Fourier Analysis of Interference Scanning Optical Probe Microscopy

Emmanuel Soubies, Wolfgang Bacsa

Fourier Analysis of Interference Scanning Optical Probe Microscopy

Emmanuel Soubies and Wolfgang Bacsa

***Abstract*—As opposed to popular far-field and near-field optical microscopy techniques, Interference Scanning Optical probe Microscopy (ISOM) operates in the intermediate-field region, where the probing distance is typically of the order of the wavelength of incident light. Specifically, ISOM enables the imaging of nanostructures through numerical inverse scattering of standing waves generated by the interference between the incident (or reflected) and scattered waves. In this work, we shed new light on this microscopy modality through an in-depth Fourier analysis. Our analysis reveals insights on the required acquisition sampling step as well as on the resolution limit of the system. Moreover, we propose two novel methods to address the associated inverse scattering problem, leveraging the intrinsic structure of the image formation model to reduce computational complexity and sensitivity to errors in model parameters. Finally, we illustrate our theoretical findings with numerical experiments.**

***Index Terms*—Fourier analysis, Intermediate-field optical microscopy, Inverse scattering, Resolution limit.**

I. Introduction

OPTICAL microscopy techniques can be classified into two categories based on whether they operate in the near-field of far-field region. Traditional far-field optical methods collect transmitted or reflected light scattered by the sample, using a lens positioned at a distance $z_{\rm im}$ much larger than the wavelength of light λ (i.e., $z_{\rm im} \gg \lambda$). At these distances, only propagative far-field waves—which carry limited low-frequency information—can be collected, resulting in the well-known diffraction limit. However, recent advancements in super-resolution techniques [1]–[3] have enabled far-field microscopy to overcome this diffraction barrier. This breakthrough was made possible through innovative light-sample interactions (e.g., patterned light, specific fluorescent probes) coupled with numerical reconstruction.

Alternatively, Near-field Scanning Optical Microscopy (NSOM) [4]–[7] bypasses the diffraction limit by probing the light in the near-field region, at distances much smaller than the wavelength of light (i.e., $z_{\rm im} \ll \lambda$). At such proximity, non-propagative evanescent waves—which carry high spatial frequencies of the sample—can be captured, enabling sub-wavelength resolution. Through its long and rich history [8]–[11], NSOM has attracted a significant attention and established itself as a powerful technique for marker-free super-resolution imaging.

E. Soubies is with IRIT, Université de Toulouse, CNRS 31000 Toulouse, France (e-mail: emmanuel.soubies@irit.fr). W. Bacsa is with CEMES-CNRS, Université de Toulouse, 31055 Toulouse, France (e-mail: wolfgang.bacsa@cemes.fr).

While extensive research has focused on near-field and far-field microscopy, the intermediate-field region (i.e., $z_{\rm im} \sim \lambda$) has received significantly less attention in the development of optical microscopy techniques. In this region, the interference between scattered and incident or reflected waves generates complex standing waves. Similar to NSOM, these interference patterns can be observed by scanning an optical fiber tip in collection mode near the surface [12]. Interference Scanning Optical probe Microscopy (ISOM) [13], [14] exploits this principle by integrating such measurements (interference patterns between reflected and scattered waves) with inverse scattering techniques to reconstruct the imaged sample. ISOM relies on a less invasive probing method than NSOM, reducing disturbances caused by the presence of the probe.

Contributions and Outline: In this work, we shed new light on ISOM through an in-depth Fourier analysis. Building on the description of standing waves near surfaces using a dipole model of scatterers [14], [15], we show that the image formation model for a distribution of scatterers can be expressed as a combination of sinusoidal modulations and convolution operations (Section II). This specific decomposition allows us to perform a Fourier analysis that reveals key insights into the required acquisition sampling step and the resolution limit of the system (Section III). Then, in Section IV, we introduce two approaches for solving the inverse scattering problem, leveraging the particular structure of the forward model to reduce computational complexity and enhance robustness to errors in model parameters. Finally, in Section V, we present numerical experiments that are in good agreement with our theoretical Fourier analysis.

Notations: We write vectors as bold lowercase letters (e.g., $\mathbf{x}$, $\mathbf{k}$), matrices as bold uppercase letters (e.g., $\mathbf{A}$) and their transpose as $\mathbf{x}^{\rm T}$, $\mathbf{k}^{\rm T}$, $\mathbf{A}^{\rm T}$. The mth component of a vector $\mathbf{x} \in \mathbb{R}^M$ is x_m or $[\mathbf{x}]_m$, while the (m,n) component of a matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$ is a_{mn} or $[\mathbf{A}]_{mn}$. We denote the Fourier transform of a function h by $\hat{h}$ or $\mathcal{F}\{h\}$, depending on the situation ($\mathcal{F}\{h\}$ is preferred when, instead of $\hat{h}$, we have a longer expression of which we want to express its Fourier transform). The symbols $*$ and $\odot$ stand for convolution and point-wise multiplication operations, respectively. The delta Dirac is denoted $\delta()$.

II. Image Formation Model

To accurately reconstruct the imaged object from standing optical waves, it is essential to establish a physical model that elucidates the formation of interference fringes.

The incident wave E_i is assumed to be planar, while the scattered waves E_s from the object are regarded as point

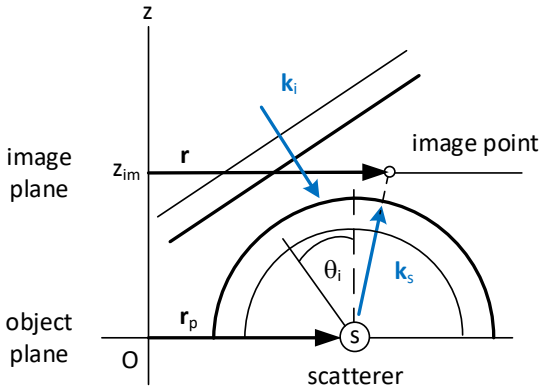

Fig. 1. Schematic of Interference Scanning Optical probe Microscopy. Object ($z = 0$) and image ($z = z_{\rm im}$) planes with position vector of the scatterer $\mathbf{r}_p$, image point $\mathbf{r}$, and wave vectors of incident ($\mathbf{k}_i$) and scattered ($\mathbf{k}_s$) waves.

dipolar waves. For a single scatterer (dipole) at position $\mathbf{r}_p \in \mathbb{R}^2$ in the object plane ($z = 0$), we have for all $(\mathbf{r}, z) \in \mathbb{R}^3$ and $t > 0$ [15],

$$E_i(\mathbf{r}, z, t) = A_i \cos(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z - \omega t), \tag{1}$$

$$E_s(\mathbf{r} - \mathbf{r}_p, z, t) = \frac{A_p \cos\left(k_0\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z^2} - \omega t\right)}{\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z^2}}, \tag{2}$$

where $\mathbf{k}_i = (\mathbf{k}_r^{\rm T}, k_z) = -k_0.(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^{\rm T}$ and A_i are the wave vector and amplitude of the incident plane wave, respectively, with θ and ϕ the angles defining the direction of $\mathbf{k}_i$ (see Fig. 1). The scalar $k_0 = 2\pi/\lambda$ corresponds to the wavenumber, with λ the wavelength of the incident plane wave, while $\omega \in [0, 2\pi)$ represents the angular frequency. Finally, A_p models the amplitude of the wave emitted by the dipole, ignoring polarization effects, so that A_i and A_p are scalars.

Let the imaged object be made of P scaterrers at positions $\{\mathbf{r}_p\}_{p=1}^P$ of the object plane ($z = 0$) and assume that the optical probe can be approximated as a point detector. Then, the recorded data at the image plane located at a distance $z = z_{\rm im}$ is an interferogram resulting from the superposition of the incident wave with the scattered waves. We denote this recorded interferogram by y. Formally, for all $\mathbf{r} \in \mathbb{R}^2$ of the image plane ($z = z_{\rm im}$), we have

$$y(\mathbf{r}) = \left\langle \left(E_i(\mathbf{r}, z_{\rm im}, t) + \sum_{p=1}^{P} E_s(\mathbf{r} - \mathbf{r}_p, z_{\rm im}, t) \right)^2 \right\rangle_t, \tag{3}$$

where $\langle f \rangle_t = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} f(t){\rm d}t$ denotes the time average operator. Injecting (1)–(2) into (3) and assuming that $A_i \gg \frac{P}{z_{\rm im}}\max_p A_p$ (roughly speaking, the amplitude A_i of the incident beam is significantly larger than that of the scattered waves A_p), we obtain (details in Appendix A)

$$y(\mathbf{r}) = \frac{A_i^2}{2} + \sum_{p=1}^{P} A_i A_p I(\mathbf{r}, \mathbf{r}_p), \tag{4}$$

with

$$I(\mathbf{r}, \mathbf{r}_p) = \frac{\cos\left(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im} - k_0\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2}\right)}{\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2}}. \tag{5}$$

corresponding to the interference between the incident wave and the scattered wave from the scatterer at position $\mathbf{r}_p$.

Remark 1. *The scattered wave expression in* (2) *is valid for* $z > \frac{1}{k_0} = \frac{\lambda}{2\pi} \simeq \frac{\lambda}{6}$. *For lower values of* z, *additional terms must be considered [15, Section 3.2]. As such, we focus in this work on the regime* $z_{\rm im} > \frac{\lambda}{6}$. *Moreover, with the described setting, we implicitly consider the imaging of flat surfaces whose thickness should not exceed the imaging distance* $z_{\rm im}$.

A handier formulation: We can equivalently write I in (5) as

$$\begin{aligned} I(\mathbf{r}, \mathbf{r}_p) &= \frac{1}{\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2}} \cos\Big(\mathbf{k}_r^{\rm T}(\mathbf{r} - \mathbf{r}_p) + k_z z_{\rm im} \\ &\qquad - k_0\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2} + \mathbf{k}_r^{\rm T}\mathbf{r}_p\Big) \\ &= \sum_{p=1}^{P} h_c(\mathbf{r} - \mathbf{r}_p) w_c(\mathbf{r}_p) - h_s(\mathbf{r} - \mathbf{r}_p) w_s(\mathbf{r}_p) \end{aligned}$$

where, for all $\mathbf{r} \in \mathbb{R}^2$,

$$w_c(\mathbf{r}) = \cos(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im}), \tag{6}$$

$$w_s(\mathbf{r}) = \sin(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im}), \tag{7}$$

$$h_c(\mathbf{r}) = \cos\left(\mathbf{k}_r^{\rm T}\mathbf{r} - k_0\sqrt{\|\mathbf{r}\|^2 + z_{\rm im}^2}\right)/\sqrt{\|\mathbf{r}\|^2 + z_{\rm im}^2}, \tag{8}$$

$$h_s(\mathbf{r}) = \sin\left(\mathbf{k}_r^{\rm T}\mathbf{r} - k_0\sqrt{\|\mathbf{r}\|^2 + z_{\rm im}^2}\right)/\sqrt{\|\mathbf{r}\|^2 + z_{\rm im}^2}. \tag{9}$$

In practice, it is more convenient to represent the imaged object as a density function $\rho : \mathbb{R}^2 \to \mathbb{R}$ rather than the sum of a finite number of $P \gg 1$ scatterers. In this context, the model writes

$$y(\mathbf{r}) = \frac{A_i^2}{2} + A_i\left(h_c * (w_c \odot \rho) - h_s * (w_s \odot \rho)\right)(\mathbf{r}) \tag{10}$$

where $\rho(\mathbf{r})$ represents the amplitude of a scatterer at position $\mathbf{r} \in \mathbb{R}^2$.

As such, we have expressed the image formation model as a combination of sinusoidal modulations and convolutions. This particular formulation will greatly facilitate the Fourier analysis conducted in Section III. Additionally, it will enable a substantial reduction of the computational cost of iterative reconstruction algorithms by leveraging the fast Fourier transform for efficient convolution calculations (see Section IV).

III. Fourier Analysis

A. Fourier Transforms of Convolution Kernels

We start the analysis with the derivation of the Fourier transforms of h_c and h_s in Proposition 1 below. The proof is provided in Appendix B.

Proposition 1. *The Fourier transforms of* h_c *and* h_s *are given by, for all* $\boldsymbol{\nu} \in \mathbb{R}^2$,

$$\widehat{h}_c(\boldsymbol{\nu}) = g_1(\|\boldsymbol{\nu} + \mathbf{k}_r\|) + g_2(\|\boldsymbol{\nu} - \mathbf{k}_r\|) \tag{11}$$

$$\widehat{h}_s(\boldsymbol{\nu}) = {\rm i}\left(g_1(\|\boldsymbol{\nu} + \mathbf{k}_r\|) - g_2(\|\boldsymbol{\nu} - \mathbf{k}_r\|)\right) \tag{12}$$

where for $q \in \{1,2\}$ and all $\nu > 0$,

$$g_q(\nu) = \begin{cases} \dfrac{\mathrm{i}\pi^{\frac{3}{2}}\sqrt{z_{\mathrm{im}}}H^{(q)}_{-\frac{1}{2}}\left(z_{\mathrm{im}}\sqrt{k_0^2-\nu^2}\right)}{(-1)^{q+1}\sqrt{2}(k_0^2-\nu^2)^{\frac{1}{4}}} & \text{if } \nu < k_0 \\ \dfrac{\pi\exp\left(-z_{\mathrm{im}}\sqrt{\nu^2-k_0^2}\right)}{\sqrt{\nu^2-k_0^2}} & \text{if } \nu > k_0 \end{cases} \tag{13}$$

with $H^{(q)}_{-\frac{1}{2}}$ the Hankel function of first ($q=1$) or second ($q=2$) kind of order $-\frac{1}{2}$.

From these expressions, we learn that most of the energy of the Fourier transform of both h_c and h_s is concentrated within two disks of radius k_0 centered respectively at $\pm\mathbf{k}_r$ (see Fig 2). Outside of these disks, the energy decays as $\exp(-z_{\mathrm{im}}\nu)/\nu$. Although the support (i.e., non-zero region) of $\widehat{h}_c$ and $\widehat{h}_s$ is unbounded, this decay implies that, for sufficiently large frequencies $\boldsymbol{\nu}$, the attenuation becomes too strong relative to the level of noise that degrades the data. This observation leads us to define the *effective support* of $\widehat{h}_c$ and $\widehat{h}_s$ as

$$\begin{aligned} \mathrm{supp}_{\mathrm{eff}}(\widehat{h}_c) &= \mathrm{supp}_{\mathrm{eff}}(\widehat{h}_c) \\ &= \left\{\boldsymbol{\nu} \in \mathbb{R}^2 : \|\boldsymbol{\nu} \pm \mathbf{k}_r\| < k_0 + \nu_{\mathrm{eff}}\right\}, \end{aligned} \tag{14}$$

where

$$\nu_{\mathrm{eff}} = \sqrt{\frac{1}{z_{\mathrm{im}}^2}\left(W\left(\frac{\pi z_{\mathrm{im}}}{\eta_{\mathrm{eff}}}\right)\right)^2 + k_0^2} - k_0 \tag{15}$$

with W the Lambert function[1] and $\eta_{\mathrm{eff}} > 0$ the Fourier magnitude below which noise is considered dominant.

The effective support of $\widehat{h}_c$ corresponds to the region within the two red circles in Fig. 2 (top). An illustration of the relationship between ν_{eff} and η_{eff} is provided in Fig. 2 (bottom). Finally, it is found that the larger z_{im}, the faster the decay $\exp(-z_{\mathrm{im}}\nu)/\nu$ and thus the smaller the effective support for a fixed level η_{eff}.

B. Proper Sampling Step

To prevent aliasing in the measurements and to assure correct reconstruction, the sampling step needs to be properly determined. A too large sampling step will create aliased data and make the reconstruction task much more challenging, affecting image quality. Conversely, a too small sampling step will slow down the acquisition.

A proper choice is to set the sampling step according to the Shannon-Nyquist sampling theorem. From (14), we get that the maximal frequency visible through the filters h_c and h_s is given by $\nu_{\max} = \|\mathbf{k}_r\| + k_0 + \nu_{\mathrm{eff}} = k_0(1+\sin\theta) + \nu_{\mathrm{eff}}$. Following Shannon-Nyquist sampling theorem, the proper sampling frequency is $\nu_{\mathrm{s}} = 2\nu_{\max}$, which corresponds to the spatial sampling step

$$\begin{aligned} \delta r = \frac{1}{\nu_{\mathrm{s}}} &= \frac{1}{2k_0(1+\sin\theta+\nu_{\mathrm{eff}}/k_0)} \\ &= \frac{\lambda}{4\pi(1+\sin\theta+\nu_{\mathrm{eff}}/k_0)} \text{ [m/rad]} \end{aligned} \tag{16}$$

that is (in meters) $\delta r = \lambda/(2(1+\sin\theta+\nu_{\mathrm{eff}}/k_0))$.

[1]The Lambert function is defined as $w = W(z)$ such that $z = we^w$.

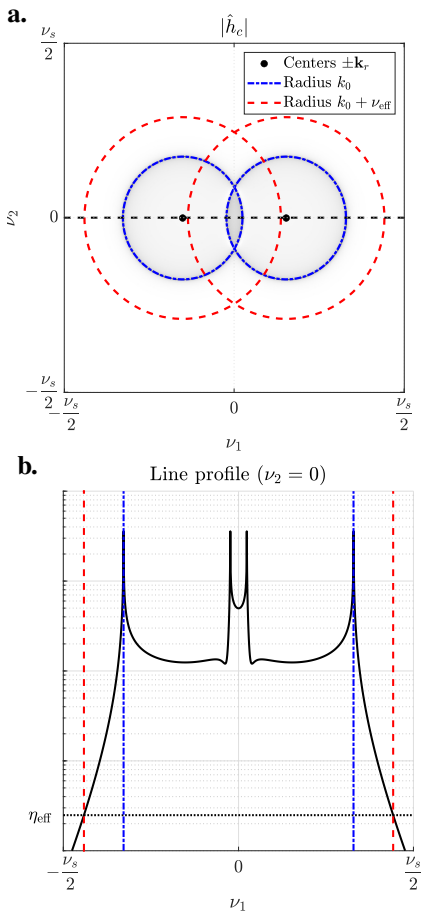

Fig. 2. Amplitude of the Fourier transform of the cosine kernel. **a.** Two-dimensional representation in the Fourier plane ((ν_1,ν_2) are the frequency coordinates). For this example, $\lambda = 500$nm, $\theta = 60°$, $\phi = 0°$, and the image plane position is set to half the incident wavelength (i.e., $z_{\mathrm{im}} = \lambda/2$). The dashed red circles represent the effective Fourier support of recorded data, which is defined in (14). The value of η_{eff} has been adjusted manually so as to match the visible Fourier frequencies of the noiseless data in Fig. 3. **b.** Line profile corresponding to $\nu_2 = 0$. The vertical dashed red lines represent the extremity of the dashed red circles in **a**, that is the frequency where the amplitude of the kernel reaches the level η_{eff}.

C. On the Resolution Limit of the System

Extending the previous Fourier analysis made for the kernels h_c and h_s to the full forward model (10) will give us insights on the frequency information contained in the data, that is, on the resolution limit of the system. At first sight, Model (10) bears strong similarities with structured illumination microscopy (SIM) [16], [17] in that it is made up of a combination of modulations and convolutions operations. As such, building upon our knowledge for SIM, we could expect from sinusoidal modulations (with frequency $\mathbf{k}_r$) followed by convolutions (with effective Fourier support $\|\mathbf{k}_r\| + k_0 + \nu_{\mathrm{eff}}$) that the recorded data carry frequency information (in direction $\pm\mathbf{k}_r$) up to $2\|\mathbf{k}_r\| + k_0 + \nu_{\mathrm{eff}}$.

Yet, this is not the case due to cancellation effects between the cosine and sine parts of Model (10), as highlighted by Proposition 2 (proof in Appendix C).

Proposition 2. *The Fourier transform of the measurements y*

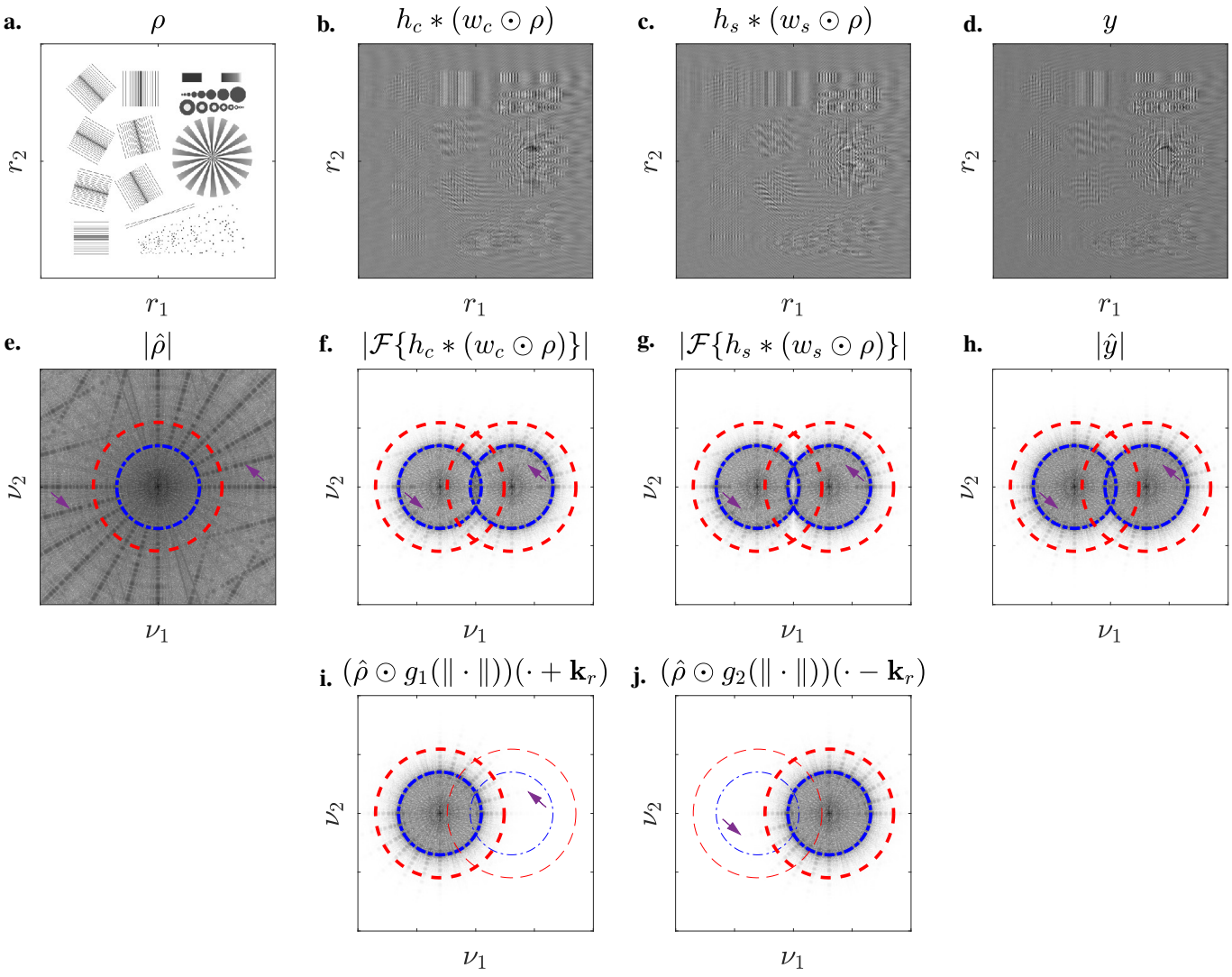

Fig. 3. Illustration of the different terms of Model (10) in spatial and Fourier domains. **a.** Object from [18]. **b.** Cosine part of (10). **c.** Sine part of (10). **d.** Measurements y. Figures **e** to **f** represent the Fourier transforms of **a** to **d**. Physical parameters are the same as those used in Fig. 2. Blue and red circles have thus the same radius as those in Fig. 2. Arrows in **e** highlight examples of high frequency content which, although being "shifted" within the effective Fourier support of the convolution kernels in the cosine (**f**) and sine (**g**) parts of the model, is not present in the final measurement y (**h**). This observation is theoretically explained with Proposition 2 and illustrated in **i** and **j** where the two main terms of (17) are displayed.

is given by, for all $\boldsymbol{\nu} \in \mathbb{R}^2$,

$$\widehat{y}(\boldsymbol{\nu}) = \frac{A_i^2}{2}\delta(\boldsymbol{\nu}) + A_i\Big(\mathrm{e}^{-\mathrm{i}k_z z_{\mathrm{im}}}\left(\widehat{\rho} \odot g_1(\|\cdot\|)\right)(\boldsymbol{\nu} + \mathbf{k}_r) + \mathrm{e}^{\mathrm{i}k_z z_{\mathrm{im}}}\left(\widehat{\rho} \odot g_2(\|\cdot\|)\right)(\boldsymbol{\nu} - \mathbf{k}_r)\Big). \quad (17)$$

We present in Fig. 3 an illustration of the cosine and sine parts of Model (10) together with their combination (i.e., the measurements y), both in spatial and Fourier domains. Taken individually, we can see that the cosine and sine parts carry high frequency information (see purple arrows). Indeed, as in SIM, the spatial sinusoidal modulations "shift" (in Fourier) this high frequency content to the effective Fourier support of the convolution kernels. However, when combining the cosine and sine parts, we observe that this high frequency information is lost and that only frequencies within the dashed red circle in Fig. 3.e are preserved. This observation is explained by Proposition 2 and illustrated in Fig. 3 (bottom row) where we get that Model (10) can be equivalently expressed as: filtering first and then modulating. Indeed, from (17), one can see that $\widehat{\rho}$ is first multiplied by $g_1(\|\cdot\|)$ (and $g_2(\|\cdot\|)$), prior to be shifted at positions $\pm\mathbf{k}_r$ and summed. As such, high frequency content is lost before modulating (i.e., before "shifting" in Fourier domain) and the frequency information of the object ρ contained in the measurements y is limited by the effective support of $g_1(\|\cdot\|)$ (and $g_2(\|\cdot\|)$). This corresponds to effective frequency (ν_{lim}) and resolution (r_{lim}) limits

$$\nu_{\mathrm{lim}} = k_0 + \nu_{\mathrm{eff}} \text{ [rad/m]}, \quad (18)$$

$$r_{\mathrm{lim}} = \frac{1}{\nu_{\mathrm{lim}}} = \frac{\lambda}{1 + \nu_{\mathrm{eff}}/k_0} \text{ [m]}. \quad (19)$$

In Fig. 4, we report the evolution of r_{lim} in (19) as a function of the image plane distance z_{im}. Given that the resolution limit

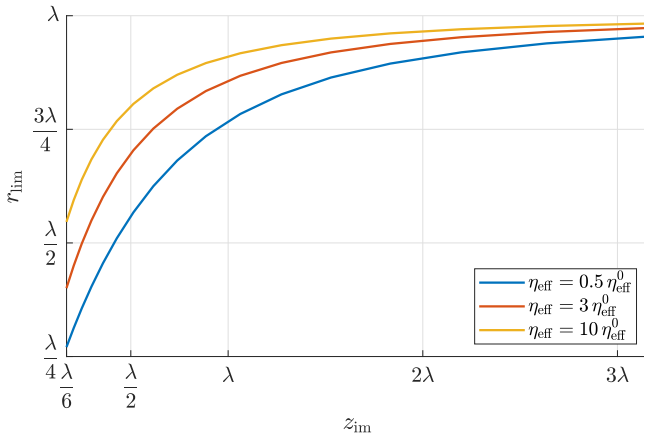

Fig. 4. Evolution of the resolution limit $r_{\rm lim}$ in (19) as a function of the image plane distance $z_{\rm im}$, all represented as fractions of the incident wavelength $\lambda = 500$nm. Physical parameters are set to the same values as for Fig. 2. The reference level $\eta^0_{\rm eff}$ is set as in Figs. 2 and 3. Each curve corresponds to a variation of the level $\eta_{\rm eff}$ with respect to this reference level.

depends on $\nu_{\rm eff}$, and thus on the choice of the level $\eta_{\rm eff}$ in (15), we report curves for different choice of $\eta_{\rm eff} \propto \eta^0_{\rm eff}$ where the reference $\eta^0_{\rm eff}$ is fixed as in Figs. 2 and 3, i.e., such that the effective Fourier support of convolution kernels includes all visible frequencies in display of the noiseless setting of Fig. 3. As $\eta_{\rm eff}$ increases, the reported curves in Fig. 4 can be associated to more and more noisy settings.

As expected, the resolution limit improves when the distance $z_{\rm im}$ decreases, due to the fact that the exponential decay $\exp(-z_{\rm im}\nu)/\nu$ of g_1 and g_2 for $\nu > k_0$ is slower for smaller values of $z_{\rm im}$. Similarly, the lower $\eta_{\rm eff}$ (in practice the lower the noise) the better the resolution. For large $z_{\rm im}$, all curves tend to λ, which corresponds to the worst resolution achievable by the system: in this regime the exponential decay becomes so fast that, even for very small levels $\eta_{\rm eff}$, g_1 and g_2 can be considered as zero beyond k_0.

IV. Solving the Inverse Scattering Problem

Only a few works have addressed the problem of image reconstruction in ISOM [14], [19]. In [14], a basic deconvolution of the interferogram y is performed using the cosine kernel h_c from (8), which provides a rough approximation of the image formation model and may result in significant artifacts. The study in [19] focuses on the localization of isolated particles by fitting ellipses to the observed elliptical fringes. In contrast, the following sections introduce two novel reconstruction methods capable of handling arbitrary samples, both relying on the complete model (10).

A. Discrete Forward Model

Let denote by $\mathbf{y} \in \mathbb{R}^M$ the acquired image made of samples of y in (10) at grid points $\mathcal{G}_{\rm im} := \{\mathbf{r}^{\rm im}_m \in \mathbb{R}^2\}_{m=1}^M$ of the image plane ($z = z_{\rm im}$). Moreover, we define $\boldsymbol{\rho} \in \mathbb{R}^M$, the vector of samples of ρ at grid points $\mathcal{G}_{\rm ob} := \{\mathbf{r}^{\rm ob}_m \in \mathbb{R}^2\}_{m=1}^M$ of the object plane ($z = 0$). These two bi-dimensional grids only differ by a translation, i.e., there exists $\mathbf{t} \in \mathbb{R}^2$ such that $\mathbf{r}^{\rm im}_m = \mathbf{r}^{\rm ob}_m + \mathbf{t}$ for all $m \in \{1, \dots, M\}$. The reason for introducing this translation can be understood from Fig. 1. Given the angle of the incident wave and the image plane distance $z_{\rm im}$, the lateral position of the (main) diffraction figure is shifted relative to the lateral position of the object (e.g., shift between the scatterer and image point in Fig. 1). As such, to keep the size of the discrete problem reasonable, it is preferable to discretize ρ on a shifted version of the acquisition grid, rather than on an enlarged version of it that would include both the object and the measurements.

Then, we get the following discrete version of (10)

$$\mathbf{y} = \frac{A_i^2}{2} + (\mathbf{H}_c\mathbf{D}_c - \mathbf{H}_s\mathbf{D}_s)\boldsymbol{\rho}, \tag{20}$$

where $\mathbf{D}_c \in \mathbb{R}^{M\times M}$ and $\mathbf{D}_s \in \mathbb{R}^{M\times M}$ are diagonal operators while $\mathbf{H}_c \in \mathbb{R}^{M\times M}$ and $\mathbf{H}_s \in \mathbb{R}^{M\times M}$ are convolution operators with zero boundary conditions (implemented through zero-padding and FFTs). More precisely, the diagonal entries of $\mathbf{D}_c$ and $\mathbf{D}_s$ are made of samples of w_c in (6) and w_s in (7) at grid points $\mathcal{G}_{\rm ob}$. Similarly, the convolution kernels of $\mathbf{H}_c$ and $\mathbf{H}_s$ are made of samples of h_c in (8) and h_s in (9) at grid points $\mathcal{G}_{\rm im}$. Finally, note that here the entries of $\boldsymbol{\rho}$ are given by $\rho_m = A_i A_m$ where A_m represents the amplitude of a scatterer at position $\mathbf{r}^{\rm ob}_m$.

Remark 2. *In addition to being central to the Fourier analysis made in Section III, the decomposition* (10) *allows us to significantly reduce the computational burden of iterative algorithms used for reconstruction. Indeed, the complexity to evaluate the resulting discrete forward model* (20) *is of the order of $O(M \log M)$, that is the same as the basic deconvolution approch of [14], which has to be compared to the $O(M^2)$ complexity that one would obtain through a direct discretization of* (4).

B. Image Reconstruction

We propose two strategies to reconstruct the imaged object $\boldsymbol{\rho}$ from the measurements $\mathbf{y}$ and discuss their advantages and drawbacks.

1) Reconstruction Based on Full Model: The most natural approach is to solve an optimization problem of the form

$$(\hat{\boldsymbol{\rho}}, \hat{a}) \in \arg\min_{\boldsymbol{\rho}\in\mathcal{C}^M, a\in\mathbb{R}_{\geq 0}} J(\boldsymbol{\rho}, a), \tag{21}$$

with

$$J(\boldsymbol{\rho}, a) := \frac{1}{2}\|\mathbf{H}\boldsymbol{\rho} + \mathbf{1}a - \mathbf{y}\|_2^2 + \mu R(\boldsymbol{\rho}).$$

Here, $\mathcal{C} = \mathbb{R}$ or $\mathcal{C} = \mathbb{R}_{\geq 0}$ (non-negativity constraint), $\mathbf{1} \in \mathbb{R}^M$ is a vector of ones, $\mathbf{H} = \mathbf{H}_c\mathbf{D}_c - \mathbf{H}_s\mathbf{D}_s$, $R : \mathbb{R}^M \to \mathbb{R}$ is a convex regularization term weighted by the parameter $\mu > 0$. The scalar $a > 0$ aims to estimate the constant term $A_i^2/2$ present in the model in order to prevent the interpretation of it through artifacts in $\boldsymbol{\rho}$.

To solve (21), we alternate between an optimization on $\boldsymbol{\rho}$ with a fixed and an optimization on a with $\boldsymbol{\rho}$ fixed. While the minimization of $J(\cdot, a)$ over $\mathcal{C}^N$ for fixed a requires to deploy

an iterative method, the minimization of $J(\boldsymbol{\rho}, \cdot)$ over $\mathbb{R}_{\geq 0}$ for fixed $\boldsymbol{\rho}$ admits a closed-form solution

$$a = \max\left(\frac{1}{M}\sum_{m=1}^{M}\left[\mathbf{y} - \mathbf{H}\boldsymbol{\rho}\right]_m, 0\right). \tag{22}$$

We implemented this reconstruction strategy within the `GlobalBioIm` framework [20] which offers a variety of regularizers R and convex optimization algorithms to minimize $J(\cdot, a)$ over $\mathcal{C}^N$ when a is fixed.

2) Reconstruction Based on Joint Deconvolution: As we shall see in Section V-B, the above reconstruction method is quite sensitive to the knowledge of model parameters $\mathbf{k}_i$ and z_{im}, which can be difficult to precisely estimate in practice. In particular, it is essential to have access to accurate modulation functions w_c and w_s to guarantee a faithful reconstruction. This is (again) akin to the importance of illumination patterns estimation in SIM [18], [21]. Differently, the effect of inaccuracies on the kernels h_c and h_s is less critical.

This motivates the reconstruction of both $\boldsymbol{\rho}_c = \mathbf{D}_c\boldsymbol{\rho}$ and $\boldsymbol{\rho}_s = \mathbf{D}_s\boldsymbol{\rho}$ through the following joint deconvolution problem

$$(\hat{\boldsymbol{\rho}}_c, \hat{\boldsymbol{\rho}}_s, \hat{a}) \in \arg\min_{\boldsymbol{\rho}_c, \boldsymbol{\rho}_s \in \mathbb{R}^M, a \in \mathbb{R}^M_{\geq 0}} \tilde{J}(\boldsymbol{\rho}_c, \boldsymbol{\rho}_s, a) \tag{23}$$

with

$$\tilde{J}(\boldsymbol{\rho}_c, \boldsymbol{\rho}_s, a) := \frac{1}{2}\|\mathbf{H}_c\boldsymbol{\rho}_c - \mathbf{H}_s\boldsymbol{\rho}_s + \mathbf{1}a - \mathbf{y}\|_2^2 + \mu R(\boldsymbol{\rho}_c, \boldsymbol{\rho}_s).$$

Then, we obtain $\hat{\boldsymbol{\rho}} = \sqrt{\hat{\boldsymbol{\rho}}_c^{\circ 2} + \hat{\boldsymbol{\rho}}_s^{\circ 2}}$ (where $(\cdot)^{\circ 2}$ denote the component-wise square). As previously, we tackle this optimization problem using alternating minimization between a and $(\boldsymbol{\rho}_c, \boldsymbol{\rho}_s)$, and implement it within the `GlobalBioIm` framework [20].

3) On the Computational Complexity: Whenever the considered optimization algorithm to solve the $\boldsymbol{\rho}$ (resp., $\boldsymbol{\rho}_c$ and $\boldsymbol{\rho}_s$) subproblem, the main computational burden lies in the application of the forward operator. For the full model reconstruction, applying $\mathbf{H}$ has a complexity of the order of $O(4M \log M + 2M)$ (using FFTs). In contrast for the joint deconvolution approach the complexity of applying either $\mathbf{H}_c$ or $\mathbf{H}_s$ is $O(2M \log M)$. While we can say that the overall rough complexity of the two proposed approaches is $O(M \log M)$, their practical speed will strongly rely on the number of iterations required to converge, which depends on the setting and cannot be generalized. In our numerical experiments (Section V), we observed that the joint deconvolution approach is generally 1.5 to 2 times slower.

V. Numerical Experiments

A. Numerical Evaluation of the Resolution Limit

We consider the phantom of Fig. 3.a and simulate measurement images $\mathbf{y}$ for $\lambda = 500\text{nm}$, $\theta = 60°$, $\phi = 0°$ and various image plane distances $z_{\text{im}} \in \{\frac{\lambda}{4}, \frac{\lambda}{2}, \lambda, 2\lambda, 3\lambda\}$. We fix the image and object grids to be the same, $\mathcal{G}_{\text{im}} = \mathcal{G}_{\text{ob}}$ (i.e., $\mathbf{t} = \mathbf{0}$), and to cover a $45\mu\text{m} \times 45\mu\text{m}$ region with a sampling step of 88nm. An example of noiseless data for $z_{\text{im}} = \frac{\lambda}{2}$ is presented on Fig 3.d. These data are then corrupted with a Gaussian noise $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma_\eta \mathbf{I})$ for σ_η equal to 0.1% or 2.5%

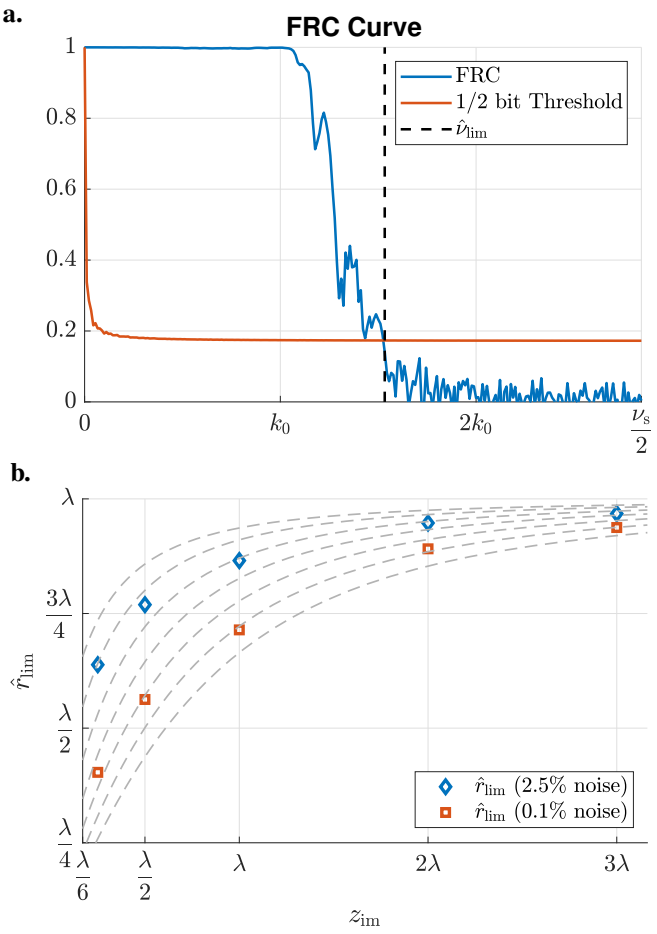

Fig. 5. Numerical assessment of the resolution limit. Parameters used to generate the measurements $\mathbf{y}$ are described in Section V-A. **a.** Example of FRC curve. **b.** Obtained 1/2 bit threshold resolution $\hat{r}_{\text{lim}}$ as a function of the image plane distance z_{im} and the level of noise. Dashed gray lines correspond to theoretical curves (19) for different level η_{eff}.

of the maximum value of the noiseless data corresponding to the case $z_{\text{im}} = \frac{\lambda}{4}$. The level of noise is thus the same for all considered distances z_{im} and the signal-to-noise ratio decreases as z_{im} increases.

Reconstruction is performed using the full forward model as described in Section IV-B1 with $\mathcal{C} = \mathbb{R}$ (no constraint) and a Tikhonov regularizer $R(\boldsymbol{\rho}) = \frac{1}{2}\|\boldsymbol{\rho}\|_2^2$. This corresponds to the simplest prior we can define in order to effectively assess the resolution achievable with the sole information contained in the data and compare it to the theoretical one derived in Section III-C. We let the regularization parameter be proportional to the noise level σ_η, that is $\mu \propto \sigma_\eta^2$. The proportionality constant being adjusted manually so as to maximize the performance. Note that this choice is driven by the Gaussian nature of the noise we added to the data. Finally, the optimization with respect to $\boldsymbol{\rho}$ is tackled through the second-order variable-metric limited-memory-bounded algorithm [22] belonging to the family of L-BFGS methods

To assess the resolution of the reconstructed image $\hat{\boldsymbol{\rho}}$, we rely on the *Fourier Ring Correlation* (FRC) [23], [24]. The FRC corresponds to a one-dimensional curve made of cross-correlations between the Fourier transforms of the recon-

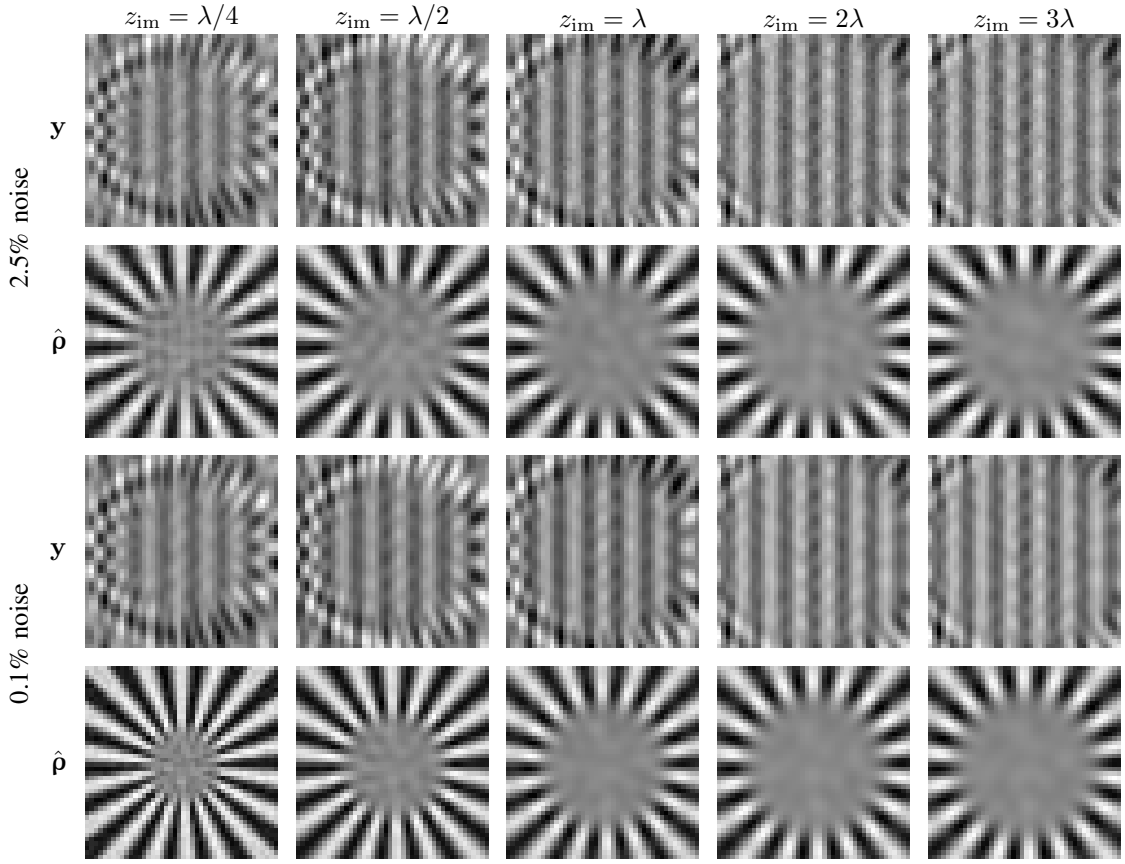

Fig. 6. Zoom (center of the star-like structure of the object) on both the measurements $\mathbf{y}$ and reconstructed object $\hat{\boldsymbol{\rho}}$ associated to each point in Fig 5.b.

structed and true images along concentric frequency bands (rings). An example is given in Fig 5.a where the x-axis represents the radius of the frequency ring and the y-axis the cross-correlation along the ring. The FRC curve is globally decreasing, which shows that low frequencies are better reconstructed than high frequencies. In particular, we can observe that the frequency content below k_0 (corresponding to blue disk in Fig 3.e) is very accurately reconstructed (FRC close to 1). The FRC start decreasing beyond k_0, that is for frequency content that lies, in the measurements, at locations where the Fourier transform of both h_c and h_s decay exponentially (see Proposition 1). Then, the standard practice (e.g., in structural biology [23] or molecule localization microscopy [24]) is to define the achieved resolution as $\hat{r}_{\rm lim} = 1/\hat{\nu}_{\rm lim}$ where $\hat{\nu}_{\rm lim}$ corresponds to the frequency where the FRC curve reaches a certain threshold. Several thresholds exists [23] and we consider the *half-bit threshold* that defines the point where there is 1/2 bit of information per pixel (cf. Fig 5.a). It is worth noting that other commonly used thresholds (e.g., 1/7 or Sobel) would lead to the same conclusions.

The obtained resolutions limits are reported in Fig. 5.b for the considered distances $z_{\rm im}$ and noise levels. Theoretical curves (19) for various levels $\eta_{\rm eff}$ are also displayed on the same graph. We observe that numerically calculated resolutions closely follow the shape of theoretical curves. As expected, the resolution decreases as $z_{\rm im}$ increases and as the level of noise increases. These behaviors are also clearly visible on the reconstructed $\hat{\boldsymbol{\rho}}$ (zooms) displayed in Fig 6.

Remark 3. *As mentioned earlier, we considered a simple Tikhonov regularizer without non-negativity constraint. This allowed us to make relevant comparisons with the theoretical resolution in* (19)*, which only rely on the information contained in the data without considering any prior. Yet, it is worth mentioning that, in practice, the consideration of more sophisticated priors as well as a non-negativity constraint can push the resolution beyond the reported ones, provided these priors/constraints are relevant for the imaged object.*

Remark 4. *In practice, noise arises from multiple sources—such as photon counting process, detector noise, and thermal drift—leading to a distribution that is more complex than a simple Gaussian. Such noises can be accounted for by replacing the least-squares data fidelity term in J or $\tilde{J}$ with the negative log-likelihood corresponding to the actual noise distribution. Approximating the negative log-likelihood with a weighted least-squares term can also serve as a practical compromise between computational efficiency and accurate noise modeling [25], [26].*

B. Robustness to Model Mismatches

We now analyze the robustness of both the full model (FM) reconstruction approach of Section IV-B1 and the joint deconvolution (JD) approach of Section IV-B2 with respect to

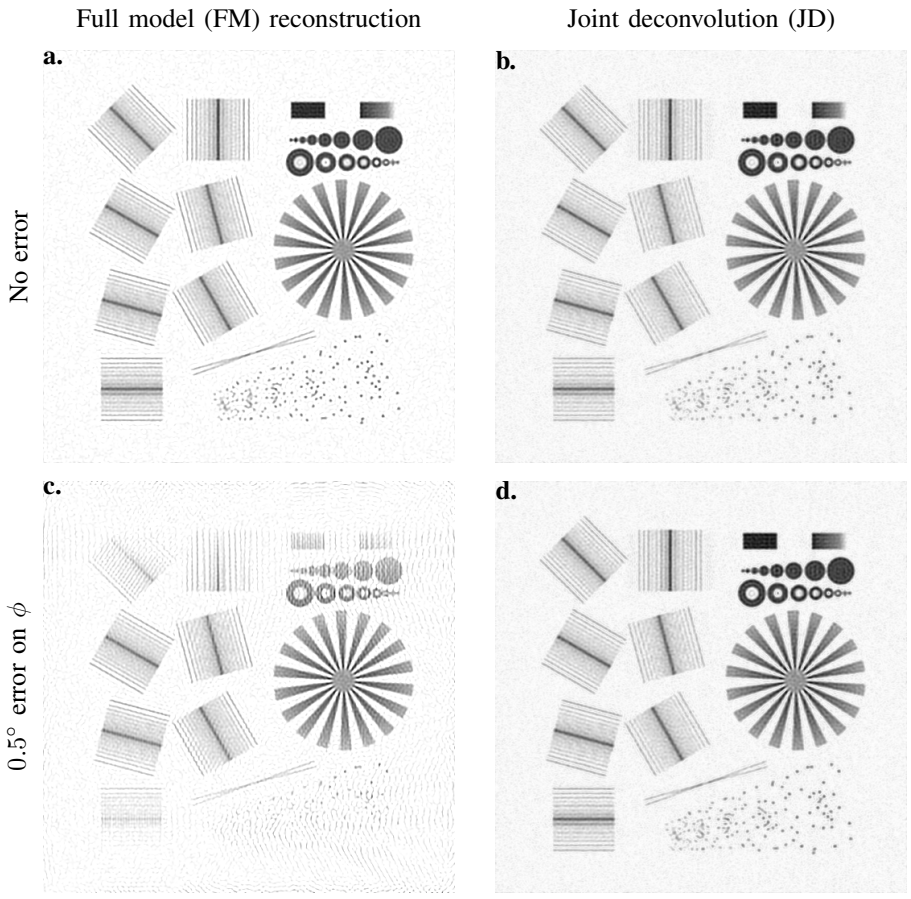

Fig. 7. Comparison of the full model (FM) and joint deconvolution (JD) reconstructions when model parameters are perfectly known (**a-b)** and when an error of 0.5° is made on ϕ (**c-d**). Simulation parameters are as in Section V-A with fixed image plane $z_{\mathrm{im}} = \frac{\lambda}{2}$ and noise level to 2.5%. Computational times (Intel® Xeon(R) w5-3423 × 24) are respectively from **a** to **d**: 19s, 28s, 18s, 39s.

errors on system parameters θ, ϕ, and z_{im}. In practice these parameters are difficult to calibrate and we cannot expect to know their exact value for a given experiment. It is thus worth analyzing the impact of errors on these parameters in terms of reconstruction quality.

To that end, we consider the same setting as in Section V-A and fix the image plane distance to $z_{\mathrm{im}} = \frac{\lambda}{2}$ and the noise level to 2.5%. Reconstructions for both FM and JD methods consider a Tikhonov regularizer. The non-negativity constraint $\mathcal{C} = \mathbb{R}_{\geq 0}$ is enforced for FM while being implicit by construction for JD.

First, we can observe on Figs 7.a and 7.b that, when model parameters are perfectly known, the JD approach provides a reconstruction with quality comparable to that of FM reconstruction. However, it is significantly more robust to model mismatches. Indeed, we can see on Figs 7.c and 7.d that, with an error of 0.5° of ϕ, the joint deconvolution achieves the same performance while the full model reconstruction is significantly degraded with strong artifacts.

The superior robustness to model mismaches of the joint deconvolution approach is also clearly visible on the graphs of Fig 8. Regardless of whether the errors are on θ, ϕ, or z_{im}, we observe that, beyond a certain error, the estimated resolution limit explodes while the PSNR drops sharply for the full model reconstruction. In contrast, the performance of the joint deconvolution approach remains stable for larger errors. Specifically, the performance of the full model reconstruction (resp. joint deconvolution) drops for parameters errors larger than 0.8° (resp., 10°) for θ, 0.5° (resp., 3°) for ϕ and 0.075λ (resp., 0.15λ) for z_{im}.

VI. Discussion and Conclusion

In this work, we conducted a Fourier analysis of ISOM, offering new perspectives on its imaging capabilities. Specifically, our analysis showed that while, in an ideal noiseless setting, the recorded image carries unlimited frequency bands, high-frequency information undergoes strong exponential attenuation. This observation led us to introduce the notion of *effective* resolution limit, balancing frequency attenuation against noise level. Numerical experiments confirmed that measured resolutions (evaluated via FRC) closely align with our theoretical predictions.

As expected, the resolution decreases when either the probing distance z_{im} or the noise level increases. Hence, for a fixed z_{im} in the intermediate-field region, the achievable resolution depends on our ability to mitigate measurement noise or effectively account for it in the reconstruction process. In this study, we used the simplest Tikhonov regularizer to

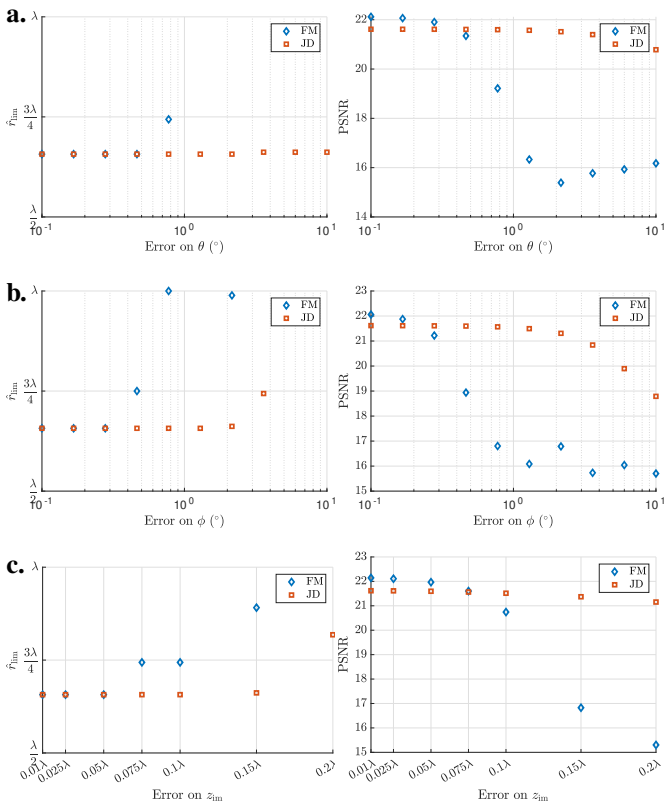

Fig. 8. Evolution of the resolution limit (estimated from FRC, see Section V-A) and PSNR with errors on θ (**a**), ϕ (**b**), and $z_{\rm im}$ (**c**). Simulation parameters are as in Section V-A with fixed image plane to $z_{\rm im} = \frac{\lambda}{2}$ and noise level to 2.5%. "Missing" points on $\hat{r}_{\rm lim}$ graphs are points that lie above λ.

validate our theoretical framework (cf. Remark 3). However, incorporating more sophisticated priors would undoubtedly enhance reconstruction performance. In particular, it is now well established that learned priors, using deep neural networks either in a Plug-and-Play [27] or an unrolling [28] framework, can significantly enhance the resolution of imaging inverse problems. Moreover, exploiting the effectiveness of deep-learning-based denoisers [29], [30] for data pre-processing could further improve performance. These are promising avenues for pushing the limits of ISOM.

From a practical standpoint, the proposed analysis offers valuable guidelines for selecting an appropriate sampling step during data acquisition. It also emphasizes that imaging distances suitable for achieving super-resolution (i.e., below $\lambda/2$) typically range between $\lambda/6$ and λ, depending on the noise level and our ability to efficiently handle it. The two proposed reconstruction methods contribute to expanding the currently limited set of reconstruction techniques available for ISOM, which typically depend on stringent assumptions regarding the model or sample properties. Last but not least, we showed the importance of estimating systems parameters accurately in order to ensure meaningful and reliable reconstructions.

Future research will include experimental validation to verify ISOM's practical resolution against our theoretical findings, as well as a comparative analysis with alternative intermediate-field techniques. Moreover, advancing acquisition and reconstruction methods for real-time nanoscale motion observation represents a promising direction for future research. Finally, the shared optical characteristics of ISOM and NSOM could enable the development of multimodal imaging techniques.

Acknowledgments

E.S. acknowledges the support from ANR Micro-Blind (grant ANR-21-CE48-0008) and ANR LabEx CIMI (grant ANR-11-LABX-0040) in the framework of the 'Programme des Investissements d'Avenir'. W.B. acknowledges the support from ANR Labex NanoX (grant no. ANR-17-EURE0009) in the framework of the 'Programme des Investissements d'Avenir'. The authors thank Sylvain Goimard, Anthony Torosyan, and Gargee Phukon for a early studies of this problem during internships.

Appendix A
Details on the Image Formation Model (4)

Let first develop the expression (3), denoting $E_i(\mathbf{r}) = E_i(\mathbf{r}, z_{\rm im}, t)$ and $E_s(\mathbf{r} - \mathbf{r}_p) = E_s(\mathbf{r} - \mathbf{r}_p, z_{\rm im}, t)$,

$$y(\mathbf{r}) = \left\langle E_i^2(\mathbf{r}) \right\rangle_t + \sum_{p=1}^{P} \left\langle E_i(\mathbf{r}) E_s(\mathbf{r} - \mathbf{r}_p) \right\rangle_t + \sum_{p,p'=1}^{P,P} \left\langle E_s(\mathbf{r} - \mathbf{r}_{p'}) E_s(\mathbf{r} - \mathbf{r}_p) \right\rangle_t . \quad (24)$$

Using the expressions of E_i and E_s in (1)–(2) we get that the first term is given by (with $T = 2\pi/\omega$)

$$\left\langle E_i^2(\mathbf{r}) \right\rangle_t = \frac{A_i^2}{T} \int_0^T \cos^2(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im} - \omega t)\, \mathrm{d}t = \frac{A_i^2}{2}.$$

For the second term, we need to solve

$$\left\langle E_i(\mathbf{r}) E_s(\mathbf{r} - \mathbf{r}_p) \right\rangle_t = \frac{A_i A_p}{T\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2}} \times \int_0^T \cos(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im} - \omega t) \cos\left(k_0 \sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2} - \omega t \right) \mathrm{d}t$$

Then, we can separate the above integral as the sum of the following two integrals

$$\int_0^T \cos\left(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im} + k_0 \sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2} - 2\omega t \right) \mathrm{d}t = 0$$

$$\int_0^T \cos\left(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im} - k_0 \sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2} \right) \mathrm{d}t = T \cos\left(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im} - k_0 \sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2} \right).$$

We thus get that

$$\left\langle E_i(\mathbf{r}) E_s(\mathbf{r} - \mathbf{r}_p) \right\rangle_t = \frac{A_i A_p \cos\left(\mathbf{k}_r^{\rm T}\mathbf{r} + k_z z_{\rm im} - k_0 \sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2} \right)}{\sqrt{\|\mathbf{r} - \mathbf{r}_p\|^2 + z_{\rm im}^2}}.$$

Similarly, we get

$$\langle E_s(\mathbf{r}-\mathbf{r}_p)E_s(\mathbf{r}-\mathbf{r}_{p'})\rangle_t = \frac{A_pA_{p'}\cos\left(k_0\left(\sqrt{\|\mathbf{r}-\mathbf{r}_p\|^2+z_{\rm im}^2}-\sqrt{\|\mathbf{r}-\mathbf{r}_{p'}\|^2+z_{\rm im}^2}\right)\right)}{\sqrt{\|\mathbf{r}-\mathbf{r}_p\|^2+z_{\rm im}^2}\sqrt{\|\mathbf{r}-\mathbf{r}_{p'}\|^2+z_{\rm im}^2}}.$$

The amplitude A_i of the incident beam is significantly larger than that of the scattered waves A_p. In particular, we assume that $A_i \gg \frac{P}{z_{\rm im}}\max_p A_p$ which, combined with the fact that $\sqrt{\|\mathbf{r}-\mathbf{r}_p\|^2+z_{\rm im}^2} \geq z_{\rm im}$, implies

$$\sum_{p=1}^{P}\frac{A_iA_p}{\sqrt{\|\mathbf{r}-\mathbf{r}_p\|^2+z_{\rm im}^2}} \gg \sum_{p=1}^{P}\frac{A_pP\max_{p'}A_{p'}}{z_{\rm im}\sqrt{\|\mathbf{r}-\mathbf{r}_p\|^2+z_{\rm im}^2}} \geq \sum_{p,p'=1}^{P,P}\frac{A_pA_{p'}}{\sqrt{\|\mathbf{r}-\mathbf{r}_p\|^2+z_{\rm im}^2}\sqrt{\|\mathbf{r}-\mathbf{r}_{p'}\|^2+z_{\rm im}^2}}.$$

As such, the third term in (24) can be neglected in front of the second one and we get (4).

Appendix B
Proof of Proposition 1

Our goal here is to get an expression of the Fourier transforms of h_c in (8) and h_s in (9). To that end, we first decompose these two functions as follows:

$$h_c(\mathbf{r}) = f_c(\|\mathbf{r}\|)\cos(\mathbf{k}_r^{\rm T}\mathbf{r}) + f_s(\|\mathbf{r}\|)\sin(\mathbf{k}_r^{\rm T}\mathbf{r})$$
$$h_s(\mathbf{r}) = f_c(\|\mathbf{r}\|)\sin(\mathbf{k}_r^{\rm T}\mathbf{r}) - f_s(\|\mathbf{r}\|)\cos(\mathbf{k}_r^{\rm T}\mathbf{r})$$

where

$$f_c(r) = \cos\left(k_0\sqrt{r^2+z_{\rm im}^2}\right)/\left(\sqrt{r^2+z_{\rm im}^2}\right), \quad (25)$$

$$f_s(r) = \sin\left(k_0\sqrt{r^2+z_{\rm im}^2}\right)/\left(\sqrt{r^2+z_{\rm im}^2}\right). \quad (26)$$

Given these decompositions, we will focus on deriving the Fourier transforms of the radial functions $f_c(\|\mathbf{r}\|)$ and $f_s(\|\mathbf{r}\|)$ and then use the Fourier convolution theorem to deduce the Fourier transforms of h_c and h_s.

The Fourier transform of a radial function is also a radial function defined through the Hankel transform. Let $r = \|\mathbf{r}\|$ and $\nu = \|\boldsymbol{\nu}\|$ with $\boldsymbol{\nu} \in \mathbb{R}^2$ the frequency vector. Then we have from [31, Sec. 6.737, Eq. 6]

$$\widehat{f}_c(\nu) = 2\pi\int_0^\infty f_c(r)J_0(r\nu)r\,{\rm d}r \quad \text{(Hankel transform)}$$
$$= \begin{cases} -\frac{\pi\sqrt{2\pi z_{\rm im}}}{(k_0^2-\nu^2)^{\frac{1}{4}}}Y_{-\frac{1}{2}}\left(z_{\rm im}\sqrt{k_0^2-\nu^2}\right) & \text{if } \nu < k_0 \\ \frac{2\sqrt{2\pi z_{\rm im}}}{(\nu^2-k_0^2)^{\frac{1}{4}}}K_{\frac{1}{2}}\left(z_{\rm im}\sqrt{\nu^2-k_0^2}\right) & \text{if } \nu > k_0 \end{cases}$$

where J_α and Y_α are the Bessel functions of first and second kind, respectively, and K_α the modified Bessel function of second kind. Similarly we get from [31, Sec. 6.737, Eq. 5]

$$\widehat{f}_s(\nu) = 2\pi\int_0^\infty f_s(r)J_0(r\nu)r\,{\rm d}r$$
$$= \begin{cases} \frac{\pi\sqrt{2\pi z_{\rm im}}}{(k_0^2-\nu^2)^{\frac{1}{4}}}J_{-\frac{1}{2}}\left(z_{\rm im}\sqrt{k_0^2-\nu^2}\right) & \text{if } \nu < k_0 \\ 0 & \text{if } \nu > k_0 \end{cases}$$

Denoting $\nu_- = \|\boldsymbol{\nu}-\mathbf{k}_r\|$ and $\nu_+ = \|\boldsymbol{\nu}+\mathbf{k}_r\|$, we get from these expression and the Fourier convolution theorem,

$$\widehat{h}_c(\boldsymbol{\nu}) = \frac{1}{2}\left(\widehat{f}_c(\nu_-)+\widehat{f}_c(\nu_+)\right) - \frac{\rm i}{2}\left(\widehat{f}_s(\nu_-)-\widehat{f}_s(\nu_+)\right)$$
$$= \begin{cases} -\frac{1}{2}\frac{\pi\sqrt{2\pi z_{\rm im}}}{(k_0^2-\nu_-^2)^{\frac{1}{4}}}\left(Y_{-\frac{1}{2}}\left(z_{\rm im}\sqrt{k_0^2-\nu_-^2}\right)\right. & \\ \qquad\left. +{\rm i}J_{-\frac{1}{2}}\left(z_{\rm im}\sqrt{k_0^2-\nu_-^2}\right)\right) & \text{if } \nu_- < k_0 \\ \frac{1}{2}\frac{2\sqrt{2\pi z_{\rm im}}}{(\nu_-^2-k_0^2)^{\frac{1}{4}}}K_{\frac{1}{2}}\left(z_{\rm im}\sqrt{\nu_-^2-k_0^2}\right) & \text{if } \nu_- > k_0 \end{cases}$$
$$+ \begin{cases} -\frac{1}{2}\frac{\pi\sqrt{2\pi z_{\rm im}}}{(k_0^2-\nu_+^2)^{\frac{1}{4}}}\left(Y_{-\frac{1}{2}}\left(z_{\rm im}\sqrt{k_0^2-\nu_+^2}\right)\right. & \\ \qquad\left. -{\rm i}J_{-\frac{1}{2}}\left(z_{\rm im}\sqrt{k_0^2-\nu_+^2}\right)\right) & \text{if } \nu_+ < k_0 \\ \frac{1}{2}\frac{2\sqrt{2\pi z_{\rm im}}}{(\nu_+^2-k_0^2)^{\frac{1}{4}}}K_{\frac{1}{2}}\left(z_{\rm im}\sqrt{\nu_+^2-k_0^2}\right) & \text{if } \nu_+ > k_0 \end{cases}$$

Using the fact that the Hankel functions of first and second kind satisfiy $H_\alpha^{(q)} = J_\alpha + (-1)^{q+1}{\rm i}Y_\alpha$ for $q \in \{1,2\}$, and that $K_{\frac{1}{2}}(z) = \sqrt{\pi/2}\exp(-z)/\sqrt{z}$, we can further simplify the above expressions as

$$\widehat{h}_c(\boldsymbol{\nu}) = \begin{cases} -\frac{\pi\sqrt{\pi z_{\rm im}}}{\sqrt{2}(k_0^2-\nu_-^2)^{\frac{1}{4}}}{\rm i}H_{-\frac{1}{2}}^{(2)}\left(z_{\rm im}\sqrt{k_0^2-\nu_-^2}\right) & \text{if } \nu_- < k_0 \\ \frac{\pi}{\sqrt{\nu_-^2-k_0^2}}\exp\left(-z_{\rm im}\sqrt{\nu_-^2-k_0^2}\right) & \text{if } \nu_- > k_0 \end{cases}$$
$$+ \begin{cases} \frac{\pi\sqrt{\pi z_{\rm im}}}{\sqrt{2}(k_0^2-\nu_+^2)^{\frac{1}{4}}}{\rm i}H_{-\frac{1}{2}}^{(1)}\left(z_{\rm im}\sqrt{k_0^2-\nu_+^2}\right) & \text{if } \nu_+ < k_0 \\ \frac{\pi}{\sqrt{\nu_+^2-k_0^2}}\exp\left(-z_{\rm im}\sqrt{\nu_+^2-k_0^2}\right) & \text{if } \nu_+ > k_0 \end{cases}$$

which corresponds to (11) with g_1 and g_2 given by (13). Doing the same for $\widehat{h}_s$ we recover (12).

Appendix C
Proof of Proposition 2

From the convolution theorem, we have $\mathcal{F}\{h_c * (w_c \odot \rho)\} = \widehat{h}_c \odot (\widehat{w}_c * \widehat{\rho})$. Combining this expression with $\widehat{w}_c = \frac{1}{2}(\mathrm{e}^{{\rm i}k_z z_{\rm im}}\delta(\cdot-\mathbf{k}_r) + \mathrm{e}^{-{\rm i}k_z z_{\rm im}}\delta(\cdot+\mathbf{k}_r))$ and (11), we get,

$$\mathcal{F}\{h_c * (w_c \odot \rho)\} = \frac{1}{2}\left(g_1(\|\cdot+\mathbf{k}_r\|) + g_2(\|\cdot-\mathbf{k}_r\|)\right) \odot \left(\mathrm{e}^{{\rm i}k_z z_{\rm im}}\widehat{\rho}(\cdot-\mathbf{k}_r) + \mathrm{e}^{-{\rm i}k_z z_{\rm im}}\widehat{\rho}(\cdot+\mathbf{k}_r)\right)$$

Similarly, with $\widehat{w}_s = \frac{1}{2{\rm i}}(\mathrm{e}^{{\rm i}k_z z_{\rm im}}\delta(\cdot-\mathbf{k}_r) - \mathrm{e}^{-{\rm i}k_z z_{\rm im}}\delta(\cdot+\mathbf{k}_r))$ and (12), we get

$$\mathcal{F}\{h_s * (w_s \odot \rho)\} = \frac{1}{2}\left(g_1(\|\cdot+\mathbf{k}_r\|) - g_2(\|\cdot-\mathbf{k}_r\|)\right) \odot \left(\mathrm{e}^{{\rm i}k_z z_{\rm im}}\widehat{\rho}(\cdot-\mathbf{k}_r) - \mathrm{e}^{-{\rm i}k_z z_{\rm im}}\widehat{\rho}(\cdot+\mathbf{k}_r)\right)$$

Combining these two expressions, together with the linearity of the Fourier transform and the fact the the Fourier transform of a constant function is the delta Dirac function, we obtain

$$\widehat{y} = \frac{A_i^2}{2}\delta + A_i\Big(\mathrm{e}^{{\rm i}k_z z_{\rm im}}\widehat{\rho}(\cdot-\mathbf{k}_r)\odot g_2(\|\cdot-\mathbf{k}_r\|) + \mathrm{e}^{-{\rm i}k_z z_{\rm im}}\widehat{\rho}(\cdot+\mathbf{k}_r)\odot g_1(\|\cdot+\mathbf{k}_r\|)\Big), \quad (27)$$

which can be further compacted as in (17).

REFERENCES

[1] L. Schermelleh, R. Heintzmann, and H. Leonhardt, "A guide to super-resolution fluorescence microscopy," *Journal of Cell Biology*, vol. 190, no. 2, pp. 165–175, 2010.

[2] L. Schermelleh, A. Ferrand, T. Huser, C. Eggeling, M. Sauer, O. Biehlmaier, and G. P. Drummen, "Super-resolution microscopy demystified," *Nature Cell Biology*, vol. 21, no. 1, pp. 72–84, 2019.

[3] K. Prakash, B. Diederich, R. Heintzmann, and L. Schermelleh, "Super-resolution microscopy: a brief history and new avenues," *Philosophical Transactions of the Royal Society A*, vol. 380, no. 2220, p. 20210110, 2022.

[4] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, "Development of a 500 Å spatial resolution light microscope: I. light is efficiently transmitted through λ/16 diameter apertures," *Ultramicroscopy*, vol. 13, no. 3, pp. 227–231, 1984.

[5] D. W. Pohl, W. Denk, and M. Lanz, "Optical stethoscopy: Image recording with resolution λ/20," *Applied Physics Letters*, vol. 44, no. 7, pp. 651–653, 1984.

[6] R. Reddick, R. Warmack, and T. Ferrell, "New form of scanning optical microscopy," *Physical Review B*, vol. 39, no. 1, p. 767, 1989.

[7] E. Betzig and J. K. Trautman, "Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit," *Science*, vol. 257, no. 5067, pp. 189–195, 1992.

[8] H. Heinzelmann and D. Pohl, "Scanning near-field optical microscopy," *Applied Physics A*, vol. 59, pp. 89–101, 1994.

[9] A. Lereu, A. Passian, and P. Dumas, "Near field optical microscopy: a brief review," *International Journal of Nanotechnology*, vol. 9, no. 3-7, pp. 488–501, 2012.

[10] B. Hecht, B. Sick, U. P. Wild, V. Deckert, R. Zenobi, O. J. Martin, and D. W. Pohl, "Scanning near-field optical microscopy with aperture probes: Fundamentals and applications," *The Journal of Chemical Physics*, vol. 112, no. 18, pp. 7761–7774, 2000.

[11] L. Gross, "Recent advances in submolecular resolution with scanning probe microscopy," *Nature chemistry*, vol. 3, no. 4, pp. 273–278, 2011.

[12] N. Umeda, Y. Hayashi, K. Nagai, and A. Takayanagi, "Scanning wiener-fringe microscope with an optical fiber tip," *Applied Optics*, vol. 31, no. 22, pp. 4515–4518, 1992.

[13] W. Bacsa and A. Kulik, "Interference scanning optical probe microscopy," *Applied Physics Letters*, vol. 70, no. 26, pp. 3507–3509, 1997.

[14] W. Bacsa, "Optical interference near surfaces and its application in subwavelength microscopy," in *Advances in Imaging and Electron Physics*. Elsevier, 2010, vol. 163, pp. 1–32.

[15] W. Bacsa, R. Bacsa, and T. Myers, *Optics near Surfaces and at the Nanometer Scale*. New York: Springer, 2020.

[16] R. Heintzmann and C. G. Cremer, "Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating," in *Optical biopsies and microscopic techniques III*, vol. 3568. SPIE, 1999, pp. 185–196.

[17] M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *Journal of Microscopy*, vol. 198, no. 2, pp. 82–87, 2000.

[18] K. Wicker, O. Mandula, G. Best, R. Fiolka, and R. Heintzmann, "Phase optimisation for structured illumination microscopy," *Optics Express*, vol. 21, no. 2, pp. 2032–2049, 2013.

[19] T. Myers, H. Ribera, and W. Bacsa, "Optical diffraction from isolated nanoparticles," *arXiv preprint arXiv:1902.08680*, 2019.

[20] E. Soubies, F. Soulez, M. T. McCann, T. an Pham, L. Donati, T. Debarre, D. Sage, and M. Unser, "Pocket guide to solve inverse problems with globalbioim," *Inverse Problems*, vol. 35, no. 10, p. 104006, 2019.

[21] E. Soubies, A. Nogueron, F. Pelletier, T. Mangeat, C. Leterrier, M. Unser, and D. Sage, "Surpassing light inhomogeneities in structured-illumination microscopy with flexsim," *Journal of Microscopy*, vol. 296, no. 1, pp. 94–106, 2024.

[22] E. Thiébaut, "Optimization issues in blind deconvolution algorithms," in *Astronomical Data Analysis II*, vol. 4847. SPIE, 2002, pp. 174–183.

[23] M. Van Heel and M. Schatz, "Fourier shell correlation threshold criteria," *Journal of Structural Biology*, vol. 151, no. 3, pp. 250–262, 2005.

[24] N. Banterle, K. H. Bui, E. A. Lemke, and M. Beck, "Fourier ring correlation as a resolution criterion for super-resolution microscopy," *Journal of Structural Biology*, vol. 183, no. 3, pp. 363–367, 2013.

[25] A. Sawatzky, C. Brune, J. Müller, and M. Burger, "Total variation processing of images with poisson statistics," in *International Conference on Computer Analysis of Images and Patterns*. Springer, 2009, pp. 533–540.

[26] A. Singh, E. Soubies, and C. Chaux, "Learning weighted least squares data term for poisson image deconvolution," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025, pp. 1–5.

[27] U. S. Kamilov, C. A. Bouman, G. T. Buzzard, and B. Wohlberg, "Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications," *IEEE Signal Processing Magazine*, vol. 40, no. 1, pp. 85–97, 2023.

[28] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.

[29] M. Elad, B. Kawar, and G. Vaksman, "Image denoising: The deep learning revolution and beyond—a survey paper," *SIAM Journal on Imaging Sciences*, vol. 16, no. 3, pp. 1594–1654, 2023.

[30] L. Qu, S. Zhao, Y. Huang, X. Ye, K. Wang, Y. Liu, X. Liu, H. Mao, G. Hu, W. Chen *et al.*, "Self-inspired learning for denoising live-cell super-resolution microscopy," *Nature Methods*, vol. 21, no. 10, pp. 1895–1908, 2024.

[31] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*. San Diego: Academic press, 2014.