

Spatio-angular fluorescence microscopy I. Basic theory (Draft: November 24, 2018)

TALON CHANDLER,^{1,*} AUTHORS TBD,^{2,*} RUDOLF OLDENBOURG^{2,*},
AND PATRICK LA RIVIÈRE¹

¹University of Chicago, Department of Radiology, Chicago, Illinois 60637, USA

²Publications Department, The Optical Society, 2010 Massachusetts Avenue NW, Washington, DC 20036, USA

³Currently with the Department of Electronic Journals, The Optical Society, 2010 Massachusetts Avenue NW, Washington, DC 20036, USA

*talonchandler@talonchandler.com

Abstract: We introduce the basic elements of a spatio-angular theory of fluorescence microscopy. We start by modeling an aplanatic microscope imaging an ensemble of in-focus fluorescent dipoles as a linear Hilbert-space operator, and we express the operator in terms of four different sets of object-space basis functions. We show that the operator takes a particularly convenient form when expressed in a basis of spatial and spherical harmonics—a form we call the dipole spatio-angular transfer function. We demonstrate our formalism by analyzing a single-view fluorescence microscope without using the monopole or scalar approximations. We show that this imaging system has an angular band limit, and we demonstrate the value of the transfer function approach by efficiently simulating the imaging process with simple phantoms. Notably, we show that information about the out-of-plane orientation of ensembles of in-focus fluorophores is recorded by fluorescence microscopes. We discuss the implications of our analysis for all quantitative fluorescence microscopy studies and lay out a path towards a complete theory.

1. Introduction

Fluorescence microscopes are widely used in the biological sciences for measuring the spatial distribution of fluorophores throughout a sample [cite]. While an unprocessed fluorescence micrograph reports the approximate distribution of fluorophores throughout a sample, all microscopes are diffraction limited, so the image is a blurred version of the true fluorophore distribution.

Restoration techniques attempt to recover the true distribution of fluorophores using the measured data and a model of the imaging process. A model of the imaging process can be obtained theoretically (by mathematically modeling the instrument under idealized conditions), experimentally (by measuring the instrument’s response to a known input), or by a combination of theory and experiment (by measuring parameters of an instrument model). In all cases the accuracy of the restored fluorophore distribution is limited by the accuracy of the imaging model. All theoretical imaging models make simplifying approximations that limit the accuracy of restorations, so it is important to verify that the approximations introduce an acceptable level of error. This work investigates the errors introduced by two common approximations in models of fluorescence microscopes—the *monopole approximation* and the *scalar approximation*.

The widely-used monopole model treats fluorophores as molecules that absorb and emit light isotropically. Despite their use in models, monopole absorbers/emitters do not exist in nature. All real fluorophores absorb and emit light anisotropically, and in this work we investigate the more realistic dipole model of fluorophores. Since the absorption and emission patterns of many widely used fluorophores including green fluorescent protein (GFP) are known to be well described by the dipole model [cite Inoue], we start by considering the dipole model before higher-order multipole models.

The scalar approximation is another widely-used approximation in fluorescence microscopy.

The scalar approximation models the propagation of light using a scalar-valued field U instead of the more realistic vector-valued electric field \mathbf{E} . Although the scalar model yields realistic results in many cases, the scalar field ignores polarization-dependent effects. The absorption and emission response of dipoles strongly depends on the polarization of the fields, so we need to use an electromagnetic model of light propagation to accurately model a microscope imaging dipoles.

This work lies at the intersection of three classes of fluorescence microscopy: (1) ensemble imaging techniques where many fluorophores are imaged in parallel and the goal is to find their concentration throughout the sample, (2) single molecule localization microscopy (SMLM) techniques where fluorophores are sparse in the sample and the goal is to find the position of each fluorophore, and (3) polarized fluorescence techniques where many fluorophores are imaged in parallel and the goal is to find their concentration and orientation. We briefly review these three classes and focus on their use of the monopole and scalar approximations.

The large majority of fluorescence microscopes are used to image ensembles of fluorophores, and many widely used restoration techniques make either or both of the monopole and scalar approximations. The Gibson-Lanni model [1] and the Born and Wolf model [2] are two examples of scalar models that are widely used to restore data collected by fluorescence microscopes, and the Richards and Wolf model [3] is another widely-used model that accounts for vectorial diffraction. All three of these models describe how optical fields propagate through an optical system, but they are agnostic to the object that is being imaged. In other words, these models are *optical* models, and when they are used to model fluorescent objects they implicitly make the monopole approximation.

The SMLM community has pioneered the use of rigorous electromagnetic models for dipole imaging systems that do not adopt the scalar or monopole approximations [4–6]. When a single molecule fluoresces, its irradiance pattern on the detector is strongly dependent on the orientation of the emitting molecule. Backlund et al. [7] have shown that ignoring the orientation of fluorophores can bias position estimates, so the most accurate SMLM experiments must jointly estimate the position and orientation of each molecule. Some SMLM experiments use polarizers to estimate the orientation of single molecules more accurately [Backer DNA paper, other Moerner papers], and these works do not adopt the scalar or monopole approximations.

A smaller community of microscopists is interested in measuring the orientation and position of ensembles of fluorophores [8] [Forkey, Goldman, Moerner, Oldenbourg]. These techniques typically use polarized illumination or polarized detection to make multiple measurements of the same object. Current angular restoration techniques use a model of the dipole excitation and emission processes [Fourkas] to recover the orientation of fluorophores using pixel-wise arithmetic, but these techniques do not perform any spatial restoration so they do not use all of the available information.

In this work we place these three communities on a common theoretical footing. We begin in Section 2 by formulating the imaging problem mathematically. We start with an abstract description of fluorescence imaging systems and extend the usual monopole imaging model to dipoles. Along the way we introduce transfer functions that will allow us to efficiently analyze spatio-angular microscopes. In Section 3 we use the mathematical tools we developed in Section 2 to analyze and simulate a single-view epi-fluorescence microscope, and we demonstrate our model with four simple phantoms. Finally, in Section 4 we discuss the results and their broader implications.

2. Theory

We begin our analysis with the abstract Hilbert space formalism of Barrett and Myers [9]. Our first task is to formulate the imaging process as a mapping between two Hilbert spaces $\mathcal{H} : \mathcal{U} \rightarrow \mathcal{V}$, where \mathcal{U} is a set that contains all possible objects, \mathcal{V} is a set that contains all possible datasets, and \mathcal{H} is a model of the instrument that maps between these two spaces. We denote (possibly

infinite-dimensional) Hilbert-space vectors in \mathbb{U} with \mathbf{f} , Hilbert-space vectors in \mathbb{V} with \mathbf{g} , and the mapping between the spaces with

$$\mathbf{g} = \mathcal{H}\mathbf{f}. \quad (1)$$

Once we have identified the spaces \mathbb{U} and \mathbb{V} , we can start expressing the mapping between the spaces in a specific object-space and data-space basis. In most cases the easiest mapping to find uses a delta-function basis—we expand object and data space into delta functions, then express the mapping as an integral transform. After finding this mapping we can start to investigate the same mapping in different bases.

The above discussion is quite abstract, but it is a powerful point of view that will enable us to unify the analysis of spatio-angular fluorescence microscopes. In Section 2.1 we will demonstrate the formalism by examining a familiar monopole imaging model, and we will demonstrate the mapping between object and data space in two different bases. In Section 2.2 we will extend the monopole imaging model to dipoles and examine the mapping in four different bases.

2.1. Monopole imaging in different bases

We start by considering a microscope that images a field of in-focus monopoles by recording the irradiance on a two-dimensional detector. This section treads familiar ground, but it serves to establish the concepts and notation that will be necessary when we extend to the dipole case.

We can represent the object as a function that assigns a real number to each point on a plane, so we identify object space as $\mathbb{U} = \mathbb{L}_2(\mathbb{R}^2)$ —the set of square-integrable functions on the two-dimensional plane. Similarly, we have a two-dimensional detector that measures a real number at each point on a plane, so data space is the same set $\mathbb{V} = \mathbb{L}_2(\mathbb{R}^2)$.

Next, we name the representations of our object and data in a specific basis. In a delta function basis the object can be represented by a function $f(\mathbf{r}_o)$ called the *monopole density*—the number of monopoles per unit area at the two-dimensional position \mathbf{r}_o . Similarly, in a delta function basis the data can be represented by a function $g'(\mathbf{r}'_d)$ called the *irradiance*—the power received by a surface per unit area at position \mathbf{r}'_d . It may seem pedantic to emphasize that these representations are in a delta function basis, but we highlight this point to make it clear that the delta function basis is not special—we can choose a different basis at will.

A reasonable starting point is to assume that the relationship between the object and the data is *linear*—this is usually true in fluorescence microscopes since fluorophores emit incoherently, so a scaled sum of fluorophores will result in a scaled sum of the irradiance patterns created by the individual fluorophores. If the mapping is linear, we can write the irradiance as a weighted integral over a field of monopoles

$$g'(\mathbf{r}'_d) = \int_{\mathbb{R}^2} d\mathbf{r}_o h'(\mathbf{r}'_d, \mathbf{r}_o) f(\mathbf{r}_o), \quad (2)$$

where $h'(\mathbf{r}'_d, \mathbf{r}_o)$ is the irradiance at position \mathbf{r}'_d created by a point source at \mathbf{r}_o .

Next, we assume that the optical system is *aplanatic*—Abbe’s sine condition is satisfied and on-axis points are imaged without aberration. Abbe’s sine condition guarantees that off-axis points are imaged without spherical aberration or coma [10], so the imaging system can be modeled within the field-of-view of the optical system as a magnifier with shift-invariant blur

$$g'(\mathbf{r}'_d) = \int_{\mathbb{R}^2} d\mathbf{r}_o h'(\mathbf{r}'_d - m\mathbf{r}_o) f(\mathbf{r}_o), \quad (3)$$

where m is a magnification factor.

We can simplify our analysis by changing coordinates and writing Eq. 3 as a convolution [9]. We define a demagnified detector coordinate $\mathbf{r}_d = \mathbf{r}'_d/m$ and a normalization factor that corresponds

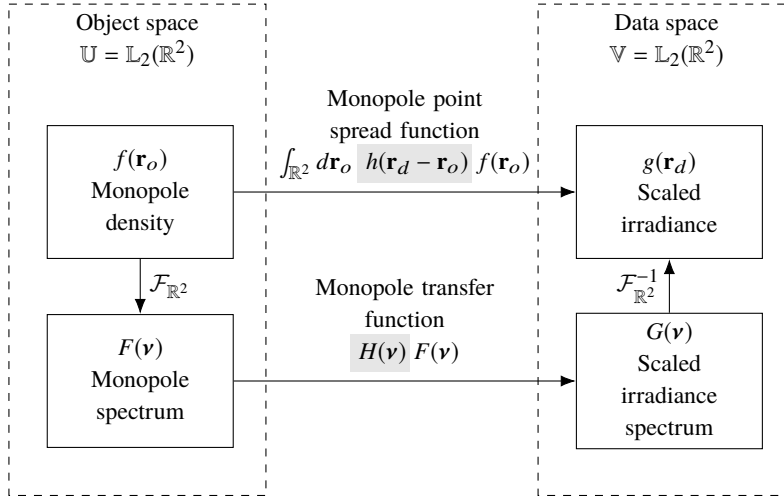


Fig. 1. The mapping between the object and data space of a monopole fluorescence microscope can be computed in two different bases—a delta function basis and a spatial harmonic basis. The change of basis can be computed with a two-dimensional Fourier transform denoted $\mathcal{F}_{\mathbb{R}^2}$.

to the total power incident on the detector plane due to a point source $P_{\text{mono}} = \int_{\mathbb{R}^2} d\mathbf{r} h'(\mathbf{m}\mathbf{r})$ where $\mathbf{r} = \mathbf{r}_d - \mathbf{r}_o$. We use these scaling factors to define the *monopole point spread function* as

$$h(\mathbf{r}_d - \mathbf{r}_o) = \frac{h'(m[\mathbf{r}_d - \mathbf{r}_o])}{P_{\text{mono}}}, \quad (4)$$

and the *scaled irradiance* as

$$g(\mathbf{r}_d) = \frac{g'(m\mathbf{r}_d)}{P_{\text{mono}}}. \quad (5)$$

With these definitions we can express the mapping between the object and the data as a familiar convolution

$$g(\mathbf{r}_d) = \int_{\mathbb{R}^2} d\mathbf{r}_o h(\mathbf{r}_d - \mathbf{r}_o) f(\mathbf{r}_o). \quad (6)$$

We will mostly be using scaled coordinates, so we have adopted a slightly unusual convention of using primes to denote unscaled coordinates.

We have chosen to normalize the monopole point spread function so that

$$\int_{\mathbb{R}^2} d\mathbf{r} h(\mathbf{r}) = 1. \quad (7)$$

The monopole point spread function corresponds to a measurable irradiance, so it is always real and positive.

The mapping between the object and the data in a linear shift-invariant imaging system takes a particularly simple form in a spatial harmonic basis. If we apply the Fourier convolution theorem to Eq. 6 we find that

$$G(\mathbf{v}) = H(\mathbf{v})F(\mathbf{v}), \quad (8)$$

where we define the *scaled irradiance spectrum* as

$$G(\boldsymbol{\nu}) = \int_{\mathbb{R}^2} d\mathbf{r} g(\mathbf{r}) \exp(-2\pi i \mathbf{r} \cdot \boldsymbol{\nu}), \quad (9)$$

the *monopole transfer function* as

$$H(\boldsymbol{\nu}) = \int_{\mathbb{R}^2} d\mathbf{r} h(\mathbf{r}) \exp(-2\pi i \mathbf{r} \cdot \boldsymbol{\nu}), \quad (10)$$

and the *monopole spectrum* as

$$F(\boldsymbol{\nu}) = \int_{\mathbb{R}^2} d\mathbf{r} f(\mathbf{r}) \exp(-2\pi i \mathbf{r} \cdot \boldsymbol{\nu}). \quad (11)$$

The monopole point spread function is normalized and real, so we know that the monopole transfer function is normalized $H(0) = 1$ and conjugate symmetric $H(-\boldsymbol{\nu}) = H^*(\boldsymbol{\nu})$ where z^* denotes the complex conjugate of z .

Notice that Eqs. 6 and 8 are expressions of the same mapping between object and data space in different bases. Figure 1 summarizes the relationship between object and data space in both bases.

We have been careful to use the term *monopole transfer function* instead of the commonly-used term *optical transfer function*. We reserve the term *optical transfer function* for optical systems—the optical transfer function maps between an input irradiance spectrum and an output irradiance spectrum in an optical system. We can use optical transfer functions to model the propagation of light through a microscope, but ultimately we are always interested in the object, not the light emitted by the object. We will find the distinction between the optical transfer function and the object transfer function to be especially valuable when we consider dipoles later in this section.

2.1.1. Monopole coherent transfer functions

Although the Fourier transform can be used to calculate the monopole transfer function directly from the monopole point spread function, there is a well-known alternative that exploits coherent transfer functions. The key idea is that the monopole point spread function can always be written as the absolute square of a scalar-valued *monopole coherent spread function* defined as

$$|c(\mathbf{r}_d - \mathbf{r}_o)|^2 = h(\mathbf{r}_d - \mathbf{r}_o). \quad (12)$$

Physically, the monopole coherent spread function corresponds to the scalar-valued field on the detector with appropriate scaling.

Since the monopole transfer function is given by the two-dimensional Fourier transform of the monopole point spread function

$$H(\boldsymbol{\nu}) = \int_{\mathbb{R}^2} d\mathbf{r} h(\mathbf{r}) \exp[-2\pi i \mathbf{r} \cdot \boldsymbol{\nu}], \quad (13)$$

we can plug in Eq. 12 and use the autocorrelation theorem to rewrite the monopole transfer function as

$$H(\boldsymbol{\nu}) = \int_{\mathbb{R}^2} d\boldsymbol{\tau} C(\boldsymbol{\tau}) C^*(\boldsymbol{\tau} - \boldsymbol{\nu}), \quad (14)$$

where we have introduced the *monopole coherent transfer function* as the two-dimensional Fourier transform of the monopole coherent spread function

$$C(\boldsymbol{\tau}) = \int_{\mathbb{R}^2} d\mathbf{r} c(\mathbf{r}) \exp[-2\pi i \mathbf{r} \cdot \boldsymbol{\tau}]. \quad (15)$$

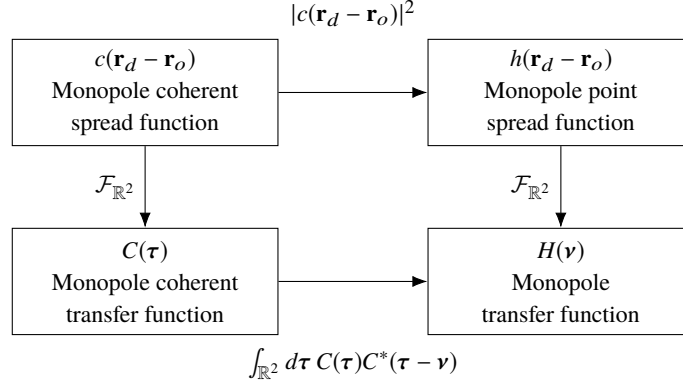


Fig. 2. The incoherent monopole transfer functions are related by a two-dimensional Fourier transform (right column). The coherent monopole transfer functions (left column) can be used to simplify the calculation of the remaining transfer functions.

Physically, the monopole coherent transfer function corresponds to the scalar-valued field in a Fourier plane of the detector with appropriate scaling.

The coherent transfer function provides a valuable shortcut for analyzing microscopes since it is often straightforward to calculate the field in a Fourier plane of the detector. A typical approach for calculating the transfer functions is to (1) calculate the field in a Fourier plane of the detector, (2) scale the field to find the monopole coherent transfer function, then (3) use the relationships in Figure 2 to calculate the other transfer functions.

2.2. Dipole imaging in different bases

Now we consider a microscope imaging a field of in-focus dipoles by recording the irradiance on a two-dimensional detector. A function that assigns a real number to each point on a plane is not sufficient to specify a field of dipoles because the dipoles can have different orientations. To represent the object we need to extend object space to $\mathbb{U} = \mathbb{L}_2(\mathbb{R}^2 \times \mathbb{S}^2)$ where \mathbb{S}^2 is the two-dimensional sphere (the usual sphere embedded in \mathbb{R}^3).

In a delta function basis the object can be represented by a function $f(\mathbf{r}_o, \hat{\mathbf{s}}_o)$ called the *dipole density*—the number of dipoles at position \mathbf{r}_o per unit area oriented along $\hat{\mathbf{s}}_o$ per unit solid angle. Similar to the monopole case, we model the mapping between the object and the irradiance in a delta function basis as an integral transform

$$g'(\mathbf{r}'_d) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o \int_{\mathbb{R}^2} d\mathbf{r}_o h'(\mathbf{r}'_d, \mathbf{r}_o, \hat{\mathbf{s}}_o) f(\mathbf{r}_o, \hat{\mathbf{s}}_o). \quad (16)$$

where $h'(\mathbf{r}'_d, \mathbf{r}_o, \hat{\mathbf{s}}_o)$ is the irradiance at position \mathbf{r}'_d created by a point source at \mathbf{r}_o with orientation $\hat{\mathbf{s}}_o$. If the optical system is aplanatic we can write the integral transform as

$$g'(\mathbf{r}'_d) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o \int_{\mathbb{R}^2} d\mathbf{r}_o h'(\mathbf{r}'_d - m\mathbf{r}_o, \hat{\mathbf{s}}_o) f(\mathbf{r}_o, \hat{\mathbf{s}}_o). \quad (17)$$

We define the same demagnified detector coordinate $\mathbf{r}_d = \mathbf{r}'_d/m$ and a new normalization factor that corresponds to the total power incident on the detector due to a spatial point source with an angularly uniform distribution of dipoles $P_{\text{dip}} = \int_{\mathbb{S}^2} d\hat{\mathbf{s}} \int_{\mathbb{R}^2} d\mathbf{r} h'(m\mathbf{r}, \hat{\mathbf{s}}_o)$. We use these scaling factors to define the *dipole point spread function* as

$$h(\mathbf{r}_d - \mathbf{r}_o, \hat{\mathbf{s}}_o) = \frac{h'(m[\mathbf{r}_d - \mathbf{r}_o], \hat{\mathbf{s}}_o)}{P_{\text{dip}}}, \quad (18)$$

and the *scaled irradiance* as

$$g(\mathbf{r}_d) = \frac{g'(\mathbf{r}_d)}{P_{\text{dip}}}. \quad (19)$$

With these definitions we can express the mapping between the object and the data as

$$g(\mathbf{r}_d) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o \int_{\mathbb{R}^2} d\mathbf{r}_o h(\mathbf{r}_d - \mathbf{r}_o, \hat{\mathbf{s}}_o) f(\mathbf{r}_o, \hat{\mathbf{s}}_o). \quad (20)$$

Eq. 20 is a key result because it represents the mapping between object space and data space in a delta function basis. The integrals in 20 would be extremely expensive to compute for an arbitrary object, but the integrals simplify to an efficient sum if the object consists of spatially fixed dipoles in a fixed orientation. In other words, Eq. 20 is ideal for simulating and analyzing objects that are spatially and angularly sparse; e.g. single fluorophores that are rigidly attached to a structure .

Similar to the monopole case, we have chosen to normalize the dipole point spread function so that

$$\int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o \int_{\mathbb{R}^2} d\mathbf{r} h(\mathbf{r}, \hat{\mathbf{s}}_o) = 1. \quad (21)$$

The dipole point spread function is a measurable quantity, so it is real and positive.

2.2.1. Spatial dipole transfer function

We can make our first change of basis by applying the Fourier-convolution theorem to Eq. 20 which yields

$$G(\mathbf{v}) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o H(\mathbf{v}, \hat{\mathbf{s}}_o) F(\mathbf{v}, \hat{\mathbf{s}}_o), \quad (22)$$

where we define the *spatial dipole transfer function* as

$$H(\mathbf{v}, \hat{\mathbf{s}}_o) = \int_{\mathbb{R}^2} d\mathbf{r} h(\mathbf{r}, \hat{\mathbf{s}}_o) \exp(-2\pi i \mathbf{r} \cdot \mathbf{v}), \quad (23)$$

and the *spatial dipole spectrum* as

$$F(\mathbf{v}, \hat{\mathbf{s}}_o) = \int_{\mathbb{R}^2} d\mathbf{r} f(\mathbf{r}, \hat{\mathbf{s}}_o) \exp(-2\pi i \mathbf{r} \cdot \mathbf{v}). \quad (24)$$

Since the dipole point spread function is normalized and real, we know that the spatial dipole transfer function is normalized $\int_{\mathbb{S}^2} d\hat{\mathbf{s}} H(0, \hat{\mathbf{s}}_o) = 1$ and conjugate symmetric $H(-\mathbf{v}, \hat{\mathbf{s}}_o) = H^*(\mathbf{v}, \hat{\mathbf{s}}_o)$.

This basis is ideal for simulating and analyzing objects that are angularly sparse and spatially dense; e.g. rod structures that contain fluorophores in a fixed orientation.

2.2.2. Dipole angular transfer function

The spherical harmonics are another set of convenient basis functions that are play the same role as complex exponentials in spatial transfer functions—see Appendix A for an introduction to the spherical harmonics. We can change basis from spherical delta functions to spherical harmonics by applying the generalized Plancharel theorem for spherical functions

$$\int_{\mathbb{S}^2} d\hat{\mathbf{s}} p(\hat{\mathbf{s}}) q^*(\hat{\mathbf{s}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} P_{\ell}^m Q_{\ell}^{m*}, \quad (25)$$

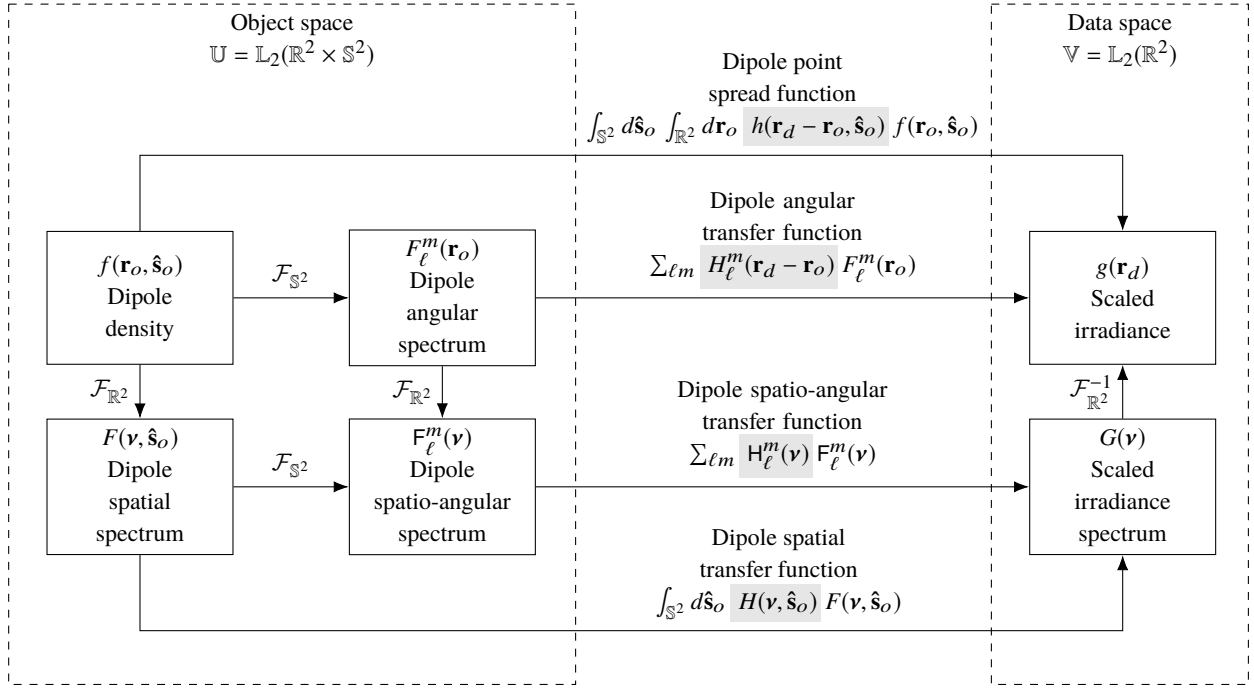


Fig. 3. The mapping between the object and data space of a dipole imaging system can be computed in four different bases—a delta function basis, a spatial-harmonic/angular-delta basis, a spatial-delta/spherical-harmonic basis, and a spatio-angular harmonic basis. The changes of basis can be computed with the two-dimensional Fourier transform denoted $\mathcal{F}_{\mathbb{R}^2}$, and the spherical Fourier transform denoted $\mathcal{F}_{\mathbb{S}^2}$.

where $p(\hat{\mathbf{s}})$ and $q(\hat{\mathbf{s}})$ are arbitrary functions on the sphere, P_ℓ^m and Q_ℓ^m are their spherical Fourier transforms defined by

$$P_\ell^m = \int_{\mathbb{S}^2} d\hat{\mathbf{s}} p(\hat{\mathbf{s}}) Y_\ell^{m*}(\hat{\mathbf{s}}), \quad (26)$$

and $Y_\ell^m(\hat{\mathbf{s}})$ are the spherical harmonic functions defined in Appendix A. Eq. 25 expresses the fact that scalar products are invariant under a change of basis (see Eq. 3.78 of [9]). The left-hand side of Eq. 25 is the scalar product of $\mathbb{L}_2(\mathbb{S}^2)$ functions in a delta function basis and the right-hand side is the scalar product of $\mathbb{L}_2(\mathbb{S}^2)$ functions in a spherical harmonic function basis. Applying Eq. 25 to Eq. 20 yields

$$g(\mathbf{r}_d) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} H_\ell^m(\mathbf{r}_d - \mathbf{r}_o) F_\ell^m(\mathbf{r}_o), \quad (27)$$

where we have defined the *dipole angular transfer function* as

$$H_\ell^m(\mathbf{r}_d - \mathbf{r}_o) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o h(\mathbf{r}_d - \mathbf{r}_o, \hat{\mathbf{s}}_o) Y_\ell^{m*}(\hat{\mathbf{s}}_o), \quad (28)$$

and the *dipole angular spectrum* as

$$F_\ell^m(\mathbf{r}_o) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o f(\mathbf{r}_o, \hat{\mathbf{s}}_o) Y_\ell^{m*}(\hat{\mathbf{s}}_o). \quad (29)$$

Since the dipole point spread function is normalized and real, we know that the dipole angular transfer function is normalized $\int_{\mathbb{R}^2} d\mathbf{r} H_0^0(\mathbf{r}) = 1$ and conjugate symmetric $H_\ell^{-m}(\mathbf{r}) = (-1)^m H_\ell^{m*}(\mathbf{r})$.

This basis is well suited for simulating and analyzing objects that are spatially sparse and angularly dense; e.g. rotating single fluorophores that are spatially fixed, or many fluorophores that are close together with varying orientations.

2.2.3. Spatio-angular dipole transfer function

We can arrive at our final basis in two ways: by applying the generalized Plancharel theorem for spherical functions to Eq. 22 or by applying the Fourier convolution theorem to Eq. 27. We follow the second path and find that

$$G(\boldsymbol{\nu}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} H_\ell^m(\boldsymbol{\nu}) F_\ell^m(\boldsymbol{\nu}), \quad (30)$$

where we have defined the *dipole spatio-angular transfer function* as

$$H_\ell^m(\boldsymbol{\nu}) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o H(\boldsymbol{\nu}, \hat{\mathbf{s}}_o) Y_\ell^{m*}(\hat{\mathbf{s}}_o), \quad (31)$$

and the *dipole spatio-angular spectrum* as

$$F_\ell^m(\boldsymbol{\nu}) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o F(\boldsymbol{\nu}, \hat{\mathbf{s}}_o) Y_\ell^m(\hat{\mathbf{s}}_o). \quad (32)$$

Since the dipole point spread function is normalized and real, we know that the dipole spatio-angular transfer function is normalized $H_0^0(0) = 1$ and conjugate symmetric $H_\ell^{-m}(-\boldsymbol{\nu}) = (-1)^m H_\ell^{m*}(\boldsymbol{\nu})$.

This basis is well suited for simulating and analyzing arbitrary samples because it exploits the band limit of the imaging system. We will see this advantage explicitly in Section 3 when we calculate a specific dipole spatio-angular transfer function and see that it has spatial and angular band limits.

Figure 3 summarizes the relationship between the four bases that we can use to compute the image of a field of dipoles. We reiterate that all four bases may be useful depending on the sample.

2.2.4. Dipole coherent transfer functions

Similar to the monopole case, there is efficient way to calculate the transfer functions using coherent transfer functions. The dipole point spread function can always be written as the absolute square of a vector-valued function called the *dipole coherent spread function*

$$|\mathbf{c}(\mathbf{r}_d - \mathbf{r}_o, \hat{\mathbf{s}}_o)|^2 = h(\mathbf{r}_d - \mathbf{r}_o, \hat{\mathbf{s}}_o). \quad (33)$$

Physically, the dipole coherent spread function corresponds to the vector-valued electric field on the detector with appropriate scaling. We need a vector-valued coherent transfer function since the polarization of the field plays a significant role in dipole imaging, so the point spread function cannot be written as an absolute square of a scalar-valued function.

Since the dipole spatial transfer function is the two-dimensional Fourier transform of the dipole point spread function

$$H(\boldsymbol{\nu}) = \int_{\mathbb{R}^2} d\mathbf{r} h(\mathbf{r}, \hat{\mathbf{s}}_o) \exp[-2\pi i \mathbf{r} \cdot \boldsymbol{\nu}], \quad (34)$$

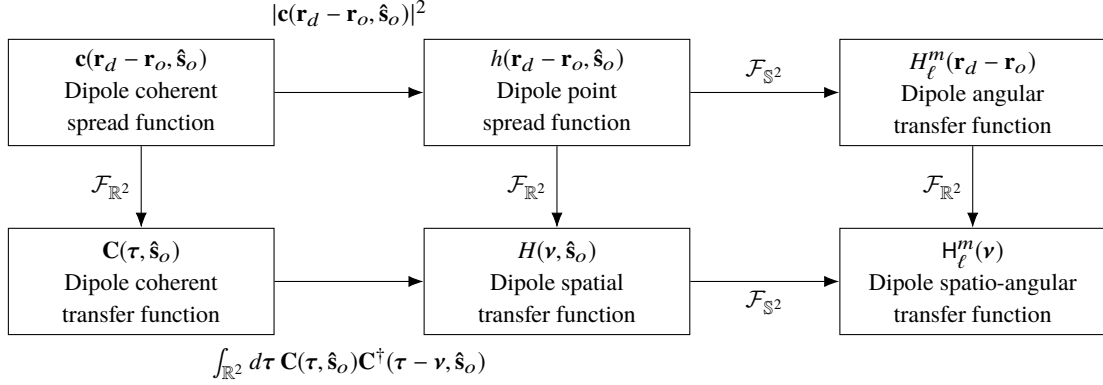


Fig. 4. There is one incoherent transfer function for each set of object-space basis functions, and these transfer functions are related by two-dimensional and spherical Fourier transforms—see center and right columns. There are an additional pair of coherent transfer functions that are useful for calculating the incoherent transfer functions—see left column.

we can plug in Eq. 33 and use the autocorrelation theorem to rewrite the spatial dipole transfer function as

$$H(\mathbf{v}) = \int_{\mathbb{R}^2} d\boldsymbol{\tau} \mathbf{C}(\boldsymbol{\tau}, \hat{\mathbf{s}}_o) \mathbf{C}^\dagger(\boldsymbol{\tau} - \mathbf{v}, \hat{\mathbf{s}}_o), \quad (35)$$

where we have introduced the *dipole coherent transfer function* $\mathbf{C}(\boldsymbol{\tau}, \hat{\mathbf{s}}_o)$ as the two-dimensional Fourier transform of the dipole coherent spread function

$$\mathbf{C}(\boldsymbol{\tau}, \hat{\mathbf{s}}_o) = \int_{\mathbb{R}^2} d\mathbf{r} \mathbf{c}(\mathbf{r}, \hat{\mathbf{s}}_o) \exp[-2\pi i \mathbf{r} \cdot \boldsymbol{\tau}]. \quad (36)$$

Physically, the dipole coherent transfer function corresponds to the vector-valued electric field created by a dipole oriented along $\hat{\mathbf{s}}_o$ in a Fourier plane of the detector with appropriate scaling. Similar to the monopole case, we can calculate the orientation-dependent fields in a Fourier plane of the detector, scale appropriately to find the dipole coherent transfer function, then use the relationships in Figure 4 to calculate the other transfer functions.

3. Results

In this section we begin to apply the tools we developed in the previous section to analyze a specific fluorescence microscope. During our initial discussion we only required the optical system to be aplanatic so that we could model it as a shift-invariant magnifier. Now we will be more specific and consider an aplanatic optical system in a *4f configuration* with an arbitrary first lens (the objective lens) and a *paraxial second lens* (the tube lens) as shown in Figure 5. A lens can be considered paraxial if the angle α between the optical axis of the lens and the marginal ray is small enough that $\sin \alpha \approx \alpha$. As a rule of thumb, non-paraxial effects only become significant when the numerical aperture of a lens exceeds 0.7 [11], but this is only a rough guideline.

We start by defining *pupil functions* for both monopole and dipole imaging systems, and we explicitly relate the pupil functions to the previously defined coherent transfer functions. After establishing the link between physical calculations and the transfer functions, we calculate the monopole and dipole transfer functions for imaging systems with a paraxial objective, and we demonstrate our approach by simulating simple phantoms.

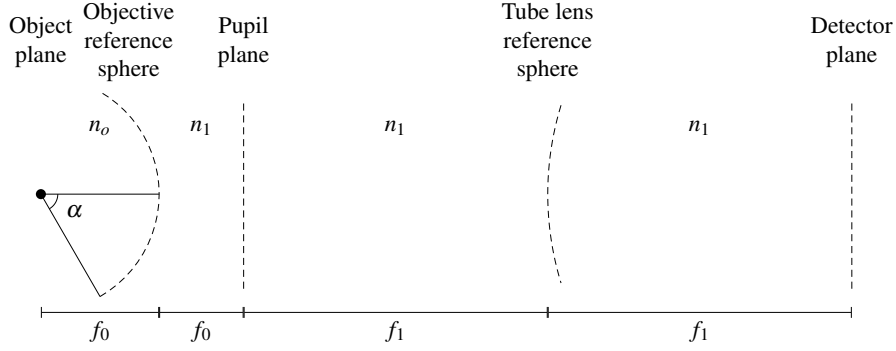


Fig. 5. Schematic of an aplanatic imaging system in a $4f$ geometry with a paraxial tube lens. Since we have an aplanatic optical system we only need to consider the image created by on-axis objects.

3.1. Monopole pupil functions

We define the *monopole pupil function* $p(\mathbf{r}_p)$ of the imaging system as the field immediately following the pupil plane created by an on-axis monopole where \mathbf{r}_p is an unscaled two-dimensional coordinate in the pupil plane. In this section we will relate the monopole pupil function to the transfer function we defined in Section 2 by following a near variant of the derivation in Barrett and Myers Section 9.7 [9].

Since monopoles emit scalar fields, the monopole pupil function is a scalar-valued function. The optical system is aplanatic, so we can write the field created at a point in the aperture \mathbf{r}_p by a monopole at position \mathbf{r}_o as

$$U_p(\mathbf{r}_p, \mathbf{r}_o) \propto p(\mathbf{r}_p) \exp \left[-2\pi i \frac{n_o}{\lambda f_0} \mathbf{r}_p \cdot \mathbf{r}_o \right]. \quad (37)$$

Eq. 37 is a restatement of the aplanatic condition for a $4f$ optical system—the field in the aperture plane can be written as the pupil function multiplied by a linear phase factor that encodes the position of the object.

Since the second lens is paraxial, we can model the relationship between the field in the aperture plane and the field on the detector with a scaled Fourier transform [12]

$$U_d(\mathbf{r}'_d, \mathbf{r}_o) \propto \int_{\mathbb{R}^2} d\mathbf{r}_p p(\mathbf{r}_p) \exp \left[-2\pi i \frac{n_o}{\lambda f_0} \mathbf{r}_p \cdot \mathbf{r}_o \right] \exp \left[-2\pi i \frac{n_1}{\lambda f_1} \mathbf{r}_p \cdot \mathbf{r}'_d \right]. \quad (38)$$

If we define $P(\boldsymbol{\tau})$ as the two-dimensional Fourier transform of the pupil function then we can rewrite Eq. 38 as

$$U_d(\mathbf{r}'_d, \mathbf{r}_o) \propto P \left(\frac{n_o}{\lambda f_0} \mathbf{r}_o + \frac{n_1}{\lambda f_1} \mathbf{r}'_d \right), \quad (39)$$

which we can simplify further by writing in terms of $m = -\frac{f_1 n_o}{f_0 n_1}$:

$$U_d(\mathbf{r}'_d - m\mathbf{r}_o) \propto P \left(\frac{n_1}{\lambda f_1} [\mathbf{r}'_d - m\mathbf{r}_o] \right). \quad (40)$$

The irradiance on the detector is the absolute square of the field so

$$h'(\mathbf{r}'_d - m\mathbf{r}_o) \propto \left| P \left(\frac{n_1}{\lambda f_1} [\mathbf{r}'_d - m\mathbf{r}_o] \right) \right|^2. \quad (41)$$

If we demagnify the coordinates with $\mathbf{r}_d = \mathbf{r}'_d/m$ and demagnify the irradiance with $h(\mathbf{r}_d - \mathbf{r}_o) \propto h'(m[\mathbf{r}_d - \mathbf{r}_o])$ we find that the monopole point spread function is related to the Fourier transform of the monopole pupil function by

$$h(\mathbf{r}_d - \mathbf{r}_o) \propto \left| P \left(-\frac{n_o}{\lambda f_o} [\mathbf{r}_d - \mathbf{r}_o] \right) \right|^2. \quad (42)$$

The monopole coherent spread function is the absolute square of the monopole point spread function so

$$c(\mathbf{r}_d - \mathbf{r}_o) \propto P \left(-\frac{n_o}{\lambda f_o} [\mathbf{r}_d - \mathbf{r}_o] \right). \quad (43)$$

Finally, the monopole coherent transfer function is the Fourier transform of the monopole coherent spread function so

$$C(\boldsymbol{\tau}) \propto p \left(\frac{\lambda f_o}{n_o} \boldsymbol{\tau} \right). \quad (44)$$

Equation 44 is the key result of this section—the monopole coherent transfer function is a scaled monopole pupil function.

3.2. Dipole pupil function

We define the *dipole pupil function* $\mathbf{p}(\mathbf{r}_p, \hat{\mathbf{s}}_o)$ of the imaging system as the electric field immediately following the pupil plane created by an on-axis dipole oriented along $\hat{\mathbf{s}}_o$. Since dipoles emit vector-valued electric fields, the dipole pupil function is a vector-valued function. Almost all of the arguments in the previous section carry over to the dipole case. Briefly, we can write the electric field created at a point in the pupil \mathbf{r}_p by a dipole at \mathbf{r}_o oriented along $\hat{\mathbf{s}}_o$ as

$$\mathbf{E}(\mathbf{r}_p, \mathbf{r}_o, \hat{\mathbf{s}}_o) = \mathbf{p}(\mathbf{r}_p, \hat{\mathbf{s}}_o) \exp \left[-2\pi i \frac{n_0}{\lambda f_0} \mathbf{r}_p \cdot \mathbf{r}_o \right]. \quad (45)$$

The second lens is paraxial, so we can find the field on the detector with a Fourier transform

$$\mathbf{E}_d(\mathbf{r}'_d, \mathbf{r}_o, \hat{\mathbf{s}}_o) \propto \int_{\mathbb{R}^2} d\mathbf{r}_p \mathbf{p}(\mathbf{r}_p, \hat{\mathbf{s}}_o) \exp \left[-2\pi i \frac{n_0}{\lambda f_0} \mathbf{r}_p \cdot \mathbf{r}_o \right] \exp \left[-2\pi i \frac{n_1}{\lambda f_1} \mathbf{r}_p \cdot \mathbf{r}'_d \right]. \quad (46)$$

Note that the Fourier transform of a vector field is the Fourier transform of its scalar-valued orthogonal components, so Eq. 46 specifies two two-dimensional Fourier transforms. We follow the same manipulations as the previous section and find that the dipole coherent transfer function is a scaled dipole pupil function

$$\mathbf{C}(\boldsymbol{\tau}, \hat{\mathbf{s}}_o) \propto \mathbf{p} \left(\frac{\lambda f_o}{n_o} \boldsymbol{\tau}, \hat{\mathbf{s}}_o \right). \quad (47)$$

3.3. Special functions

We adopt and generalize Bracewell's notation [13] for several special functions which will simplify our calculations. First, we define a *rectangle function* as

$$\Pi(x) = \begin{cases} 1 & \text{if } |x| < \frac{1}{2}, \\ 0 & \text{else.} \end{cases} \quad (48)$$

We also define the n^{th} -order jinc function as

$$\text{jinc}_n(r) = \frac{J_{n+1}(\pi r)}{2r}, \quad (49)$$

where $J_{n+1}(r)$ is the $(n+1)^{th}$ -order Bessel function of the first kind.

In Appendix B we derive the following two-dimensional Fourier transform relationships between the jinc functions and the weighted rectangle functions

$$i^n \begin{Bmatrix} \exp(in\phi_r) \\ \cos(n\phi_r) \\ \sin(n\phi_r) \end{Bmatrix} \text{jinc}_n(r) \xrightarrow{\mathcal{F}_{\mathbb{R}^2}} (2\nu)^n \begin{Bmatrix} \exp(in\phi_\nu) \\ \cos(n\phi_\nu) \\ \sin(n\phi_\nu) \end{Bmatrix} \Pi(\nu), \quad (50)$$

where the entries inside the curly braces are to be taken one at a time.

Finally, we define the n^{th} -order chat function as the two-dimensional Fourier transform of the squared n^{th} -order jinc function

$$\text{jinc}_n^2(r) \xrightarrow{\mathcal{F}_{\mathbb{R}^2}} \text{chat}_n(\nu). \quad (51)$$

In Appendix B we show that the zero and first order chat functions can be written in closed form as

$$\text{chat}_0(x) = \frac{1}{2} \left[\cos^{-1} |x| - |x| \sqrt{1-x^2} \right] \Pi\left(\frac{x}{2}\right), \quad (52)$$

$$\text{chat}_1(x) = \frac{1}{2} \left[\cos^{-1} |x| - |x|(3-2x^2)\sqrt{1-x^2} \right] \Pi\left(\frac{x}{2}\right). \quad (53)$$

3.4. Monopole transfer functions

Our first step towards the monopole transfer functions is to calculate the monopole pupil function and coherent transfer function. Several works [14, 15] have modeled an aplanatic fluorescence microscope imaging monopole emitters with the scalar pupil function

$$p(\mathbf{r}_p) \propto \tilde{C}\left(\frac{r_p}{f_o}\right) \Pi\left(\frac{r_p}{2f_o \sin \alpha}\right), \quad (54)$$

where

$$\tilde{C}(x) = (1-x^2)^{-1/4} = 1 + \frac{x^2}{4} + \frac{5x^4}{32} + \dots. \quad (55)$$

The $\tilde{C}(x)$ function models the radial dependence of the field and ensures that power is conserved on either side of an aplanatic objective, and the rectangle function models the aperture stop of the objective. Applying Eq. 44 and collecting constants we find that the coherent monopole transfer function is

$$C(\tau) \propto \tilde{C}\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) \Pi\left(\frac{\tau}{\nu_c}\right), \quad (56)$$

where $\text{NA} = n_o \sin \alpha$ and $\nu_c = 2\text{NA}/\lambda$. This coherent transfer function models objectives with an arbitrary numerical aperture, but for our initial analysis we restrict ourselves to the paraxial regime. We retain only zero and first order radial terms to find that

$$C(\tau) \stackrel{(p)}{\propto} \Pi\left(\frac{\tau}{\nu_c}\right), \quad (57)$$

where (p) indicates that we have used the paraxial approximation for the objective lens.

We can find the monopole coherent spread function by taking the inverse Fourier transform of the monopole coherent transfer function

$$c(\mathbf{r}) \stackrel{(p)}{\propto} \text{jinc}_0(v_c r). \quad (58)$$

The monopole point spread function is the (normalized) absolute square of the monopole coherent spread function so,

$$h(\mathbf{r}) \stackrel{(p)}{=} \frac{4}{\pi} \text{jinc}_0^2(v_c r) \quad (59)$$

which is the well-known *Airy disk*.

Finally, we can calculate the monopole transfer function as the two-dimensional Fourier transform of the monopole point spread function (or the autocorrelation of the coherent transfer function) and find that

$$H(\nu) \stackrel{(p)}{=} \frac{4}{\pi} \text{chat}_0\left(\frac{\nu}{\nu_c}\right). \quad (60)$$

3.5. Dipole transfer functions

To calculate the dipole transfer function we proceed similarly to the monopole case—we find the pupil function, scale to find the coherent dipole transfer function, then calculate the remaining transfer functions.

Backer and Moerner [4] have calculated the dipole pupil function for a high-NA objective as

$$\mathbf{p}(\mathbf{r}_p, \hat{\mathbf{s}}_o) \propto \begin{bmatrix} \tilde{C}_0\left(\frac{r_p}{f_o}\right) + \tilde{C}_2\left(\frac{r_p}{f_o}\right) \cos(2\phi_p) & \tilde{C}_2\left(\frac{r_p}{f_o}\right) \sin(2\phi_p) & \tilde{C}_1\left(\frac{r_p}{f_o}\right) \cos \phi_p \\ \tilde{C}_2\left(\frac{r_p}{f_o}\right) \sin(2\phi_p) & \tilde{C}_0\left(\frac{r_p}{f_o}\right) - \tilde{C}_2\left(\frac{r_p}{f_o}\right) \cos(2\phi_p) & \tilde{C}_1\left(\frac{r_p}{f_o}\right) \sin \phi_p \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix} \Pi\left(\frac{r_p}{2f_o \sin \alpha}\right). \quad (61)$$

where

$$\tilde{C}_0(x) = \frac{1}{2}(\sqrt{1-x^2} + 1)(1-x^2)^{-1/4} = 1 + \frac{x^4}{32} + \frac{x^6}{32} + \dots, \quad (62)$$

$$\tilde{C}_1(x) = x(1-x^2)^{-1/4} = x + \frac{x^3}{4} + \frac{5x^5}{32} + \dots, \quad (63)$$

$$\tilde{C}_2(x) = \frac{1}{2}(\sqrt{1-x^2} - 1)(1-x^2)^{-1/4} = -\frac{x^2}{4} - \frac{x^4}{8} - \frac{11x^6}{128} - \dots. \quad (64)$$

Similar to the monopole case, the dipole pupil function conserves power and has a cutoff at the objective aperture, but the dipole pupil function is vector-valued to model the complete electric field in the pupil plane. The fields in the aperture plane have a negligible $\hat{\mathbf{z}}$ component which is a consequence of our assumption that the tube lens is paraxial.

Scaling the dipole pupil function using Eq. 47 yields the dipole coherent transfer function

$$\mathbf{C}(\boldsymbol{\tau}, \hat{\mathbf{s}}_o) \propto \begin{bmatrix} \tilde{C}_0\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) + \tilde{C}_2\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) \cos(2\phi_\tau) & \tilde{C}_2\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) \sin(2\phi_\tau) & \tilde{C}_1\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) \cos \phi_\tau \\ \tilde{C}_2\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) \sin(2\phi_\tau) & \tilde{C}_0\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) - \tilde{C}_2\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) \cos(2\phi_\tau) & \tilde{C}_1\left(\frac{2\text{NA}}{n_o} \frac{\tau}{\nu_c}\right) \sin \phi_\tau \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix} \Pi\left(\frac{\tau}{\nu_c}\right). \quad (65)$$

We restrict our analysis to the paraxial regime by dropping all but the zero and first order radial terms to find that

$$\mathbf{C}(\boldsymbol{\tau}, \hat{\mathbf{s}}_o) \stackrel{(p)}{\propto} \begin{bmatrix} 1 & 0 & \frac{2\text{NA}}{n_o} \frac{\tau}{v_c} \cos \phi_\tau \\ 0 & 1 & \frac{2\text{NA}}{n_o} \frac{\tau}{v_c} \sin \phi_\tau \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix} \Pi\left(\frac{\tau}{v_c}\right). \quad (66)$$

Under the paraxial approximation the transverse component of the dipoles (s_x, s_y) creates purely transverse fields in the pupil plane and the axial component of the dipole (s_z) creates purely radial fields in the pupil plane. The paraxial approximation may seem crude compared to Backer and Moerner's numerical results, but the approximation will allow us to calculate the transfer functions in closed form so that we can gain an intuition for the limits of the microscope. We also note that many existing works in ensemble polarized fluorescence microscopy make stronger approximations than ours. For example, Fourkas only considers the total irradiance in the back focal plane while ignoring the propagation of fields to the detector [16].

The dipole coherent spread function is the inverse Fourier transform of the dipole coherent transfer function. Applying Eq. 50 in reverse yields

$$\mathbf{c}(\mathbf{r}, \hat{\mathbf{s}}_o) \stackrel{(p)}{\propto} \begin{bmatrix} \text{jinc}_0(v_c r) & 0 & \frac{\text{NA}}{n_o} i \cos \phi \text{jinc}_1(v_c r) \\ 0 & \text{jinc}_0(v_c r) & \frac{\text{NA}}{n_o} i \sin \phi \text{jinc}_1(v_c r) \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix}. \quad (67)$$

Notice that the radial component of the dipole coherent spread function has a $\pi/2$ phase shift relative to the transverse component. This phase factor arises because the Fourier transform of a real and odd function is purely imaginary.

3.5.1. Paraxial dipole point spread function

The dipole point spread function is the (normalized) absolute square of coherent dipole spread function

$$h(\mathbf{r}, \hat{\mathbf{s}}_o) \propto \mathbf{c}(\mathbf{r}, \hat{\mathbf{s}}_o) \mathbf{c}^\dagger(\mathbf{r}, \hat{\mathbf{s}}_o). \quad (68)$$

Plugging in the paraxial dipole coherent spread function and normalizing yields

$$h(\mathbf{r}, \hat{\mathbf{s}}_o) \stackrel{(p)}{=} N \left[\text{jinc}_0^2(v_c r) \sin^2 \vartheta + \left(\frac{\text{NA}}{n_o} \right)^2 \text{jinc}_1^2(v_c r) \cos^2 \vartheta \right], \quad (69)$$

where $\sin^2 \vartheta = s_x^2 + s_y^2$, $\cos^2 \vartheta = s_z^2$, and the normalization factor is

$$N = 6v_c^2 \pi^{-3/2} \left[2 + \left(\frac{\text{NA}}{n_o} \right)^2 \right]^{-1}. \quad (70)$$

Notice that the transverse and radial fields are out of phase on the detector, so the total irradiance is the sum of the contributions from the transverse and radial components. In Figure 6 we plot the dipole point spread function for several dipole orientations and numerical apertures, and in Figure 7 we compare the monopole point spread function to the dipole point spread function. The paraxial monopole and dipole models are equivalent when the sample consists of only transverse dipoles which is clear if we notice that Eq. 69 reduces to an Airy disk when $\vartheta = \pi/2$.

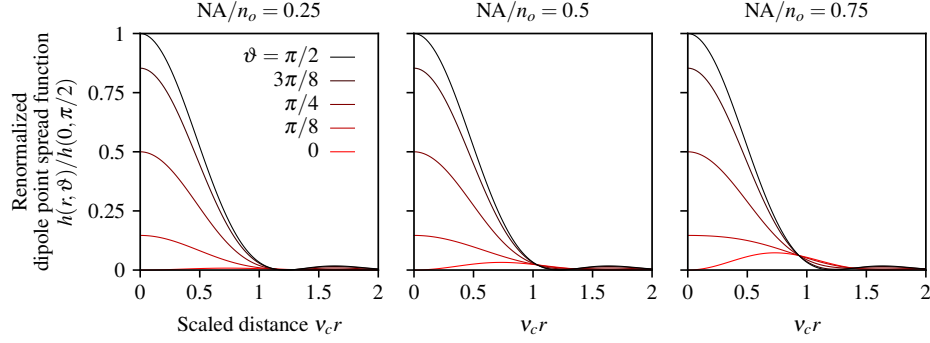


Fig. 6. Renormalized paraxial dipole point spread function as a function of the scaled radial coordinate $v_c r$, the dipole inclination angle ϑ , and NA/n_o . For small numerical apertures (left) the irradiance pattern created by axial dipoles (**red**) is small compared to transverse dipoles (**black**), but the relative contribution of axial dipoles increases with the numerical aperture (see **red** lines from left to right).

To demonstrate the paraxial dipole point spread function we simulate a set of equally spaced dipoles with varying orientation

$$f_{(ph1)}(r_x, r_y, \vartheta, \varphi) = \sum_{j=0}^3 \sum_{k=0}^3 \delta(r_x - j) \delta(r_y - k) \delta(\cos \vartheta - \cos \vartheta_j) \delta(\varphi - \varphi_k), \quad (71)$$

where $\vartheta_j = j \frac{\pi}{6}$, $\varphi_k = k \frac{\pi}{4}$, the subscript $(ph1)$ indicates that this is the first phantom, and the spatial coordinates are expressed in μm . To find the irradiance pattern created by the phantom Eq. 71 we plug it into Eq. 20 and use the sifting property to find that

$$g_{(ph1)}(r_x, r_y) = \sum_{j=0}^3 \sum_{k=0}^3 h\left(\sqrt{(r_x - j)^2 + (r_y - k)^2}, \vartheta_j\right). \quad (72)$$

In Figure 8 we plot the phantom and scaled irradiance for an imaging system with $\text{NA} = 0.75$, $\lambda = 500 \text{ nm}$, and $n_o = 1.33$. We sample and plot the scaled irradiance at $20\times$ the Nyquist rate $\Delta x = 1/[20(2v_c)]$, so the irradiance patterns are free of aliasing. The output demonstrates that the irradiance pattern depends on the dipole inclination, but not its azimuth.

3.6. Paraxial dipole spatial transfer function

The dipole spatial transfer function is the spatial Fourier transform of the dipole point spread function (or the complex autocorrelation of the dipole coherent transfer function). Applying the Fourier transform to Eq. 69 we find that

$$H(\mathbf{v}, \vartheta) \stackrel{(p)}{=} \frac{N}{v_c^2} \left[\text{chat}_0\left(\frac{v}{v_c}\right) \sin^2 \vartheta + \left(\frac{\text{NA}}{n_o}\right)^2 \text{chat}_1\left(\frac{v}{v_c}\right) \cos^2 \vartheta \right]. \quad (73)$$

In Figure 9 we plot the dipole spatial transfer function for several dipole orientations dipole orientations and numerical apertures. We find that the dipole spatial transfer function is negative for axial dipoles at high spatial frequencies, especially for larger numerical apertures. The negative spatial transfer corresponds to a contrast inversion for high-frequency patterns of axial dipoles because the irradiance minimum corresponds to the position of the dipole.

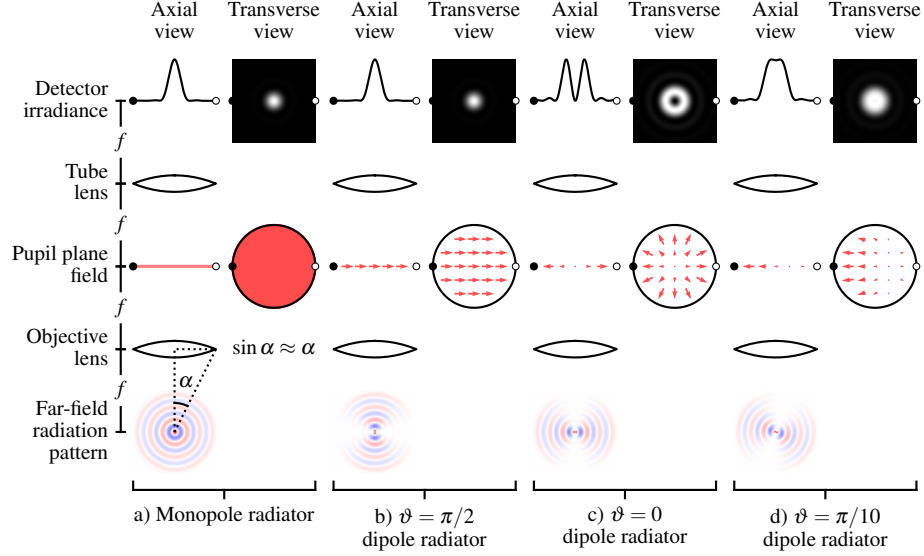


Fig. 7. Comparison of paraxial models for monopole radiators a) and dipole radiators b)–d). a) Monopole radiators fill the pupil plane with a uniform scalar field which gives rise to an Airy disk on the detector. b) A transverse dipole radiator also creates an Airy disk, but the pupil plane is filled with a uniform vector field. c) An axial dipole radiator creates a radial electric field pattern in the back focal plane which creates a $\text{jinc}_1^2(r)$ (or higher-order Airy disk) on the detector. d) Dipoles that are not transverse or axial still create radially symmetric irradiance patterns under the paraxial approximation. Fields from transverse dipoles are real and even while fields from axial dipoles are real and odd which causes a relative $\pi/2$ phase shift for the fields on the detector. This phase shift means that the fields from transverse and axial components of the dipole do not interfere, which causes radially symmetric irradiance patterns.

To demonstrate the spatial dipole transfer function we simulate a set of equally spaced disks with varying diameter containing fluorophores with varying orientation

$$f_{\text{(ph2)}}(r_x, r_y, \vartheta) = \sum_{j=0}^3 \sum_{k=0}^3 \frac{1}{D_k^2} \Pi \left(\frac{1}{D_k} \sqrt{(r_x - j)^2 + (r_y - k)^2} \right) \delta(\cos \vartheta - \cos \vartheta_j) \quad (74)$$

where $D_k = 0.15(1 + k) \mu\text{m}$ and $\vartheta_j = j \frac{\pi}{6}$. Notice that we have scaled the disks so that the total number of fluorophores in each disk is constant.

We can calculate the irradiance pattern by taking the spatial Fourier transform for each orientation in the phantom, filtering the result with the dipole spatial transfer function, integrating over the orientations, then taking the inverse spatial Fourier transform. Since our phantom is angularly sparse, the integral becomes a sum and we find that

$$g_{\text{(ph2)}}(r_x, r_y) = \mathcal{F}_{\mathbb{R}^2}^{-1} \left\{ \sum_j H(\nu, \vartheta_j) \mathcal{F}_{\mathbb{R}^2} \{ f_{\text{(ph2)}}(r_x, r_y, \vartheta_j) \} \right\}. \quad (75)$$

In Figure 10 we plot the phantom and scaled irradiance with the same imaging parameters as the previous section. The small disks create irradiance patterns that are similar to the point sources in the previous section, while larger disks create increasingly uniform irradiance patterns that hide the orientation of the fluorophores.

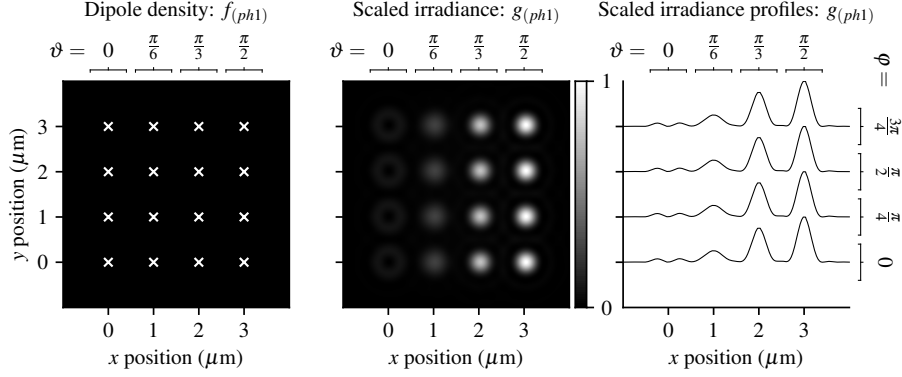


Fig. 8. **Left:** A spatially and angularly sparse phantom—uniformly spaced single dipoles with varying orientations (increasing ϑ from left to right and increasing φ from bottom to top). White crosses mark the positions of the dipoles. **Center:** Scaled irradiance for an imaging system with $\text{NA} = 0.75$, $\lambda = 500$ nm, and $n_o = 1.33$ sampled at $20\times$ the Nyquist rate. **Right:** x profiles through the scaled irradiance. The response is independent of the azimuth angle and strongly dependent on the inclination angle.

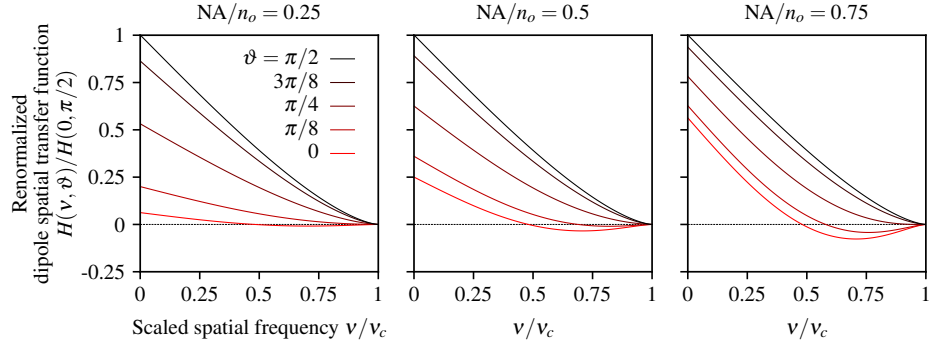


Fig. 9. Dipole spatial transfer function as a function of the scaled spatial frequency v/v_c , the dipole inclination angle ϑ , and NA/n_o . For small numerical apertures (left) the dipole spatial transfer function for axial dipoles (**red**) is small compared to transverse dipoles (**black**), but the relative contribution of axial dipoles increases with the numerical aperture (see **red** lines from left to right). The spatial dipole transfer function of axial dipoles is negative at high spatial frequencies because the central minimum of the axial dipole point spread function corresponds to the position of the dipole. Equivalently, a high-spatial-frequency pattern of axial dipoles will generate an irradiance pattern where the minimum irradiance corresponds to the peak of the axial dipole density.

3.6.1. Paraxial dipole angular transfer function

To calculate the angular dipole transfer function we take the spherical Fourier transform of the dipole point spread function

$$H_l^m(\mathbf{r}) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o h(\mathbf{r}, \hat{\mathbf{s}}_o) Y_l^{m*}(\hat{\mathbf{s}}_o). \quad (76)$$

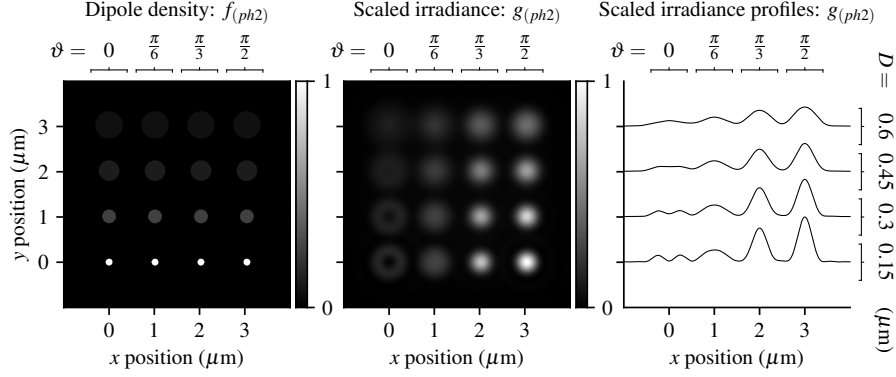


Fig. 10. **Left:** A spatially dense and angularly sparse phantom—uniformly spaced disks with varying size (increasing D from bottom to top) and dipole orientation (increasing ϑ from left to right) **Center:** Scaled irradiance for an imaging system with $\text{NA} = 0.75$, $\lambda = 500$ nm, and $n_o = 1.33$ sampled at $20\times$ the Nyquist rate. **Right:** x profiles through the scaled irradiance. Larger disks generate increasingly uniform irradiance patterns with fewer details that may indicate the orientation of fluorophores.

After evaluating the integrals and normalizing, the angular dipole transfer function is

$$H_l^m(\mathbf{r}) \stackrel{(p)}{=} \frac{N}{3} \left[2\text{jinc}_0^2(\nu_c r) + \left(\frac{\text{NA}}{n_o} \right)^2 \text{jinc}_1^2(\nu_c r) \right] \Lambda_0 \delta_{\ell 0} \delta_{m 0} + \frac{N}{3} \left[-2\text{jinc}_0^2(\nu_c r) + 2 \left(\frac{\text{NA}}{n_o} \right)^2 \text{jinc}_1^2(\nu_c r) \right] \Lambda_2 \delta_{\ell 2} \delta_{m 0}, \quad (77)$$

where $\Lambda_\ell = \sqrt{4\pi/(2\ell + 1)}$.

In Figure 11 we plot the dipole angular transfer function for both spherical harmonic terms and several numerical apertures. Note that the dipole angular transfer function can be negative because the spherical harmonics contain negative values. The $\ell = 0$ term shows that angularly uniform distributions of dipole create spatial irradiance patterns that are nearly identical to the Airy disk, while the $\ell = 2$ term shows a negative pattern because of the large contribution of the transverse negative dipoles in the Y_2^0 spherical harmonic.

To demonstrate the dipole angular transfer function we simulate a set of equally spaced fluorophore distributions with varying orientation and angular distributions

$$f_{(\text{ph3})}(r_x, r_y, \vartheta) = \sum_{j=0}^3 \sum_{k=0}^3 \delta(r_x - j) \delta(r_y - k) f_{(\text{cone})}(\vartheta, \varphi; \vartheta_j, 0, \Delta_k), \quad (78)$$

where

$$f_{(\text{cone})}(\hat{\mathbf{s}}_o; \hat{\mathbf{s}}'_o, \Delta) = f_{(\text{cone})}(\vartheta, \varphi; \vartheta', \varphi', \Delta) = \frac{1}{2\pi(1 - \cos \Delta)} \Pi \left(\frac{\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}'}{\cos \Delta} \right), \quad (79)$$

is an angular double cone distribution with central direction $\hat{\mathbf{s}}'$ and cone half-angle Δ ; $\vartheta_j = j\frac{\pi}{6}$; and $\Delta_k = k\frac{\pi}{6}$. Notice that when $\Delta = 0$ the angular double cone reduces to a single direction, and when $\Delta = \pi/2$ the angular double cone reduces to an angularly uniform distribution.

Our first step towards the irradiance pattern is to calculate the dipole angular spectrum of the phantom. In Appendix C we calculate the spherical Fourier transform of the double cone

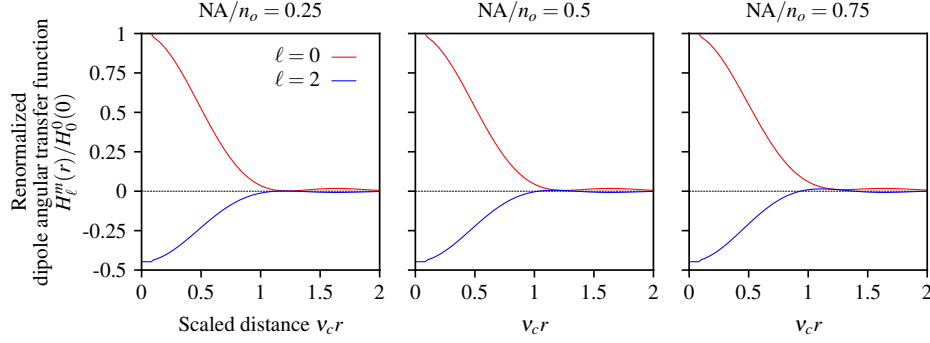


Fig. 11. Paraxial dipole angular transfer function in terms of a scaled radial detection coordinate $v_c r$, the spherical harmonic degree ℓ , and NA/n_o . Angularly uniform distributions of dipoles $\ell = 0$ generate a spatial pattern that is nearly identical to an Airy disk. $\ell = 2$ distributions have a negative response because $Y_2^0(\hat{s})$ is negative for transverse directions. As the numerical aperture increases, the relative contribution of positive axial dipoles in the $\ell = 2$ distribution increases.

distribution $F_{\ell,(\text{cone})}^m(\vartheta', \varphi'; \Delta)$ which we can use to express the dipole angular spectrum as

$$F_{\ell,(\text{ph3})}^m(r_x, r_y, \vartheta) = \sum_{j=0}^3 \sum_{k=0}^3 \delta(r_x - j) \delta(r_y - k) F_{\ell,(\text{cone})}^m(\vartheta_j, 0, \Delta_k). \quad (80)$$

To find the irradiance we plug the dipole angular spectrum into Eq. 27 and use the sifting property we find that

$$g_{(\text{ph3})}(r_x, r_y) = \sum_{\ell m} \sum_{j=0}^3 \sum_{k=0}^3 H_{\ell}^m \left(\sqrt{(r_x - j)^2 + (r_y - k)^2} \right) F_{\ell,(\text{cone})}^m(\vartheta_j, 0, \Delta_k). \quad (81)$$

In Figure 12 we plot the phantom and scaled irradiance with the same imaging parameters as the previous sections. For small cone angles the irradiance patterns are similar to the point sources in the previous sections, while larger cone angles create increasingly uniform irradiance patterns that hide the angular information about the distributions.

3.6.2. Paraxial dipole spatio-angular transfer function

We can calculate the dipole spatio-angular transfer function by taking the spatial Fourier transform of the dipole angular transfer function (or the spherical Fourier transform of the dipole spatial transfer function) to find that

$$H_{\ell}^m(\mathbf{r}) \stackrel{(p)}{=} \frac{N}{3v_c^2} \left[2\text{chat}_0 \left(\frac{v}{v_c} \right) + \left(\frac{\text{NA}}{n_o} \right)^2 \text{chat}_1 \left(\frac{v}{v_c} \right) \right] \Lambda_0 \delta_{\ell 0} \delta_{m 0} + \frac{N}{3v_c^2} \left[-2\text{chat}_0 \left(\frac{v}{v_c} \right) + 2 \left(\frac{\text{NA}}{n_o} \right)^2 \text{chat}_1 \left(\frac{v}{v_c} \right) \right] \Lambda_2 \delta_{\ell 2} \delta_{m 0}. \quad (82)$$

In Figure 13 we plot the dipole spatio-angular transfer function for both spherical harmonic terms and several numerical apertures. The $\ell = 0$ term shows that angularly uniform distributions of dipole create spatial irradiance patterns that are nearly identical to the Airy disk, with high frequencies increasingly suppressed as the numerical aperture increases. The $\ell = 2$ term shows

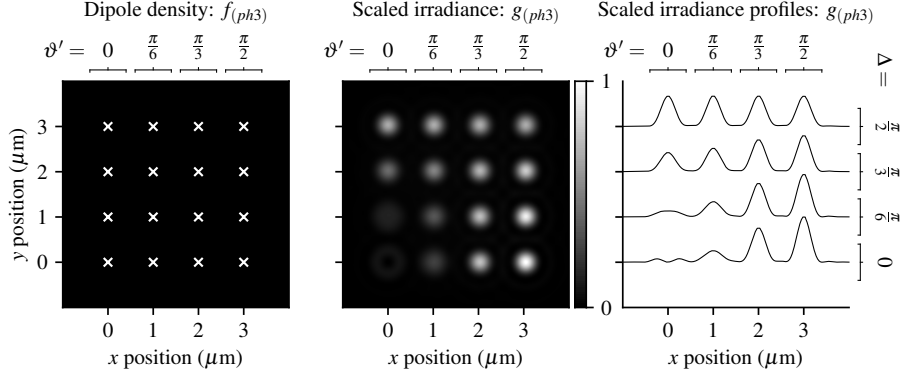


Fig. 12. **Left:** A spatially sparse and angularly dense phantom—uniformly spaced double cone distributions of fluorophores with varying central direction (increasing ϑ' from left to right) and varying cone half-angle (increasing Δ from bottom to top). **Center:** Scaled irradiance for an imaging system with $\text{NA} = 0.75$, $\lambda = 500$ nm, and $n_o = 1.33$ sampled at 20 \times the Nyquist rate. **Right:** x profiles through the scaled irradiance. Small cone angles have irradiance patterns that vary with the central direction, while larger cones angles have increasingly uniform irradiance patterns that hide angular information.

a negative pattern because of the large contribution of the transverse negative dipoles in the Y_2^0 spherical harmonic, and as the numerical aperture increases the relative contribution of the positive axial dipoles increases and the $\ell = 2$ term becomes less negative.

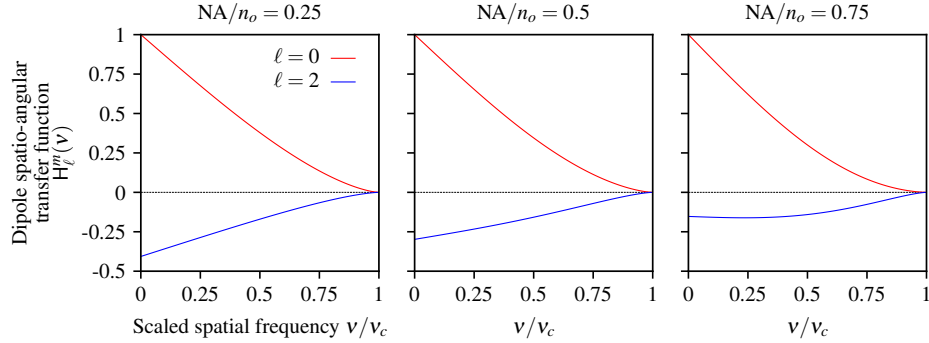


Fig. 13. Spatio-angular dipole transfer function as a function of the scaled spatial frequency v/v_c , the spherical harmonic degree ℓ , and NA/n_o . When the numerical aperture is small the transverse dipoles contribute the most to the signal which gives rise to a positive $\ell = 0$ component and a negative $\ell = 2$ component. As the numerical aperture increases, the relative contribution of axial dipoles increases and the $\ell = 2$ component becomes less negative.

To demonstrate the spatio-angular transfer function, we simulate a set of equally spaced disks of fluorophores with varying radius and angular distributions

$$f_{(\text{ph4})}(r_x, r_y, \vartheta, \varphi) = \sum_{j=0}^3 \sum_{k=0}^3 \frac{1}{D_k^2} \Pi \left(\frac{1}{D_k} \sqrt{(r_x - j)^2 + (r_y - k)^2} \right) f_{(\text{cone})} \left(\vartheta, \varphi; \frac{\pi}{2}, 0, \Delta_j \right), \quad (83)$$

where $D_k = 0.15(1 + k) \mu\text{m}$, and $\Delta_j = j \frac{\pi}{6}$.

Our first step towards calculating the irradiance pattern is to calculate the dipole spatio-angular spectrum given by the spatial Fourier transform of the dipole angular spectrum

$$F_{\ell,(\text{ph4})}^m(\nu_x, \nu_y) = \mathcal{F}_{\mathbb{R}^2} \left\{ \sum_{j=0}^3 \sum_{k=0}^3 \frac{1}{D_k^2} \Pi \left(\frac{1}{D_k} \sqrt{(r_x - j)^2 + (r_y - k)^2} \right) F_{\ell, \text{cone}}^m \left(\frac{\pi}{2}, 0, \Delta_j \right) \right\}. \quad (84)$$

We can plug the dipole spatio-angular spectrum into Eq. 30 to find the scaled irradiance as

$$g_{(\text{ph4})}(r_x, r_y) = \mathcal{F}_{\mathbb{R}^2}^{-1} \left\{ \sum_{\ell m} H_{\ell}^m(\nu_x, \nu_y) F_{\ell,(\text{ph4})}^m(\nu_x, \nu_y) \right\}. \quad (85)$$

In Figure 14 we plot the phantom and scaled irradiance with the same imaging parameters as the previous sections. Small cone angles and small disks create relatively unique irradiance patterns, while increasing the cone angle or disk size creates increasingly irradiance patterns.

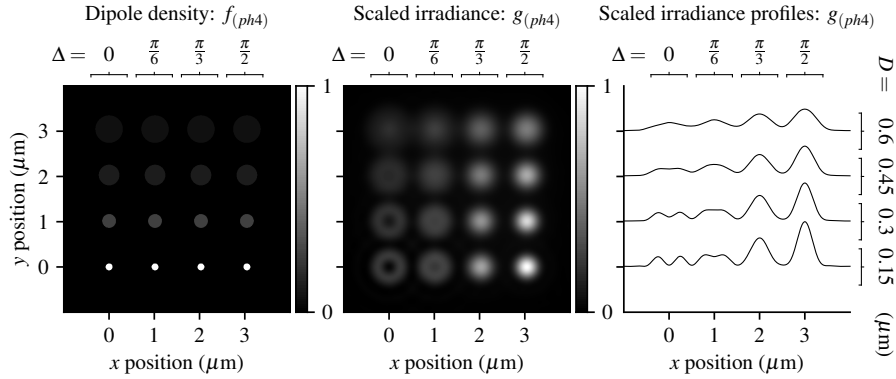


Fig. 14. **Left:** A spatially and angularly dense phantom—uniformly spaced disks with varying size (increasing D from bottom to top) and double cone half angle (increasing Δ from left to right) **Center:** Scaled irradiance for an imaging system with $\text{NA} = 0.75$, $\lambda = 500 \text{ nm}$, and $n_o = 1.33$ sampled at $20\times$ the Nyquist rate. **Right:** x profiles through the scaled irradiance.

4. Discussion

4.1. Reconstructions

- Very ill-posed reconstruction problem by dimensionality.
- Better-posed reconstruction problem by adding polarized illumination/detection formulated as $\mathcal{H} : \mathbb{L}_2(\mathbb{R}^2 \times \mathbb{S}^2) \rightarrow \mathbb{L}_2(\mathbb{R}^3 \times \mathbb{S}^1)$ and/or multiple views $\mathcal{H} : \mathbb{L}_2(\mathbb{R}^2 \times \mathbb{S}^2) \rightarrow \mathbb{L}_2(\mathbb{R}^2 \times \mathbb{S}^1)^V$.
- Mention SVD for future papers

4.2. Alternative transfer functions

Throughout this work we have used the spherical harmonic functions as a basis for functions on the sphere, but there are other basis functions that can be advantageous in some cases. Backer and [TODO Brasselet, others] have used the second moments as basis functions for the sphere because they arise naturally when computing the dipole point spread function. Mathematically,

Backer and Moerner [4] use an alternative to the dipole angular transfer function that uses the second moments as basis functions so their forward model becomes

$$g(\mathbf{r}_d) = \sum_{j=1}^6 H_j(\mathbf{r}_d - \mathbf{r}_o) F_j(\mathbf{r}_o), \quad (86)$$

where

$$H_j(\mathbf{r}_d - \mathbf{r}_o) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o h(\mathbf{r}_d - \mathbf{r}_o, \hat{\mathbf{s}}_o) Z_j(\hat{\mathbf{s}}_o), \quad (87)$$

$$F_j(\mathbf{r}_o) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}}_o f(\mathbf{r}_o, \hat{\mathbf{s}}_o) Z_j(\hat{\mathbf{s}}_o), \quad (88)$$

and $Z_j(\hat{\mathbf{s}}) = \{s_x^2, s_y^2, s_z^2, s_x s_y, s_y s_z, s_x s_z\}$ are the second moments. This formulation is similar to the angular transfer function approach because it exploits the spatial sparsity of sample, but it does not require a cumbersome expansion of the dipole point spread function onto spherical harmonics.

However, the spherical harmonics provide several advantages over the second moments. First, the spherical harmonics form a complete basis for functions on the sphere, while the second moments span a much smaller function space. The usual approach to extending the span of the second moments is to use the fourth (or higher) moments, but these are completely new basis functions. Second, the spherical harmonics are orthonormal which will allow us to easily deploy invaluable tools from linear algebra—linear subspaces, rank, decompositions, etc.—to analyze and compare microscope designs. Finally, using the spherical harmonics provides access to a set of fast algorithms. The naive expansion of an arbitrary discretized N point spherical function onto spherical harmonics (or second moments) requires a $\mathcal{O}(N^2)$ matrix multiplication, while pioneering work by Driscoll and Healy [cite] showed that the forward discrete fast spherical harmonic transform can be computed with a $\mathcal{O}(N(\log N)^2)$ algorithm and its inverse can be computed with a $\mathcal{O}(N^{3/2})$ algorithm. To our knowledge no similarly fast algorithms exist for expansion onto the higher-order moments.

TODO: Cite Tournier/Descoteux for an easy change of basis between spherical harmonics and second moments

4.3. What determines the angular bandwidth?

Spatial imaging systems have a spatial bandwidth which characterizes is the highest spatial frequency that the system can image. Similarly, angular imaging systems have an angular bandwidth that characterizes the highest angular frequency, but in the angular case there are two different types of angular bandwidths that we call *ell* and *m* bandwidths. The ℓ -bandwidth can be interpreted in a similar way to the spatial bandwidth—it characterizes the smallest angular features that the imaging system can measure. The m -bandwidth does not have a direct analog in the spatial domain—it characterizes the angular uniformity of the imaging system. If the ℓ and m bandwidths are equal then the imaging system can be said to be *rotation invariant*.

The spatial bandwidth of a fluorescence microscope is well known to be $\nu_c = \frac{2\text{NA}}{\lambda}$. In other words, we can increase the spatial resolution of a fluorescence microscope by increasing the NA of the instrument or by choosing a fluorophore with a shorter emission wavelength. Similarly, the angular bandwidth of a fluorescence microscope depends on both the instrument and the sample.

The microscope we considered in this work has an *ell* bandwidth of $\ell = 2$ and an *m* bandwidth of $m_c = 0$, so it is not a rotation invariant imaging system. In future work we will consider several approaches to improving the angular bandwidths in detail, but we briefly mention that non-paraxial microscopes, microscopes with polarizers on the illumination or detection paths,

and multiview microscopes all have higher angular bandwidths than the microscope considered here.

The angular bandwidth is also sample dependent. Even though they are just a model, monopoles emit light isotropically so they have an ℓ bandwidth of 1, while dipoles have an ℓ bandwidth of 2, and higher-order excitation and detection moments will have even higher bandwidths. Multi-photon excitation and other non-linear methods can also increase the ℓ bandwidth.

4.4. Towards more realistic models

The theoretical model we presented in this work is an extreme simplification of a real microscope. We have ignored the effects of thick samples, refractive-index mismatch, aberration, scattering, rotation of fluorophores, finite fluorescence lifetimes, and interactions between fluorophores among others. Because of this long list of unknown effects, real experiments will likely require extensions of the models developed here.

TODO: Discuss/plan future papers. SVD? Defocus/3D imaging: $\mathcal{H} : \mathbb{L}_2(\mathbb{R}^3 \times \mathbb{S}^2) \rightarrow \mathbb{L}_2(\mathbb{R}^3)$.

TODO: Mention and cite rotating dipoles like Lew (model start point, rotation, end point): $\mathcal{H} : \mathbb{L}_2(\mathbb{R}^3 \times \mathbb{S}^2 \times \mathbb{SO}^3 \times \mathbb{S}^2) \rightarrow \mathbb{V}$.

TODO: Cite connection between dipole pupil and Jones pupil. We can use the Jones/dipole pupil as a way to extend our model in a straightforward way. Add phase aberration (like the Gibson-Lanni model), polarization aberrations, amplitude masks, phase masks, etc.

TODO: Cite Piestun Green’s tensor engineering.

5. Conclusion

Most models of fluorescence microscopes use a monopole model to describe the object and the scalar approximation to model the imaging system. A more complete description of any fluorescence microscope requires a dipole model of the object and vector optics to model the propagation of light through the microscope. We developed several transfer functions that simplify the mapping between the dipole density and the irradiance pattern on the detector, and we demonstrated these transfer functions by efficiently simulating a paraxial fluorescence microscope.

References

1. S. F. Gibson and F. Lanni, “Diffraction by a circular aperture as a model for three-dimensional optical microscopy,” *J. Opt. Soc. Am. A* **6**, 1357–1367 (1989).
2. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Elsevier Science Limited, 1980).
3. B. Richards and E. Wolf, “Electromagnetic diffraction in optical systems, ii. structure of the image field in an aplanatic system,” *Proc. Royal Soc. Lond. A: Math. Phys. Eng. Sci.* **253**, 358–379 (1959).
4. A. S. Backer and W. E. Moerner, “Extending single-molecule microscopy using optical Fourier processing,” *J. Phys. Chem. B* **118**, 8313–8329 (2014).
5. M. A. Lieb, J. M. Zavislan, and L. Novotny, “Single-molecule orientations determined by direct emission pattern imaging,” *J. Opt. Soc. Am. B* **21**, 1210–1215 (2004).
6. L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, 2006).
7. M. P. Backlund, M. D. Lew, A. S. Backer, S. J. Sahl, and W. E. Moerner, “The role of molecular dipole orientation in single-molecule fluorescence microscopy and implications for super-resolution imaging,” *ChemPhysChem* **15**, 587–599 (2014).
8. S. B. Mehta, M. McQuilken, P. J. La Rivière, P. Occhipinti, A. Verma, R. Oldenbourg, A. S. Gladfelter, and T. Tani, “Dissection of molecular assembly dynamics by tracking orientation and position of single molecules in live cells,” *Proc. Natl. Acad. Sci. U.S.A.* **113**, E6352–E6361 (2016).
9. H. Barrett and K. Myers, *Foundations of image science*, Wiley series in pure and applied optics (Wiley-Interscience, 2004).
10. M. Mansuripur, *Classical Optics and Its Applications* (Cambridge University Press, 2009).
11. M. Gu, *Advanced Optical Imaging Theory*, Springer Series in Optical Sciences (Springer, 2000).
12. J. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1996), 2nd ed.
13. R. Bracewell, *Fourier Analysis and Imaging* (Springer US, 2004).

14. P. N. Petrov, Y. Shechtman, and W. E. Moerner, “Measurement-based estimation of global pupil functions in 3d localization microscopy,” *Opt. Express* **25**, 7945–7959 (2017).
15. M. P. Backlund, Y. Shechtman, and R. L. Walsworth, “Fundamental precision bounds for three-dimensional optical localization microscopy with poisson statistics,” *Phys. Rev. Lett.* **121**, 023904 (2018).
16. J. T. Fourkas, “Rapid determination of the three-dimensional orientation of single molecules,” *Opt. Lett.* **26**, 211–213 (2001).
17. N. Schaeffer, “Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations,” *Geochem. Geophys. Geosystems* **14**, 751–758 (2013).
18. I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products* (Elsevier/Academic Press, Amsterdam, 2007), seventh ed.
19. J. Mertz, *Introduction to Optical Microscopy* (W. H. Freeman, 2009).

A. Spherical harmonics

The spherical harmonic function of degree ℓ and order $-\ell \leq m \leq \ell$ is defined as [17]

$$Y_\ell^m(\vartheta, \varphi) = \sqrt{\frac{2\ell+1}{4\pi}} \sqrt{\frac{(\ell-|m|)!}{(\ell+|m|)!}} P_\ell^m(\cos \vartheta) \exp(im\varphi), \quad (89)$$

where $P_\ell^m(\cos \theta)$ are the associated Legendre polynomials with the Condon-Shortley phase

$$P_\ell^m(x) = (-1)^m (1-x^2)^{|m|/2} \frac{d^{|m|}}{dx^{|m|}} P_\ell(x), \quad (90)$$

and $P_\ell(x)$ are the Legendre polynomials defined by the recurrence

$$P_0(x) = 1, \quad (91)$$

$$P_1(x) = x, \quad (92)$$

$$\ell P_\ell(x) = (2\ell-1)xP_{\ell-1}(x) - (\ell-1)P_{\ell-2}(x). \quad (93)$$

The spherical harmonics are orthonormal, which means that

$$\int_{\mathbb{S}^2} d\hat{\mathbf{s}} Y_\ell^m(\hat{\mathbf{s}}) Y_{\ell'}^{m'*}(\hat{\mathbf{s}}) = \delta_{\ell\ell'} \delta_{mm'}, \quad (94)$$

where $\delta_{\ell\ell'}$ denotes the Kronecker delta. The spherical harmonics form a complete basis so an arbitrary function on the sphere $f(\hat{\mathbf{s}})$ can be expanded into a sum of weighted spherical harmonic functions

$$f(\hat{\mathbf{s}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} F_\ell^m Y_\ell^m(\hat{\mathbf{s}}). \quad (95)$$

We can find the spherical harmonic coefficients F_ℓ^m for a given function using Fourier’s trick—multiply both sides by $Y_\ell^{m*}(\hat{\mathbf{s}})$, integrate over the sphere, and exploit orthogonality to find that

$$F_\ell^m = \int_{\mathbb{S}^2} d\hat{\mathbf{s}} f(\hat{\mathbf{s}}) Y_\ell^{m*}(\hat{\mathbf{s}}). \quad (96)$$

The coefficients F_ℓ^m are called the *spherical Fourier transform* of a spherical function.

We briefly show how two properties of spherical functions propagate to the spherical Fourier transform. First, the spherical harmonic coefficients of spherical functions that can be written entirely as a function of the inclination angle ϑ are zero when $m \neq 0$ because

$$\int_{\mathbb{S}^2} d\hat{\mathbf{s}} f(\vartheta) Y_\ell^{m*}(\hat{\mathbf{s}}) = \int_0^{2\pi} d\phi \exp[-im\varphi] \int_0^\pi d\vartheta \sin \vartheta f(\vartheta) P_\ell^m(\cos \vartheta), \quad (97)$$

$$= \delta_{m0} \int_0^\pi d\vartheta \sin \vartheta f(\vartheta) P_\ell^m(\cos \vartheta). \quad (98)$$

Second, the spherical harmonic coefficients of spherical functions that are symmetric under inversion $f(\hat{\mathbf{s}}) = f(-\hat{\mathbf{s}})$ are zero when ℓ is odd because $Y_\ell^m(-\hat{\mathbf{s}}_o) = (-1)^\ell Y_\ell^m(\hat{\mathbf{s}})$, so if ℓ is odd then

$$\int_{\mathbb{S}^2} d\hat{\mathbf{s}} f(\hat{\mathbf{s}}) Y_{2n+1}^{m*}(\hat{\mathbf{s}}) = \int_{\mathbb{S}^2/2} d\hat{\mathbf{s}} f(+\hat{\mathbf{s}}) Y_{2n+1}^{m*}(+\hat{\mathbf{s}}) + \int_{\mathbb{S}^2/2} d\hat{\mathbf{s}} f(-\hat{\mathbf{s}}) Y_{2n+1}^{m*}(-\hat{\mathbf{s}}), \quad (99)$$

$$= \int_{\mathbb{S}^2/2} d\hat{\mathbf{s}} f(\hat{\mathbf{s}}) Y_{2n+1}^{m*}(+\hat{\mathbf{s}}) - \int_{\mathbb{S}^2/2} d\hat{\mathbf{s}} f(\hat{\mathbf{s}}) Y_{2n+1}^{m*}(\hat{\mathbf{s}}) = 0. \quad (100)$$

B. Relationships between special functions

Our first task is to show that

$$i^n \begin{Bmatrix} \exp(in\phi_r) \\ \cos(n\phi_r) \\ \sin(n\phi_r) \end{Bmatrix} \text{jinc}_n(r) \xrightarrow{\mathcal{F}_{\mathbb{R}^2}} (2\nu)^n \begin{Bmatrix} \exp(in\phi_\nu) \\ \cos(n\phi_\nu) \\ \sin(n\phi_\nu) \end{Bmatrix} \Pi(\nu). \quad (101)$$

Writing the inverse Fourier transform in polar coordinates yields

$$= 2^n \int_0^{1/2} d\nu \nu^{n+1} \int_0^{2\pi} d\phi_\nu \begin{Bmatrix} \exp(in\phi_\nu) \\ \cos(n\phi_\nu) \\ \sin(n\phi_\nu) \end{Bmatrix} \exp[2\pi i \nu r \cos(\phi_\nu - \phi_r)]. \quad (102)$$

The azimuthal integral can be evaluated in terms of an n^{th} order Bessel function (see Barrett and Myers 4.111 for the complex case [9])

$$= 2^n 2\pi i^n \begin{Bmatrix} \exp(in\phi_r) \\ \cos(n\phi_r) \\ \sin(n\phi_r) \end{Bmatrix} \int_0^{1/2} d\nu \nu^{n+1} J_n(2\pi \nu r). \quad (103)$$

We can use the following identity (see GR 6.561-5 [18])

$$\int_0^1 du u^{n+1} J_n(au) = a^{-1} J_{n+1}(a) \quad (104)$$

with a change of variable $u = 2\nu$ to find the final result

$$= 2^n 2\pi i^n \begin{Bmatrix} \exp(in\phi_r) \\ \cos(n\phi_r) \\ \sin(n\phi_r) \end{Bmatrix} \int_0^1 \frac{du}{2} \left(\frac{u}{2}\right)^{n+1} J_n(\pi u r) = i^n \begin{Bmatrix} \exp(in\phi_r) \\ \cos(n\phi_r) \\ \sin(n\phi_r) \end{Bmatrix} \frac{J_{n+1}(\pi r)}{2r} = i^n \begin{Bmatrix} \exp(in\phi_r) \\ \cos(n\phi_r) \\ \sin(n\phi_r) \end{Bmatrix} \text{jinc}_n(r). \quad (105)$$

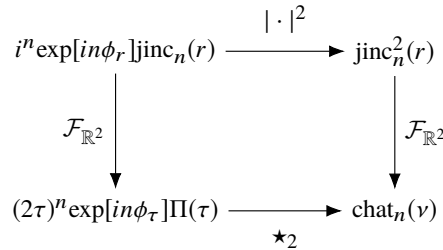


Fig. 15. The relationships between special functions. The chat functions are defined as the two-dimensional Fourier transform of the squared jinc functions, and they can be calculated with the two-dimensional complex autocorrelations (denoted by \star_2) of the complex-weighted rectangle functions.

We can use the relationship in Eq. 101 to express the chat functions in terms of a complex autocorrelation—see the diagram in Figure 15. Starting with the definition of the n^{th} -order chat function

$$\text{chat}_n(\nu) = \int_{\mathbb{R}^2} d\mathbf{r} \text{jinc}_n^2(|\mathbf{r}|) \exp[-2\pi i \mathbf{r} \nu], \quad (106)$$

we can rewrite the integrand in terms of the absolute square of a simpler function with a known Fourier transform

$$\text{chat}_n(\nu) = \int_{\mathbb{R}^2} d\mathbf{r} |t_n(\mathbf{r})|^2 \exp[-2\pi i \mathbf{r} \nu]. \quad (107)$$

$$t_n(\mathbf{r}) = i^n \exp[in\phi_r] \text{jinc}_n(r). \quad (108)$$

Now we can apply the autocorrelation theorem to rewrite the Fourier transform as

$$\text{chat}_n(\nu) = \int_{\mathbb{R}^2} d\boldsymbol{\tau} T_n(\boldsymbol{\tau}) T_n^*(\boldsymbol{\tau} - \nu), \quad (109)$$

where the function to be autocorrelated can be found with the help of Eq. 101

$$T_n(\boldsymbol{\tau}) = \int_{\mathbb{R}^2} d\mathbf{r} t_n(\mathbf{r}) \exp[-2\pi i \mathbf{r} \cdot \boldsymbol{\tau}] = (2\tau)^n \exp[in\phi_\tau] \Pi(\tau). \quad (110)$$

It will be more convenient to set up the autocorrelation in Cartesian coordinates

$$T_n(\boldsymbol{\tau}) = 2^n (\tau_x + i\tau_y)^n \Pi\left(\sqrt{\tau_x^2 + \tau_y^2}\right). \quad (111)$$

Plugging Eq. 111 into Eq. 109 gives

$$\text{chat}_n(\nu) = 4^n \int_{\mathbb{R}^2} d\boldsymbol{\tau} (\tau_x^2 + \tau_y^2 - \nu\tau_x)^n \Pi\left(\sqrt{\tau_x^2 + \tau_y^2}\right) \Pi\left(\sqrt{(\tau_x - \nu)^2 + \tau_y^2}\right). \quad (112)$$

We can interpret the autocorrelation as an integral over a region of overlap between a circle centered at the origin and a circle shifted to the right by ν (a geometric lens). Using the construction in Figure 16 we can express this region as

$$\begin{aligned} \text{chat}_n(\nu) = 4^{n+1} & \left[\int_0^{1/2} \tau d\tau \int_0^{\cos^{-1}\nu} d\phi_\tau (\tau^2 - \nu\tau \cos \phi_\tau)^n - \right. \\ & \left. \int_0^{\nu/2} d\tau_x \int_0^{\frac{\tau_x}{\nu} \sqrt{1-\nu^2}} d\tau_y (\tau_x^2 + \tau_y^2 - \nu\tau_x)^n \right] \Pi\left(\frac{\nu}{2}\right). \end{aligned} \quad (113)$$

For $n = 0$:

$$\text{chat}_0(\nu) = 4 \left[\int_0^{1/2} \tau d\tau \int_0^{\cos^{-1}\nu} d\phi_\tau - \int_0^{\nu/2} d\tau_x \int_0^{\frac{\tau_x}{\nu} \sqrt{1-\nu^2}} d\tau_y \right] \Pi\left(\frac{\nu}{2}\right), \quad (114)$$

$$\text{chat}_0(\nu) = \frac{1}{2} \left[\cos^{-1} \nu - \nu \sqrt{1 - \nu^2} \right] \Pi\left(\frac{\nu}{2}\right), \quad (115)$$

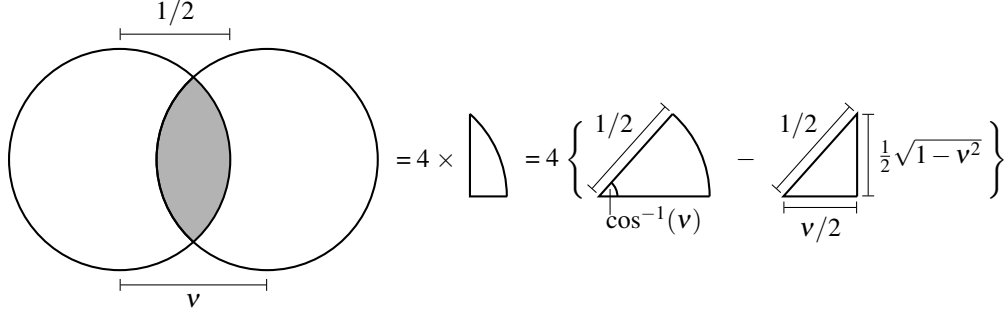


Fig. 16. Geometric construction for evaluating the autocorrelation. We need to integrate over the overlapping region of two circles with radius $1/2$ and distance v between their centers. The region is four times the difference in area between a sector of angle $\arccos(v)$ and radius $1/2$ and a right triangle with base $v/2$ and hypotenuse $1/2$.

which is a well-known result [12, 13, 19]. For $n = 1$:

$$\text{chat}_1(v) = 16 \left[\int_0^{1/2} \tau d\tau \int_0^{\cos^{-1} v} d\phi_\tau (\tau^2 - v\tau \cos \phi_\tau) - \right. \quad (116)$$

$$\left. \int_0^{v/2} d\tau_x \int_0^{\frac{\tau_x}{v} \sqrt{1-v^2}} d\tau_y (\tau_x^2 + \tau_y^2 - v\tau_x) \right] \Pi\left(\frac{v}{2}\right), \quad (117)$$

$$\text{chat}_1(v) = \frac{1}{2} \left[\cos^{-1} v - v(3 - 2v^2)\sqrt{1 - v^2} \right] \Pi\left(\frac{v}{2}\right). \quad (118)$$

C. Spherical Fourier transform of a double cone

In this appendix we will evaluate the spherical Fourier transform of a normalized double-cone angular distribution with central direction $\hat{\mathbf{s}}'$ and cone half-angle Δ

$$f_{(\text{cone})}(\hat{\mathbf{s}}; \hat{\mathbf{s}}', \Delta) = \frac{1}{2\pi(1 - \cos \Delta)} \Pi\left(\frac{\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}'}{\cos \Delta}\right). \quad (119)$$

The spherical Fourier transform is

$$F_{\ell(\text{cone})}^m(\hat{\mathbf{s}}', \Delta) = \int_{\mathbb{S}^2} d\hat{\mathbf{s}} f_{(\text{cone})}(\hat{\mathbf{s}}; \hat{\mathbf{s}}', \Delta) Y_\ell^{m*}(\hat{\mathbf{s}}). \quad (120)$$

The limits of integration will be difficult to find unless we change coordinates to exploit the axis of symmetry $\hat{\mathbf{s}}'$. Since the spherical function is rotationally symmetric about $\hat{\mathbf{s}}'$ we can rotate the function so that the axis of symmetry is aligned with $\hat{\mathbf{z}}$ and multiply by $\sqrt{\frac{4\pi}{2l+1}} Y_\ell^m(\hat{\mathbf{s}}')$ to account for the rotation

$$F_{\ell(\text{cone})}^m(\hat{\mathbf{s}}', \Delta) = \sqrt{\frac{4\pi}{2l+1}} Y_\ell^m(\hat{\mathbf{s}}') \int_{\mathbb{S}^2} d\hat{\mathbf{s}} f_{(\text{cone})}(\vartheta; \hat{\mathbf{z}}, \Delta) Y_\ell^0(\hat{\mathbf{s}}). \quad (121)$$

In this coordinate system the double cone is independent of the azimuthal angle, so we can evaluate the azimuthal integral and express the function in terms of an integral over ϑ

$$F_{\ell(\text{cone})}^m(\hat{\mathbf{s}}', \Delta) = 2\pi Y_\ell^m(\hat{\mathbf{s}}') \int_0^\pi d\vartheta \sin \vartheta f_{(\text{cone})}(\vartheta; \hat{\mathbf{z}}, \Delta) P_\ell(\cos \vartheta). \quad (122)$$

The function $f_{\text{(cone)}}(\vartheta; \hat{\mathbf{z}}, \Delta)$ is only non-zero on the intervals $\vartheta \in [0, \Delta]$ and $\vartheta \in [\pi - \Delta, \pi]$ so

$$F_{\ell(\text{cone})}^m(\hat{\mathbf{s}}', \Delta) = \frac{Y_{\ell}^m(\hat{\mathbf{s}}')}{2(1 - \cos \Delta)} \left[\int_0^{\Delta} d\vartheta \sin \vartheta P_{\ell}(\cos \vartheta) + \int_{\pi-\Delta}^{\pi} d\vartheta \sin \vartheta P_{\ell}(\cos \vartheta) \right]. \quad (123)$$

Applying a change of coordinates with $u = \cos \vartheta$ yields

$$F_{\ell(\text{cone})}^m(\hat{\mathbf{s}}', \Delta) = \frac{Y_{\ell}^m(\hat{\mathbf{s}}')}{2(1 - \cos \Delta)} \left[\int_{\cos \Delta}^1 d\vartheta P_{\ell}(u) + \int_{-1}^{-\cos \Delta} d\vartheta P_{\ell}(u) \right]. \quad (124)$$

The Legendre polynomials $P_{\ell}(u)$ are even (odd) on the interval $[-1, 1]$ when ℓ is even (odd), so the pair of integrals will be identical when ℓ is even and cancel when ℓ is odd. For even ℓ

$$F_{\ell(\text{cone})}^m(\hat{\mathbf{s}}', \Delta) = \frac{Y_{\ell}^m(\hat{\mathbf{s}}')}{1 - \cos \Delta} \int_{\cos \Delta}^1 d\vartheta P_{\ell}(u). \quad (125)$$

The integral evaluates to (see GR 7.111 [18])

$$\int_{\cos \Delta}^1 d\vartheta P_{\ell}(u) = \begin{cases} 1 - \cos \Delta, & \ell = 0, \\ \sin \Delta P_{\ell}^{-1}(\cos \Delta), & \text{else,} \end{cases} \quad (126)$$

where $P_l^{-1}(\cos \Delta)$ is the associated Legendre polynomial with order $m = -1$, not an inverse Legendre polynomial. Bringing everything together

$$F_{\ell(\text{cone})}^m(\hat{\mathbf{s}}', \Delta) = \begin{cases} \sqrt{1/(4\pi)}, & \ell = 0, \\ 0, & \ell \text{ odd,} \\ Y_{\ell}^m(\hat{\mathbf{s}}') \cot\left(\frac{\Delta}{2}\right) P_l^{-1}(\cos \Delta), & \ell > 0 \text{ even.} \end{cases} \quad (127)$$