Stokes-based analysis for the estimation of 3D dipolar emission

Isael Herrera, Luis Alemán-Castañeda, Sophie Brasselet, Miguel Alonso

Stokes-based analysis for the estimation of 3D dipolar emission.

ISAEL HERRERA,[1] LUIS A. ALEMAN-CASTANEDA,[1] SOPHIE BRASSELET, [1] MIGUEL A. ALONSO,[1,2,*]

[1]*Institut Fresnel, Aix Marseille Univ, CNRS, Centrale Med, Marseille, France;*
[2]*The Institute of Optics, University of Rochester, Rochester, NY, USA.*
[*]*miguel.alonso@fresnel.fr*

Abstract: We provide a general description of the measurement capabilities of systems that probe the 3D state of polarization of light emitted by a dipole or a collection of dipoles. This analysis is based on a generalization of the Stokes parameters for 3D polarization, and its goal is to provide insight into what constitutes a good measurement system under specific circumstances, through the definition of appropriate merit functions. Three cases are considered: the general case of arbitrary states of 3D polarization, the special case of 3D linear full or partial polarization states, and the even more specific case of linear dipoles that wobble with rotational symmetry around a central direction. Note that the latter two cases are of interest in fluorescence microscopy. The analysis presented here is illustrated by applying it to two different approaches used commonly in orientation microscopy: PSF engineering and radiometric measurements.

1. Introduction

In addition to intensity and phase, the local description of an optical field involves polarization, which is the characterization of the geometry of the oscillations of the electric field vector. For paraxial light, these oscillations are essentially constrained to a plane normal to the main direction of propagation. The statistical description of polarization is based on three parameters that can be inferred from intensity measurements, and that determine the main orientation of the oscillations, their average ellipticity/handedness, and the uniformity of oscillation (known as the degree of polarization). On the other hand, the characterization of an optical field that is not constrained to the paraxial regime requires a description of the oscillations of its three Cartesian components. This is the case, for example, of fields in focal regions or at the vicinity of sources. The geometric description of polarization becomes more complex, and up to eight parameters are required to achieve the same level of characterization [1–3]. The measurement of these parameters can be referred to as nonparaxial polarimetry. Many approaches have been used to produce arbitrary 3D polarization states [4–8], typically involving highly focused beams and/or evanescent wave components. Characterizing these 3D polarization states generally requires analyzing the field that is either scattered by nano-probes [5–7] or emitted by fluorescent molecules excited by the unknown field [4, 8].

Some techniques used in fluorescence microscopy under high numerical aperture (NA) imaging conditions can also be regarded as a form of nonparaxial polarimetry, in which the polarization state is directly linked to the orientation and wobbling of a fluorescent molecule [9]. These molecules usually behave like linear dipoles, whose direction determines both their excitability by an external field and the dipolar pattern of the light they radiate. Further, since they are attached to structures of biological interest, their 3D direction and wobble provide useful information about the sample's structural properties. Methods like single-molecule orientation localization microscopy (SMOLM) enable the 3D spatial localization of emitters while also providing information about their 3D orientational behavior. They rely on approaches such as polarization channel splitting [10–12] and/or point spread function (PSF) engineering, in which the information about the dipole's 3D localization and orientation is encoded in the shape of the single molecule's PSF [13–17].

In all of the cases described above, intensity measurements of different projections of the radiation from a particle can be used to construct what is referred to as the 3×3 second moment matrix [1], described in Section 2, which characterizes the second order statistics of the field oscillations, and hence its 3D polarization. Prior work based on the Fisher information has been proposed for the specific case of fluorescence microscopy [18, 19], for which the matrix can be assumed to be real, causing a reduction in the number of parameters being estimated. The Fisher information and the corresponding Cramér-Rao bounds have also been used to compare different SMOLM techniques [15, 20–22].

The goal of this work is to provide a general description that is based on the eight parameters necessary for the characterization of the 3D polarization state of the field emitted or scattered by a molecule/nanoparticle, and to define merit functions that qualify the appropriateness of a technique to measure 3D polarization. This approach is based on the decomposition of the matrix into a basis composed of the Gell-Mann matrices in order to define generalizations of the Stokes parameters [2, 3]. A general description of the high-NA microscopy systems used for these measurements is given in Section 3, which also provides conditions for reducing estimation bias between the transverse and longitudinal components of the dipole. This section also includes a forward model for the measured intensity in terms of the parameters mentioned earlier and gives insights into what constitutes a good system for measuring unambiguously different aspects of polarization. Section 4 proposes a merit function for the system's ability to measure 3D polarization based on the Fisher information matrix. We also show that this merit function can be written as the product of four factors: the measured number of signal photons, a measure of the pixel localization of the signal, a measure of the orthonormality of the basis of intensity components associated with each Stokes parameter, and a measure of the "monomodality" of the pixel detection. The general results of Section 4 are applied in Section 5 to the case of fluorescence microscopy, where the number of parameters is reduced (due to the absence of spin). The merit function is adapted to this dimensionality reduction, and is also expressed in terms of angular parameters that are more easily associated with fluorophore orientation. A further simplification is performed for the specific (and commonly used) case of fluorophores whose wobbling is assumed to be statistically rotationally symmetric about a main direction. Finally, Section 7 illustrates these ideas by applying the merit functions to two different types of measurement technique [12, 17, 23, 24], and using them for optimizing free parameters in these methods.

Let us stress that the results found here can be applied either to fluorescence microscopy, where the goal is to obtain information about the fluorophore's orientation and dynamics, or to scattering microscopy, where the particles act as probes to measure the 3D state of polarization of an incident field. Since our focus is to better understand the measurement of 3D polarization/orientation, we ignore the estimation of the spatial localization of the particle.

2. The second moment matrix

The definition and description of the second moment matrix and its decomposition into generalizations of the Stokes parameters has been treated in several prior works, including two recent reviews [3, 9]. Nevertheless, it is convenient to start by giving a brief description of these concepts. Consider an electric dipolar field emanating from some molecule/nanoparticle with known position. This dipole is determined by the normalized vector $\hat{\mu} = (\mu_x, \mu_y, \mu_z)^{\mathrm{T}}$, with $^{\mathrm{T}}$ denoting a transpose (so that vectors are assumed to be columns). For the case of a fixed fluorophore, this dipole has a linear orientation specified by two spherical angles η (the off-plane or polar angle) and ξ (the in-plane or azimuthal angle), as shown in Fig. 1(a). However, fluorophores often wobble within the detection time, as illustrated in Fig. 1(b), so more parameters are needed to describe the statistics of their oscillations. In the more general case of 3D polarimetry, the field scattered by a sub-wavelength particle can be described by an elliptic dipolar oscillation with

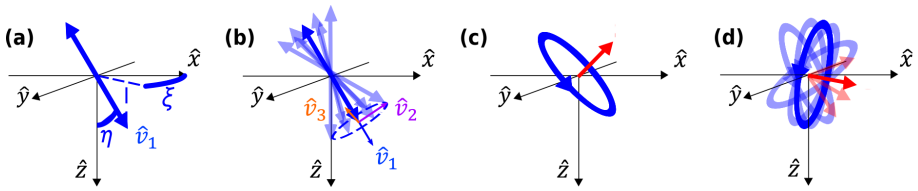

Fig. 1. (a) In-plane and off-plane angles, ξ and η, for a fixed real dipole. (b) Wobbling dipole, whose directional statistics can be described in terms of a principal direction, $\hat{v}_1$, and two minor ones, $\hat{v}_{2,3}$, which form an orthonormal triad. (c) Fully polarized 3D scattered field, described by an ellipse of polarization with arbitrary ellipticity and orientation in 3D. Here, the red arrow represents the spin density vector, which is normal to the plane of the ellipse and whose length is proportional to the area enclosed by the ellipse. (d) Partially polarized 3D scattered field, described by a fluctuating polarization ellipse.

some ellipticity and handedness and with arbitrary 3D orientation, as shown in Fig. 1(c). The normal to the ellipse following the right-hand rule determines the direction of the spin (shown as a red arrow in the figure), whose magnitude is proportional to the ellipse's area. This ellipse can also present fluctuations within the detection time, as illustrated in Fig. 1(d). Simple sets of well-chosen intensity measurements of the radiated field allow characterizing the second-order statistics of the fluctuating dipolar radiation described earlier. These statistics are encoded in the second moment matrix, defined as [13, 14, 18, 25]

$$\mathbb{m} = \langle \hat{\mu}\hat{\mu}^\dagger \rangle_T = \frac{1}{T}\int_{t_0}^{t_0+T} \hat{\mu}\hat{\mu}^\dagger \mathrm{d}t, \tag{1}$$

where T is the integration time of the detection and † denotes a transpose conjugate. If the dipole is fixed during this integration time, the second moment matrix factorizes as an outer product; if on the other hand the dipole fluctuates within the detection time, this matrix is not factorizable and contains information about the statistics of this fluctuation. For the case of fluorescing molecules that wobble, the second moment matrix characterizes not only the main direction of the molecule but also its directional fluctuations [14]. We assume that the molecule's wobbling takes place at a timescale much shorter than the fluorescence lifetime, since otherwise there could be appreciable coupling between absorption and emission events causing the measurements to be related to higher-order moments [26, 27]. Note that the matrix $\mathbb{m}$ is explicitly Hermitian and non-negative-definite, and that it has a trace of unity given the normalization of $\hat{\mu}$. In the general case of complex dipoles, the second moment matrix includes eight independent parameters, three of which describe spin density of the emission and correspond to the imaginary parts of the off-diagonal elements. For the case of fluorescent molecules, on the other hand, the dipole emission is typically along a line so that $\hat{\mu}$ is real, as is $\mathbb{m}$, and the second moment matrix contains only five independent parameters. In either case, as described later, the quantities measured by polarized imaging are typically proportional to the elements of this matrix.

Decomposition in terms of Gell-Mann matrices. To simplify the treatment, it is common to reorganize the second moment matrix as a vector. Many authors [13, 14, 25, 28] do so by simply composing a vector with the different elements, e.g. $\vec{m} = (m_{xx}, m_{yy}, m_{zz}, m_{xy}, m_{yz}, m_{xz})^\mathrm{T}$. Another option is to decompose the matrix in a complete orthogonal basis of Hermitian 3×3

matrices. One such basis is that of the Gell-Mann matrices [29], which can be written as

$$\Theta_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Theta_{12} = \begin{pmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1/\sqrt{3} & 0 \\ 0 & 0 & -2/\sqrt{3} \end{pmatrix},$$

$$\Theta_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \Theta_{22} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \Theta_{23} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

$$\Theta_{31} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \quad \Theta_{32} = \begin{pmatrix} 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 \end{pmatrix}, \quad \Theta_{33} = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2}$$

These matrices are orthogonal under the inner product of the trace:

$$\mathrm{Tr}(\Theta_{mn}\Theta_{m'n'}) = 2\delta_{mm'}\delta_{nn'}, \tag{3}$$

with $\delta_{mm'}$ being the Kronecker symbol. They are also traceless, meaning that each of them is orthogonal to the 3×3 identity matrix $\mathbb{I}$. Therefore, together with $\mathbb{I}$, the Gell-Mann matrices form a complete orthogonal basis for Hermitian 3×3 matrices. Given its normalization, the second moment matrix can be expressed in terms of the identity and the Gell-Mann matrices as

$$\mathbb{m} = \frac{1}{3}\mathbb{I} + \frac{1}{\sqrt{3}}\sum_k s_k\Theta_k, \qquad s_k = \frac{\sqrt{3}}{2}\mathrm{Tr}(\mathbb{m}\Theta_k), \tag{4}$$

where we introduce the shorthand k to represent the two indices mn, so that k runs over the values $11, 12, 21, 22, 23, 31, 32, 33$. The coefficients s_k are called here the normalized Stokes-Gell-Mann (NSG) parameters, in analogy with the normalized Stokes parameters used in paraxial polarization, which are the coefficients for the decomposition of a normalized polarization matrix in terms of Pauli matrices; in fact, the numbering scheme used here (and in other recent work [3, 9]) for the Gell-Mann matrices is chosen to highlight the similarity with the Stokes parameters of paraxial polarization [3], as can be seen from writing explicitly all parameters in Eq. (4):

$$\begin{aligned} &s_{11} = \sqrt{3}\,\frac{m_{xx} - m_{yy}}{2}, \quad s_{12} = \frac{m_{xx} + m_{yy} - 2m_{zz}}{2}, \\ &s_{21} = \sqrt{3}\,\mathrm{Re}(m_{yz}), \quad s_{22} = \sqrt{3}\,\mathrm{Re}(m_{xz}), \quad s_{23} = \sqrt{3}\,\mathrm{Re}(m_{xy}), \\ &s_{31} = -\sqrt{3}\,\mathrm{Im}(m_{yz}), \quad s_{32} = \sqrt{3}\,\mathrm{Im}(m_{xz}), \quad s_{33} = -\sqrt{3}\,\mathrm{Im}(m_{xy}). \end{aligned} \tag{5}$$

That is, the parameters s_{1n} encode the diagonal elements of the matrix, while s_{2n} and s_{3n} encode, respectively, the real and imaginary parts of the off-diagonal ones. As mentioned earlier, for a typical fluorescent dipole, $\mathbb{m}$ is real and only the first five NSG parameters can be different from zero. For complex dipoles, however, the parameters s_{3n} are needed, and in fact the three-vector $(s_{31}, s_{32}, s_{33})^{\mathrm{T}}$ is proportional to the spin density vector of the dipole, and it points in a direction normal to the main plane of oscillation of the electrons [30], as shown in Fig. 1(c,d). Note that the numerical factors used in the definition of the NSG parameters differ by a factor of $\sqrt{3}/2$ from those used in the standard Stokes parameters. These normalization factors were chosen such that the following inequality is satisfied:

$$\sum_k s_k^2 \leq 1, \tag{6}$$

where the upper bound of unity is achieved only if $m_{ij} = \mu_i \mu_j^*$, i.e. in the absence of fluctuations/wobbling. Based on this inequality, we group the NSG parameters as an eight-component vector, $\vec{s} = (s_{11}, s_{12}, s_{21}, s_{22}.s_{23}, s_{31}, s_{32}, s_{33})^{\mathrm{T}}$, where the last three components can be ignored when considering fluorophores described by linear dipoles. That is, $\vec{s}$ is constrained to the interior and surface of a unit hypersphere, $|\vec{s}| \leq 1$, in an abstract 8D space (or a 5D space in the absence of dipole spin). As discussed elsewhere, this vector is actually constrained to a far smaller region within the hypersphere [3].

Eigenvalues of the second moment matrix. The second moment matrix can also be parametrized in terms of its three eigenvalues and the direction of its three normalized eigenvectors as

$$\mathbb{m} = \sum_{j=1}^{3} \lambda_j \hat{v}_j \hat{v}_j^{\dagger}, \tag{7}$$

where we choose, without loss of generality, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$, with $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The three eigenvectors are orthonormal, and in the absence of dipole spin they can be chosen as real (as shown in Fig. 1(b)). They can then be parametrized in terms of three angles η, ξ and β as

$$\begin{aligned}
\hat{v}_1 &= (\sin\eta\cos\xi, \sin\eta\sin\xi, \cos\eta)^{\mathrm{T}}, \\
\hat{v}_2 &= (\cos\eta\cos\xi, \cos\eta\sin\xi, -\sin\eta)^{\mathrm{T}}\cos\beta + (-\sin\xi, \cos\xi, 0)^{\mathrm{T}}\sin\beta, \\
\hat{v}_3 &= -(\cos\eta\cos\xi, \cos\eta\sin\xi, -\sin\eta)^{\mathrm{T}}\sin\beta + (-\sin\xi, \cos\xi, 0)^{\mathrm{T}}\cos\beta.
\end{aligned} \tag{8}$$

That is, η and ξ are the polar and azimuthal spherical angles of $\hat{v}_1$ and β is an angle determining the orientation of $\hat{v}_2$ (and hence $\hat{v}_3$) within the plane normal to $\hat{v}_1$ (Fig. 1(b)). In cases involving dipole spin, on the other hand, the eigenvectors can be complex and require more parameters. Different parametrizations have been discussed elsewhere [3] and are not central to our study.

Whether there is spin or not, a measure of the degree of 3D polarization of the emitted light or of fluorophore rigidity (i.e. lack of wobble) can be defined as the magnitude of the NSG vector [2, 3, 31, 32]:

$$P = |\vec{s}| = \sqrt{\frac{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_1 - \lambda_3)^2}{2}}, \tag{9}$$

The upper limit $P = 1$ is only achieved for $\lambda_2 = \lambda_3 = 0$, which corresponds to a factorizable second moment matrix, describing a fully polarized field or a non-wobbling dipole. In this limiting case we can identify $\hat{\mu} = \hat{v}_1$. The opposite limiting case, $P = 0$, holds only for $\lambda_1 = \lambda_2 = \lambda_3 = 1/3$ which indicates complete statistical isotropy, for example by a freely wobbling dipole that explores equally all directions. Several alternative measures have been proposed, particularly in the context of 3D polarization [3, 33–36]. Within the context of molecular rigidity, Zhang *et al.* [16, 18, 19] defined the rotational mobility or rotational constraint as

$$\gamma = \lambda_1 - \frac{\lambda_2 + \lambda_3}{2}. \tag{10}$$

Note that γ coincides with P for $\lambda_2 = \lambda_3$. This is the case, for example, of the common model in which the dipole is assumed to oscillate uniformly within a directional cone of angle δ, for which $P = \gamma = \cos^2(\delta/4)\cos(\delta/2)$. Also, in the very specific case of a real dipole oscillating around the z axis such that the off-diagonal elements of the second moment matrix vanish (e.g. when the wobble has rotational symmetry about this direction), $\gamma = -s_{12}$.

As shown in Appendix B the five NSG parameters not associated with spin are all proportional to the real or imaginary parts of spherical harmonics of order $\ell = 2$, leading to a simple proportionality relation between these parameters and the dipole orientation and wobbling characteristics, particularly the second order parameters of the dipole's wobbling distribution.

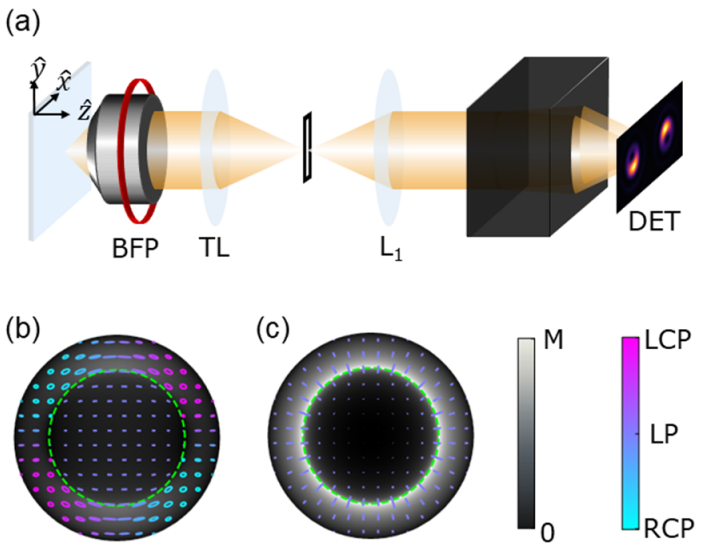

Fig. 2. (a) Schematic of the optical setup used to measure 3D polarization. Components: objective lens (OBJ), tube lens (TL) and relaying lens (L_1) for accessing the back-focal plane (BFP). The black box represents a system (possibly consisting of elements to manipulate the phase, polarization and/or amplitude) to concentrate the light onto one or several detectors or detector arrays (DET). (b,c) Polarization distribution and intensity at the BFP for a transverse dipole aligned with the x direction (b), and a longitudinal dipole aligned with the z direction (c), where the green circle corresponds to $u = n_0$, and the polarization ellipticity and handedness are encoded in the color of the ellipses, which go from right-hand circular polarization (RCP) to left-hand circular polarization (LCP) passing through linear polarization (LP).

3. Measurement system

We begin this section with a review of the mathematical description of a generic microscopy setup used to measure the 3D polarization of the field emitted or scattered by a particle [3, 9, 37, 38], as illustrated in Fig. 2(a). Since in this work we focus on the characterization of $\mathbb{m}$, we assume that the position of the particle is known. (In applications such as SMOLM, the particle's 3D localization must be estimated as well.) Let the particle be embedded in a medium with refractive index n_0. Before being collected by the objective, the light emitted or scattered by the particle typically traverses a flat interface with another medium with refractive index n_1, e.g. a glass coverslip that is index-matched to the microscope objective.

Light distribution at the pupil plane. The field distribution at the back-focal (or pupil) plane of the objective is given by $\widetilde{\mathbf{E}}$. This field is well collimated and hence only the two transverse components are needed. (Our convention in this work is to use boldface for vectors with only x and y components.) The pupil coordinate $\mathbf{u}$ is given for convenience in normalized units, so that its maximum value corresponds to the NA. The expression for the field $\widetilde{\mathbf{E}}(\mathbf{u})$ at the pupil plane is simply proportional to $\mathbb{G}(\mathbf{u})\,\hat{\mu}$, where $\mathbb{G}$ is the 2×3 Green tensor relating each component of the dipole to a paraxial field distribution at the pupil plane:

$$\mathbb{G}(\mathbf{u}) = \begin{pmatrix} g_0(u) + g_2(u)\cos 2\varphi & g_2(u)\sin 2\varphi & g_1(u)\cos\varphi \\ g_2(u)\sin 2\varphi & g_0(u) - g_2(u)\cos 2\varphi & g_1(u)\sin\varphi \end{pmatrix}, \tag{11}$$

with $u = |\mathbf{u}|$ being the radial pupil coordinate and φ the azimuthal one, and where

$$g_{0,2}(u) = \frac{n_1\sqrt{\gamma_1(u)}}{2n_0}\left[t_\mathrm{p}(u) \pm \frac{t_\mathrm{s}(u)}{\gamma_0(u)}\right], \quad g_1(u) = \frac{n_1\sqrt{\gamma_1(u)}}{n_0^2\,\gamma_0(u)}u\,t_\mathrm{p}(u), \tag{12}$$

with $\gamma_i(\mathbf{u}) = \sqrt{1 - u^2/n_i^2}$, and where t_p and t_s are the Fresnel transmission coefficients given by

$$t_\mathrm{s}(u) = \frac{2n_0\gamma_0(u)}{n_0\gamma_0(u) + n_1\gamma_1(u)}, \quad t_\mathrm{p}(u) = \frac{2n_0\gamma_0(u)}{n_1\gamma_0(u) + n_0\gamma_1(u)}. \tag{13}$$

Note that in the typical case of $n_1 > n_0$, light can be collected not only for $u \le n_0$ but also within the annular pupil region $n_1 > u > n_0$, which corresponds to evanescent components of the dipole radiation that couple into propagating ones at the interface between the two media. In the context of fluorescence microscopy, this light is referred to as supercritical angle fluorescence (SAF) [39, 40]. For $u > n_0$ some of the quantities in the equations above become imaginary or complex. Also, while here we are not considering the effect of variations in particle position, an increase in the distance between the particle and the interface causes an exponential decay in the optical power within this outer pupil region; the power within the region $u > n_0$ is appreciable only when this distance is smaller than the wavelength.

Dependence of power collection on dipole orientation. It is interesting to consider how $\Phi = \int_0^{2\pi}\int_0^{\mathrm{NA}} |\widetilde{\mathbf{E}}(\mathbf{u})|^2 u\,\mathrm{d}u\,\mathrm{d}\varphi$, which quantifies the fraction of the power captured by the microscope objective, depends on the orientation of the dipolar radiation emitted or scattered by the particle. Note that strong polarization-dependent variations in the total number of measured photons hinders the ability to measure 3D polarization. It turns out that considering linear dipoles is sufficient for this characterization, so let us then assume for now that the dipole direction $\hat{\mu}$ is real. It is easy to see by symmetry arguments that this power depends only on the angle η between the dipole and the z axis, following the simple sinusoidal relation $\Phi = \Phi_\perp \sin^2\eta + \Phi_z \cos^2\eta$, where $\Phi_\perp$ and Φ_z correspond to the two extreme cases of a dipole oriented transversely to the axis of the microscope, say, in the x direction, and one oriented parallel to the microscope axis. These powers are calculated as

$$\Phi_\perp = \int_0^{2\pi}\int_0^{\mathrm{NA}} |\mathbb{G}(\mathbf{u})\hat{x}|^2 u\,\mathrm{d}u\,\mathrm{d}\varphi = 2\pi\int_0^{\mathrm{NA}} \left[|g_0|^2 + |g_2|^2\right] u\,\mathrm{d}u, \tag{14a}$$

$$\Phi_z = \int_0^{2\pi}\int_0^{\mathrm{NA}} |\mathbb{G}(\mathbf{u})\hat{z}|^2 u\,\mathrm{d}u\,\mathrm{d}\varphi = 2\pi\int_0^{\mathrm{NA}} |g_1|^2 u\,\mathrm{d}u. \tag{14b}$$

These integrals can be solved in closed form but the resulting expressions are too long to provide any useful insight. Figure 3(a) shows the relative difference between $\Phi_\perp$ and Φ_z as a function of the NA and the index mismatch at the interface. We can see that this difference vanishes near $\mathrm{NA} = n_0$.

The closed-form expressions for these two power components become considerably simpler for $\mathrm{NA} = n_0$. They are given in Appendix A. From these relations it can be seen that the two powers match if $n_1/n_0 = 1.33$. This balance is not achieved exactly for the common case in which the first medium is water ($n_0 = 1.33$) and the second is glass ($n_1 = 1.515$), but the relative difference between $\Phi_\perp$ and Φ_z is only of about 5%, so for the sake of simplification we can assume that the two powers are nearly equal. (For these media, exactly balancing these powers would require expanding the NA slightly above n_0, but the solution would then depend significantly on the distance between the particle and the interface.)

The balance or near-balance between the transverse and longitudinal power components can be understood in the following way: on the one hand, as shown in Fig. 2(b) a transverse dipole fills the center of the pupil, which corresponds to smaller angles of incidence at the interface and hence to higher transmission, but it also populates regions of the edge of the pupil where

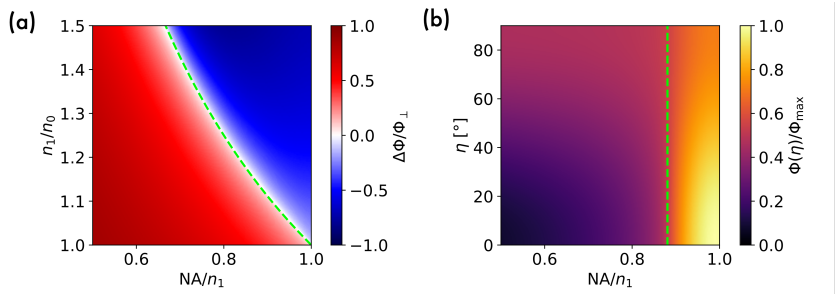

Fig. 3. (a) Relative difference $(\Phi_\perp - \Phi_z)/\Phi_\perp$ between the power captured by the microscope for a transverse and a longitudinal dipole, as a function of the normalized NA with respect the immersion liquid's refractive index and the index mismatch between the embedding medium and the immersion liquid. (b) Total power for a non-wobbling emitter with an off-plane angle η for different numerical apertures given an index mismatch corresponding to oil/water. (a-b) The green line corresponds to NA = n_0. In both, the dipole is assumed to be at the interface.

it is mostly made up of s-polarization components with weaker transmission coefficients; on the other hand, as shown in Fig. 2(c) a longitudinal dipole radiates mostly into the edges of the pupil, and hence at high angles of incidence, but with p-polarization that is better transmitted. Note that these conditions for achieving (near) independence of the power with respect to dipole orientation also apply if the dipole has spin or if there are fluctuations. In other words, under these conditions the power is (essentially) independent of the NSG parameters. As mentioned earlier, this power is then also independent of the particle's position.

Polarization splitting and detection. As shown in Fig. 2(a) the field is typically made to pass through another system between the pupil and the final detection. This system can contain elements such as: i) a pupil mask that modifies the amplitude, phase, and/or polarization at the Fourier plane; ii) a polarization-splitting device that separates the field into different polarization components that are to be detected separately; iii) a focusing system to form an image or at least to concentrate light onto the detectors; and iv) a series of detectors or detector arrays. Let the different polarization channels (typically two or four) be identified by the index σ and, for each channel, let the different pixels be labeled by the discrete variable $\mathbf{x}$. We assume for now that the pixels are sufficiently small as for the field variation within each of them to be negligible; the effect of more extended pixels is discussed towards the end of this section. The measured intensity can then be written as

$$I(\mathbf{x},\sigma) = \mathbf{h}^\dagger(\mathbf{x},\sigma)\mathbb{m}\mathbf{h}(\mathbf{x},\sigma), \tag{15}$$

where $\mathbf{h}(\mathbf{x},\sigma)$ is the Green function from the source to the polarization component of the detector in question. Consider as an example a type of imaging system used in SMOLM [10–17] where a Fourier mask is used to shape the PSF and where the light emerging from this mask is separated into polarization components $\mathbf{p}_\sigma$ before reaching the image. For this system we can write $\mathbf{h}(\mathbf{x},\sigma) = \mathbb{H}^\dagger(\mathbf{x})\mathbf{p}_\sigma$, where $\mathbb{H}(\mathbf{x})$ is the 2×3 matrix whose three columns give the paraxial vector field distribution at the image plane, prior to polarization filtering, due to dipoles in the x, y and z directions, respectively. This matrix is given by

$$\mathbb{H}(\mathbf{x}) = \frac{1}{\lambda}\int \mathbb{J}(\mathbf{u})\mathbb{G}(\mathbf{u})\exp\left(\mathrm{i}\frac{2\pi}{\lambda}\mathbf{u}\cdot\mathbf{x}\right)\mathrm{d}^2u, \tag{16}$$

where λ is the wavelength and $\mathbb{J}(\mathbf{u})$ is the transmission function for a potentially-birefringent

pupil mask. While this type of system is important, the analysis that follows applies to a broader range of systems.

If the system is unitary, meaning that no amplitude masks are used and that losses are minimal so that essentially all the light that passes through the pupil is captured by the detectors, then

$$\langle |\mathbf{h}(\mathbf{x},\sigma)|^2 \rangle_{\mathbf{x},\sigma} \approx \Phi. \tag{17}$$

where $\langle\cdot\rangle_{\mathbf{x},\sigma}$ denotes a sum over all detector pixels of all polarization channels.

Measured intensity in terms of NSG parameters. The intensity distribution created by the dipole can be written in terms of the NSG parameters as

$$I(\mathbf{x},\sigma) \propto \mathcal{I}_0(\mathbf{x},\sigma) + \vec{s}\cdot\vec{\mathcal{I}}(\mathbf{x},\sigma), \tag{18}$$

with $\vec{\mathcal{I}} = (\mathcal{I}_{11},\mathcal{I}_{12},\mathcal{I}_{21},\mathcal{I}_{22},\mathcal{I}_{23},\mathcal{I}_{31},\mathcal{I}_{32},\mathcal{I}_{33})$, for

$$\mathcal{I}_0(\mathbf{x},\sigma) = \frac{\kappa}{3}|\mathbf{h}(\mathbf{x},\sigma)|^2, \quad \mathcal{I}_k(\mathbf{x},\sigma) = \frac{\kappa}{\sqrt{3}}\mathbf{h}^\dagger(\mathbf{x},\sigma)\Theta_k\mathbf{h}(\mathbf{x},\sigma), \tag{19}$$

where κ is a normalization constant. The total number of signal photons captured by the microscope is given by $\mathcal{N}_\mathrm{s} = \langle I\rangle_{\mathbf{x},\sigma} \propto \langle\mathcal{I}_0\rangle_{\mathbf{x},\sigma} + \vec{s}\cdot\langle\vec{\mathcal{I}}\rangle_{\mathbf{x},\sigma}$. Note that by choosing NA = n_0 (so that $\mathcal{N}_\mathrm{s}$ is roughly independent of dipole orientation) $\langle\mathcal{I}_k\rangle_{\mathbf{x},\sigma}$ essentially vanishes in order for the dependence on $\vec{s}$ to be suppressed. We will then assume this choice of NA in what follows. It is convenient to choose the normalization constant κ so that $\langle\mathcal{I}_0\rangle_{\mathbf{x},\sigma} = 1$. We henceforth refer to the intensity distribution $I(\mathbf{x},\sigma)$ as the PSF even in cases where the system is not an imaging one.

In real conditions, the measurements include some background intensity I_b, so the model for the intensity (in photons per pixel) becomes

$$I(\mathbf{x},\sigma) = I_\mathrm{b} + \mathcal{N}_\mathrm{s}[\mathcal{I}_0(\mathbf{x},\sigma) + \vec{s}\cdot\vec{\mathcal{I}}(\mathbf{x},\sigma)], \tag{20}$$

where $\mathcal{N}_\mathrm{s}$ is the number of signal photons arriving from the dipole to the detector.

Reciprocal interpretation and an identity for the PSF components. The second expression in Eqs. (19) can also be written as

$$\mathcal{I}_k(\mathbf{x},\sigma) = \frac{\kappa}{\sqrt{3}}\mathrm{Tr}[\,\mathbf{h}(\mathbf{x},\sigma)\mathbf{h}^\dagger(\mathbf{x},\sigma)\Theta_k\,]. \tag{21}$$

That is, to within numerical prefactors, this relation for the PSF components $\mathcal{I}_k$ is mathematically similar to the definition of the Stokes-Gell-Mann parameters in Eq. (4), applied to a fully polarized 3D field $\mathbf{h}(\mathbf{x},\sigma)$. That is, if we were to reverse the direction of light propagation and consider the detector pixel $\mathbf{x}$ for the polarization channel σ as a source of light with polarization $\mathbf{p}_\sigma$, and if this light were to travel back throught the microscope all the way to the particle's position, it would produce an electric field proportional to $\mathbf{h}$. The NSG parameters for this light distribution at the particle would then be $(\sqrt{3}/2\kappa)\mathrm{Tr}[\mathbf{h}\,\mathbf{h}^\dagger\Theta_k]/\mathrm{Tr}[\mathbf{h}\mathbf{h}^\dagger] = \mathcal{I}_k/(2\mathcal{I}_0)$. Because this back-propagated field would be fully polarized, the norm of these NSG parameters would be unity, leading to the following identity that holds for any $\mathbf{x}$ and σ:

$$4\mathcal{I}_0^2(\mathbf{x},\sigma) = \sum_k \mathcal{I}_k^2(\mathbf{x},\sigma) = |\vec{\mathcal{I}}(\mathbf{x},\sigma)|^2, \tag{22}$$

which is analogous to Eq. (6) except that it is a strict equality. Note that the sum above is over all eight values of k, including those associated with spin. Equation (22) and the reciprocal picture that lead to it are useful for visualizing what constitutes a good system for measuring $\mathbb{m}$: the subset of pixels within all polarization channels that capture the light should each be able to couple onto a different (nearly fully polarized) 3D state of polarization at the dipole location,

so that together they cover all the states that one desires to explore or measure. In particular, if we were to consider the case of an imaging system in fluorescent microscopy where the dipoles are known a priori to have no spin, we know that the measured intensity will not include the components $\mathcal{I}_{3n}$. Nevertheless, Eq. (22) still includes these components. That is, even though we might not want to measure spin because we know it is not present, the system will typically have some ability to measure it, and this might come at the cost of weakening the remaining components $\mathcal{I}_k$ linked to parameters that we do want to measure. An ideal system to measure orientation and wobble of spinless fluorophores would then have to be designed such that $\mathcal{I}_{3n} = 0$ for all pixels over both polarization channels. In other words, in the reverse propagation picture all detector pixels should couple to 3D linear polarization states at the dipole location. This reciprocal interpretation also illustrates the advantages of polarization splitting and of using small pixel sizes in the detection. If the polarization components were not separated before being imaged, the light "emitted" from each pixel $\mathbf{x}$ would be "unpolarized" in the paraxial sense, resulting in a partially-polarized light distribution at the dipole location. That is, the left-hand sides of Eqs. (15), (19) and (21) would be independent of σ because the corresponding right-hand sides would include a sum in σ; in particular Eq. (21) would then resemble the calculation of the NSG parameters for a (nonfactorizable) 3D partially polarized field. Similarly, if the pixels were large enough so that the detected field varies appreciably within each of them, in the reciprocal interpretation they would behave as extended, spatially-incoherent sources that, even when fully polarized, can lead to a partially-polarized field at the dipole location. In this case Eq. (21) would resemble the expression for the NSG parameters of a partially polarized field because in the right-hand side, the product $\mathbf{h}(\mathbf{x},\sigma)\mathbf{h}^\dagger(\mathbf{x},\sigma)$ would be replaced by an integral of this product in the spatial coordinate over the area of the pixel centered at $\mathbf{x}$. Whether because the polarization was not filtered or because large pixels were used, the resulting partial polarization at the dipole in the reciprocal picture implies that the equality in Eq. (22) must be replaced by the inequality

$$4\,\mathcal{I}_0^2(\mathbf{x}) \geq |\vec{\mathcal{I}}(\mathbf{x},\sigma)|^2. \tag{23}$$

The components $\mathcal{I}_k$ are then weaker when compared to $\mathcal{I}_0$, making the estimation more difficult.

4. Fisher information in terms of the PSF components

For a given value of the NSG parameters, the normalized probability for a photon to hit a given pixel $\mathbf{x},\sigma$ can be written as

$$\mathcal{P}(\mathbf{x},\sigma|\vec{s}) = \frac{I(\mathbf{x},\sigma)}{\langle I(\mathbf{x},\sigma)\rangle_{\mathbf{x},\sigma}} \approx \frac{I_\mathrm{b}/\mathcal{N}_\mathrm{s} + \mathcal{I}_0(\mathbf{x},\sigma) + \vec{s}\cdot\vec{\mathcal{I}}(\mathbf{x},\sigma)}{\mathcal{N}/\mathcal{N}_\mathrm{s}}, \tag{24}$$

where $\mathcal{N} = \langle I_\mathrm{b}\rangle_{\mathbf{x},\sigma} + \mathcal{N}_\mathrm{s}$ is the total measured number of photons (since I_b is the average number of background photons per pixel), and we used the relations $\langle\mathcal{I}_0\rangle_{\mathbf{x},\sigma} = 1$ and $\langle\mathcal{I}_k\rangle_{\mathbf{x},\sigma} \approx 0$, the latter resulting from using NA = n_0. That is, the use of NA = n_0 makes this probability approximately a linear function of $\vec{s}$, which simplifies the analysis that follows.

Fisher information matrix. The elements of the Fisher information matrix $\boldsymbol{\mathcal{J}}^{(\vec{s})}$ for the NSG parameters are given by

$$\mathcal{J}^{(\vec{s})}_{kk'} = \mathcal{N}\left\langle\frac{\partial_{s_k}\mathcal{P}\,\partial_{s_{k'}}\mathcal{P}}{\mathcal{P}}\right\rangle_{\mathbf{x},\sigma} \approx \mathcal{N}_\mathrm{s}\left\langle\frac{\mathcal{I}_k\mathcal{I}_{k'}}{I_\mathrm{b}/\mathcal{N}_\mathrm{s} + \mathcal{I}_0 + \vec{s}\cdot\vec{\mathcal{I}}}\right\rangle_{\mathbf{x},\sigma} = \mathcal{N}_\mathrm{s}\left\langle\!\left\langle\mathcal{I}_k\mathcal{I}_{k'}\right\rangle\!\right\rangle, \tag{25}$$

where $\langle\!\langle\mathcal{I}_k\mathcal{I}_{k'}\rangle\!\rangle$ denotes an inner product between two PSF components $\mathcal{I}_k$ and $\mathcal{I}_{k'}$, defined as the sum over all pixels and polarization channels with a non-negative weight function given by $1/(I_\mathrm{b}/\mathcal{N}_\mathrm{s} + \mathcal{I}_0 + \vec{s}\cdot\vec{\mathcal{I}})$ (which depends on the NSG parameters to be retrieved and on the level of background intensity).

Trace constraint and its upper bound. The inner-product form in Eq. (25) of the Fisher information matrix $\boldsymbol{\mathcal{J}}^{(\vec{s})}$ means that it is a non-negative definite real 8×8 matrix. The trace of the matrix can be seen from Eq. (23) to be

$$\mathrm{Tr}\boldsymbol{\mathcal{J}}^{(\vec{s})} = \sum_k \mathcal{N}_\mathrm{s} \left\langle\!\left\langle \mathcal{I}_k^2 \right\rangle\!\right\rangle \leq 4\mathcal{N}_\mathrm{s} \left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle . \tag{26}$$

That is, the trace of the matrix is bound by $\left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle$. The equality in the last part of the expression is achieved by using polarization splitting and sufficiently small pixels. Let us assume for now this is the case, and study this result in two limiting situations:

- Consider first the ideal case in which there is no background ($I_\mathrm{b} = 0$). If we consider a highly unpolarized situation (for example, a fluorophore wobbling freely) in which $\vec{s} \approx \vec{0}$ then $\left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle \approx \langle \mathcal{I}_0 \rangle_{\mathbf{x},\sigma} = 1$ so the trace of the Fisher information matrix is simply $4\mathcal{N}_\mathrm{s}$. Typically, as the magnitude of $\vec{s}$ increases, the intensity distribution $\mathcal{I}_0 + \vec{s} \cdot \vec{\mathcal{I}}$ develops minima that approach zero and lead to regions with large values of the weight function, making the trace of the Fisher information matrix grow, to the point that for $|\vec{s}| = 1$ some elements of this matrix can diverge.

- Consider now the opposite limiting case in which the background I_b dominates. The Fisher information then takes the approximate form of an inner product with constant weight, namely

$$\mathcal{J}_{kk'}^{(\vec{s})} \approx \frac{\mathcal{N}_\mathrm{s}^2}{I_\mathrm{b}} \langle \mathcal{I}_k \mathcal{I}_{k'} \rangle_{\mathbf{x},\sigma} = \mathcal{N}_\mathrm{s}\, r_\mathrm{s/b}\, N_\mathrm{pixel}\, \langle \mathcal{I}_k \mathcal{I}_{k'} \rangle_{\mathbf{x},\sigma} , \tag{27}$$

  where $r_\mathrm{s/b} = \mathcal{N}_\mathrm{s}/(I_\mathrm{b} N_\mathrm{pixel})$ is the ratio of signal to background photons detected, with $N_\mathrm{pixel} = \langle 1 \rangle_{\mathbf{x},\sigma}$ being the total number of pixels over all polarization channels being used in the estimation of $\mathbb{m}$. (In imaging systems with a wide field of view capable of measuring polarization for many particles simultaneously, N_pixel is the number of pixels of the window(s) used for a specific particle.) The trace of this matrix then reduces to

$$\mathrm{Tr}\boldsymbol{\mathcal{J}}^{(\vec{s})} \approx 4\mathcal{N}_\mathrm{s}\, r_\mathrm{s/b}\, N_\mathrm{pixel} \left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma} . \tag{28}$$

  Recall that $\mathcal{I}_0$ corresponds to the shape of the intensity distribution over the detector pixels for a fully unpolarized (in the 3D sense) dipole, and hence to the intensity distribution with maximum possible extension for the corresponding measurement method. Therefore, $\mathcal{I}_0$ determines the region within the detectors where the intensity distribution (or PSF) for a dipole with any orientation/polarization is constrained. Recall also that we chose to normalize this distribution according to $\langle \mathcal{I}_0 \rangle_{\mathbf{x},\sigma} = 1$. The factor $\left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma}$ in the expression for the trace is then a measure of pixel localization: the more restricted $\mathcal{I}_0$ is to a small set of pixels, the larger $\left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma}$ is. This means that a large value for $\left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma}$ probably also leads to a higher signal-to-noise ratio, and in the case of SMOLM, to good transverse spatial localization (not considered here). Note, however, that $\left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma}$ can be made larger artificially by using larger pixels. (Think, for example, of an extreme case consisting of only two pixels with $\mathcal{I}_0 = 1/2$ each, leading to $\left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma} = 2 \times 1/2^2 = 1/2$; joining the two pixels into a single one would give $\left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma} = 1$ but this would certainly not improve the measurement.) An appropriate merit function of pixel localization that compensates for this effect can then be defined as $r_\mathrm{s/b}\, N_\mathrm{pixel} \left\langle \mathcal{I}_0^2 \right\rangle_{\mathbf{x},\sigma}$, which can be written simply as

$$M_\mathrm{PL} = \left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle . \tag{29}$$

Determinant inequality and a measure of orthonormality. According to the Cramér-Rao bounds, if the measurement is unbiased, a measured intensity distribution results in a probability density for the unknown parameters $\vec{s}$ that is approximately a Gaussian centered at the expected parameter values and with an ellipsoidal cross-section aligned with the eigenvectors of $\mathcal{J}^{(\vec{s})}$, where the width in each of these directions is proportional to the inverse of the square root of the corresponding eigenvalue. The total cross-section of the Gaussian is then proportional to the inverse of the square root of the determinant of the Fisher information matrix. We can then use $\mathrm{Det}\mathcal{J}^{(\vec{s})}$ (or any monotonic function of it) as the merit function that must be maximized in order to guarantee a precise measurement scheme. The trace of a matrix is the sum of its eigenvalues, while the determinant is the product of these eigenvalues. Therefore, for fixed trace, the determinant is maximized when all eigenvalues are equal, and this leads to a matrix proportional to the identity. That is, the constraint on the trace imposes an upper bound for the determinant which is only met if $\mathcal{J}^{(\vec{s})}$ is proportional to the 8×8 identity matrix. We then have the following two inequalities:

$$\mathrm{Det}\mathcal{J}^{(\vec{s})} \leq \left[\frac{\mathrm{Tr}\mathcal{J}^{(\vec{s})}}{8}\right]^8 \leq \left(\frac{\mathcal{N}_\mathrm{s}}{2}\left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle\right)^8. \tag{30}$$

The upper bound to the first inequality is achieved only if the PSF components $\mathcal{I}_k$ constitute an orthonormal set under the inner product $\langle\!\langle \mathcal{I}_k \mathcal{I}_{k'} \rangle\!\rangle$, while the second inequality approaches an equality when polarization splitting is used and the pixels are small enough that the field does not change significantly within them. We then define merit functions for the orthonormality of the basis and for the "monomodality" of the pixel detection (in polarization or spatial field variation), respectively, as

$$M_\mathrm{ON} = \frac{8\left[\mathrm{Det}\mathcal{J}^{(\vec{s})}\right]^{1/8}}{\mathrm{Tr}\mathcal{J}^{(\vec{s})}} = \frac{8\left(\mathrm{Det}\mathbb{C}\right)^{1/8}}{\mathrm{Tr}\mathbb{C}} \leq 1, \tag{31}$$

$$M_\mathrm{MM} = \frac{\mathrm{Tr}\mathcal{J}^{(\vec{s})}}{4\mathcal{N}_\mathrm{s}\left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle} = \frac{\mathrm{Tr}\mathbb{C}}{4\left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle} \leq 1, \tag{32}$$

where $\mathbb{C}$ is the matrix of inner products of the PSF components, whose elements are $C_{kk'} = \langle\!\langle \mathcal{I}_k \mathcal{I}_{k'} \rangle\!\rangle$.

Complete merit function. The complete merit function $M_8(\vec{s})$ is defined as the eighth root of the determinant of the Fisher information matrix. It can be written as the product of the number of signal photons and the measures of pixel localization, basis orthonormality, and pixel monomodality, according to

$$M_8(\vec{s}) = \left[\mathrm{Det}\mathcal{J}^{(\vec{s})}\right]^{1/8} = \frac{1}{2}\mathcal{N}_\mathrm{s}\, M_\mathrm{PL}\, M_\mathrm{ON}\, M_\mathrm{MM}. \tag{33}$$

The inverse of the square root of this measure gives the geometric mean of the standard deviation in the estimation of $\vec{s}$. Since in general this measure depends on $\vec{s}$ (except in the limit of very strong background intensity), any optimization based on it should involve its value over a sample of representative states of 3D polarization of interest. Note that the second form of this merit function as a product of several factors simply has the goal of helping understand the different aspects of what makes a good measurement, as will be discussed in the examples in Section 7; the merit function can be computed directly without necessarily calculating these factors.

5. Merit function for the case of fluorophores

Let us now consider the specific case of (possibly wobbling) linear fluorophores, whose radiation has no spin. If a priori one knows that the measured PSFs should not include the components

$\mathcal{I}_{3n}$ associated with spin, then we only need to consider a 5×5 submatrix of $\boldsymbol{\mathcal{J}}^{(\vec{s})}$ denoted here as $\overline{\boldsymbol{\mathcal{J}}}^{(\vec{s})}$. If the sum in k is limited to $11, 12, 21, 22, 23$, then Eq. (22) must be replaced with the inequality $4\mathcal{I}_0^2(\mathbf{x}, \sigma) \geq \sum_k \mathcal{I}_k^2(\mathbf{x}, \sigma)$, the equality being attainable only if the measurement system is designed so that the three PSF components $\mathcal{I}_{3n}(\mathbf{x}, \sigma)$ vanish for all pixels $\mathbf{x}$ in all polarization channels σ. An optimal system then requires that $\overline{\boldsymbol{\mathcal{J}}}^{(\vec{s})}$ be proportional to the 5×5 identity, and this leads to the upper bound

$$\mathrm{Det}\overline{\boldsymbol{\mathcal{J}}}^{(\vec{s})} \leq \left(\frac{4}{5}\mathcal{N}_\mathrm{s} \left\langle\!\left\langle \mathcal{I}_0^2 \right\rangle\!\right\rangle\right)^5. \tag{34}$$

We can then define the merit function

$$M_5(\vec{s}) = \left[\mathrm{Det}\overline{\boldsymbol{\mathcal{J}}}^{(\vec{s})}\right]^{1/5} = \frac{4}{5}\mathcal{N}_\mathrm{s}\, M_\mathrm{PL}\, \overline{M}_\mathrm{ON} M_\mathrm{MM}. \tag{35}$$

where the orthonormality measure now is limited to the five relevant PSF components normalized by their upper bound:

$$\overline{M}_\mathrm{ON} = \frac{5\left[\mathrm{Det}\overline{\boldsymbol{\mathcal{J}}}^{(\vec{s})}\right]^{1/5}}{\mathrm{Tr}\boldsymbol{\mathcal{J}}^{(\vec{s})}} = \frac{5\left(\mathrm{Det}\overline{\mathbb{C}}\right)^{1/5}}{\mathrm{Tr}\mathbb{C}} \leq 1. \tag{36}$$

Note that the quantities in the denominator do not carry a bar because they include all elements, so that $\overline{M}_\mathrm{ON}$ accounts also for how much of the measuring capabilities are wasted on the spin components of the dipole oscillation, assumed to be nonexistent. It is useful to see the form these measures take in terms of the fluorophore orientation parameters. As mentioned earlier, orientation is specified by the spherical angles η and ξ for the leading eigenvector that indicates the main direction of the fluorophore. The angle β indicates the direction of the remaining eigenvectors within the plane orthogonal to the first one, and the level of wobble is characterized by the three eigenvalues λ_j whose sum is unity. While there are other ways to parametrize the eigenvalues [41], let us use $t, \tau \in [0, 1]$ defined as

$$t = \lambda_1 - \lambda_2, \quad \tau = 2(\lambda_2 - \lambda_3). \tag{37}$$

such that

$$\lambda_1 = \frac{2 + 4t + \tau}{6}, \quad \lambda_2 = \frac{2 - 2t + \tau}{6}, \quad \lambda_3 = \frac{1 - t - \tau}{3}. \tag{38}$$

The NSG parameters can then be written in terms of the parameters $t, \tau, \eta, \xi, \beta$ by first constructing $\mathbb{m}$ according to Eqs. (7) and (8), and then substituting this construction in Eq. (4). These equations are given in Appendix B. A Jacobian matrix between the vector $\vec{s}$ (with five components) and these parameters can now be defined as the matrix of derivatives of each element with respect to each parameter:

$$\boldsymbol{\Delta} = \left[(\partial_t, \partial_\tau, \partial_\eta, \partial_\xi, \partial_\beta)\vec{s}\right]^\mathrm{T}. \tag{39}$$

The determinant of this matrix gives a very simple expression

$$\begin{aligned}\mathrm{Det}\boldsymbol{\Delta} &= \frac{\partial(s_{11}, s_{12}, s_{21}, s_{22}, s_{23})}{\partial(t, \tau, \eta, \xi, \beta)} = \frac{9}{16} t\tau(2t + \tau)\sin\eta \\ &= \frac{9}{4}(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)\sin\eta.\end{aligned} \tag{40}$$

Let us refer to the set of five parameters that determine the directional wobble of the fluorophore as $\vec{\Omega} = (t, \tau, \eta, \xi, \beta)$. (Note that this is just a shorthand notation, since unlike $\vec{s}$, $\vec{\Omega}$ does not define a meaningful vector space.) The Fisher information matrix for the parameters $\vec{\Omega}$ can be calculated then by simply multiplying $\overline{\mathcal{J}}^{(\vec{s})}$ by the Jacobian matrix on both sides:

$$\mathcal{J}^{(\vec{\Omega})} = \mathbf{\Delta}^{\dagger} \overline{\mathcal{J}}^{(\vec{s})} \mathbf{\Delta}. \tag{41}$$

The determinant of this matrix is then related to the merit function defined earlier by

$$\mathrm{Det}\mathcal{J}^{(\vec{\Omega})} = \left[\frac{9}{16} t\tau(2t+\tau) \sin\eta \right]^2 M_5^5. \tag{42}$$

Note that, in the ideal case in which the five relevant PSF elements are orthonormal while the three associated with spin are zero ($\overline{M}_{\mathrm{ON}} = 1$), this matrix reduces to

$$\mathcal{J}^{(\vec{\Omega})} \approx \frac{4}{5} \mathcal{N}_{\mathrm{s}} M_{\mathrm{PL}} M_{\mathrm{MM}} \mathbf{\Delta}^{\dagger} \mathbf{\Delta}. \tag{43}$$

The resulting expression is easily calculated but too long for reproducing here. It is worth mentioning that this matrix is not diagonal, implying some inherent coupling in the estimation between the different parameters in $\vec{\Omega}$ due to their definitions, even though there was no coupling in the estimation between the NSG parameters.

Assumption of symmetric wobble around a main dipole direction. Finally, let us study the form of the Fisher information matrix for the more restricted case corresponding to the common assumption (not necessarily justified physically) of uniform wobble within a cone around a main direction. Mathematically, this assumption implies $\lambda_2 = \lambda_3$, so that $t = P = \gamma$, $\tau = 0$, and β is irrelevant. In this case we are only interested in a smaller submatrix of the Fisher information matrix, whose rows and columns are associated with the parameters $t(= P)$, η and ξ,

$$\mathcal{J}^{(P,\eta,\xi)} = \overline{\mathbf{\Delta}}^{\dagger} \overline{\mathcal{J}}^{(\vec{s})} \overline{\mathbf{\Delta}}, \tag{44}$$

where $\overline{\Delta}$ is a 5×3 rectangular submatrix of the Jacobian, defined as

$$\overline{\mathbf{\Delta}} = \left[(\partial_t, \partial_\eta, \partial_\xi) \vec{s} \right]^{\mathrm{T}} \Big|_{t \to P, \tau \to 0}. \tag{45}$$

In the ideal case when the PSF basis achieves orthonormality such that $\overline{M}_{\mathrm{ON}} = 1$, the Fisher information matrix in Eq. (44) greatly simplifies to

$$\mathcal{J}^{(P,\eta,\xi)} \approx \frac{4}{5} \mathcal{N}_{\mathrm{s}} M_{\mathrm{PL}} M_{\mathrm{MM}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3P^2 & 0 \\ 0 & 0 & 3P^2 \sin^2\eta \end{pmatrix}. \tag{46}$$

The fact that this matrix is diagonal means that there is no coupling between the parameters, so that the standard deviations (the inverse square roots of the diagonal elements of the Fisher information matrix) satisfy

$$\sigma_P = \sigma_t \geq \sqrt{\frac{5}{4\mathcal{N}_{\mathrm{s}} M_{\mathrm{PL}} M_{\mathrm{MM}}}}, \tag{47a}$$

$$\sigma_\eta \geq \frac{1}{P} \sqrt{\frac{5}{12\mathcal{N}_{\mathrm{s}} M_{\mathrm{PL}} M_{\mathrm{MM}}}}, \tag{47b}$$

$$\sigma_\xi \geq \frac{1}{P \sin\eta} \sqrt{\frac{5}{12\mathcal{N}_{\mathrm{s}} M_{\mathrm{PL}} M_{\mathrm{MM}}}}. \tag{47c}$$

A global directional measure was defined as [17]

$$\sigma_{\mathrm{Dir}} = P^2 \sigma_P \sin\eta\, \sigma_\eta \sigma_\xi . \tag{48}$$

This measure remains finite over all values of the parameters, despite the fact that σ_ξ naturally diverges for $\eta \to 0$ and both σ_ξ and σ_η diverge for $P \to 0$. We can now see that the lower bound for this measure follows

$$\sigma_{\mathrm{Dir}} \geq \frac{5}{24}\sqrt{5}\,(\mathcal{N}_{\mathrm{s}} M_{\mathrm{PL}} M_{\mathrm{MM}})^{-3/2} \approx \frac{1}{2}(\mathcal{N}_{\mathrm{s}} M_{\mathrm{PL}} M_{\mathrm{MM}})^{-3/2}. \tag{49}$$

This idealized case helps motivate a definition of a merit function for the general case in which the PSF components are not orthonormal (and $\overline{M}_{\mathrm{ON}}$ is not unity) according to

$$M_3 = \frac{\left[\mathrm{Det}\mathcal{J}^{(P,\eta,\xi)}\right]^{1/3}}{(3P^2\sin\eta)^{2/3}}. \tag{50}$$

In the ideal case discussed earlier in which the three parameters are decoupled, $M_3 = 1/(3\sigma_{\mathrm{Dir}})^{2/3}$. When $\overline{M}_{\mathrm{ON}} = 1$, this measure is equivalent to M_5, but as will be shown in the examples in the next section, M_3 can be different from zero in cases when M_5 vanishes. One can show that, in either of the limits $P \to 0$ or $\sin\eta \to 0$, M_3 reduces to a finite value, although this value can depend on the angular parameters even for $P \to 0$. Note that, for the purpose of understanding the different factors, we can also write this merit function as a product of four factors:

$$M_3(\vec{s}) = \frac{4}{5}\mathcal{N}_{\mathrm{s}}\, M_{\mathrm{PL}}\, \overline{\overline{M}}_{\mathrm{ON}} M_{\mathrm{MM}}, \tag{51}$$

where $\overline{\overline{M}}_{\mathrm{ON}}$ is a measure of basis orthonormality when applied to this type of situation, whose calculation is not direct but through solving this equation.

6. Estimation of background and number of signal photons

The measures just discussed are defined to account for the linear independence between the PSF components associated with the NSG parameters. However, the accuracy of a method will also depend on how different these components are to $\mathcal{I}_0$ and even to a constant background. Consider first the relation between a constant intensity background and the PSF components $\mathcal{I}_k$. If, as discussed in Sec. 3, the NA is chosen so that the detected intensity is independent of polarization, $\langle \mathcal{I}_k \rangle_{\mathbf{x},\sigma} \approx 0$ and so the PSF components are approximately orthogonal to the background in the limit of large I_{b}. In the opposite limit, if there is no background, this orthogonality is clearly not an issue. For intermediate cases in which the weight function differs from a constant, there could be some slight correlations between background and $\mathcal{I}_k$, but these are expected to be small.

The orthogonality between $\mathcal{I}_0$ and $\mathcal{I}_k$ is not automatically guaranteed, even though it is expected that the inner products will not be large given that the latter PSF components contain positive and negative values in roughly equal amounts, while $\mathcal{I}_0 \geq 0$. Note that due to their non-negativity, $\mathcal{I}_0$ and I_{b} are more significantly correlated under the inner product. This correlation can introduce bias in the measurements, affecting mostly the estimation of the degree of polarization P, and hence the level of wobbling of a molecule. This correlation can be reduced by using a window that is larger than the extent of $\mathcal{I}_0$.

7. Examples

In this section, we apply the formalism proposed here to two different types of measurement techniques. The first is a PSF engineering technique proposed recently called *coordinate and*

height superresolution imaging with dithering and orientation (CHIDO) [17], which allows in principle full 3D polarimetry. The second is a radiometric technique consisting on separation into four polarization components together with segmentation of the pupil, offering more limited information.

7.1. CHIDO

This PSF-shaping technique relies on the use at the pupil plane of a stress-induced spatially-varying birefringent mask, referred to as a stress-engineered optic (SEO) [42] described by the Jones matrix

$$\mathbb{J}(\mathbf{u}) = \cos\left(\frac{cu}{2\mathrm{NA}}\right)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \mathrm{i}\,\sin\left(\frac{cu}{2\mathrm{NA}}\right)\begin{pmatrix} \cos\varphi & -\sin\varphi \\ -\sin\varphi & -\cos\varphi \end{pmatrix}, \tag{52}$$

where c is a parameter proportional to the amount of stress and that determines the strength of the increase in birefringence with distance from the center of the pupil. This mask is followed by separation into two circular polarization channels: $\mathbf{p}_\sigma = (1, \sigma\mathrm{i})^{\mathrm{T}}/\sqrt{2}$ for $\sigma = \pm 1$.

Figure 4(a) shows M_8, M_5, and M_3 for one signal photon, averaged over all the valid values of $\vec{s}$ for each case (through Monte Carlo estimation of the integrals by using 500 values), for different values of c and I_b. The standard deviations of each measure are also shown. For sufficiently low background, all three merit functions peak at $c \approx 2\pi$. However, as the background grows, the curves drop and tilt, and ultimately M_5 and M_3 peak at around $c \approx 0.5\pi$, which is also one of the main peaks of M_8 except that this merit function takes slightly higher values at $c = 0$ (no SEO). This last peak for M_8 is due to the fact that, in the absence of a birefringent mask, the spin components are easier to estimate than the remaining NSG parameters given the circular polarization channel splitting. As the background grows further, the curves maintain their general shape but simply scale down approximately inversely proportionally to I_b. Note that the standard deviations for M_8 and M_5 tend to zero as the background grows, since the dependence of these functions of the NSG parameters decreases; this not the case, though, for M_3.

The reason why larger/smaller values of c are preferable in the presence of smaller/larger background can be visualized from Fig. 4(b), which shows M_PL, M_ON and M_MM for the two extreme cases in (a): no background, and a strong background equal to four times the peak value of $\mathcal{I}_0$ without SEO. Since we assumed small pixels and we are using polarization splitting, M_MM is essentially unity in both cases. Similarly, the shape of M_ON does not change much with background and the main peak is for c between 2π and 2.5π, for which the modification of the field distribution at the pupil by the SEO makes the PSF components $\mathcal{I}_k$ sufficiently different, at the cost of spreading them spatially. The factor that changes more importantly with background is M_PL, which not only drops with growing background, but for a given non-zero background it decreases with c due to the dilution of the photons over more pixels. The PSF basis $\mathcal{I}_k$ over the two polarization channels is shown in Fig. 4(c) for $c = \pi/2$ and $c = 2\pi$, and the corresponding elements of the Fisher information matrix for these values of c in the absence of background and for a fully unpolarized dipole ($P = 0$) are represented in Fig. 4(d). We see that the PSFs become considerably larger for larger c and this causes a drop in M_PL, but in the absence of background, the more complex structure of the components $\mathcal{I}_k$ allows getting closer to orthonormality.

7.2. Pupil splitting radiometric

We now consider what is called a radiometric approach, in which the total intensity in each channel is integrated, and hence the estimation is based only on the ratios of a small number of measurement intensities (making the data analysis computationally simple). In fluorescence microscopy, this approach relies on splitting the pupil into different channels in order to estimate the off-plane angle η, in combination with linear polarization projections for estimating the

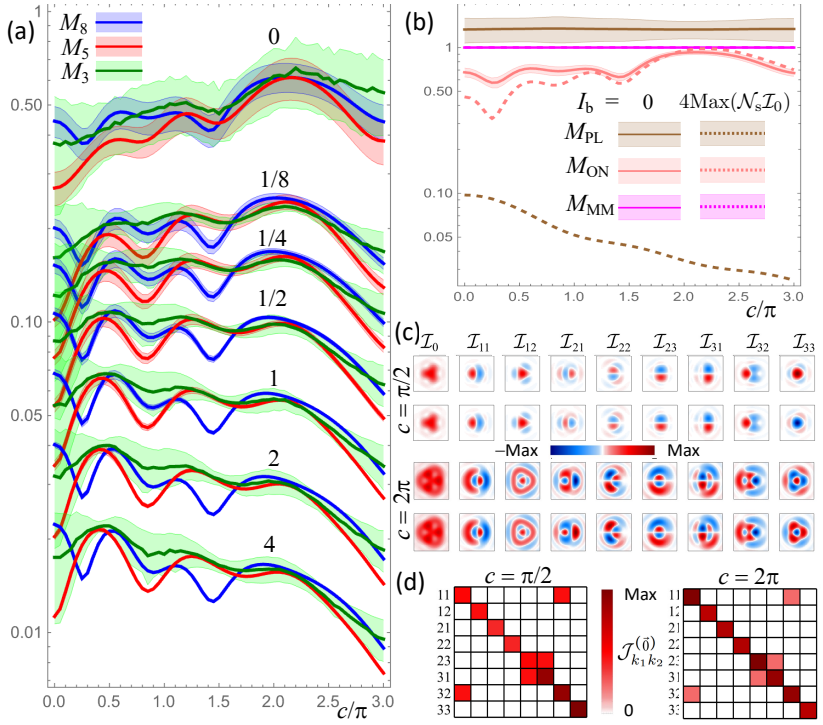

Fig. 4. (a) Averages (solid lines) and standard deviations (pale color bands) for the measures M_8, M_5 and M_3 in terms of the stress-birefringence parameter c, for several levels of background I_b indicated on the right in units of the maximum of the PSF component $\mathcal{I}_0$ for $c = 0$ (no birefringence) times the number of signal photons. (b) Averages (solid and dashed lines) and standard deviations (pale color bands) for the measures of pixel localization M_{PL}, orthonormality M_{ON}, and monomodality M_{MM}, in terms of the stress-birefringence parameter c, for the two extreme cases of background intensity I_b in part (a): no background (solid lines) and 4 times the peak intensity (dashed lines). (c) PSF basis over both polarization channels for $c = \pi/2$ (top two rows) and $c = 2\pi$ (bottom two rows), where the shown areas represent a square in object space of length 1.2λ. Note that each case was normalized to its maximum and minimum values; the PSFs for $c = 2\pi$ would be considerably fainter than those for $c = \pi/2$ due to the larger area they occupy. (d) Elements of the Fisher matrix for $c = \pi/2$ (left), and $c = 2\pi$ (right), both for $\vec{s} = \vec{0}$ and $I_b = 0$, and where each was normalized to its maximum.

in-plane angle ξ. Techniques of this type have been proposed that employ three [23, 24] or four [12] channels. Here, we explore a version of this approach that is a combination of these prior versions, as illustrated in Fig. 5(a): the pupil is split into two channels through an annular mirror, the central part having a radius $\mathrm{NA_c}$, and the outer annular region extending up to $\mathrm{NA} = n_0$. The central channel is subsequently split into two linear polarization channels, corresponding to horizontal and vertical polarizations, while the outer annular channel is split into two linear polarization channels for polarizations at ±45°. The light captured by these four channels is integrated into four "pixels" that compose the PSF.

Given this reduced number of pixels, it is not possible to recover information about all five NSG parameters required for characterizing general wobbling linear dipoles: only s_{11}, s_{12} and s_{23} can be estimated, as shown in Figs. 5(b) and (c). This means that $M_8 = M_5 = 0$. Nevertheless, M_3 can differ from zero, so that if we make the assumption of linear dipoles wobbling symmetrically around a main direction ($\lambda_2 = \lambda_3$), it is possible to estimate the parameters η, ξ and $P = \gamma$, albeit with a sign ambiguity for the off-plane angle [12] (as can be seen from the equations in Appendix B for s_{11}, s_{12} and s_{23} with $\tau = 0$). Figure 5(d) shows plots of the average and standard deviation of M_3 as functions of the cutoff numerical aperture $\mathrm{NA_c}$ under different levels of background intensity (expressed in fractions of the average signal intensity per pixel). In this case we see that the merit function peaks at $\mathrm{NA_c} \approx 0.8 n_0$ regardless of background. This is because for this value of the cutoff NA, $\mathcal{I}_{12}$ is significant and approximately orthogonal to $\mathcal{I}_0$ as shown in Fig. 5(b), where we can also see that the only two other nonzero PSF elements, $\mathcal{I}_{11}$ and $\mathcal{I}_{23}$, are orthogonal to each other and to both $\mathcal{I}_0$ and $\mathcal{I}_{12}$. Note also that M_3 drops to zero in the limiting cases when $\mathrm{NA_c}$ goes to zero or n_0, since in this case we only have two channels. The values of the relevant elements of the resulting Fisher information matrix are shown in Fig. 5(c) for the case of a fully wobbling dipole ($P = 0$) in the absence of background intensity. Figure 5(e) shows the factors contributing to M_3. Note in particular that $M_{\mathrm{MM}} \approx 0.3$, which is considerably smaller than unity despite the use of polarization splitting, given the radiometric nature of this approach in which the light in each channel is integrated (so that the pixels are "large").

8. Concluding remarks

We studied the characterization of the 3D state of polarization of a localized source, such as a fluorophore or a sub-wavelength scatterer, by using a high NA optical system. This work is based on the use of the eight NSG parameters, which are generalizations of the Stokes parameters in 2D polarization, to characterize the second moment matrix even if spin is present. For convenience, we chose the conditions under which the detected power is independent of the orientation of the polarization state, which for a standard microscope corresponds approximately to using an NA equal to the refractive index of the embedding medium. Based on the Fisher information matrix, we proposed definitions of merit functions that quantify the ability to measure the second moment matrix in several specific situations: full 3D polarimetry (including spin), general wobbling linear fluorophores, and linear fluorophores assumed to wobble symmetrically around a main direction. In all cases, we provided an analysis of the factors that contribute to these merit functions, which include the level of orthonormality of the PSF elements, the concentration of the measurement into a small set of pixels, and the coupling of each detector pixel to a specific polarization state.

An important conceptual aspect of the analysis presented here comes from the relation in Eqs. (22) and its generalization in (23). These relations can be interpreted in terms of the reversibility of light. We can imagine a situation in which light emerges from each pixel of the detector and travels towards the fluorophore. By the time it gets there it will have a 3D polarization specified by normalized Stokes-Gell-Mann parameters proportional to $\mathcal{I}_k$. This allows understanding which polarization components couple naturally to each detector pixel. A good measurement is then one in which we get a well-balanced coupling of all polarizations that are relevant to the measurement, so we can sacrifice sensitivity to some polarization states (for

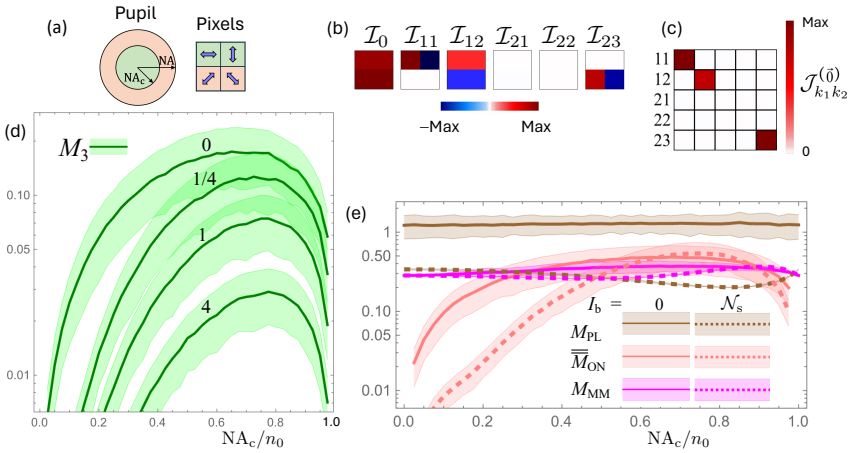

Fig. 5. (a) Schematic of the radiometric technique: splitting of the BFP into two orthogonal linearly polarized channels at the detector (DET), namely, 0 and 90° for the inner section (low NA), and ±45° for the outer annulus (high NA). (b) Stokes basis and (c) Fisher information matrix for $\vec{s} = \vec{0}$ and $\mathrm{NA_c} = 0.8n_0$. (d) Average (solid lines) and standard deviation (pale color bands) for the measure M_3 in terms of the cutoff numerical aperture, for several levels of background I_b indicated on the right in units of the average signal intensity per pixel. (e) Averages (solid and dashed lines) and standard deviations (pale color bands) for the measures of pixel localization M_PL, orthonormality $\overline{\overline{M}}_\mathrm{ON}$, and monomodality M_MM, in terms of the cutoff numerical aperture, for the two extreme cases of background intensity I_b in part (d): no background (solid lines) and 4 times the average signal intensity (dashed lines).

example, those with spin) in order to boost the ability to measure others (e.g. linear polarizations).

We illustrated the use of these merit functions in the design of measurement systems by considering two examples, a system capable of measuring all eight NSG parameters and one specific to linear dipoles wobbling with axially-symmetric statistics. In both cases, a physical parameter of the system can be optimized based on the merit functions, although the optimal value can depend on the level of background.

Note that the spatial localization of the sources was not discussed here. In applications like SMOLM, this localization must be estimated jointly with the 3D polarization, which would imply using the determinant of the Fisher information matrix for all parameters, including the three spatial variables. While the estimation of position and direction are often intrinsically coupled [43], basic estimates of the bounds to these measures can be obtained by considering them separately [21], and techniques have been proposed that allow the estimation of the localization without precise knowledge of direction/polarization [17].

Appendix A: Power dependence on polar orientation

When the numerical aperture is chosen at the limit with SAF, namely $\mathrm{NA} = n_0$, the expressions for the captured powers for a transverse and a longitudinal dipole can be simplified to the following

expressions:

$$\Phi_\perp = \frac{16\pi\alpha\left(3\nu^4+2\right)\nu^5}{a^7b^4} + \frac{16\pi\nu}{15a^6b^4}\Big[(\nu-1)(6\nu^{10}+6\nu^9+19\nu^8-21\nu^7 + 47\nu^6+26\nu^5+32\nu^4-26\nu^3-6\nu^2-4\nu-4) - b\left(6\nu^{10}+\nu^8+58\nu^6-2\nu^4+11\nu^2+1\right)\Big], \tag{53a}$$

$$\Phi_z = \frac{32\pi\alpha\nu^7\left(\nu^4-2\nu^2+2\right)}{a^7b^4} - \frac{32\pi\nu^3}{3a^6b^4}\Big[(\nu-1)\nu^2(\nu^6+\nu^5-\nu^4-10\nu^3 + 4\nu+2) - b(\nu^8-6\nu^6+6\nu^4-3\nu^2-1)\Big], \tag{53b}$$

where $\nu = n_1/n_0$, $\alpha = \operatorname{arctanh}\left(\frac{\sqrt{\nu^4-1}}{\nu^2}\right) - \operatorname{arctanh}\left(\frac{\nu^2-1}{\nu^3+1}\sqrt{\nu^2+1}\right)$, $a = \sqrt{\nu^2+1}$, and $b = \sqrt{\nu^2-1}$. These two captured powers can be shown to agree when $\nu \approx 1.33$.

Appendix B: NSG parameters in terms of fluorophore orientation parameters

The five relevant NSG parameters for a spinless dipole such as a wobbling fluorophore can be written in terms of the orientation angles η, ξ, β and the eigenvalues λ_i through the combinations $\gamma = \lambda_1 - (\lambda_2+\lambda_3)/2$ and $\tau = 2(\lambda_2-\lambda_3)$ as

$$s_{11} = \frac{\sqrt{3}}{16}\left\{\left[\tau(\cos 2\eta+3)\cos 2\beta + 4\gamma\sin^2\eta\right]\cos 2\xi - 4\tau\cos\eta\sin 2\beta\sin 2\xi\right\}, \tag{54a}$$

$$s_{12} = -\frac{1}{8}\left[3\tau\sin^2\eta\cos 2\beta + \gamma(3\cos 2\eta+1)\right], \tag{54b}$$

$$s_{21} = \frac{\sqrt{3}}{4}\left[(2\gamma-\tau\cos 2\beta)\cos\eta\sin\xi - 2\tau\cos\beta\sin\beta\cos\xi\right]\sin\eta, \tag{54c}$$

$$s_{22} = \frac{\sqrt{3}}{4}\left[(2\gamma-\tau\cos 2\beta)\cos\eta\cos\xi + \tau\sin\xi\sin 2\beta\right]\sin\eta, \tag{54d}$$

$$s_{23} = \frac{\sqrt{3}}{16}\left\{4\tau\cos\eta\cos 2\xi\sin 2\beta + \left[\tau(\cos 2\eta+3)\cos 2\beta + 4\gamma\sin^2\eta\right]\sin 2\xi\right\}. \tag{54e}$$

Interestingly, in the non-wobbling limit corresponding to $\gamma = 1$ and $\tau = 0$, these parameters turn out to be linked to spherical harmonics with first index $\ell = 2$:

$$s_{11} = \frac{\sqrt{3}}{2}\sin^2\eta\cos 2\xi = 2\sqrt{\frac{2\pi}{5}}\,\mathrm{Re}[Y_{2,2}(\eta,\xi)], \tag{55a}$$

$$s_{12} = -\frac{1}{4}(3\cos 2\eta+1) = 2\sqrt{\frac{\pi}{5}}\,Y_{2,0}(\eta,\xi), \tag{55b}$$

$$s_{21} = \frac{\sqrt{3}}{2}\sin 2\eta\sin\xi = -2\sqrt{\frac{2\pi}{5}}\,\mathrm{Im}[Y_{2,1}(\eta,\xi)], \tag{55c}$$

$$s_{22} = \frac{\sqrt{3}}{2}\sin 2\eta\cos\xi = -2\sqrt{\frac{2\pi}{5}}\,\mathrm{Re}[Y_{2,1}(\eta,\xi)], \tag{55d}$$

$$s_{23} = \frac{\sqrt{3}}{2}\sin^2\eta\sin 2\xi = 2\sqrt{\frac{2\pi}{5}}\,\mathrm{Im}[Y_{2,2}(\eta,\xi)]. \tag{55e}$$

Let us now consider the situation of a wobbling fluorescent dipole within an angular distribution function $f(\theta,\phi)$ where (θ,ϕ) give the orientation of the dipole in the (x,y,z) frame during the averaging detection time. The new NSG parameters become the average of the fixed-dipole

parameters, due to the linear relation between the contributions of all dipoles contributing to the fluorescence signal:

$$s_{11} = 2\sqrt{\frac{2\pi}{5}} \int\int \mathrm{Re}[Y_{2,2}(\theta,\phi)]\, f(\theta,\phi)\, \sin\theta d\theta d\phi, \tag{56a}$$

$$s_{12} = 2\sqrt{\frac{\pi}{5}} \int\int Y_{2,0}(\theta,\phi)\, f(\theta,\phi)\, \sin\theta d\theta d\phi, \tag{56b}$$

$$s_{21} = -2\sqrt{\frac{2\pi}{5}} \int\int \mathrm{Im}[Y_{2,1}(\theta,\phi)]\, f(\theta,\phi)\, \sin\theta d\theta d\phi, \tag{56c}$$

$$s_{22} = -2\sqrt{\frac{2\pi}{5}} \int\int \mathrm{Re}[Y_{2,1}(\theta,\phi)]\, f(\theta,\phi)\, \sin\theta d\theta d\phi, \tag{56d}$$

$$s_{23} = 2\sqrt{\frac{2\pi}{5}} \int\int \mathrm{Im}[Y_{2,2}(\theta,\phi)]\, f(\theta,\phi)\, \sin\theta d\theta d\phi. \tag{56e}$$

Assuming a decomposition of the distribution function on spherical harmonics $f(\theta,\phi) = \sum_{l,m} f_{l,m} Y_{l,m}(\theta,\phi)$ with $-l \leq m \leq l$. these expressions, by orthogonality of the spherical harmonics, become:

$$s_{11} = 2\sqrt{\frac{2\pi}{5}}\, \mathrm{Re}[f_{2,2}], \tag{57a}$$

$$s_{12} = 2\sqrt{\frac{\pi}{5}}\, f_{2,0}, \tag{57b}$$

$$s_{21} = -2\sqrt{\frac{2\pi}{5}}\, \mathrm{Im}[f_{2,1}], \tag{57c}$$

$$s_{22} = -2\sqrt{\frac{2\pi}{5}}\, \mathrm{Re}[f_{2,1}], \tag{57d}$$

$$s_{23} = 2\sqrt{\frac{2\pi}{5}}\, \mathrm{Im}[f_{2,2}]. \tag{57e}$$

As a consequence, the only accessible information of the wobbling distribution function of the molecule is limited to its second order terms ($l = 2, -2 \leq m \leq 2$). The terms $f_{l=2,m}$ (and hence the NSG parameters) can be expressed easily supposing a distribution of cylindrical symmetry along z written as $\sum_l f_{l,0} Y_{l,0}(\theta,\phi)$ (with here only $l = 2$ pertinent), rotated by the angles (η,ξ). This rotation uses the Wigner D-matrix elements $D^l_{m',m}(\eta,\xi)$:

$$Rot_{(\eta,\xi)}(Y_{l,m=0}(\theta,\phi)) = \sum_{m'=-l}^{l} D^l_{m',m=0}(\eta,\xi) Y_{l,m'}(\theta,\phi) \tag{58}$$

Therefore :

$$f_{l,m} = f_{l,0}\, D^l_{m,0}(\eta,\xi). \tag{59}$$

$D^l_{m,0}(\eta,\xi)$ is related to the associated Legendre polynomials functions P^m_l with:

$$D^l_{m,0}(\eta,\xi) = \sqrt{\frac{(l-m)!}{(l+m)!}} P^m_l(\cos\eta)\, \exp(-im\xi) \tag{60}$$

As a consequence, each NSG parameter can be written as a product of functions involving three independent variables $(f_{l,0},\eta,\xi)$, proportional to:

$$f_{l=2,m} = f_{2,0} \sqrt{\frac{(2-m)!}{(2+m)!}} P^m_2(\cos\eta)\, \exp(-im\xi) \tag{61}$$

with : $m = 2$ for $s_{11} \propto \mathrm{Re}(\mathrm{f}_{2,2})$ and $s_{23} \propto \mathrm{Im}(\mathrm{f}_{2,2})$, $m = 0$ for $s_{12} \propto Re(f_{2,0})$, $m = 1$ for $s_{21} \propto \mathrm{Im}(\mathrm{f}_{2,1})$ and $s_{22} \propto \mathrm{Re}(\mathrm{f}_{2,1})$. This expression emphasizes the effect of wobbling on the norm of the Stokes vector $P = |\vec{s}|$. Since all of the NSG parameters are proportional to $f_{2,0}$, this $(l = 2, m = 0)$ term of the spherical decomposition of the distribution function is the determinant factor quantifying wobbling. This term, which is nothing else than the order parameter of the distribution function to the order 2, is expressed as:

$$f_{2,0} = \int \int Y_{2,0}(\theta, \phi)\, f(\theta, \phi)\, \sin\theta d\theta d\phi, \tag{62}$$

with $Y_{2,0}(\theta, \phi) = \frac{1}{4}\sqrt{\frac{5}{\pi}}(2\cos^2\theta - 1)$. In the case where $f(\theta, \phi)$ is a symmetric cone distribution of width δ, this wobbling factor is equal to $f_{2,0} = 2(\cos^3\frac{\delta}{2} - \cos\frac{\delta}{2})$, which can be related to the eigenvalue of the moment matrix defined in [28].

Funding. Agence Nationale de la Recherche (ANR-20-CE42-0003, ANR21-CE24-0014, France 2030 Investment Plan IDEC ANR-21-ESRE-0002, Investissements d'Avenir CENTURI ANR-16-CONV-0001, France BioImaging National Infrastructure ANR-10-INBS-04).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. P. Soleillet, "Sur les paramètres caractérisant la polarisation partielle de la lumière dans les phénomènes de fluorescence," in *Annales de physique,* vol. 10 (1929), pp. 23–97.
2. J. C. Samson, "Descriptions of the polarization states of vector processes: Applications to ULF magnetic fields," Geophys. J. Int. **34**, 403–419 (1973).
3. M. A. Alonso, "Geometric descriptions for the polarization of nonparaxial light: a tutorial," Adv. Opt. Photon. **15**, 176–235 (2023).
4. L. Novotny, M. R. Beversluis, K. S. Youngworth, and T. G. Brown, "Longitudinal field modes probed by single molecules," Phys. Rev. Lett. **86**, 5251–5254 (2001).
5. T. Bauer, S. Orlov, U. Peschel, *et al.*, "Nanointerferometric amplitude and phase reconstruction of tightly focused vector beams," Nat. Photonics **8**, 23–27 (2014).
6. T. Bauer, P. Banzer, E. Karimi, *et al.*, "Observation of optical polarization möbius strips," Science **347**, 964–966 (2015).
7. S. Tsesses, E. Ostrovsky, K. Cohen, *et al.*, "Optical skyrmion lattice in evanescent electromagnetic fields," Science **361**, 993–996 (2018).
8. E. Otte, K. Tekce, S. Lamping, *et al.*, "Polarization nano-tomography of tightly focused light landscapes by self-assembled monolayers," Nat. Commun. **10**, 4308 (2019).
9. S. Brasselet and M. A. Alonso, "Polarization microscopy: from ensemble structural imaging to single-molecule 3D orientation and localization microscopy," Optica **10**, 1486–1510 (2023).
10. J. T. Fourkas, "Rapid determination of the three-dimensional orientation of single molecules," Opt. Lett. Vol. 26, Issue 4, pp. 211-213 **26**, 211–213 (2001).
11. C. V. Rimoli, C. A. Valades-Cruz, V. Curcio, *et al.*, "4polar-storm polarized super-resolution imaging of actin filament organization in cells," Nat. Commun. **13**, 301 (2022).
12. M. Sison, C. A. Valades Cruz, C. S. Senthil Kumar, *et al.*, "4polar3D smolm : single molecule orientation and localization microscopy using a simple pupil diaphragm and ratiometric polarization splitting," (in preparation) .
13. F. Aguet, S. Geissbühler, I. Märki, *et al.*, "Super-resolution orientation estimation and localization of fluorescent dipoles using 3-d steerable filters," Opt. Express **17**, 6829–6848 (2009).
14. A. S. Backer and W. E. Moerner, "Extending single-molecule microscopy using optical fourier processing," The J. Phys. Chem. B **118**, 8313–8329 (2014).
15. M. P. Backlund, M. D. Lew, A. S. Backer, *et al.*, "Simultaneous, accurate measurement of the 3D position and orientation of single molecules," Proc. Nat. Acad. Sci. **109**, 19087–92 (2012).
16. O. Zhang, J. Lu, T. Ding, and M. D. Lew, "Imaging the three-dimensional orientation and rotational mobility of fluorescent emitters using the tri-spot point spread function," Appl. Phys. Lett. **113** (2018).
17. V. Curcio, L. A. Alemán-Castañeda, T. G. Brown, *et al.*, "Birefringent fourier filtering for single molecule coordinate and height super-resolution imaging with dithering and orientation," Nat. Commun. **11** (2020).
18. O. Zhang and M. D. Lew, "Fundamental limits on measuring the rotational constraint of single molecules using fluorescence microscopy," Phys. Rev. Lett. **122** (2019).

19. O. Zhang and M. D. Lew, "Quantum limits for precisely estimating the orientation and wobble of dipole emitters," Phys. Rev. Res. **2** (2020).
20. A. Agrawal, S. Quirin, G. Grover, and R. Piestun, "Limits of 3D dipole localization and orientation estimation for single-molecule imaging: towards Green's tensor engineering," Opt. Express **20**, 26667–26680 (2012).
21. O. Zhang and M. D. Lew, "Single-molecule orientation localization microscopy i: fundamental limits," J. Opt. Soc. Am. A **38**, 277 (2021).
22. O. Zhang and M. D. Lew, "Single-molecule orientation localization microscopy II: a performance comparison," J. Opt. Soc. Am. A **38**, 288 (2021).
23. J. Hohlbein and C. G. Hübner, "Three-dimensional orientation determination of the emission dipoles of single molecules: The shot-noise limit," The J. Chem. Phys. **129**, 094703 (2008).
24. R. Börner, N. Ehrlich, J. Hohlbein, and C. G. Hübner, "Single molecule 3d orientation in time and space: A 6d dynamic study on fluorescently labeled lipid membranes," J. Fluoresc. **26**, 963–975 (2016).
25. T. Ding, T. Wu, H. Mazidi, *et al.*, "Single-molecule orientation localization microscopy for resolving structural heterogeneities between amyloid fibrils," Optica **7**, 602–607 (2020).
26. S. Stallinga, "Effect of rotational diffusion in an orientational potential well on the point spread function of electric dipole emitters," J. Opt. Soc. Am. A **32**, 213 (2015).
27. E. Munger, M. Sison, and S. B. let, "Influence of the excitation polarization on single molecule 3d orientation imaging," Opt. Commun. **541**, 129480 (2023).
28. A. S. Backer and W. E. Moerner, "Determining the rotational mobility of a single molecule from a single image: a numerical study," Opt. Express **23**, 4255–4276 (2015).
29. M. Gell-Mann, "Symmetries of baryons and mesons," in *The Eightfold Way,* (CRC Press, 2018), pp. 216–233.
30. K. Y. Bliokh, A. Y. Bekshaev, and F. Nori, "Extraordinary momentum and spin in evanescent waves," Nat. Commun. **5** (2014).
31. R. Barakat, "Degree of polarization and the principal idempotents of the coherency matrix," Opt. Commun. **23**, 147–150 (1977).
32. T. Setälä, M. Kaivola, and A. T. Friberg, "Degree of polarization in near fields of thermal sources: Effects of surface waves," Phys. Rev. Lett. **88** (2002).
33. J. J. Gil, J. M. Correas, P. A. Melero, and C. Ferreira, "Generalized polarization algebra," Monografias del Semin. Matematico Garcia de Galdeano **31**, 161–167 (2004).
34. J. Ellis, A. Dogariu, S. Ponomarenko, and E. Wolf, "Degree of polarization of statistically stationary electromagnetic fields," Opt. Commun. **248**, 333–337 (2005).
35. M. R. Dennis, "A three-dimensional degree of polarization based on rayleigh scattering," J. Opt. Soc. Am. A **24**, 2065 (2007).
36. G. Björk, H. de Guise, A. B. Klimov, *et al.*, "Classical distinguishability as an operational measure of polarization," Phys. Rev. A **90** (2014).
37. L. A. Alemán-Castañeda, S. Y.-T. Feng, R. Gutiérrez-Cuevas, *et al.*, "Using fluorescent beads to emulate single fluorophores," JOSA A **39**, C167–C178 (2022).
38. R. Gutiérrez-Cuevas, L. A. Alemán-Castañeda, I. Herrera, *et al.*, "Vectorial phase retrieval in super-resolution polarization microscopy," APL Photonics **9** (2024).
39. T. Ruckstuhl, J. Enderlein, S. Jung, and S. Seeger, "Forbidden light detection from single molecules," Anal. Chem. **72**, 2117–2123 (2000).
40. N. Bourg, C. Mayet, G. Dupuis, *et al.*, "Direct optical nanoscopy with axially localized detection," Nat. Photonics **9**, 587–593 (2015).
41. W. Zhou, T. Wu, and M. D. Lew, "Fundamental limits in measuring the anisotropic rotational diffusion of single molecules," The J. Phys. Chem. A (2024).
42. A. K. Spilman and T. G. Brown, "Stress birefringent, space-variant wave plates for vortex illumination," Appl. optics **46**, 61–66 (2007).
43. J. Enderlein, E. Toprak, and P. R. Selvin, "Polarization effect on position accuracy of fluorophore localization," Opt. Express **14**, 8111–8120 (2006).