Characterizing birefringent pupils in widefield fluorescence microscopy using phase retrieval

R. Gutiérrez-Cuevas[1, 2, *]

[1] *Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, UMR 7249, 13397 Marseille Cedex 20, France*
[2] *Institut Langevin, ESPCI Paris, Université PSL, CNRS, 75005 Paris, France*

(Dated: October 4, 2022)

Here we present the theory and the different strategies for retrieving the pupil of in fluorescence microscopy where the effects of birefringence play an important role. We also study the effect of separating the signal into different polarization channels.

I. INTRO

General intro about fluorescence microscopy, SMOLM and PSF engineering. Finish by saying something along the lines of: However, any aberration or misalignment in the system can affect the final shape of the PSFs and thus lead to an inaccurate estimate of the parameters.

A common solution to this problem is to perform a set of calibration measurements with a known source and at varying focal planes [1, 2]. From these measurements and an accurate model for the propagation of the light emitted by the source to the camera, the aberrations can be estimated through phase retrieval algorithms. Initially, only scalar models with point sources were used which allowed the successful use of both iterative and nonlinear optimization routines for estimating the aberrations [1–3]. However, more accurate models [4, 5] take into account the vectorial nature of light which is essential to describe the propagation of the emitted radiation through the interface between the suspension medium and the immersion liquid used for the high-NA microscope objective [6]. Moreover, it is also necessary to consider that the most common sources, such as fluorescent nanobeads, emit incoherent light and their size can thus lead to a noticeable blurring of the PSFs. Given these more accurate assumptions, iterative algorithms cannot be directly applied and have to be adapted which can make them unstable, although some success has been demonstrated [4]. A more natural approach is to use nonlinear optimization [5, 7] since it provides us with the freedom to incorporate other unknown parameters into the estimation, such as the photobleaching amplitudes or the background illumination.

Up until now, all works have assumed a scalar pupil to characterize the aberrations or to design new PSFs. However, SMOLM requires the use of birefringent elements to encode the orientation information of the emitting dipole into the shape of the two polarization components of the PSF. Therefore, it is important to update the model of propagation for an accurate characterization or design of SMOLM systems. In this work, we propose a characterization technique and algorithm to estimate the polarization aberrations of the system [8] as well as other unknown (or poorly known) parameters. This is done by applying a nonlinear optimization algorithm to a physical model where the aberrations (or widow we want to design) are now represented by a Jones matrix. The implementation is done with the neural network framework `PyTorch` [] which performs automatically all the gradient computation []. Additionally, we show that, in general, it is necessary to introduce polarization diversity in the measurements in order to properly characterize the polarization response. This is similar to the introduction of phase diversity by taking images at various focal planes. Additionally, we implement the most useful method for taking into account the blurring to the size of the nanobead discussed in []. It should also be noted that all the results obtained in this work were achieved with the software `torchPSFstack` which is freely accessible at [].

II. MODELING THE PSFS FOR FLUORESCENCE MICROSCOPY

A. Field at the back-focal plane

The first step is to have the most accurate model for the propagation of light through our microscope. As shown in Fig. **??**, the source is embedded in a medium of index of refraction n_i at a distance $d_{cs} > 0$ from the interface with immersion oil of index of refraction n_f, the interface is given by the coverslip which is assumed to be index matched to the immersion oil. For a dipolar source, the green tensor at the back-focal plane (BFP) of a high-NA objective can be computed through Richards-Wolf diffraction theory for aplanatic systems [6**?**]. An important step in this calculation is to take into account the interface between the two media since it introduces spherical aberration and the coupling of evanescent waves when $n_f > n_i$, known as supercritical angle fluorescence (SAF), which can contribute to more than half of the radiation reaching the detector [9–11].

Following [6, 12] and assuming that light propagates along the positive $\hat{z}$ direction, the Green tensor at the

* rodrigo.gutierrez-cuevas@fresnel.fr

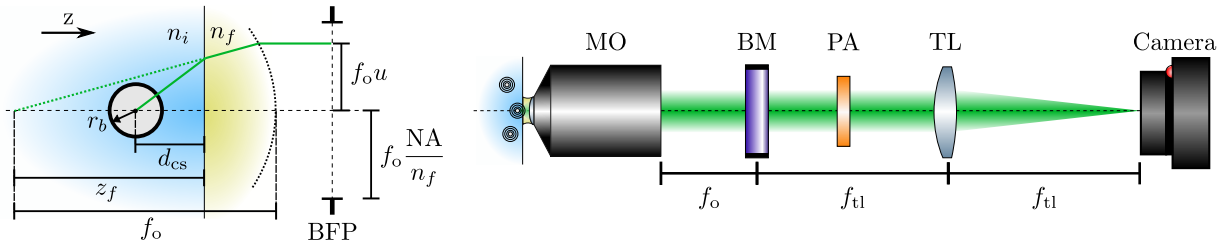

FIG. 1. .

BFP,

$$\mathbf{G}_0(\mathbf{u}) = e^{-ikn_f \boldsymbol{\rho}_o \cdot \mathbf{u}} \exp\left[ikn_i d_{cs}\sqrt{1-\left(\frac{n_f u}{n_i}\right)^2}\right] \times \exp\left(ikn_f z_f \sqrt{1-u^2}\right)\mathbf{g}(\mathbf{u}), \quad (1)$$

where

$$\mathbf{g}(\mathbf{u}) = \begin{pmatrix} g_{xx}(\mathbf{u}) & g_{xy}(\mathbf{u}) & g_{xz}(\mathbf{u}) \\ g_{yx}(\mathbf{u}) & g_{yy}(\mathbf{u}) & g_{yz}(\mathbf{u}) \end{pmatrix} \quad (2)$$

with the explicit form of its components given in the Appendix, $\boldsymbol{\rho}_o = (x_o, y_o)$ denotes the transverse location of the dipole, $k = 2\pi/\lambda$ is the wavenumber with λ being the wavelength, and z_f is the location of the focal plane from the interface, $z_f < 0$ ($z_f > 0$) if the focal plane is in the medium with index of refraction n_i (n_f). The vector $\mathbf{u} = (u_x, u_y)$ denotes the normalized coordinates at the BFP, see Fig. **??**. The maximum value of $u = \|\mathbf{u}\|$ is limited by the NA through $u_{\max} = \mathrm{NA}/n_f$. The matrix elements g_{ij} include the effect of the Fresnel coefficients of the boundary and depend on $\mathbf{u}$ but not on the location of the dipole or the focal plane. For a dipole oriented along the unit vector $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)$, the electric field distribution at the BFP is given by

$$\mathbf{E}_0(\mathbf{u}) = \mathbf{g}(\mathbf{u}; \mathbf{r}_0) \cdot \boldsymbol{\mu}. \quad (3)$$

Therefore, the three columns of the Green tensor represent the field distribution produced by a dipole along each of the three coordinate axes.

B. Propagation from the BFP to the image plane

In order to encode information about the position and orientation of the dipole into the shape of the PSF, it is necessary to include a mask into the path of the emitted light at the BFP with the possible help a relay system as shown in Fig. 1. The most general case is that of a birefringent mask represented by a 2×2 Jones matrix [],

$$\mathbf{J}_{\mathrm{M}}(\mathbf{u}) = \begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix}. \quad (4)$$

Note that a scalar mask case with apodization is regained if $q_j = 0$ for $j = 1, 2, 3$, and a pure phase mask if also $q_0 = 1$. After multiplying the Green tensor by this Jones matrix, the field can then be propagated form the BFP to the image plane through

$$\mathbf{G}_{\mathrm{IP}}(\boldsymbol{\rho}) = \iint \mathbf{J}_{\mathrm{M}}(\mathbf{u}) \cdot \mathbf{G}_0(\pm\mathbf{u}) e^{-ik\frac{n_f \boldsymbol{\rho}}{M}\cdot\mathbf{u}} d\mathbf{u}, \quad (5)$$

where M is the total magnification of the system. Note that the normalized coordinate is defined at the location of the birefringent mask. Assuming an incoherent source, the PSF is then given by

$$I_{\mathrm{IP}}(\boldsymbol{\rho}) = \|\mathbf{G}_{\mathrm{IP}}\|^2 = \sum_{i=x,y}\sum_{j=x,y,z} |G_{\mathrm{IP},ij}(\mathbf{u})|^2. \quad (6)$$

III. MODELING FOR THE PHASE RETRIEVAL

A. Polarization aberrations

The decomposition of polarization aberrations by a Jones matrix can be done in the following manner

$$\mathbf{J}_{\mathrm{A}}(\mathbf{u}) = e^{i2\pi W(\mathbf{u})} \begin{pmatrix} q_0(\mathbf{u}) + iq_3(\mathbf{u}) & q_2(\mathbf{u}) + iq_1(\mathbf{u}) \\ -q_2(\mathbf{u}) + iq_1(\mathbf{u}) & q_0(\mathbf{u}) - iq_3(\mathbf{u}) \end{pmatrix}, \quad (7)$$

which allows separating the scalar aberrations, contained in W from the vectorial correction given by the q's. Since aberrations tend to be described by smooth functions, it generally suffices to decompose the various elements of the Jones matrix using the Zernike polynomials which constitute a complete basis on the unit disk. Therefore, we write

$$W(\mathbf{u}) = {\sum_l}' c_l^{(W)} Z_l(\mathbf{u}/u_{\max}), \quad (8)$$

$$q_j(\mathbf{u}) = \sum_l c_l^{(j)} Z_l(\mathbf{u}/u_{\max}), \quad (9)$$

where $j = 0, ..., 3$, and we used a single index notation for the basis. Note that $\sum'$ in the expression for W indicates that the terms corresponding to piston and defocus should be excluded, this is automatically handled by the software []. The piston term only fixes a global phase which cannot be determined from intensity measurements while the defocus term is redundant with

the more accurate defocus parameter z_f in Eq. 1. This Zernike expansion is inspired by the Nijboer-Zernike theory [13–15] where a scalar mask would be separated into real and imaginary parts before decomposing in terms of Zernike polynomials. Note that other models can be used, such as pixel by pixel decomposition, but a Zernike decomposition with enough terms should be able to handle most cases found in microscopes. Moreover, computationally there is no advantage to considering a pixel decomposition since the number of DFT (the most costly operation) would be the same. It should also be noted that a scalar mask can easily be modeled by just setting the coefficients $c_l^{(j)} = 0$ for all l and $j = 1, 2, 3$.

B. Best focus and the distance to the coverslip

In general the position chosen as the nominal focus (or best focus) z_f and distance to the coverslip of the emitter d_{cs} are not perfectly known. Therefore, it is worth considering them as part of the optimization parameters that will be estimated along with the coefficients of the Zernike decomposition of the birefringent mask. The most obvious solution is to directly use z_f and d_{cs} as parameters, however both phase terms have similar effect when the radial coordinate u is below the SAF radiation and this might cause the nonlinear optimization to fall into a local minimum which does not provide the real values. Moreover, the estimation of the distance of the best focus z_f will generally depend on the distance of the nanobead from the coverslip. For example, if the paraxial approximation is used then the distance to the best focus will be $z_f = -n_f d_{\text{cs}}/n_i$. A similar result will hold even if we consider the spherical aberration induced by the interface, the only thing that changes is the factor in front of d_{cs}. Therefore, it is best to consider instead

$$\mathbf{G}_0(\mathbf{u}) = D_{\delta,\alpha}(\mathbf{u})\mathbf{g}_0(\mathbf{u}), \tag{10}$$

where

$$D_{\delta,\alpha}(\mathbf{u}) = \exp\left\{\mathrm{i}2\pi n_f \delta\left[\frac{n_i}{n_f}\sqrt{1-\left(\frac{n_f u}{n_i}\right)^2} - \alpha\sqrt{1-u^2}\right]\right\}, \tag{11}$$

with $\delta = d_{\text{cs}}/\lambda$ and $z_f = -\alpha d_{\text{cs}}$. Note that it has been assumed that the source is placed along the optical axis so that $\boldsymbol{\rho}_0 = (0, 0)$. Any deviation from this assumption will be corrected be the scalar tilts given by the tilt Zernike polynomials in the phase term W of $\mathbf{J}_A$. The parameter δ now predominantly controls the amount of exponential decay for the SAF radiation while α controls the defocus of the system. For fluorescing nanobead samples it is safe to assume that they are all fixed to the coverslip so that their distance to the coverslip can be taken as their radius and can be taken off the optimization routine. However, the choice of the best focus is quite subjective and should therefore always be included.

C. Phase and polarization diversity

In phase retrieval algorithms for optical microscopes, it is common practice to assume that we have access to a stack of intensity images at varying focal distances separated by Δz_ζ from the location of the best focus. This varying focal distances are taken into account by multiplying the Green tensor by the phase factor

$$D^\zeta(\mathbf{u}) = \exp\left[\mathrm{i}kn_f\Delta z_\zeta\sqrt{1-u^2}\right] \tag{12}$$

This additional information, referred to as phase diversity, helps the algorithm converge to an appropriate solution without falling into local minima and helps discriminate between the right and left phase vortices. These measurements are sufficient when the aberrations are taken as scalar, however when birefringence effects need to be taken into account it is necessary to implement a method to also provides information about the polarization state of the PSFs for each focal distance. This supplementary information can be obtained by inserting a polarization analyzer right after the birefringence mask (see Fig. 1). This polarization analyzer can be composed of a combination of waveplates and polarizers where at least one element can rotate in order to change the polarization projection of the output. This polarization diversity is modeled by a set of constant Jones matrices $\mathbf{P}^{(p)}$ that is applied after all other birefringent masks. Therefore, the stack of Green tensors at the BFP is given by

$$\mathbf{G}_{\text{BFP}}^{(\zeta,p)}(\mathbf{u}) = D^{(\zeta)}(\mathbf{u})\mathbf{P}^{(p)} \cdot \mathbf{J}_{\text{A}}(\mathbf{u}) \cdot \mathbf{J}_{\text{M}}(\mathbf{u}) \cdot \mathbf{G}_0(\mathbf{u}). \tag{13}$$

In this model $\mathbf{J}_{\text{M}}$ represents any known birefringent mask into the nominal Green tensor which can have approximately known parameters that can be incorporated into the optimization routine. It is worth noting that while experimentally the polarization diversity happens at the BFP, computationally it is better to perform it at the image plane in order to avoid the computation of unnecessary DFTs.

D. Modeling the total measured intensity

To model the intensity measured by the camera, the Green tensor is first propagated to the image plane via,

$$\mathbf{G}_{\text{IP}}^{(\zeta,p)}(\boldsymbol{\rho}) = \iint \mathbf{G}_{\text{BFP}}^{(\zeta,p)}(\mathbf{u}) e^{-\mathrm{i}k\frac{n_f\boldsymbol{\rho}}{M}\cdot\mathbf{u}}\mathrm{d}\mathbf{u}. \tag{14}$$

Here, it will be assumed that the source emits fully unpolarized light. This is the case for fluorescing nanobeads which are commonly used to characterize or test fluorescence microscopes since they have a stronger signal than that emitted by single fluorescing molecules. In this case

the measured intensity is given by the incoherent sum of the PSFs produced by dipoles oriented along each of the three Cartesian axes which is the same as the squared Frobenius norm of the Green tensor at the image plane

$$I_{\mathrm{IP}}^{(\zeta,p)}(\boldsymbol{\rho}) = \left\| \mathbf{G}_{\mathrm{IP}}^{(\zeta,p)} \right\|^2 = \sum_{i=x,y} \sum_{j=x,y,z} |G_{\mathrm{IP,ij}}^{(\zeta,p)}(\mathbf{u})|^2. \quad (15)$$

Depending on the size of the source it might be necessary to perform one of the blurring operations described in []. If a three-dimensional blurring is needed the computation of supplementary quantities will need to added to the forward model, for example, for the exact hard-sphere model we would also need to compute the total intensity for other values of $d_c s$. The downside is that this would slow the algorithm considerably. However, as long as their diameter is smaller than 30nm one can safely skip this step. More details about the blurring models implemented can be found in the Supplementary Information.

As a last step for computing the measured intensity, we also consider the effect of photobleaching of the fluorescing nanobeads and the background illumination. The photobleaching causes the number of photons emitted by the nanobead to diminish with time. Its effect can be taken into account by implementing an overall amplitude factor $\mathcal{A}^{(p,\zeta)}$ which depends on both the phase and polarization diversities. The background illumination is then added incoherently to the photobleached PSF stack. The simplest model is to assume that the background illumination is constant across each intensity image and determined by the term $\mathcal{B}^{(p,\zeta)}$. Therefore, the final total measured intensities of the ZP stack are given by

$$I_{\mathrm{tot}}^{(\zeta,p)}(\boldsymbol{\rho}) = a^{(p,\zeta)} I_{\mathrm{IP}}^{(\zeta,p)}(\boldsymbol{\rho}) + b^{(p,\zeta)}. \quad (16)$$

It is possible to assume a more complicated model for the background illumination, such as a quadratic expansion [16].

E. Cost function

The last piece we need to consider is the choice of a cost function that compares how good our modeled PSFs $I_{\mathrm{tot}}^{(\zeta,p)}$ are compared to the measured ones $I_{\mathrm{exp}}^{(\zeta,p)}$. Under the absence of noise, any choice of cost function that has a minimum when the two quantities are the same should provide the same result. However, noise is ever present in experimental measurements and thus needs to be taken account. If the noise in the camera follows a Poisson distribution then the log-likelihood cost function,

$$\mathcal{C}_{\mathrm{LL}} = \sum_{\zeta,p} \iint w(\boldsymbol{\rho}) \left\{ I_{\mathrm{exp}}^{(\zeta,p)}(\boldsymbol{\rho}) \log \left[I_{\mathrm{tot}}^{(\zeta,p)}(\boldsymbol{\rho}) \right] - I_{\mathrm{tot}}^{(\zeta,p)}(\boldsymbol{\rho}) \right\} \mathrm{d}^2\boldsymbol{\rho}, \quad (17)$$

Whereas if it follows a Gaussian distribution then the sum of the squared difference is the most appropriate,

$$\mathcal{C}_{\mathrm{SS}} = \sum_{\zeta,p} \iint w(\boldsymbol{\rho}) \left[I_{\mathrm{exp}}^{(\zeta,p)}(\boldsymbol{\rho}) - I_{\mathrm{tot}}^{(\zeta,p)}(\boldsymbol{\rho}) \right]^2 \mathrm{d}^2\boldsymbol{\rho}. \quad (18)$$

A window function w has been introduced to represent the use of a finite region at the image plane and to exclude bad pixels if there are any. It is also worth making a technical note. In order for the choice of cost function to make sense we need to use as $I_{\mathrm{exp}}^{(\zeta,p)}$ the values that actually follows the assumed distribution. In general this means that the images should not be denoised and that the offset of the camera should be removed. A detailed description for the forward model can be found in the Supplementary Information as well as other considerations that need to be taken into account before launching the optimization routine.

IV. NUMERICAL EXPERIMENTS

To exemplify the implementation of the phase retrieval algorithm, two main phase masks will be considered: (2) a stress-engineered optical (SEO) element [] and (3) an s-plate []. Both have been used in practice to measure the position and orientation of dipolar emitters []. All the modelled data stacks are computed using `pyPSFstack` which allows computing the PSF stack with phase and polarization diversities, and can compute the three-dimensional blurring due to the size of the fluorescing nano-bead. It can also simulate the effect of the background illumination and noise. The polarization diversity will be taken as the Jones matrix given by a quarter-wave plate at various angles followed by a horizontal or vertical linear polarizer. This choice is inspired by the setup used in [] where the SEO is followed by a quarter-wave plate and then a Wollaston prism separates the two linearly-polarized components.

A. Polarization diversity VS more phase diversity

Let us first compare the use of polarization diversity when characterizing the pupil from

The data is modelled assuming the source is a 20nm fluorescing nano-bead, with an average of 10000 photons hitting our detectors per PSF to which we add 50 photons per pixel as the background illumination. Noise following a Poisson distribution is then added to the images. Under these conditions, an image defocused by more than 500nm past the nominal focal plane will be indistinguishable from noise. Therefore, images are taken from -500nm to 500nm with a step size of 100nm.

B. Incorporating blurring due to size to the pupil retrieval

V. CHARACTERIZING THE POLARIZATION ABERRATIONS IN AN EXPERIMENTAL SETUP

ACKNOWLEDGMENTS

S. Paine for useful discussions.

Appendix A: Expressions for the Green tensor at the BFP

As outlined in [6], a closed-form for the Green tensor at the BFP for a dipolar source placed close to an interface can be obtained. In particular the components of the **g** tensor in Eq. (**??**) are given by

$$\mathbf{g}(\mathbf{u}) = \frac{1}{(1-u^2)^{1/4}} \begin{pmatrix} \cos^2\phi\sqrt{1-u^2}\Phi_2 + \sin^2\phi\Phi_3 & \cos\phi\sin\phi(\sqrt{1-u^2}\Phi_2 - \Phi_3) & -u\cos\phi\Phi_1 \\ \cos\phi\sin\phi(\sqrt{1-u^2}\Phi_2 - \Phi_3) & \sin^2\phi\sqrt{1-u^2}\Phi_2 + \cos^2\phi\Phi_3 & -u\sin\phi\Phi_1 \end{pmatrix}. \tag{A1}$$

where

$$\Phi_1(u) = t^p(u)\frac{n_i k_i\sqrt{1-u^2}}{n_m\sqrt{k_m^2 - k_i^2u^2}} \tag{A2a}$$

$$\Phi_2(u) = t^p(u)\frac{n_i}{n_m} \tag{A2b}$$

$$\Phi_3(u) = t^s(u)\frac{k_i\sqrt{1-u^2}}{\sqrt{k_m^2 - k_i^2u^2}} \tag{A2c}$$

with

$$t^s(u) = \frac{2k_{m,z}}{k_{m,z} + k_{i,z}}, \tag{A3a}$$

$$t^p(u) = \frac{n_m}{n_i}\frac{2n_i^2k_{m,z}}{n_i^2k_{m,z} + n_m^2k_{i,z}}. \tag{A3b}$$

being the Fresnel coefficients for p and s polarized light.

[1] B. M. Hanser, M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, Optics Letters **28**, 801 (2003).
[2] B. M. HANSER, M. G. L. GUSTAFSSON, D. A. AGARD, and J. W. SEDAT, Journal of Microscopy **216**, 32 (2004).
[3] N. A. Clark, *Microscope Characterization using Phase Retrieval Applied to Determine the Spatial Distribution of Membrane-Associated Proteins in Hematocytes*, Ph.D. thesis, University of Rochester (2012).
[4] N. H. Thao, O. Soloviev, and M. Verhaegen, Journal of the Optical Society of America A **37**, 16 (2019).
[5] B. Ferdman, E. Nehme, L. E. Weiss, R. Orange, O. Alalouf, and Y. Shechtman, Optics Express **28**, 10179 (2020).
[6] L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, 2006).
[7] J. R. Fienup, Applied Optics **21**, 2758 (1982).
[8] E. W. Hansen, in *Polarization Considerations for Optical Systems*, edited by R. A. Chipman (SPIE, 1988).
[9] E. H. Hellen and D. Axelrod, Journal of the Optical Society of America B **4**, 337 (1987).
[10] D. Axelrod, Traffic **2**, 764 (2001).
[11] D. Axelrod, Biophysical Journal **104**, 1401 (2013).
[12] M. A. Lieb, J. M. Zavislan, and L. Novotny, Journal of the Optical Society of America B **21**, 1210 (2004).
[13] A. J. E. M. Janssen, Journal of the Optical Society of America A **19**, 849 (2002).
[14] J. J. M. Braat, P. Dirksen, A. J. E. M. Janssen, and A. S. van de Nes, Journal of the Optical Society of America A **20**, 2281 (2003).
[15] J. J. Braat, P. Dirksen, A. J. Janssen, S. van Haver, and A. S. van de Nes, Journal of the Optical Society of America A **22**, 2635 (2005).
[16] A. Aristov, B. Lelandais, E. Rensen, and C. Zimmer, Nature Communications **9** (2018), 10.1038/s41467-018-04709-4.