Three dimensional single molecule localization using a phase retrieved pupil function

Sheng Liu,1 Emil B. Kromann,2,3 Wesley D. Krueger,1 Joerg Bewersdorf,2,3 and Keith A. Lidke1,*

1Department of Physics and Astronomy, University of New Mexico, NM 87106, USA
2Department of Cell Biology, Yale University School of Medicine, New Haven, CT 06520, USA
3Department of Biomedical Engineering, Yale University, New Haven, CT 06511, USA

*klidke@unm.edu

Abstract: Localization-based superresolution imaging is dependent on finding the positions of individual fluorophores in a sample by fitting the observed single-molecule intensity pattern to the microscope point spread function (PSF). For three-dimensional imaging, system-specific aberrations of the optical system can lead to inaccurate localizations when the PSF model does not account for these aberrations. Here we describe the use of phase-retrieved pupil functions to generate a more accurate PSF and therefore more accurate 3D localizations. The complex-valued pupil function contains information about the system-specific aberrations and can thus be used to generate the PSF for arbitrary defocus. Further, it can be modified to include depth dependent aberrations. We describe the phase retrieval process, the method for including depth dependent aberrations, and a fast fitting algorithm using graphics processing units. The superior localization accuracy of the pupil function generated PSF is demonstrated with dual focal plane 3D superresolution imaging of biological structures.

OCIS codes: (100.6640) Superresolution; (180.2520) Fluorescence microscopy; (100.5070) Phase retrieval; (100.6890) Three-dimensional image processing.

References and links
1. M. J. Rust, M. Bates, and X. Zhuang, “Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM),” Nat. Methods 3, 793–796 (2006).
2. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, “Imaging intracellular fluorescent proteins at nanometer resolution,” Science 313, 1642–1645 (2006).
3. S. T. Hess, T. P. K. Girirajan, and M. D. Mason, “Ultra-high resolution imaging by fluorescence photoactivation localization microscopy,” Biophys. J. 91, 4258–4272 (2006).
4. M. Heilemann, S. van de Linde, M. Schüttelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer, “Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes,” Angew. Chem. Int. Ed. 47, 6172–6176 (2008).
5. J. Folling, V. Belov, D. Riedel, A. Schöne, A. Egner, C. Eggeling, M. Bossi, and S. W. Hell, “Fluorescence Nanoscopy with Optical Sectioning by Two-Photon Induced Molecular Switching using Continuous-Wave Lasers,” Chem. Phys. Chem 9, 321–326 (2008).
6. R. P. J. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grünwald, S. Stallinga, and B. Rieger, “Measuring image resolution in optical nanoscopy,” Nat.Methods 10, 557–562 (2013).
7. S. Stallinga and B. Rieger, “Accuracy of the gaussian point spread function model in 2D localization microscopy,” Opt. Express 18, 24461-24476 (2010).
8. M. F. Juette, T. J. Gould, M. D. Lessard, M. J. Mlodzianski, B. S. Nagpure, B. T. Bennett, S. T. Hess, and J. Bewersdorf, “Three-dimensional sub100 nm resolution fluorescence microscopy of thick samples,” Nat. Methods 5, 527–529 (2008).

9. S. Ram, P Prabhat, J. Chao, E. S. Ward, and R. J. Ober, “High accuracy 3D quantum dot tracking with multifocal plane microscopy for the study of fast intracellular dynamics in live cells,” Biophys. J. 95, 6025–6043 (2008).

10. B. Huang, W. Wang, M. Bates, and X. Zhuang, “Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy,” Science 319, 810–813 (2008).

11. S. R. P. Pavani, M. A. Thompson, J. S. Bitenc, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun, and W. E. Moerner, “Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function,” Proc. Natl. Acad. Sci. USA. 106, 2995–2999 (2009).

12. H. Kirshner, F. Aguet, D. Sage, and D. Unser, “3-D PSF fitting for fluorescence microscopy: implementation and localization application,” J. Microsc. 249, 13–25 (2012).

13. M. J. Mlodzianski, M. F. Juette, G. L. Beane, and J. Bewersdorf, “Experimental characterization of 3D localization techniques for particle-tracking and super-resolution microscopy,” Opt. Express 17, 8264–8277 (2009).

14. A. G. York, A. Ghitani, A. Vaziri, M. W. Davidson, and H. Shroff, “Confined activation and subdiffractional localization enables whole-cell PALM with genetically expressed probes,” Nat. Methods 8, 327–333 (2011).

15. E. Toprak, H. Balci, B. H. Blehm, and P. R. Selvin, “Three-dimensional particle tracking via bifocal imaging,” Nano Lett. 7, 2043-2045 (2007).

16. K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg, “Optimized localization analysis for single-molecule tracking and super-resolution microscopy,” Nat. Methods 7, 377-381 (2010).

17. B. Richards and E. Wolf, “Electromagnetic diffraction in optical systems. II. structure of the image field in an aplanatic system,” Proc. R. Soc. Lond. A Math. Phys. Sci. 253, 358-379 (1959).

18. B. M. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, “Phase retrieval for high-numerical-aperture optical systems,” Opt. Lett. 28, 801–803 (2003).

19. B. M. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, “Phase-retrieved pupil functions in wide-field fluorescence microscopy,” J. Microsc. 216, 32-48 (2004).

20. E. B. Kromann, T. J. Gould, M. F. Juette, J. E. Wilhjelm, and J. Bewersdorf, “Quantitative Pupil Analysis in STED Microscopy Using Phase Retrieval,” Opt. Lett. 37, 1805–1807 (2012).

21. M. R. Foreman, C. L. Giusca, P. Török, and R. K. Leach, “Phase–retrieved pupil function and coherent transfer function in confocal microscopy,” J. Microsc. 251, 99–107 (2013).

22. S. Quirin, S. R. P. Pavani, and R. Piestun, “Optimal 3D single-molecule localization for superresolution microscopy with aberrations and engineered point spread functions,” Proc. Natl. Acad. Sci. USA. 109, 675–679 (2011).

23. J. W. Goodman, Introduction to Fourier Optics, Third Edition (Roberts & Company Publishers, 2005), pp.31–50.

24. R. W. Gerchberg and W. O. Saxton, “A practical algorithm for the determination of the phase from image and diffraction plane pictures,” Optik 35, 237–246 (1972).

25. J. C. Wyant and K. Creath, “Basic wavefront aberration theory for optical metrology,” in Applied Optics and Optical Engineering, VOL. XI (Academic Press, 1992), pp.27–39.

26. M. Sambridge and K. Mosegaard, “Monte Carlo methods in geophysical inverse problems,” Reviews of Geophysics 40, 1–29 (2002).

27. S. F. Gibson and F. Lanni, “Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy,” J. Opt. Soc. Am. A 19, 1601–1613 (1991)

28. D. Axelrod, “Fluorescence excitation and imaging of single molecules near dielectric coated and bare surfaces: a theoretical study,” J. Microsc. 247, 147–160 (2012).

29. D. Axelrod, “Evanescent Excitation and Emission in Fluorescence Microscopy,” Biophys. J. 7, 1401–1409 (2013).

30. J. Enderlein, T. Ruckstuhl, and S. Seeger, “Highly efficient optical detection of surface-generated fluorescence,” Appl. Opt. 38, 724–732 (1999).

31. J. Enderlein, I. Gregor, and T. Ruckstuhl, “Imaging properties of supercritical angle fluorescence optics,” Opt. Express 19, 8011–8018 (2011).

32. S. Stallinga, “Compact description of substrate-related aberrations in high numerical-aperture optical disk readout,” Appl. Opt. 44, 849–858 (2005).

33. B. Huang, S. A. Jones, B. Brandenburg, and X. Zhuang, “Whole-cell 3D STORM reveals interactions between cellular structures with nanometer-scale resolution,” Nat. Methods 5, 1047–1052 (2008).

34. NVIDIA, “Compute unified device architecture (CUDA),” (2007), http://www.nvidia.com/object/cuda_home_new.html.

35. C. S. Smith, N. Joseph, B. Rieger, and K. A. Lidke, “Fast, single-molecule localization that achieves theoretically minimum uncertainty,” Nat. Methods 7, 373-375 (2010).

36. F. Huang, S. L. Schwartz, J. M. Byars, and K. A. Lidke, “Simultaneous multiple-emitter fitting for single molecule super-resolution imaging,” Biomed. Opt. Express 2, 1377–1393 (2011).

37. F. Aguet, D. Van De Ville, and M. Unser, “A maximum-likelihood formalism for sub-resolution axial localization of fluorescent nanoparticles,” Opt. Express 13, 10503-10522 (2005).
1. Introduction

Single molecule localization based superresolution operates on the principle of localizing the individual fluorescent molecules that label a sample [1–5]. In the image analysis process, the actual position of each fluorescent molecule is estimated with sub-diffraction precision. The resulting coordinate estimates are then combined to produce an image with given sufficient labeling density and localization precision [6], with sub-diffraction resolution. For 2D imaging (such as in total internal reflection mode), simple PSF models, (such as a 2D Gaussian model), have been found to be sufficient for imaging under typical conditions [7]. 3D imaging, such as that facilitated by dual focal plane [8,9] astigmatic imaging [10], and techniques that make use of more complex engineered PSFs [11, 12] require a 3D PSF model, and therefore an accurate description of the 3D PSF is crucial to optimizing the localization accuracy in image processing.

The 3D PSF model can be based on the experimental measurement of, for example, a sub-diffraction sized fluorescent bead [8,13,14]. In this case, the system aberrations and fluorescence emission properties are inherently included. However, noise in the measurements can introduce inaccuracies into the model and, because the PSF can only be measured at a finite number of points, a continuous description of the PSF throughout 3D space requires interpolation between the discrete measurement points. The measured PSF is also difficult to modify in order to take into account additional aberrations, such as depth dependent spherical aberration caused by refractive index mismatch. Storage and use of a large calibration PSF stack is computationally cumbersome and can limit the speed of analysis – a very practical concern when data sets may contain images of more than $10^6$ fluorophores.

Alternatively, the PSF model can be based purely on a theoretical model. One widely used theoretical model is the 3D Gaussian model [10, 15]. The Gaussian model requires little computational overhead and therefore offers significant speed advantages in the localization algorithm. On the other hand, the Gaussian model only provides a good approximation to the 3D PSF within a limited spatial range near the focus. Furthermore, the Gaussian model also assumes that the optical system is free of aberrations, which in practice is rarely the case. More realistic theoretical PSF models are described by taking into account the vector property of light [7,16,17] and considering the fluorophore as a dipole emitter. The vectorial models provide a complete description of the PSF patterns from dipole emission and can be used to model fixed or freely rotating dipoles. However, these models exclude system specific aberrations that are important in practice.

To overcome the limitations of the PSF modeling approaches described above, a phase retrieval method for wide field microscopy was developed to retrieve the pupil function of the optical system [18,19]. The phase retrieval process makes use of images acquired with the optical system in question, and therefore inherently includes system specific aberrations. A realistic PSF computed from the retrieved pupil function has been used in 3D image deconvo-
olution [19], and more recently, modified phase retrieval methods were investigated specifically for STED microscopy [20] and confocal microscopy [21].

A phase retrieval method was also recently described for use in 3D superresolution imaging using an engineered ‘double-helix’ PSF [22]. In that work, a phase retrieval process was used to model the PSF between the discrete measurement positions along the optical axis. The method, however, relies on local interpolation between z planes separated by about 100 nm, thus effectively using a different pupil function for each z plane and retaining the need to store and access experimentally acquired PSFs during localization.

In this paper, we describe a process for using phase retrieved pupil functions for 3D superresolution imaging with no need of measured PSFs during localization. We describe the phase retrieval process, PSF generation, and 3D single molecule localization. We also demonstrate how to account for depth-dependent spherical aberration induced by sample media refractive index mismatch and elucidate the role of supercritical angle fluorescence in phase retrieval and PSF generation. Finally, we demonstrate improved accuracy of the approach compared to unaberrated PSF models by performing superresolution using a dual focal plane geometry, imaging both immobilized beads and cellular structures.

2. Phase retrieval of the pupil function

According to scalar diffraction theory [23], the relationship between the pupil function and the point spread function is [19],

$$I_{PSF}(x,y,z) = \left| \int \int P(k_x, k_y) e^{2\pi i (k_x x + k_y y)} e^{2\pi i k_z z} dk_x dk_y \right|^2,$$

where $P(k_x, k_y)$ is the pupil function in $k$ space with a length $k = n/\lambda$, $n$ is the refractive index of the immersion medium, and $I_{PSF}(x,y,z)$ is the 3D intensity PSF in Cartesian space. The defocus phase is described by the $\exp(2\pi i k_z z)$ factor, where $k_z = (k^2 - k_x^2 - k_y^2)^{1/2}$.

Our approach for phase retrieval and PSF generation is based on that described by Hanser et al. [19] where an iterative algorithm is used to retrieve $P(k_x, k_y)$ from a series of images of a small fluorescent bead collected with various amounts of known defocus. As in [19], we make the assumption that the fluorescent bead used for phase retrieval measurements contains a large collection of randomly oriented fluorophores and therefore the dipole nature of the emission of individual fluorophores results in a radial modification of the Optical Transfer Function (OTF).

As in [19], we account for these effects with an OTF rescaling performed after calculating the integral in Eq. (1). OTF rescaling will be discussed further in a later section.

In this section, we describe the process of obtaining the pupil function from an experimentally measured PSF. The phase retrieval process consists of four steps: 1) PSF measurement; 2) Data preprocessing; 3) Phase retrieval; and 4) Expansion into Zernike polynomials. Each of these steps is described in the following subsections.

2.1. Data collection (PSF measurement)

Images of a single, isolated bead are collected and used for phase retrieval. The bead is approximately centered in the field of view and placed in focus. During data acquisition, the stage is moved through a range of $\pm 1$ μm about the focus position in 100 nm steps. At each axial position (which we take to be the z axis), 100 images are collected, each measuring $128 \times 128$ pixels. See section 5 for further details.

2.2. Data preprocessing

For each z position, we compute the average of the collected images, yielding a single image at each sampled z position. The resulting images are cropped to a size $R_1 \times R_1$, where $R_1 \approx 40$
pixels, corresponding to 4 μm. This cropping preserves the core structure while eliminating those parts of the image that are indistinguishable from noise. The in-focus image is fitted to a 2D Gaussian intensity distribution and the position is used to define the center of the cropped frame for all images [Fig. 1(a)].

After cropping, a mean background value is subtracted from each of the images. The mean background value is calculated for each image by computing the mean value of the pixels at each edge of the corresponding cropped image and choosing the smallest of the resulting four values. After the mean background has been subtracted, any resulting negative pixel values are set to zero, and a binary circular mask with a diameter of $R_1$ pixels is multiplied onto each image, effectively setting all pixel values outside the circle to zero [Fig. 1(b)]. Finally, the size of each image is restored to the full $128 \times 128$ pixels by padding with zero-valued pixels [Fig. 1(c)].

In this application of superresolution imaging, we use a dual focal plane setup (see section 5.1) and the preprocessing steps described above are applied to the recorded PSF images at both planes separately.

Fig. 1. Preprocessing of raw PSF data. All images show the PSF at $z = -0.6$ μm. On the left side of each image, the pixel intensities have been linearly scaled, while the right side shows the same image with intensity values on a log scale. (a) The raw PSF, cropped to $40 \times 40$ pixels. (b) Subtraction of background from (a) and application of the circular mask (see section 2.2). (c) PSF from (b) after extension to $128 \times 128$ pixels. This is the measured PSF used in phase retrieval algorithm at $i$th iteration where the $i \leq i_0$th iteration (see section 2.3). (d) PSF with additional processing (see section 2.3), which is used as measured PSF in phase retrieval algorithm at the $i$th iteration, where $i > i_0$. 
2.3. Phase retrieval process

The phase retrieval process is an iterative process that retrieves the pupil function of the optical system from a set of measured PSF images. The algorithm was originally proposed by Gerchberg and Saxton (Gerchberg-Saxton algorithm [24]), and further implemented in high numerical aperture microscope system by Hanser et al. in [18, 19]. In our dual focal plane system, we apply a similar procedure to the measured PSF for each plane separately, thereby obtaining a pupil function for each focal plane. As detailed in [19], the phase retrieval algorithm takes as input the numerical aperture of the objective lens \( NA \), the emission wavelength \( \lambda \), the refractive index of the objective immersion medium \( n \), the actual CCD pixel size \( \Delta d \), and the lateral magnification \( M \).

At each iteration of the procedure, the magnitude of the amplitude PSF is replaced by the square root of the measured intensity PSF. However, the magnitude of the generated PSF model never reaches zero at any point in space. This conflicts with our data preprocessing scheme for the measured PSF, in which negative pixel values and extended pixel values are set to zero. To prevent mis-convergence due to the relatively large number of zero-valued pixels following the data preprocessing described in the previous section, the zero-valued pixels in the extended PSF are set to the value of the corresponding pixels in the phase retrieved pupil function generated PSF (PR-PSF) from the previous iteration. This approach has the advantage of preserving the parts of the measured and processed PSF images [Fig. 1(d)] that would otherwise be set to zero. Because the PR-PSF is still very different from the measured PSF during the first few iterations, this additional preprocessing step can cause the calculation to diverge if applied too early in the iterative scheme. We applied the modification only after a certain number of iterations \( i_0 \). We then determined the optimal value of \( i_0 \) through a process described in section 2.5.

2.4. Zernike decomposition and compensation for \( x, y, z \) shifts

Zernike polynomials are a set of orthogonal polynomials that are widely used for describing aberrations in optical systems. We decompose the PR pupil function into Zernike polynomials, which has the following advantages: First, the fitted coefficients of Zernike polynomials indicate particular aberrations present in the optical system; and second, the diffraction integral can be written as a sum of one dimensional integrals, simplifying the PSF calculation (see Appendix A). We decompose both the magnitude and phase parts of the pupil function into the first 49 Zernike polynomials (the polynomials are ordered as in [25]).

The first few coefficients of the phase-fitted Zernike polynomials are used to estimate how well the measured PSF is centered on the cropped region and how well the calculated in-focus PSF image matches with the measured in-focus PSF.

To estimate the lateral shift of the PSF relative to the center of the cropped region, we calculate:

\[
x_{\text{shift}} = C_2 \times \frac{M \lambda}{2\pi \text{NA} \cdot \Delta d},
\]

\[
y_{\text{shift}} = C_3 \times \frac{M \lambda}{2\pi \text{NA} \cdot \Delta d},
\]

where \( C_2 \) and \( C_3 \) are the 2nd and 3rd Zernike coefficients respectively. The axial shift was calculated by fitting the defocus phase \( 2\pi k_z z_{\text{shift}} \) to the 4th Zernike polynomial. The first-order Taylor expansion of the defocus phase can be expressed by the 4th Zernike polynomial plus a constant phase, so that the axial shift is found by minimizing the integral,

\[
f = \iint (2\pi k_z z_{\text{shift}} - W_4 - \phi_0)^2 dk_x dk_y,
\]
where \( W_4 \) is the 4th Zernike polynomial and \( \phi_0 \) is a constant phase. We found that the first order approximation is sufficient to minimizing defocus aberration, so that the 4th Zernike coefficient is less than 0.01. The lateral shift was minimized by shifting the original PSF images accordingly before cropping. The axial shift was minimized by redefining the \( z \) positions of the measured PSF. After correcting for \( x, y, z \) shift, the phase retrieval method described above was repeated, resulting in a Zernike expanded pupil function with lateral shift and defocus near zero.

### 2.5. Parameter optimization

Many of the parameters required by the phase retrieval process can be obtained by direct measurement or from product specifications; however, the optimal values of the parameters \( i_0, R_1, n \) and \( \lambda \), as defined in the previous sections, may depend on the specifics of a particular sample and the experimental conditions. The optimal values of \( i_0 \) and \( R_1 \) are primarily affected by the spectra of the sample beads and the effective pixel size on the sample plane; \( n \) can be affected by the type and temperature of the immersion oil; and \( \lambda \) is primarily affected by the finite width of the beads’ emission spectra and the used emission filters. Given that we ultimately desire minimum error in 3D localization, we optimize \( i_0, R_1, n \) and \( \lambda \) by minimizing the \( z \) localization error of the same bead data used in the phase retrieval process. The optimization is performed by the following steps: 1) Phase retrieval of the pupil function; 2) 3D localization of the bead data recorded at the \( z \) positions from -0.8 \( \mu \)m to 0.6 \( \mu \)m, in 0.2 \( \mu \)m intervals, and the bead data used in localization were preprocessed and averaged from the original data (section 5.4); 3) Calculation of the sum of square error (SSE) of the found \( z \) positions compared with the actual \( z \) positions, which is the set stage positions in \( z \) plus founded \( z \) shift (see section 2.4); 4) Update of the optimization parameters using a Markov chain Monte Carlo (MCMC) method [26], and return to step 1. The complete optimization typically requires about 30 iterations.

### 3. PSF generation

Under scalar diffraction theory, the PSF can be generated from Eq. (1), or when using Zernike polynomials, from Eq. (27) (see Appendix A). This PSF, directly calculated from the pupil function (PR-PSF), does not incorporate effects that originate from fluorophore dipole emission and index mismatch aberration. In this section, we describe modifications to the PR-PSF that address these issues as well as derive the aberration phase caused by index mismatch and clarify the role of supercritical angle fluorescence.

#### 3.1. OTF rescaling accounts for dipole PSF broadening

Scalar diffraction theory does not account for fluorophore dipole emission, and results in a PSF with sharper structures than the measured PSF. Although the pupil function could be modified to correspond to a static dipole of arbitrary orientation [19], we proceed under the assumption that the fluorophore is freely rotating and appears as a collection of randomly oriented dipoles. In order to account for the effects of dipole emission, a 2D empirical function is applied to the PR-PSF, which operates like a Gaussian filter.

\[
I_{\text{PSF}} = |\mathcal{F}\{P(k_x,k_y)e^{2\pi ik_zz}\}|^2 \otimes \mathcal{F}\{G(k_x,k_y)\},
\]

where \( \mathcal{F}\{f(k_x,k_y)\} \) denotes a 2D Fourier transform, and symbol \( \otimes \) denotes a convolution. \( G(k_x,k_y) \) is the empirical rescaling function. This modified PR-PSF (mPR-PSF) is smoother than the PR-PSF and the full width of half maximum (FWHM) of the in-focus PSF is wider, which matches with the vectorial PSF model of a fluorophore with isotropic dipole emission [7].
mPR-PSF could be generated from PR-PSF that either comes from PR pupil function or Zernike fitted pupil function, however, as detailed in section 6.3.3 and Appendix A, we used the latter case for mPR-PSF calculation.

\( G(k_x, k_y) \) is a 2D Gaussian function fitted to the OTF ratio, which is defined as the measured 2D in-focus OTF divided by the retrieved 2D in-focus OTF. The fitting parameters are the magnitude \( f \) and the width of the Gaussian function in both the \( x \) and \( y \) axis, \( \sigma_x \) and \( \sigma_y \). Typically, we find \( \sigma_x \neq \sigma_y \). Because the covariance, \( \sigma_{xy} \), only results in a negligible term in \( \mathcal{F} \{ G(k_x, k_y) \} \), we didn’t include it in the Gaussian function.

### 3.2. Index mismatch aberration and supercritical angle fluorescence

Due to their high numerical aperture, oil immersion objectives are often used even when imaging cellular samples mounted in a water based medium. Imaging away from the cover glass will induce additional spherical aberration that is more pronounced as the distance from the cover glass increases. However, the resulting aberration can be accounted for with a depth-dependent modification of the pupil function.

The beads used for the phase retrieval process are adhered to the cover glass. Therefore, light coming from the bead would travel a negligible distance in the sample medium (reflection and transmission still occur at the interface) and consequently there would be no aberration caused by refractive index mismatch. Therefore, the phase retrieval result does not include such aberration. We adopt Gibson-Lanni’s model [27] for estimating the aberrations induced by refractive index mismatch. We simplify this approach by using a two medium optical system consisting of the sample medium and the immersion medium, with refractive indices \( n_2 \) and \( n_1 \), respectively, and assuming the refractive index of the cover glass and the objective lens are the same as that of the immersion medium. In Fig. 2(a), point \( P \) is the designed focal position of the objective lens, which is at the cover glass–sample interface, and \( t^*_x, t^*_y \) are the designed thicknesses of the cover glass and immersion medium, respectively. In order to image a point source emitter inside the sample medium, the cover glass (dashed red lines) was moved closer to the objective lens, while the thickness of the cover glass remained unchanged (\( t_x = t^*_x \)). For a single emitter \( O \) (red dot, Fig. 2(a)) located at a distance \( z \) from the interface, the optical path difference of the corresponding ray coming from emitter \( O \) and designed focus \( P \) is,

\[
OPD = n_2|OA| + n_1|AB| - n_1|PC|,
\]

and the phase aberration at the pupil will be

\[
\Phi_{aber} = \frac{2\pi}{\lambda} \left[n_2 \cos(\theta_2) - (t^*_x - t_x) n_1 \cos(\theta_1) \right].
\]

For high numerical aperture, for example \( NA = 1.45 \), and with \( n_2 = 1.33, n_1 = 1.52 \), the \( \cos(\theta_2) \) term could be imaginary, given that \( \theta_1 \) satisfies the relation \( n_1 \sin(\theta_1) \leq NA \). This occurs when \( \theta_1 \) is larger than the critical angle. Eq. (6) is still valid and the resulting imaginary phase will give rise to an exponential decay term in the \( z \) direction, which is physically the result of evanescent waves from near field dipole emission. If the fluorophore is very close to the cover glass (within about \( \lambda/2 \)), the evanescent wave will penetrate into the cover glass and transmit at \( \theta_1 \) above the critical angle, a phenomenon known as supercritical angle fluorescence (SAF) [28–31]. If \( \theta_1 \) satisfies the relation \( n_1 \sin(\theta_1) \leq NA \), the SAF will be captured by the objective lens. Since our measured PSF for the phase retrieval process is taken from the sample with the beads adhered to the cover glass in aqueous solution, we assume that the transmission distribution at the cover glass–sample interface is recovered from the phase retrieval algorithm and also includes the transmission pertaining to the SAF effect. Because of the decay term, the numerical aperture will appear smaller when imaging depth increases, up to about
1.5 \(\lambda\), where the numerical aperture converges to the refractive index of sample medium. No additional modification beyond the usage of the complex aberration phase in Eq. (6) is needed to correctly account for SAF effects. See Appendix B for the derivation of this result.

Fig. 2. Schematic illustration of aberration caused by refractive index mismatch. (a) Geometric ray tracing from an on-axis point (blue) at the designed position and an emitter (red) located at an arbitrary \(z\) position. \(t_g^*\) and \(t_i^*\) are the designed thicknesses of the cover glass and the immersion medium respectively, while \(t_g\) and \(t_i\) are the corresponding thicknesses when the objective lens is moved away from the designed position. We assume the refractive index of the cover glass, the immersion medium, and the objective lens are the same and equal to \(n_1\), while \(n_2\) is refractive index of sample medium. (b) In the paraxial range, the image of a fluorophore located at reference position \(z_o\) is at the designed focus position \(P\).

Whole-cell 3D imaging may require imaging depths up to tens of microns from the cover glass surface. Due to refractive index mismatch, the stage position that produces the best focus can be far from the actual depth of the emitter [32]. The quantity \((t_i^* - t_i)\) is the actual movement of the stage in \(z\), which is known and can be denoted as the stage position, \(z_{\text{stage}}\). For 3D emitter localization, it is convenient to introduce a reference position, \(z_o\), that is near the stage position and is defined as,

\[ z_o \equiv \frac{n_2}{n_1} z_{\text{stage}}, \]  

and a relative \(z\) position, \(z'\),

\[ z' \equiv z - z_o. \]  

Figure 2(b) illustrates the physical meaning of the reference position, equivalent to the actual focal plane in [33]. With the fluorophore O located at the reference position, its paraxial image will be at the designed position \(P\). After introducing the reference position, Eq. (6) can be rewritten in terms of \(z_o\) and \(z'\):

\[ \varphi_{\text{aber}} = \frac{2\pi}{\lambda} \left[ z_o n_2 \cos(\theta_2) - z_o \frac{n_1^2}{n_2^2} \cos(\theta_1) + z' n_2 \cos(\theta_2) \right]. \]  

In Eq. (9), the third term can be interpreted as a relative defocus phase, while the first two terms represent the aberration phase introduced by the refractive index mismatch at a given stage position. Therefore, for accurate 3D whole-cell localization, \(\varphi_{\text{aber}}\) should be added to the pupil function (see Appendix A), and the actual fitting parameter should be \(z'\).
3.3. Calculation of the PSF using Graphics Processing Units

In order to make our mPR-PSF based localization algorithm practical for localizing thousands or even millions of emitters, we implemented the PSF calculation on Graphics Processing Unit (GPU) hardware using NVIDIA’s CUDA Architecture [34]. Our previous 2D localization algorithms have shown approximately two orders of magnitude speed improvement using GPU implementations as compared to CPU implementations [35,36]. In this case, the PSF generation itself is the predominant computational demand and only the PSF generation is implemented on the GPU.

In the diffraction integral, we take advantage of the Zernike expanded pupil function to express the PR-PSF calculation as a one-dimensional integral (see Appendix A). The integral along the radial dimension is computed numerically using a quadrature method based on rectangle rules. The calculation is parallelized in CUDA such that one ‘block’ calculates a 2D PR-PSF with a particular defocus and one ‘thread’ calculates one pixel of the 2D PR-PSF. To make more efficient use of the GPU, many PR-PSF calculations are done in parallel, which in the localization algorithm corresponds to several individual fits, each needing 16 2D mPR-PSF images per iteration (See section 4). The OTF rescaling is performed in a second kernel call that performs a real space convolution of the PR-PSF to produce the mPR-PSF and truncates edge pixels to match the data region used in localization.

4. 3D localization algorithm

Using the mPR-PSF obtained via the algorithm described in section 3.1, it is straightforward to model the signal from a single emitter (µ) for any position in space (x, y, z), any photon count (I), and any background level (bg):

\[ \mu(x,y,z,I,bg) = I \cdot \text{PSF}_0(x,y,z) + bg. \]  \hspace{1cm} (10)

The localization algorithm described in this section is particular to a dual focal plane setup, but can be easily modified for other optical configurations. It is based on the maximum likelihood estimator (MLE) assuming a Poisson noise model [35–37] and finds the MLE using a Newton-Raphson method to iteratively update the fitting parameters. Under a Poisson process, the likelihood function for a dual focal plane method is,

\[ L(\theta) = \prod_{q} \frac{\mu_{1q}^{N_{1q}} e^{-\mu_{1q}}}{N_{1q}!} \cdot \prod_{q} \frac{\mu_{2q}^{N_{2q}} e^{-\mu_{2q}}}{N_{2q}!}, \]

\[ \theta = \theta(x_1, y_1, z_1, I_1, bg_1, bg_2), \]  \hspace{1cm} (11)

where \( N_{1q} \) and \( N_{2q} \) are the expected photon counts at the \( q \)th pixel in plane 1 and plane 2 respectively, while \( \mu_{1q}, \mu_{2q} \) are the expected photon counts of a single emitter model at \( q \)th pixel. The fitting parameter \( \theta \) includes the \( x, y, z \) position of a single fluorophore in plane 1, the expected total photon counts, \( I_1 \), for a single fluorophore in plane 1, and the background photon counts in plane 1 (\( bg_1 \)) and plane 2 (\( bg_2 \)). The corresponding values for plane 2 are easily computed once the values for plane 1 have been determined: The \( x_2, y_2 \) position in plane 2 can be calculated from \( x_1 \) and \( y_1 \) by means of a linear transform matrix, and the \( z_2 \) position in plane 2 is related to \( z_1 \) by a known plane separation. The expected total photon count in plane 2 (\( I_2 \)) is related to \( I_1 \) by a ratio factor \( I_{\text{ratio}} \). The background levels in plane 1 and plane 2, \( bg_1 \) and \( bg_2 \), are fitted separately because they could come from auto-fluorescence or out-of-focus emission, and thus can vary independently.

To maximize this likelihood function, we use the Newton-Raphson method to find the zero
of the objective function,
\[
f(\theta|D) = \frac{\Delta \ln(L(\theta))}{\Delta \theta} = \sum_q N_{1q} \frac{\mu_{1q}}{\Delta \mu_{1q}} + \sum_q N_{2q} \frac{\mu_{2q}}{\Delta \mu_{2q}}
\]

where the symbol \( \Delta \) represents forward difference on discrete functions. The fitting parameter \( \theta \) is updated through,
\[
\theta_{n+1} = \theta_n - \frac{f(\theta_n|D)}{f'(\theta_n|D)}.
\]

Since the mPR-PSF can only be described by pixelized images, the derivatives of the log likelihood function, \( \ln(L(\theta)) \), are calculated numerically.

To ensure a proper convergence, the Newton-Raphson method requires an adequate initial guess of the parameters in question. Therefore, before the Newton-Raphson iterative calculation, we perform some basic calculations on the fitting data to ensure an appropriate initial guess of the parameters. The fitting subregions measure 16 × 16 pixels with an effective pixel size of about 106 nm on the sample plane. Each subregion usually contains a single fluorophore, located near the center. The initial \( x, y \) position of the fluorophore to be localized is calculated from the center of mass (COM) of a 5 × 5 pixel area at the center of the subregion in plane 1, which reduces the effect of fluorophores near the border of the subregion. The initial background value is estimated by calculating the mean value of the pixels at each border of the subregion and choosing the smallest of the four values. The background estimation is done for plane 1 and plane 2 separately. The initial intensity is estimated by summing over the subregion for plane 1 after subtracting the initial plane 1 background estimate. We computed the initial \( z \) localization estimate using a modified method based on [38]. We compute the center intensity ratio, \( Z_{est} \), which is the ratio of center intensities in plane 1 and plane 2, where the center intensity in each plane is calculated by summing over an 11 × 11 pixel region at the center of the subregion in each plane after subtracting the estimated background value. We found that this center intensity ratio is linear in \( z \) and can therefore be used as an initial \( z \) estimator. Using crimson bead data for \( Z_{est} \) calibration, we obtained the following \( z \) dependence:
\[
Z_{est} = p_1z + p_2,
\]

where \( p_1, p_2 \) are coefficients of linear fit.

The \( z \) estimation for whole-cell imaging is more complicated. As introduced in section 3.2, with a given reference position \( z_o \), the actual fitting parameter in \( z \) is \( z' \). Thus, the \( z \) localization for whole-cell imaging consists of two steps: finding \( z' \) for a given \( z_o \), and recovering the actual \( z \) position using Eq. (7) and Eq. (8).

The initial \( z' \) estimation is the same as described above. However, different \( z_o \) values will yield very different \( p_1 \) and \( p_2 \) values. Using results from simulated data, we found the linear relationships between \( z_o \) and \( p_1, p_2 \):
\[
p_1 = S_2z_o + S_1, \quad p_2 = S_4z_o + S_3,
\]

where the \( S_i \) are coefficients of linear fits.

For the dual focal plane method, the transform matrix that links the \( x, y \) position between the two planes can also affect the fitting result. The transform matrix consists of a series of linear transforms: scaling, translation, and rotation. In order to obtain the correct transform matrix, a set of bead data is taken by moving a single bead along \( x, y \) in equal increments across the full imaging area. Then the \( x, y \) positions of bead in both planes were found by performing a
2D Gaussian fitting of a \(20 \times 20\) pixel subregion around the center of the bead image. Having obtained this set of coordinates in \(x, y\), we were able to calculate the transform matrix using the \textit{fminsearch} function in Matlab (MathWorks Inc.).

5. Experimental methods

5.1. Dual focal plane setup

A schematic illustration of our dual focal plane setup is shown in Fig. 3. The emission beam was collimated through lens L1, and divided into two optical paths using a beam splitter (BS). The transmission path was a 4f configuration, with \(f=125\) mm (L1, L2). A mirror, M2, was placed at the Fourier plane. In the reflection path, lens L3 (\(f=1000\) mm) was placed about 90 mm before the imaging lens L2 in order to create a second focal plane at the CCD. A filter wheel was placed before L1 for the selection of different emission spectra. We used a band-pass filter (FF01-692/40-25, Semrock) for imaging bead samples and cell samples labeled with Alexa 647 fluorophores. The camera was an EMCCD camera (iXon 860, Andor Technologies PLC.), with a CCD size of \(128 \times 128\) pixels and a pixel size of 16 \(\mu\)m (the actual pixel size is 24 \(\mu\)m, we used a 1.5 magnifier before L1, which resulted in a effective pixel size of 16 \(\mu\)m).

The microscope was an inverted microscope (IX71, Olympus America Inc.). We used a dichroic mirror (650 nm, Semrock) and a TIRF objective (UAPON 150XOTIRF, Olympus America Inc.) with \(NA=1.45\) for recording data used in Fig. 4, 5, 6, 9, 10, while a TIRF objective (UAPON 100XOTIRF, Olympus America Inc.) with \(NA=1.49\) was used for obtaining data for Fig. 11 and Fig. 12. Our setup was slightly modified during the paper preparation: L1 and L2 had \(f=100\) mm, M1 was removed, M2 was placed at 45° relative to the incoming beam, the camera was replaced by an EMCCD camera (iXon 897, Andor, Andor Technologies PLC.), with a CCD size of \(512 \times 512\) pixels and a pixel size of 16 \(\mu\)m. The new setup has a relative magnification close to 1 between both planes, so that the overall magnifications used through out the paper were not the same, as specified in each result.

5.2. Sample preparation

For the phase retrieval process, we used 20 nm diameter beads (FluoSpheres F-8782, crimson, Invitrogen) and 40 nm diameter beads (FluoSpheres F-8789, dark red, Invitrogen) diluted to \(1 : 10^9\) in deionized water. We added 5 \(\mu\)L of this diluted bead solution to 500 \(\mu\)L 1X PBS (phosphate-buffered saline) in a single well of a 8-well chambered coverslip (155411, LabTek, Nunc) and added 10 \(\mu\)L 1M NaCl to improve adherence of the beads to the coverslip surface. For measuring PSFs with refractive index mismatch aberration, we added dilute 40 nm diameter beads solution to 1X PBS medium with fixed RBL cells.

For whole-cell imaging, we used RBL cells prepared in an 8-well chambered coverslip kept at 37°C. We added 4 \(\mu\)L IgE-Alexa Fluor 647 (A-20006, Invitrogen) conjugate (with a dye to protein ratio of 2.4) to each well. The cells were kept at 37°C overnight, then rinsed with warm 1X PBS once, and fixed with 4% paraformaldehyde (PFA) in 1X PBS for 15 min. After fixation, the samples were rinsed with 10 mM Tris twice, allowing 5 minutes for each rinse. Immediately before imaging, the Tris buffer was replaced by a STORM imaging buffer (450 \(\mu\)L 10% (w/v) glucose in 50 mM Tris, 10 mM NaCl, pH 8.5; 50 \(\mu\)L oxygen scavenger buffer [14040U catalase (C9322–1G, Sigma Aldrich), 1688U glucose oxidase (G2133–50KU, Sigma Aldrich) in 50 mM Tris, 10 mM NaCl, pH 8.5]; 8 \(\mu\)L 1M MEA, pH 8.5) and capped with 150 \(\mu\)L paraffin oil to prevent the entry of additional oxygen.

RBL cells in an 8-well chambered coverslip were also used for microtubule imaging. In this case, the labeling steps were: 1) cells were rinsed with 1X PBS at 37°C and extracted with 0.1% Triton X-100, 3% BSA (bovine serum albumin) in 1X PBS for 10 s at room temperature (RT); 2) cells were fixed with 4% PFA (paraformaldehyde) in 1X PBS for 15 minutes at RT.
and rinsed 4 times with 10 mM Tris for 5 minutes each time; 3) cells were permeabilized with 0.1% Triton X-100, 3% BSA in 1X PBS for 15 minutes at RT; 4) cells were incubated with 10 µg/ml anti-alpha and anti-beta tubulin primary antibody (T6074, T8328, Sigma Aldrich) for an hour at 37°C, and rinsed 3 times with 1% BSA in 1X PBS for 5 minutes each time; 5) cells were incubated with 10 µg/ml secondary antibody (107299, Jackson ImmunoResearch) conjugated with Alexa Fluor 647 (dye/protein ratio of 2.52) for an hour at 37°C, and rinsed 3 times with 1% BSA in 1X PBS for 5 minutes each time; 6) cells were post fixed with 4% PFA for 15 minutes at RT, and rinsed 4 times with 10 mM Tris for 5 minutes each time; 7) cells were rinsed with 1X PBS and kept in 1X PBS, ready for imaging. Antibodies were diluted in 0.1% Triton X-100, 1% BSA in 1X PBS. The imaging buffer preparation was the same as for whole cell imaging.

5.3. Data acquisition

The bead sample was excited with a 637 nm laser (laser diode, HL63133DG, Thorlabs, with home build collimation optics) at an excitation intensity of 0.1 kW/cm². Data was acquired at a frame rate of 50 Hz, 100 frames per z position. The z = 0 µm position was set at the in-focus position of cover glass at plane 1, and data used in generating results in section 6 were acquired at z positions either from -1 µm to 1 µm, or from -2 µm to 2 µm, in increments of 100 nm by moving the sample with a piezo-driven nano-stage (Nano-LPS100, Mad City Labs Inc.) along z-axis.

The cell sample was first excited with a 637 nm laser at low intensity (0.03 kW/cm²), in order to find a region of interest. Then the excitation intensity was increased to 1.1 kW/cm², and the sample was illuminated slightly out of TIRF for a few minutes, allowing most of the fluorophores to switch to the dark state. At this point, data acquisition was started at a frame rate of 50 Hz, with 2500 frames per z position. The whole-cell image data was taken from the
cover glass to the top of the cell, with a \( z \) step size of 1 \( \mu m \). During data acquisition, a 405 nm laser (DL-405-020, CrystaLaser) was used to accelerate switching the fluorophores from the dark state to the active state. The laser excitation intensity was gradually adjusted from 0 to 6 W/cm\(^2\) to ensure a proper density of excited fluorophores.

### 5.4. Data preprocessing for 3D localization

Our 3D localization algorithm relies on that the photon counts at each pixel are Poisson distributed over time (expected photon counts in section 4). Thus, a CCD offset was subtracted from original data, and the results were multiplied by a fitted gain factor, as detailed in [35, 39]. After converting pixel values of original data into expected photon counts, we applied the method described in [36] to identify the single emitters in each frame. The method basically performed two filtering steps to reduce the Poisson noise and smooth out the data, then found the pixel coordinates of local maxima and set them as the centers of fitting regions. Each fitting region measured 16 \( \times \) 16 pixels, corresponding to a size of 2.87 \( \mu m^2 \) at the sample plane. Then, all the fitting regions and the pixel coordinates of the corresponding upper left corner of each fitting region are fed into the 3D localization algorithm.

In the case of 3D localization of bead data, the pixel coordinates of each fitting region center were set to a constant, and each PSF was centered in the fitting region, by shifting the single frames or the averaged bead data in the \( x-y \) plane with the same amount used in the phase retrieval process. The averaged bead data were generated by averaging over 100 frames at each axial position.

### 5.5. Drift correction

We used drift correction for constructing a superresolution image from image data acquired from a sample with fluorescence labeled microtubules. The data was collected in 27 consecutive sequences, where each sequence consisted of 1000 frames. For each sequence, a 2D image was reconstructed from all the accepted fitting results, by replacing each fitting point with a circularly symmetric Gaussian, with a sigma four times smaller than the radius of the Airy pattern. A reference image was generated by averaging over all 27 reconstructed images. The shifts between all 27 reconstructed images and the reference image were calculated by using the function \texttt{findshift} in the DIPimage toolbox [40]. The \texttt{findshift} function utilizes the method of gradient-based shift estimation [41]. The \( x, y \) shifts were calculated by reconstructing the images in the \( x-y \) plane, while \( z \) shifts were obtained from the images reconstructed in the \( x-z \) plane.

### 6. Results and discussion

#### 6.1. Phase retrieval result

Figure 4 shows the pupil function obtained from the phase retrieval process and Zernike fitting, as detailed in section 2, and the mPR-PSF generated from the Zernike fitted pupil function with OTF rescaling (section 3.1). mPR-PSFs match well with measured PSFs in both planes of the dual focal plane setup, and are free of noise and artifacts. Fitting the pupil functions with Zernike modes, eliminated the noise visible in the PR pupil functions. The pupil functions for plane 2 are slightly smaller than for plane 1, due to a slightly larger magnification. We found that the optimization procedure (see section 2.5) yielded the following values for the unknown parameters in the phase retrieval process of the data shown in Fig. 4: \( i_0 = 7, R_1 = 44, n = 1.5132, \) and \( \lambda = 670 \) nm. However, the optimization parameters vary slightly for different measured PSF data sets under the same experimental conditions, so we typically choose the phase retrieval result that produces the best fitting accuracy (section 6.3.1) for 3D localization.
Fig. 4. Pupil function and mPR-PSF for the dual focal plane setup. These images were obtained using the methods detailed in section 2 and section 3. mPR-PSFs were generated from Zernike fitted pupil function. Measured PSF images at $z$ positions as indicated were used for phase retrieval. For this particular data, the calculated defocus was 3.9 nm. The plane separation was 334.9 nm, $\lambda = 670$ nm, and $NA = 1.46$, while the magnifications at plane 1 and plane 2 were 147 and $147 \times 1.148$, respectively. The $x$-$z$ sections in the right-most column shows the measured PSF and mPR-PSF in the $x$-$z$ plane. The bottom images show the corresponding pupil functions in $k$ space. Here, the Zernike fitted pupil functions are constructed from 49 Zernike polynomials. The scale bar is 500 nm. The support of pupil function in $k$-space is $NA/\lambda$. 

---

Received 24 Sep 2013; revised 4 Nov 2013; accepted 7 Nov 2013; published 21 Nov 2013

(C) 2013 OSA
6.2. **PSF with index mismatch aberration**

![Image of measured PSF, mPR-PSF, and S-PSF at various depths](image)

**Fig. 5.** Measured PSFs and mPR-PSFs at various depths in the sample medium. The data were acquired at \( z \) positions from -2 to 2 \( \mu \text{m} \), with 100 nm step size. The images were generated by summing over the 100 frames taken at each \( z \) step. The displayed depths are the \( z \) positions found in the sample medium using the method presented in section 4. The mPR-PSF was retrieved from the measured PSF at the cover glass where the depth is 0 \( \mu \text{m} \), and the aberration phase is given by Eq. (9). The S-PSF was generated from an unaberrated pupil function.

Figure 5 shows different PSF models with refractive index mismatch aberration. Comparing the leftmost column to the center column in Fig. 5, it is clear that the mPR-PSF with aberration phase [Eq. (6)] matches the measured PSF quite well. The depth values listed in Fig. 5 indicate the bead’s distance from the cover glass in the sample medium. These values were found by fitting the measured PSFs using the method described in section 4. The refractive index of the sample medium and immersion medium were 1.33 and 1.52 respectively. For a bead at the cover glass, the measured and mPR-PSF show small spherical aberration. This spherical aberration is partially canceled by the refractive index mismatch aberration at the depth of \( \sim 1 \mu \text{m} \), so that the axial dimension of the PSFs are smaller at a depth of 0.87 \( \mu \text{m} \). The rightmost column was generated using the scalar PSF (S-PSF) model, which is calculated from the Fourier transform of an unaberrated pupil function. The scalar PSFs are sharper, as they do not include OTF rescaling (see section 3.1). A comparison of all three columns shows that the mPR-PSF matches the measured PSF well, and thus ensures better localization accuracy than the S-PSF.
In calculating both the mPR-PSF and the S-PSF, we used $NA = 1.46, \lambda = 670$ nm, a lateral magnification of $M = 149.5$ at both plane 1 and plane 2, and a CCD pixel size of $\Delta d = 16 \, \mu m$.

### 6.3. Localization of beads and simulated data

#### 6.3.1. Comparison of bead data localizations using various PSF models

Figure 6 shows the 3D localization results for one set of bead data, where beads were immobilized on a cover slip and imaged repeatedly at various stage positions. This data was also used to generate the phase retrieval results shown in Fig. 4. Figure 6(a) compares $z$ localization of bead data using various PSF models. The blue crosses and the superimposed red line show the fitting results obtained by using mPR-PSF (Fig. 4) as the PSF model for localization of the collected single frame data and the averaged data respectively, where the averaged data is the collected data after averaging over 100 frames at each $z$ position. The green dashed line shows the fitting results obtained by using the S-PSF model with OTF rescaling (see section 3.1) (mS-PSF), while the brown dashed line shows the fitting results obtained by using only the S-PSF model. Both the green and brown dashed lines show the fitting results on averaged data. The plots show that our localization algorithm based on the mPR-PSF model performs well within a $z$ range of $\sim 1.4 \, \mu m$, compared to the S-PSF model, which produces large deviations [Fig. 6(b-d)] in fitting accuracy. It is interesting to note that the S-PSF model plus OTF rescaling (mS-PSF) greatly improves the fitting accuracy, especially in $z$, which is also true for the PR-PSF model (not shown). This indicates that both the mS-PSF and mPR-PSF models are closer to the realistic PSF. The mPR-PSF model performs by far the best in fitting accuracy. Over a depth of $\sim 1.4 \, \mu m$, the average $z$-position deviates by less than 10 nm from the set $z$-positions [Fig. 6(b)], which is well below the typical axial localization precision achieved with single molecules; the localization precision in $z$ and $x, y$ are around 15 nm and 4 nm respectively [Fig. 6(e,f)], under an expected photon count of around 4000 photons at plane 1. It is important to note that because the beads adhere to the cover glass, the mPR-PSF model is only correct for analyzing data with samples located at or near the cover glass when imaging samples with refractive index mismatch. Thus, when imaging samples far from the cover glass, the mPR-PSF needs to be modified to account for aberrations caused by refractive index mismatch (see section 3.2 and section 6.2).

#### 6.3.2. Theoretical localization precision

The localization accuracy can be estimated by calculating the Cramér-Rao lower bound (CRLB) \[37, 42\],

$$\text{var}(\theta_i) \geq F_{ii}^{-1},$$

where $F$ is the Fisher information matrix. For a discrete likelihood function, each element in $F$ is given by

$$F_{ij} = E[\frac{\Delta \ln(L(\theta|D))}{\Delta \theta_i} \frac{\Delta \ln(L(\theta|D))}{\Delta \theta_j}],$$

$$\theta = \theta(x_1, y_1, z_1, I_1, bg_1, bg_2).$$

Assuming there is no correlation between any two different pixels and under a Poisson process, we have

$$F_{ij} = \sum_q \frac{1}{\mu_{1q}} \frac{\Delta \mu_{1q}}{\Delta \theta_i} \frac{\Delta \mu_{1q}}{\Delta \theta_j} + \sum_q \frac{1}{\mu_{2q}} \frac{\Delta \mu_{2q}}{\Delta \theta_i} \frac{\Delta \mu_{2q}}{\Delta \theta_j}. \tag{18}$$

Figure 7 shows the theoretical estimation error from the CRLB (solid line) of the 3D localization based on the mPR-PSF and the standard deviation (circles) of the found positions of simulated data using the same mPR-PSF model, which are in good agreement.
**Fig. 6.** 3D localization of bead data. Data was acquired at 100 frames per z position, from -1 µm to 1 µm. (a) z positions found from the bead data: using the mPR-PSF model on the collected single frame data (blue crosses); using the mPR-PSF model (red solid line), the mS-PSF model (green dashed line), and the S-PSF model (brown dashed line) on the averaged PSF data. (b-d) Deviations in the found x, y, z positions of the averaged PSF data from the true positions. Color representation is the same as in (a), and deviations beyond 50 nm are not shown in the plots. (e-f) Fitting precision in x, y, z, using the mPR-PSF model on the averaged PSF data.

**Fig. 7.** Fitting precision and CRLB for simulated data. Simulated data were generated from the mPR-PSF shown in Fig. 4, to which Poisson noise was added. $I_1 = 4000, b_{g1} = 4.5, b_{g2} = 3.5, I_{ratio} = 0.7, z = [-1, -0.9, ..., 0.7] \mu m$, 100 simulated single emitters per z position.
6.3.3. Localization speed

Figure 8 shows the speed of GPU-based localization for different numbers of Zernike coefficients and Newton-Raphson iterations. The plots were generated by fitting 1000 simulated single emitters at random $x$, $y$, $z$ positions. We used the following procedure: First, localization was performed using only the first Zernike coefficient and 25 iterations to generate an initial guess for the second fit. We systematically varied the number of Zernike coefficients and the number of Newton-Raphson iterations for the second fit, and calculated the corresponding total fitting times and the averaged SSE (sum of square errors) over all localizations. The SSE is defined as,

$$SSE_{fit} = \sum_q (N_q - \mu_q)^2,$$

where $N_q$ and $\mu_q$ are pixel values of the simulated data and the fitted model, respectively. The fitting time plots show that the total fitting time increases almost linearly with both the number of Zernike coefficients and Newton-Raphson iterations. From the SSE plots, we found that more than 25 Zernike coefficients (Fig. 8(a), fixed at 15 iterations) or more than 5 iterations (Fig. 8(b), fixed at 9 Zernike coefficients) do not decrease the SSE value significantly. We also noticed that the SSE drops significantly when using 9 Zernike coefficients while still being $\sim 2$ times faster than when using 25 Zernike coefficients. We, therefore, chose 15 iterations and 9 Zernike coefficients as the preferred parameters for the second fit, which yielded a fitting time of 15.5 s for 1000 localizations.

6.4. 3D whole-cell reconstruction

Figure 9 shows 3D whole-cell reconstructions using various PSF models. From the four images, it is obvious that using the mPR-PSF combined with the aberration phase [Eq. (6)] gives the best results [Fig. 9(a)]: the $z$ localization is continuous throughout the whole cell, and the number of accepted fits is greater than for the other three images [Fig. 9(b-d)] under the same rejection threshold. It is important to note that, with refractive index mismatch aberration, the...
actual $z$ position should always be positive, a negative $z$ position has no physical meaning under supercritical angle fluorescence. We therefore rejected all the fits with negative $z$ positions. The second best fitting result uses the mPR-PSF without the aberration phase. In this case, the $z$ fitting is continuous up to 3 $\mu$m inside the sample medium: above this, the $z$ fitting creates a separate band at each reference $z$ position, $z_0$ [Eq. (7)], and the whole cell appears a bit stretched along the $z$ axis. The separate bands imply that mPR-PSF alone doesn’t match well to the realistic PSF at depth greater than 3 $\mu$m. The bottom two images show the fitting results using the S-PSF model with no aberration: one [Fig. 9(c)] takes the reference $z$ position as defined in Eq. (7), and one [Fig. 9(d)] assumes the reference $z$ position is the same as the stage position. For the S-PSF model, the $z$ fitting is discontinuous throughout the whole cell, and the separate bands are much thinner compared to Fig. 9(b). This indicates that the S-PSF model only matches to a realistic PSF in a small $z$ range at each reference $z$ position. Only about 7% of the total fits are accepted, compared to 60% in Fig. 9(a). Furthermore, when Eq. (7) is not taken into account, the cell is stretched to occupy the entire axial distance through which the stage was scanned, while the actual size of the cell is much smaller. These observations illustrate that a realistic PSF model is essential in 3D localization.

6.5. 3D reconstruction of microtubules

Figure 10 shows the 3D localization of microtubules in a RBL cell using different PSF models. We see that the mPR-PSF model produces a relatively continuous and complete reconstruction of the labeled structures. In Fig. 10(b), the microtubule shows an artifactual step from using the S-PSF model, while it is continuous when using the mPR-PSF model. However, the $z$ localization precision [Fig. 10(c)] appears to be better under the S-PSF model. This probably arises from the localization bias towards certain $z$ positions under the S-PSF model. Another weakness of the S-PSF model is that it leads to a large divergence in the localization algorithm, thus resulting in a substantial number of unaccepted localizations. The empty region in Fig. 10(a) (the white box) from the S-PSF model exemplifies the detrimental effect of this large divergence. The $z$ localizations and photon counts before rejection [Fig. 10(d)] show that most fitting results under the S-PSF model diverge significantly from the mPR-PSF results.

6.6. Supercritical angle fluorescence effect

Figure 11(a,b) show the theoretical pupil function calculated from Eq. (34) (see Appendix B), at $d = 0$ and $d = \lambda$, respectively. Fig. 11(c) shows the Zernike-fitted PR pupil function, while Fig. 11(d) is Fig. 11(c) multiplied by the same decay term as in Fig. 11(b). In Fig. 11(a,c), with a single emitter located at the cover glass, both the theoretical and retrieved pupil functions have a bright ring near the edge due to supercritical angle fluorescence (SAF). However, Fig. 11(a) has greater contrast, and the ring is brighter towards the edge. This indicates a higher transmission loss in SAF during the experiment, so that the effect is weakened. In Fig. 11(b,d), with a single fluorophore at depth $d = \lambda$, both pupil functions become smaller due to the decay term introduced from index mismatch aberration (see section 3.2). For the theoretical model [Fig. 11(b)], the intensity decreases towards the edge, while there is still a partial bright ring left in the Zernike fitted PR pupil function [Fig. 11(d)]. This is probably caused by the noisy PR pupil function and the limitations of Zernike expansion in approximating effect of SAF. It is important to notice that here we used an objective lens with $NA = 1.49$, and a magnification of 100, while we, in Fig. 4, used an objective lens with $NA = 1.46$ and a magnification of 147. Comparing Fig. 11(c) with the Zernike fitted pupil function in Fig. 4, the SAF effect is more obvious in the former case. We assume that the phase retrieval process can retrieve the SAF effect, but under different optical conditions, the SAF effect may vary in strength. However, the following Figs. illustrate that a pupil function with a reduced SAF effect can also give
Fig. 9. 3D whole cell reconstruction using various PSF models. The sample consists of RBL cells labeled with IgE-Alexa 647 conjugate. Data was taken by scanning through the entire cell along the $z$ axis, with a step size of 1 µm, and recording 2500 frames per step. The basic reconstruction method is described in section 4. All units are in µm. The color map indicates the actual $z$ position and the color bars have the same color scale. (a) Reconstruction using the mPR-PSF and correcting for index-mismatch aberration. (b) Reconstruction using the mPR-PSF without index-mismatch correction. (c) Reconstruction using the S-PSF model. (d) Reconstruction using the S-PSF model and assuming the reference plane is the same as the stage position, without applying Eq. (7). In calculating PSF models, we used $NA = 1.46, \lambda = 690$ nm, a lateral magnification of $M = 148.6$ at plane 1 and $M = 148.6 \times 1.135$ at plane 2, and a CCD pixel size of $\Delta d = 16$ µm. The refractive index of sample medium and immersion medium are 1.33 and 1.53 respectively.
Fig. 10. 3D reconstruction of microtubules in an RBL cell. Data was acquired at a fixed $z$ position near the cover glass. (a) 2D scatter plots of the 3D localization using the mPR-PSF (left) and S-PSF (right) as PSF models. Color represents the $z$ position. (b) 2D scatter plot of the x-z projection of a single microtubule (the red box in (a)). (c) Probability plot of the $z$ localization of a small region (the red boxes) of the microtubule in (b). (d) Probability plots of $z$ localization (left) and expected photon counts from plane 1 (right) of the raw fitting results of the region in the white box in (a). The raw fitting results are the total localizations before rejection and drift correction. In calculating both the mPR-PSF and the S-PSF, we used $NA = 1.46$, $\lambda = 682.7$ nm, a refractive index of immersion medium $n = 1.56$, a lateral magnification of $M = 149.5$ at both plane 1 and plane 2, and a CCD pixel size of $\Delta \xi = 16$ µm.
acceptable fitting accuracy.

Figure 12 shows the Zernike expansion of the pupil function to different Zernike orders \( n \) [25], where the number of Zernike coefficients is \((n+1)^2\). The SAF effect is more apparent with higher order, and beyond \( n = 7 \), the pupil functions are quite similar, which indicates that using \( n = 7 \) is sufficient to approximate the SAF effect. However, in 3D localization, using a large number of Zernike coefficients will significantly decrease the localization speed. We found that increasing Zernike order doesn’t linearly increase fitting accuracy. Thus, it is better to use the lowest Zernike order that does not sacrifice fitting accuracy (see section 6.3.3).

Fig. 11. Pupil function magnitude with supercritical effect. (a)-(b) Theoretical calculation from Eq. (34), with depth \( d = 0 \) and \( d = \lambda \). (c) Zernike fitted PR pupil function. (d) Pupil function generated by multiplying the decay term in (b) to (c). The PSF data used in phase retrieval process were acquired at \( z \) positions from -2 to 2 \( \mu \text{m} \). In generating both theoretical and Zernike fitted pupil functions, we used \( \lambda = 670 \text{ nm} \), \( NA = 1.49 \), a lateral magnification of \( M = 100 \), and a CCD pixel size of \( \Delta d = 16 \mu \text{m} \). The refractive index of sample medium and immersion medium is 1.33 and 1.52 respectively.

Fig. 12. Zernike expansion of the magnitude of PR pupil function. It uses the set of Zernike coefficients in Fig. 11. \( n \) is the order of Zernike coefficients, and each image shows the Zernike expansion up to \( n \)th order. The number of Zernike coefficients is \((n+1)^2\). In generating Zernike fitted pupil functions, we used \( \lambda = 670 \text{ nm} \), \( NA = 1.49 \), a lateral magnification of \( M = 100 \), and a CCD pixel size of \( \Delta d = 16 \mu \text{m} \). The refractive index of immersion medium is 1.52.

7. Conclusion

In this paper, we have demonstrated a new 3D localization method based on a modified phase-retrieved PSF, and successfully applied this method to 3D whole-cell imaging. The mPR-PSF
is constructed from an experimentally obtained PSF, and therefore contains the aberrations of the optical system. Our PSF model is continuous and accurate in \( z \) over a range of \( \sim 1.4 \) \( \mu \text{m} \) and allows for fast on-the-fly modeling of the realistic PSF, thereby eliminating the need to store and access a large experimental data set during localization.

The demonstrated biological applications show that our localization algorithm can accurately localize a sufficient amount of fluorophores and generate a complete and accurate reconstruction of biological structures over large depths, something which is not possible with the tested simpler PSF models. We also showed that for 3D whole cell imaging, it is necessary to account for the aberration caused by refractive index mismatch. The results again illustrate that an accurate PSF model is essential for 3D localization. By implementing the PSF calculation on GPU, we were able to speed up the localization algorithm nearly 100 fold, thereby allowing to obtain 100,000 localizations in less than half an hour.

**Appendix A: Derivation of the diffraction integral expressed in Zernike polynomials**

In scalar diffraction theory, PSF of an imaging system can be expressed as an integral over all plane wave components:

\[
U(x, y, z) = \int \int P(k_x, k_y) e^{2\pi i (k_x x + k_y y)} e^{2\pi i k_z z} dk_x dk_y,
\]

where \( U(x, y, z) \) is the PSF at the sample plane and \( P(k_x, k_y) \) is the pupil function at the back focal plane of objective lens [18]. Converting Eq. (20) to polar coordinates, we obtain:

\[
U(r, \phi, z) = \int \int P(k_r, \theta) e^{2\pi i k_r \cos(\theta - \phi)} e^{2\pi i \sqrt{k_r^2 - k_z^2} r} dk_r d\theta.
\]

If we now let \( \theta - \phi \rightarrow \theta \):

\[
U(r, \phi, z) = \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} P(k_r, \theta + \phi) e^{2\pi i k_r r \cos \theta} d\theta.
\]

In anticipation of carrying out the \( \theta \) integration first, we expand \( P(k_r, \theta) \) in terms of the Zernike polynomials. The phase retrieval process returns the pupil function in the form \( A e^{i\Phi} \), but rather than expanding \( A \) and \( \Phi \) separately, we change the pupil function into the form as \( U + iV \), where \( U = A \cos(\Phi) \) and \( V = A \sin(\Phi) \). This is an exact transform, as long as \( \Phi \) is within \([ -\pi, \pi ]\), which is the case for our imaging system. The expansion of the pupil function is then:

\[
P(k_r, \theta) = \sum_{a=0}^{N} \sum_{m=0}^{n} R_n^m \left[ C_a \cos(m\phi + a\theta) + C_b \sin(m\phi + m\theta) \right],
\]

where \( C_a, C_b \) are the \( a \)th and \( b \)th complex coefficients of the Zernike polynomials. For a given \( n \) and \( m \), \( a = n^2 + 2(n - m) \) and \( b = n^2 + 2(n - m - 1) \). \( N \) is the maximum order of Zernike polynomials, and we have found that \( N = 6 \) is sufficient for our purposes, so there are \( (N+1)^2 = 49 \) coefficients. \( R_n^m \) is the radial part of the Zernike polynomials, and is given by:

\[
R_n^m = \sum_{s=0}^{n-m} (-1)^s \frac{(2n-m-s)!}{s!(n-s)!(n-m-s)!} \rho^{s+2(n-m-s)},
\]

where:

\[
\rho = \frac{k_r}{NA/\lambda}.
\]

Then applying trigonometric identities we have:

\[
P(k_r, \theta) = \sum_{a=0}^{N} \sum_{m=0}^{n} R_n^m \left[ C_a \cos(m\phi) \cos(m\theta) - C_a \sin(m\phi) \sin(m\theta) + C_b \sin(m\phi) \cos(m\theta) + C_b \cos(m\phi) \sin(m\theta) \right].
\]
Replacing $P(k_r, \theta)$ back into integral [Eq. (22)], we can use the identities,
\[
\int_0^{2\pi} \cos(m\theta)e^{2\pi ik_r \cos\theta} d\theta = 2\pi i^m J_m(2\pi k_r r),
\]
\[
\int_0^{2\pi} \sin(m\theta)e^{2\pi ik_r \cos\theta} d\theta = 0.
\]
to evaluate the $\theta$ integral, leaving a one dimensional integral in $k_r$:
\[
U(r, \phi, z) = 2\pi \int_0^{NA/\lambda} dk_r k_r e^{2\pi i z \sqrt{k_r^2 - k_0^2}} \sum_{n,m} R_{nm}^m J_m(2\pi k_r r) [C_a \cos(m\phi) + C_b \sin(m\phi)].
\]
(27)
The above integral is easily modified to include the aberration phase [Eq. (6)], so
\[
U(r, \phi, z) = 2\pi \int_0^{NA/\lambda} dk_r k_r e^{i\phi_{aber}} \sum_{n,m} R_{nm}^m J_m(2\pi k_r r) [C_a \cos(m\phi) + C_b \sin(m\phi)].
\]
(28)

Equation (27) and Eq. (28) can then be solved by numerical integration methods.

**Appendix B: Supercritical angle fluorescence effect**

Oil immersion objectives have the advantage of high photon collection efficiency. However, by using oil immersion objectives, a phenomenon called supercritical angle fluorescence (SAF) needs to be considered in the PSF model. In the results section 6.6, we show that the PR pupil function can retrieve the SAF effect as shown in the theoretical model. In this section we give the derivation of the theoretical model of the SAF effect.

In TIRF microscopy, oil immersion objectives with a high numerical aperture (NA) are used to enable total internal reflection (TIR) at the cover glass-sample interface, resulting in an evanescent field reaching into the sample medium. However, if we consider the inverse of TIR, the evanescent waves that originate from fluorescence dipole emission can penetrate the interface and become a propagating wave at an angle greater than the critical angle. This effect is called supercritical angle fluorescence (SAF) [28–31]. For simplicity, we consider a two layer system consisting of the sample medium and the immersion medium, with refractive indices of $n_2$ and $n_1$ respectively, where $n_2 < n_1$.

We consider a single fluorophore located at a distance $z = d$ away from the interface of the two medium system. It is convenient to use the plane wave decomposition of the fluorescence dipole emission, and for simplicity, instead of the complete expression of dipole emission, we assume all its plane wave components have the same magnitude. For a plane wave originating from the fluorophore, its propagation and transmission can be expressed in the matrix form [43],
\[
\begin{bmatrix}
E_r \\
E_i
\end{bmatrix} = \frac{1}{\tau_{21}} \begin{bmatrix}
e^{ik_2 d} & 0 \\
0 & e^{-ik_2 d}
\end{bmatrix} \begin{bmatrix}
1 \\
\rho_{21}
\end{bmatrix} \begin{bmatrix}
C_a \\
C_b
\end{bmatrix},
\]
\]
(29)
where $e^{ik_2 d}$ is the propagation phase of the wave propagating from the fluorophore to the interface, and $k_2 = 2\pi n_2/\lambda \cos \theta_2$, $\rho_{21}$ and $\tau_{21}$ are the Fresnel reflection and transmission coefficients at the interface. $E_i$, $E_r$ and $E_t$ represent the incident, reflected and transmitted waves. From the above Eqs. we can obtain the relationship between $E_i$, $E_r$, $E_t$ and $E_i$,
\[
E_t = \tau_{21} e^{ik_2 d} E_i,
\]
(30)
where $\tau_{21}$ is equivalent to the pupil function given all plane wave components have the same magnitude. For real optical systems, $\tau_{21}$ could be retrieved from the phase retrieval process.
Under SAF fluorescence, the Fresnel transmission coefficients [43] are still valid:

\[
\tau_\perp = \frac{2i\gamma}{1 + i\gamma}, \\
\tau_\parallel = \frac{2i\gamma^{22}_{21}}{i\gamma + \frac{2n_2}{n_1}}, \\
\gamma = \sqrt{(n_1 \sin \theta_1)^2 - n_2^2} \left/ n_1 \cos \theta_1 \right.,
\]

(31)

where \(\tau_\perp, \tau_\parallel\) are the transmission coefficients for perpendicular and parallel polarization. Hence \(\tau_{21} = (\tau_\perp + \tau_\parallel)/2\), assuming the incident wave is unpolarized. \(k_{2z}\) becomes imaginary,

\[
k_{2z} = \frac{2\pi}{\lambda} \sqrt{(n_1 \sin \theta_1)^2 - n_2^2} = i/\delta.
\]

(32)

Then the transmitted wave becomes,

\[
E_t = \frac{\tau_\perp + \tau_\parallel e^{-d/\delta}}{2} E_i,
\]

(33)

where the propagation term in Eq. (30) becomes a decay term, \(e^{-d/\delta}\), with a skin depth \(\delta\). Since \(\delta\) is comparable to the wavelength of the emission light [Eq. (32)], the magnitude of the evanescent wave \((e^{-d/\delta} E_i)\) will decay to zero with \(d \gg \lambda\) before the wave is transmitted through the sample-immersion interface. Thus, SAF only exists in the near interface region. In order to illustrate this confined fluorescence, we include the decay term in the pupil function. Then, the pupil function with the SAF effect can be expressed as,

\[
P = \begin{cases} 
\frac{(\tau_\perp + \tau_\parallel)/2}{e^{-d/\delta} (\tau_\perp + \tau_\parallel)/2} & n_1 \sin \theta_1 \leq n_2 \\
\end{cases} \begin{cases} 
& n_2 < n_1 \sin \theta_1 \leq NA.
\end{cases}
\]

(34)

Acknowledgment

S. L. and K. A. L were supported by NIH grant R21GM104691 and the New Mexico Spatiotemporal Modeling Center NIH P50GM085273. K. A. L was additionally supported by NSF CAREER award #0954836. W.D.K was supported by NIH grant R01NS071116. This work was additionally supported by the Wellcome Trust (095927/A/11/Z). E. B. K. was supported by: The Denmark-America Foundation (Coloplast Stipend), Civilingeniør Frants Allings Legat, Knud Højgaards Fond, Reinholdt W. Jorcks Fond, BergNielsens Legat, Ingeniør Alexandre Haynman og Hustru Nina Haynmans Fond. J. B. discloses significant financial interest in Vutara Inc.