shows those events classified as Atto532 and the right panel those as Atto565 ( $< 6\%$ ). **(b)** Gauge measurement for tubulin labeled with Atto565. Again classification was performed and the right panel contains those events correctly identified as Atto565 while the left panel contains the cross-talk ( $< 6\%$ ). Scale bars: (a,b)  $1\mu\text{m}$ .

## Supplementary Note 1:

# Theory and Implementation of the Position Estimators

## 1 Theory

As described in detail below, in the present setup, four channels are created in which the fields through the objective lenses  $a$  and  $b$  (Fig. 1a) interfere with different relative phases. This is achieved by different phase delays of the  $p$  (parallel to plane of incidence,  $E_x$ ) and  $s$ -polarized light in the interferometer arms. Moreover, our multi-color measurements use a dichroic mirror to make the intensity ratio between the  $p$  and the  $s$ -polarized channels dye specific and use this to identify the dye species during the measurement.

### 1.1 Dipole Orientation

Unless the emitters rotate fast compared to the timescale of the fluorescence emission during a single *off-on-off* cycle, both the intensity and polarization of the detected fields will depend on the azimuthal and polar orientations. For example, an  $x$ -oriented dipole would be detected almost exclusively in the  $p$ -polarized channels, a  $y$ -oriented one in the  $s$ -polarized channel and a dipole oriented along the  $z$ -axis would create equal intensities in both channels. To avoid the concomitant orientation-dependent bias, we introduced  $\lambda/4$  plates with the crystal optical axis oriented at  $\pi/4$  against the  $x$ -axis in the exit pupil of the lens. Linearly polarized light thus exhibits equal  $p$ -polarized and  $s$ -polarized parts before the beam splitter. Let  $\mathbf{a}(r, \phi, \theta_p, \phi_p, z)$  be the field created by a dipole located on the optic axis at  $(0, 0, z)$  and oriented along  $\mathbf{p} = (\sin \theta_p \cos \phi_p, \sin \theta_p \sin \phi_p, \cos \theta_p)$  at position  $\mathbf{r} = (r \cos \phi, r \sin \phi)$  in the exit pupil of the lens. Here,  $0 \leq r \leq \sin \alpha$  and  $\alpha$  is the semi-aperture angle of the lens. The field in the exit pupil after passing the  $\lambda/4$  plate is then given by

$$\mathbf{A} = e^{i\pi/4} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \mathbf{a} \quad (1)$$

Using the cylindrical symmetry of the lens and the mirror symmetry about the dipole's projection into the image plane, one can show that the two polarization components are related by

$$A_s(\phi_p + \phi, \phi_p) = ie^{-2i\phi_p} A_p(\phi_p - \phi, \phi_p). \quad (2)$$

Thus, the  $s$ -polarized field is the  $p$ -polarized field mirrored about the polar dipole orientation and multiplied by a constant phase that depends on the dipole orientation. Therefore, the intensity ratio in the two channels as well as the ratio and  $z$ -dependence of the (cylindrically symmetric) weighted moments introduced below are independent of the polar orientation of the dipole.

We note that the polar orientation changes the form of  $\mathbf{a}$  but not its phase. Nevertheless, it generates images with the center of mass shifted away from the geometric focus for moderate tilts and ring-like images for  $z$ -oriented dipoles [1]. For completely static dipoles, a minor influence on the analysis described below

can therefore not be excluded and a careful quantification of this effect will be the subject of further studies. We also note that the effect is not limited to the axial position estimation, and should be assessed for 2-dimensional SMS experiments as well.

## 1.2 Coherent detection

If the two arms denoted by  $a$  and  $b$  are identical and if we ignore a constant phase for the  $p$ -polarization, the components of  $\mathbf{A}$  are given by

$$A_{\kappa a/b}(r, \phi, \theta_p, \phi_p, z) = A(r, \phi, \theta_p, \phi_p) \exp(\pm ikz\sqrt{1-r^2}) \quad (3)$$

where  $\kappa = s, p$  denotes the polarization component. In the absence of aberrations,  $A$  is a real function of the dipole angles and the position in the back-aperture. For large focal lengths, the  $z$ -component of the electric field in the image plane can be neglected and the fields generated through the lenses  $a$  and  $b$ ,  $E_{\kappa a/b}$ , are linearly polarized along  $\kappa$  and given by the Fourier transform of  $A_{\kappa a/b}$ . Taking into account the phase delay of  $\varphi_{\kappa a/b}$  in the two arms and the reflection at the beam-splitter, the electric fields in the images of beam-paths 1 and 2 are given by

$$\begin{aligned} E_{\kappa 1} &= [E_{\kappa a} \exp(i\pi/2) \exp(i\varphi_{\kappa a}) + E_{\kappa b} \exp(i\varphi_{\kappa b})] / \sqrt{2} \\ E_{\kappa 2} &= [E_{\kappa a} \exp(i\varphi_{\kappa a}) + E_{\kappa b} \exp(i\varphi_{\kappa b}) \exp(i\pi/2)] / \sqrt{2}. \end{aligned} \quad (4)$$

Because the imaging system is aplanatic, the lateral dependence of the field only depends on the axial position of the emitter and its lateral position  $(x, y)$  relative to the point of detection. We denote the phase differences between both arms for each polarization as,

$$\Delta\varphi_{\kappa} = \varphi_{\kappa b} - \varphi_{\kappa a} \quad (5)$$

and remark that they can be adjusted independent of beam-path and polarization using the mirrors of the cavity or the beam-splitter and the compensator in arm  $a$ . The setup is adjusted such that the phase of  $E_{\kappa a/b}(x, y, 0)$  is equal for all  $(x, y)$  and over the whole emission spectrum. This can be ensured by proper alignment of the triangular cavity and by using the Babinet Soleil compensator in arm  $b$  respectively. We note that due to equation (3), we ideally have  $E_{\kappa a}(x, y, z) = E_{\kappa b}(x, y, -z)$ . However, this symmetry may be broken in the presence of aberrations induced for example by the sample, uncompensated variations in the cover glass thickness when imaging with water immersion lenses or simply by a refractive index mismatch between the immersion medium and the sample. For the success of our  $z$ -estimation this symmetry is not mandatory because we use a gauge procedure to compensate for deviations. Extracting the  $\Delta\varphi_{\kappa}$ -dependence, the intensities in the four channels can be reduced to the following formula

$$I_{\kappa 1/2} = |E_{\kappa a}|^2/2 + |E_{\kappa b}|^2/2 \pm 2|E_{\kappa a}^* E_{\kappa b}| \cos[\varphi_{\kappa}(z) - \Delta\varphi_{\kappa}] \quad (6)$$

where the  $z$ -dependent phase is given by

$$\begin{aligned} \cos[\varphi_{\kappa}] &= -\Im[E_{\kappa a}^* E_{\kappa b}] / (|E_{\kappa a}| |E_{\kappa b}|) \\ \sin[\varphi_{\kappa}] &= -\Re[E_{\kappa a}^* E_{\kappa b}] / (|E_{\kappa a}| |E_{\kappa b}|) \end{aligned} \quad (7)$$

The fields and the phase depend on  $(x, y, z)$  and we use its  $z$ -dependence to estimate the axial position of our emitter.

### 1.3 Moment-based estimator

The setup resembles of course that of a simple interferometric position monitor. Indeed if lenses of very low numerical aperture are used, the square root in equation (3) can be set to unity and we have

$$[E_{\kappa a}^* E_{\kappa b}](x, y, z) = [E_{\kappa a}^* E_{\kappa b}](x, y, 0) \exp(-2ikz) \quad (8)$$

and thus only the phase is  $z$ -dependent in equations (7) with

$$\varphi_{\kappa}(x, y, z) = 2kz \quad (9)$$

It is immediately obvious that this will only allow estimation in a layer of thickness  $\lambda/2$  due to the periodic nature of the intensity variations. All approaches which do not take the high numerical aperture into account are at best similarly limited and deliver false estimates in the worst case.

In order to extract unambiguous  $z$ -estimates we have to use the fact that our objective lenses have high focusing angles. Therefore the  $z$ -dependent oscillation of the intensity in the back-aperture will depend on the distance from the optic axis and the same will be true in the Fourier transform, i.e. in the image of the emitter on the camera. While the exact dependence is not readily derived in closed form, it can be inferred from equation (3) that the central part of  $A$  and thus the outer rings of  $E$  oscillate faster. In order to use this fact we define the weighted moment operator  $\mathbf{M}_l$  acting on an intensity distribution  $I(\mathbf{r})$  by

$$\mathbf{M}_l[I](z) = \int dx dy \exp[-R^2/2\sigma^2] R^l I(x, y, z). \quad (10)$$

with  $R = \sqrt{x^2 + y^2}$ . We will use the moments  $M_{l\kappa 1/2} = \mathbf{M}_l[I_{\kappa 1/2}]$  of an emitter's intensity distribution as parameters for the determination of its  $z$ -position. We note that our particular choice of weighted moments has proven successful but may not be the optimum choice for all applications. While we believe that cylindrical symmetry should be maintained as it ensures a certain robustness against tilt and rotation of the emitters, we are certain that alternative weighting functions can be used in equation (10) to derive valid estimators. For instance, one could imagine ring-like weights with different diameters exploiting the very same features of the detection PSFs as our approach. A more detailed analysis would be a worthwhile subject for further studies.

Applying the moment operator  $\mathbf{M}_l$  on the intensities defined in equation (6), we can write

$$M_{l\kappa 1/2}(z) = \mathbf{M}_l[|E_{\kappa a}|^2 + |E_{\kappa b}|^2]/2 \pm |\mathbf{M}_l[E_{\kappa a}^* E_{\kappa b}]| \cos[\varphi_{l\kappa} - \Delta\varphi_{\kappa}]. \quad (11)$$

where  $\varphi_{l\kappa}$  is determined analogous to  $\varphi_{\kappa}$ . While all quantities now vary with  $z$ , the envelope and modulation strength vary due to the aberrations introduced by the square-root factor in equation (3) and thus quite slowly while the phase function exhibits a much stronger dependence on  $z$  with a period of approximately half a wavelength. Importantly, the period of the phase functions depends on  $l$  which allows us to resolve the ambiguity occurring in layers thicker than  $\lambda/2$ .

This is in agreement with former theoretical studies [2] which show that the PSF contains enough information to uniquely identify positions within a much thicker layer.

$z$ -estimation will be based on a gauge measurement and the amplitudes in (11) may slightly change due to spectral changes effecting a change in relative transmission in the  $s$  and  $p$  polarized channels and due to aberrations and the averaging over a larger or smaller emission spectrum affecting the modulation depth. Robustness can thus be increased by using relative values

$$m_{l\kappa}(z) = [M_{l\kappa 1}(z) - M_{l\kappa 2}(z)]/M_{l\kappa}(z) = b_{l\kappa}(z) \cos[\varphi_{l\kappa}(z) - \Delta\varphi_\kappa] \quad (12)$$

where the total moment for each polarization direction is defined as

$$M_{l\kappa}(z) = M_{l\kappa 1}(z) + M_{l\kappa 2}(z) \quad (13)$$

and the relative modulation strength,  $b_{l\kappa}(z)$ , is implicitly defined. Note that the  $m_{l\kappa}$  remain unchanged if the detection efficiency of  $s$  versus  $p$ -polarized light varies. In experiments,  $\Delta\varphi_s - \Delta\varphi_p$  is tuned to  $\pi/2$  to maximize the information content [2], while the absolute phase  $\Delta\varphi = \Delta\varphi_s$  is often unknown and may drift between experiments. The  $z$ -Information is predominantly contained in the phase  $\varphi_{l\kappa}$  which is independent of the polarization and varies fast with  $z$ . It is therefore advantageous to further decrease the amount of parameters by defining

$$\begin{aligned} c_l(z) &= m_{ls}(z)/\rho_l(z) = \cos[\varphi_l(z) - \Delta\varphi] \\ s_l(z) &= m_{lp}(z)/\rho_l(z) = \sin[\varphi_l(z) - \Delta\varphi] \end{aligned} \quad (14)$$

where

$$\rho_l(z) = [m_{ls}(z)^2 + m_{lp}(z)^2]^{1/2}. \quad (15)$$

The quantities  $c_l(z)$  and  $s_l(z)$  will be used for  $z$ -estimation (see below) and are called *reduced moments* in the rest of this document. The impact of employing further moments in contrast to solely analyzing the intensity can be seen in Supplementary Fig. 2. We note explicitly that the dependence of  $\varphi_l(z)$  on  $z$  deviates slightly from a linear relationship even in an ideal setup. Importantly, as long as these deviations are the same for both gauge and real measurement, they will merely lead to a slightly changing localization precision over the  $z$ -range.

## 1.4 Multicolor Experiments

Analogous to the method described for multi-color GSDIM microscopy earlier [3], we separate different dye species  $\nu$  based on a characteristic emission ratio into the detection channels. Because our  $z$ -estimator is based on four instead of three channels (which are sufficient for mere  $z$  estimation), multi-color detection can be added to it by a simple modification which ensures that the relative intensities of the  $s$ - and  $p$ -polarized channels (the ratio of the sums of the two channels of each polarization is  $z$ -independent) varies between molecular species. Here, we add a dichroic mirror at an angle where the transmission characteristics strongly depend on the polarization direction. As the dichroic mirror does not split the images but just attenuates one of the polarizations more, the two channels are naturally aligned. Consequently, no algorithm for

alignment of the two channels has to be applied to the data. Lateral chromatic aberrations are neglected here, since the spectral shift between the used dyes is only about 30 nm. However, if necessary the chromatic aberrations can be corrected as already shown in former studies [4]. The edge of the dichroic mirror is chosen at a position such that  $s$  and  $p$  polarized light is transmitted with significantly different relative efficiencies for the dye species involved in the measurement. Note that  $z$ -estimation as described above is not affected by this measure as it is designed to explicitly takes this into account. Now let  $p_s(\nu)$  and  $p_p(\nu) = 1 - p_s(\nu)$  be the probability that a photon emitted by a dye of species  $\nu$  is detected in the  $s$  or  $p$ -polarized channels, respectively. The ratio of  $p_s$  and  $p_p$  is then independent of the dye's orientation due to the  $\lambda/4$  plates and given by the ratio of the total transmission efficiency for the two polarization directions, which is specific for the dye's emission spectrum.

We can then develop a maximum-likelihood estimator for the species  $\nu$  if we are given the probabilities  $p[n_s, n_p; \nu]$  of species  $\nu$  producing an event  $(n_s, n_p)$  where the  $n_\kappa$  are the  $M_{0\kappa}$  calculated from the detected intensity distributions and normalized and discretized such that they estimate the actual number of  $s$  and  $p$  polarized photons registered by the detector. We have

$$\sum_{n_s, n_p} p[n_s, n_p; \nu] = 1 \quad (16)$$

Assuming that all species have equal  $a$ -priori probability to be present at a given position, we estimate the species of an event in a mixed sample by

$$\nu_E(n_s, n_p) = \arg \max_{\nu} p[n_s, n_p; \nu]. \quad (17)$$

where the  $n_\kappa$  are now the values calculated from the detected event. Obviously the estimation is less reliable in the overlap region and one can calculate a confidence level for each value pair  $(n_s, n_p)$  given by

$$c(n_s, n_p) = p[n_s, n_p; \nu_E(n_s, n_p)] / \sum_{\nu} p[n_s, n_p; \nu]. \quad (18)$$

which essentially gives the probability of picking the correct type under the assumption that both have equal probability of occurring. Cross-talk can be reduced by excluding events that are likely to produce false assignments (events for which the confidence level is low). To this end we introduce a confidence threshold  $C$ . If, indeed, all species have the same  $a$ -priori probability of occurring, the confidence level gives the probability of a correct estimate and obviously discarding events with  $c(n_s, n_p) < C$  will thus reduce the cross-talk. The cross-talk matrix and thus the cross-talk in real measurements is readily estimated from the gauge measurements:

$$c_{\nu\mu} = \sum_{n_s, n_p} \theta[c(n_s, n_p) - C] \delta_{\nu\nu_E(n_s, n_p)} p[n_s, n_p; \mu]. \quad (19)$$

Inhomogeneous broadening usually widens the distributions  $p[n_s, n_p; \nu]$  beyond the theoretically expected values when assuming a fixed expected ratio of  $s$  and  $p$  polarized light. Similar to the  $z$ -determination, one can attempt to assess these effects theoretically but it has proven much more reliable to use gauge

measurements. Supplementary Fig. 7a depicts  $p[n_s, n_p; \nu]$  extracted from the gauge measurement used for Fig. 4 of the manuscript. To minimize the effect of shot-noise on the estimator the histograms (consisting of 800x800 bins) were smoothed with a Gaussian (FWHM 24 bins). The resulting estimator and confidence level is depicted in 7b and c, respectively. Regions where the smoothed and added histograms from the gauge measurement dropped below a value of 0.125 were also excluded. The cross-talk estimation from this gauge measurement is obtained by simply classifying the events contained in the gauge measurement with the estimator and quantifying the false assignments. In Fig. 4 of the manuscript we used a confidence threshold of  $C = 80\%$  resulting in  $< 6\%$  false assignments for both Atto565 and Atto532, while about 40 % of the events are discarded due to the threshold.

In order to check the reliability of this method and of our cross-talk estimates, we performed cross-validation of additional gauge measurements on two Atto532 and three Atto565 samples. We extracted 6 estimators by pairing every Atto532 measurement with every Atto565 measurement and for each of them we checked the cross talk it produces when classifying the gauge measurements at 80 % confidence threshold. The cross-talk values found were  $5\% \pm 1\%$  for Atto532 (number of estimates  $n = 12$ ) and  $8\% \pm 1\%$  for Atto565 ( $n = 18$ ). The small variation confirms the validity of our approach and the difference to the values found for the gauge measurement shown in Fig. 4 of the manuscript is rooted in the fact that the setup had been re-aligned, slightly changing the detection efficiencies of the dyes.

## 2 Implementation of the Position Estimators

The estimation of the axial position is implemented in a collection of MATLAB macros as follows: The acquired frames are background corrected by a Gaussian filter and the positions of events are estimated independently in each channel as described in former studies [5]. In short: relevant image segments are identified by applying a count threshold to the frame. The quadratic segment is then multiplied with an initially centered, normalized Gaussian of equal FWHM as the PSF. In each iteration the center of mass of the product of the Gaussian with the background-corrected image is calculated and used as the center for the Gaussian in the next iteration. An event is listed if this iteration converges, i.e. if in less than 100 iterations the displacement of two consecutive iterations becomes less than  $10^{-5}$  of the pixel size, otherwise the event is discarded.

Prior to actual measurements, we determine the three affine transformations that register channel 2-4 with channel 1. This is done by recording a fluorescent bead distribution, applying the localization algorithm, cutting out four quadrants of each camera frame and determine the affine transformations that best co-localize the identified positions. To this end we apply three starting transformations to channels 2-4 which have to be a fair estimates of the actual transformations. For each bead (event in channel 1) we then identify the three positions (in channels 2-4) that correspond to it by identifying the events which lie closest to it after back-transformation. Assuming that the assignment is correct for a majority of beads, the correct transformations are then found by minimizing the mean square distance and iterating the whole process if necessary (bad initial guesses and/or dense bead distributions). We remark that the

transformations may be calculated from the actual data itself but using beads to obtain master transformations turned out to be more reliable and accurate due to the better SNR of the images. The relative movement of the four channels between calibration measurements was controlled and turned out to be negligible.

When analyzing the data, quadruples of events in the four channels are identified as above, the center of the event is calculated as the weighted mean of the centers in the four channels. Using the common center, the weighted moments from equation (10) are calculated and recorded for each channel. The data is corrected for the detection efficiency of the individual channels which are determined without the dichroic used for multi-color imaging and the z-dependent reduced moments  $c_{lk}$  and  $s_{lk}$  defined by equation (14) and the phases  $\varphi_{lk}$ , for  $l = 0, 3$  are computed for each event  $k$ .

Our estimator is then based on a reverse-lookup in a calibration curve acquired by scanning fluorescent beads along  $z$  (over 1.2 microns around the focal plane with steps of 40nm) and interpolating the recorded moments along  $z$  (resulting in a calibration curve  $c_{lG}(z), s_{lG}(z)$  with steps of 0.4nm). As mentioned above, the absolute phase common to all moments defined in equation (14) can differ between gauge curve and experiments. The gauge curve can be corrected for this by applying a common phase shift to both  $\varphi_{0,G}$  and  $\varphi_{3,G}$  and using the fact that each absolute phase will result in a characteristic 2d-histogram over the axes  $\varphi_0$  and  $\varphi_3$  to determine the correct shift: We use the phase shift which maximizes the product of the 2d-histograms determined from the gauge measurement with the histogram from the actual measurement.

The z-estimate is then obtained by minimizing the distance of an event  $k$  to the z-gauge curve:

$$D_k^2(z) = \sum_{l=0,3} w_l^2 \{ (c_{lG}(z) - c_{lk})^2 + (s_{lG}(z) - s_{lk})^2 \}, \quad (20)$$

The weights  $w_j$  are chosen according to the experimental uncertainty of the two phase parameters, in our case  $w_0$  and  $w_3$  are 1/0.35 respectively.

Potential common drifts (of all 4 channels together) in space are corrected by cross-correlating packages of 10,000 frames.

If the setup is correctly aligned, the axial position  $z = 0$  is defined by the  $z$  plane in which  $\varphi_0 = \varphi_3$  for all wavelengths but the PSF still scales with the wavelength. Because the mean wavelength of the dye,  $\lambda$ , will usually differ from the mean wavelength of the gauge measurement,  $\lambda_G$ ,  $z$ -positions have to be corrected by scaling the phase-focus-centered  $z$  values with the ratio of the mean wavelengths  $\lambda/\lambda_G$ .

To visualize the data we use color coding of the  $z$  information. To this end we calculate the mean  $z$  value for each pixel of our xy-histogram images and assign a color based on a  $z$ -colormap. The colormap is created by identifying four  $z$  positions with specific colors and by linearly interpolating between them. The  $z$  positions of the specific colors are defined by the 0/0.33/0.66/1-quantiles of the cumulative  $z$  distribution to maximize the dynamical contrast. The color intensity reflects the event counts in a bin. For visibility purpose we saturated some pixels of the 2D-histogram images.



## Supplementary Note 2:

### 4Pi-SMS microscope system parts list

EMCCD camera:	IXON-Plus DU-897, Andor Technology
Objective lenses:	63X PL APO N.A. 1.20 water immersion, Leica
$\lambda/4$ -plates:	Achromatic Quartz $\text{MgF}_2$ $\lambda/4$ -plate, Bernhard Halle Nachfl.
Babinet-Soleil compensator:	custom-made Babinet-Soleil compensator, Bernhard Halle Nachfl.
Cavity dichroics:	zt620-140rb 22.5°, AHF Analysentechnik
Multicolor dichroic:	555dcxr, Chroma
Notch filter:	Single Notch Filter 532 nm - U grade, AHF Analysentechnik
Bandpass filter	HQ 560/40, AHF Analysentechnik
Bandpass filter (multicolor)	Brightline 582/75-25, Semrock
Polarizing beamsplitter	Broadband Polarizing Cube Beamsplitters, Newport
Beamsplitter	Broadband Non-Polarizing Beamsplitter Cube, Newport
Lenses:	Lead-free 400-700 nm cemented achromatic lenses, CVI Melles Griot (L1,L2,L3: f=350,140,300)
Apertures:	Rectangular aperture, Owis

## References

- [1] Engelhardt, J. *et al.* Molecular orientation affects localization accuracy in superresolution far-field fluorescence microscopy. *Nano Letters* **11**, 209–213 (2011).
- [2] v. Middendorff, C., Egner, A., Geisler, C., Hell, S. W. & Schönle, A. Isotropic 3D nanoscopy based on single emitter switching. *Optics Express* **16** (2008).
- [3] Bossi, M. *et al.* Multicolor far field fluorescence nanoscopy through isolated detection of distinct molecular species. *Nano Letters* **8**, 2463–2468 (2008).
- [4] Pertsinidis, A., Zhang, Y. & Chu, S. Subnanometre single-molecule localization, registration and distance measurements. *Nature* **466**, 647–651 (2010).
- [5] Egner, A. *et al.* Fluorescence nanoscopy in whole cells by asynchronous localization of photoswitching emitters. *Biophys. J.* **93**, 3285–3290 (2007).