On the Superresolution Capacity of Imagers Using Unknown Speckle Illuminations

Jérôme Idier, *Member, IEEE*, Simon Labouesse, Marc Allain, Penghuan Liu, Sébastien Bourguignon, and Anne Sentenac, *Member, IEEE*

***Abstract*—Speckle-based imaging consists of forming a super-resolved reconstruction of an unknown sample from low-resolution images obtained under random inhomogeneous illuminations (speckles). In a blind context, where the illuminations are unknown, we study the *intrinsic* capacity of speckle-based imagers to recover spatial frequencies outside the frequency support of the data, with minimal assumptions about the sample. We demonstrate that, under physically realistic conditions, the covariance of the data has a super-resolution power corresponding to the squared magnitude of the imager point spread function. This theoretical result is important for many practical imaging systems such as acoustic and electromagnetic tomographs, fluorescence and photoacoustic microscopes, or synthetic aperture radar imaging. A numerical validation is presented in the case of fluorescence microscopy.**

***Index Terms*—Multi-illumination imaging, high resolution, cut-off frequency, second-order statistics, optical microscopy, photoacoustic imaging, synthetic aperture radar.**

I. Introduction

In MOST active wave imaging systems, the recorded data z can be modeled as the convolution of a point spread function (PSF) h with the product of the sample ρ with an illumination E, plus some additive noise ε:

$$z = h \otimes (\rho E) + \varepsilon \tag{1}$$

where $\otimes$ stands for the convolution operator, either in two or three spatial dimensions. This simple model applies to imaging configurations as diverse as microwave scanners or anechoic chambers [1], radar remote sensing [2] or fluorescence microscopy [3].

The shape of the point spread function h depends on the imager geometry, *e.g.*, the numerical aperture (NA) of the microscope objective, or the size of the antenna array in radar imaging. It accounts for the wave propagation from the sample to the detector. In most configurations, free-space propagation prevents the wavefield high frequencies from reaching the detector. As a result, h has necessarily a bounded Fourier support $\mathcal{D}_{\mathrm{PSF}}$. For instance, in a microwave scanner, $\mathcal{D}_{\mathrm{PSF}}$ is a hollow sphere of radius $1/\lambda$ (where λ is the illumination wavelength) when the field scattered by the sample is recorded under all possible directions, or a cap of sphere when the observation is performed only over a small solid angle [4]. Similarly, in two (or three) dimensional fluorescence microscopy, $\mathcal{D}_{\mathrm{PSF}}$ is a disk (or a solid torus) of radius $2\mathrm{NA}/\lambda$ [3].

When the illumination is homogeneous throughout the target, solely the sample frequency components in $\mathcal{D}_{\mathrm{PSF}}$ can be restored from the data by linear methods, which limits fundamentally the image resolution. To improve the latter, synthetic imaging using multiple illuminations has been developed. Its main principle is to use several known inhomogeneous illuminations $E_m, m = 1, \ldots, M$, to probe the sample. The frequency mixing of E_m with ρ causes a down-modulation of the sample high spatial frequencies into the frequency-support $\mathcal{D}_{\mathrm{PSF}}$. Using appropriate data processing, sample frequencies beyond $\mathcal{D}_{\mathrm{PSF}}$ can be recovered, yielding a much better resolution. This idea is at the core of many imaging configurations such as Synthetic Aperture Radar (SAR) [2], diffraction tomography [5], and Structured Illumination fluorescence Microscopy (SIM) [6], [7], among others.

In all these imaging modalities, the standard numerical or analog process that forms the super-resolved image from the stack of low resolution data assumes the precise knowledge, and thus the tight control, of the different illuminations E_m. The super-resolution capacity of the process is then both theoretically and practically demonstrated. However, the full control of the illumination patterns is a major constraint for the experimental implementation and in some cases proves impossible. The case of thick samples imaged with three-dimensional SIM is a classical example since samples are likely to introduce distorsions on the excitation pattern [8], [9]. Hence, some groups have developed reconstruction algorithms able to handle some uncertainty about the illuminations [8]–[11]. Following a less conventional option, others have advocated using of totally uncontrolled illuminations of speckle type [12]. This recent blind approach could dramatically simplify the experimentation by further relaxing the constraints on controlling the illumination patterns. Examples of implementations can be found in optical

Manuscript received December 16, 2016; revised May 12, 2017 and October 25, 2017; accepted November 4, 2017. Date of publication November 8, 2017; date of current version February 8, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jeffrey A. Fessler. This work was supported in parts by the GdR 720 ISIS, the Agence Nationale de la Recherche under Grant ANR-12-BS03-0006, and the ITMO Cancer of the "Plan Cancer 2014-2019." *(Corresponding author: Marc Allain)*

J. Idier, P. Liu, and S. Bourguignon are with the CNRS and Ecole Centrale Nantes, Laboratoire des Sciences du Numérique de Nantes (LS2N, CNRS UMR 6004), F-44321 Nantes, France (e-mail: jerome.idier@ls2n.fr; penghuan.liu@irccyn.ec-nantes.fr; sebastien.bourguignon@ls2n.fr).

S. Labouesse, M. Allain, and A. Sentenac are with the Institut Fresnel, Aix Marseille University, CNRS, 13013 Marseille, France (e-mail: simon.labouesse@fresnel.fr; marc.allain@fresnel.fr; anne.sentenac@fresnel.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2017.2771729

microscopy [13]–[17] and photoacoustic imaging [18], [19]. The proposed inversion schemes take advantage of the nonnegativity of the sample ρ, and on statistical information on ρ and/or the illuminations E_m. In particular, some of them introduce sparsity information on ρ [14], or on the products ρE_m [17]. Generally speaking, the stack of low resolution speckle data yielded reconstructed images with significantly better resolution than that provided by a standard imager using homogeneous illumination. However, many questions remain unanswered about the theoretical resolution that one can expect from such a system, in particular with respect to the speckle statistics.

To the best of our knowledge, this paper provides the first comprehensive mathematical understanding of the super-resolution (SR) capacity of synthetic imaging using speckle illuminations in a blind way. Our analysis is very general and basically holds when the data can be modeled by Eq. (1) with $\rho \geq 0$. Pivotal SR results are for instance provided for two popular microscopy modalities, namely optical *fluorescence microscopy* and optical *coherent imaging*.

Fluorescence microscopy is an incoherent imaging modality, for which the quantities involved in (1) have the following physical interpretation: ρ is the fluorescence density distribution, *i.e.*, a real-valued, nonnegative, quantity; the incoherent PSF h is real-valued, nonnegative, and the support of its Fourier transform is a disk or a torus in two and three dimensions, respectively [3]; the illumination E is a speckle intensity pattern produced by a coherent light beam, *i.e.*, a random real-valued, nonnegative, quantity (see for instance [13]); and z is an intensity measurement plagued by real-valued instrumental noise ε. In addition, in the low counting-rate regime, one may consider photon counting fluctuations in the observation model: in this case, the quantity $h \otimes (\rho E)$ in (1) is connected to the mean of the counting statistics. The case of intensity measurement plagued by both photon counting fluctuations and electronic noise is specifically addressed in Appendix A.

In coherent imaging, such as tomographic diffraction microscopy [4], we have the following correspondence for the model (1): ρ is the relative permittivity contrast distribution, a complex-valued function in general, although our mathematical analysis is restricted to real nonnegative ρ, *i.e.*, it is restricted to lossless dielectric objects; the coherent PSF h is complex-valued, and the support of its Fourier transform is a sphere cap [4]; the illumination E is a complex-valued random field, *e.g.*, a circular Gaussian random field if it arises from a (scalar) electric field stemming from a fully developed speckle produced by coherent light [20]; the recorded data z is the scattered electric field plagued[1] by circular complex-valued instrumental noise ε, such that z is also a circular complex-valued random field.

In the sequel, the term *super-resolution* is understood as the ability to recover spatial frequencies of the sample that cannot be obtained with either a constant illumination in incoherent imaging, or a single plane-wave with normal incidence in coherent imaging. Following standard results [23, ch. 6], we recall that the incoherent (intensity) PSF is obtained by squaring the magnitude of the coherent (complex electric field) PSF of the optical system. As a result, transmitted spatial frequencies with an incoherent illumination span twice the domain transmitted with a coherent illumination. However, incoherent illuminations do not provide the permittivity contrast, but its squared magnitude, which prevents any direct comparison in terms of spatial resolution between coherent and incoherent optical systems, see *e.g.*, [23, Sec. 6.5] for details. In contrast, the present work shows that a double spatial resolution can be obtained in *both* cases thanks to random illuminations.

We finally note that the model (1) also encompasses some other situations, namely *microwave imaging* [24] (the measured data are complex fields, ρ is the complex permittivity; the noise being mainly an electronic fluctuation, it can be assumed Gaussian for both the real and imaginary parts) and *photoacoustic imaging* [18], [19] (the measured data are real-valued B-mode images corrupted by real-valued Gaussian noise, and ρ represents the optical absorption). Hereafter, we consider the complex-valued setting, since the real setting can be deduced straightforwardly as a particular case where the imaginary parts of the relevant quantities vanish.

The article is organized as follows. The next section provides the mathematic assumptions required in our SR analysis. Section III establishes the expression of the first two moments of the data. In Section IV, the dependency between the latter expressions and the spatial frequency components of the sample ρ is further examined. Clear conclusions about the SR capacity of the system are obtained if the speckle illuminations are "sufficiently" correlated, in the sense that their spectral power density lies within the frequency support of the PSF. Such conclusions constitute the main contribution of this paper. The opposite case of uncorrelated speckles is also considered. Section V deals with the practical question of a computational scheme to reconstruct the unknown scene. A two-dimensional simulation of an optical fluorescence microscope using correlated speckle illuminations is provided, and it supports that the expected SR ratio can be obtained from the data empirical second-order statistics. Finally, Section VI discusses the practical consequences of the obtained results, and evokes possible extensions and remaining points to address.

II. Model and Assumptions for the SR Analysis

We consider M images $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_M)$ of the same sample that have been acquired using M different speckle illuminations. Each image $\boldsymbol{z}_m = (z_m(\boldsymbol{r}_1), \ldots, z_m(\boldsymbol{r}_N))$ is a set of N pixels, each of which being indexed by a spatial coordinate vector $\boldsymbol{r}_n$. In practice, vector $\boldsymbol{r}_n$ spans a finite d-dimensional rectangular grid $\mathcal{G}$, common to all images, d being equal to two or three. Without loss of generality, we consider the spatial sampling rate to be normalized to unity in each direction. By convention, we also consider that $\boldsymbol{z}_m$ are column vectors obtained by scanning the image grid $\mathcal{G}$ in an arbitrary order. Hence, for all $m \in \{1 \cdots M\}$ and $\boldsymbol{r} \in \mathcal{G}$, the observation model reads

$$z_m(\boldsymbol{r}) = y_m(\boldsymbol{r}) + \varepsilon_m(\boldsymbol{r}), \tag{2}$$

[1]This measurement z is usually obtained by the Fourier transform of measured real valued intensities in an off-axis interferometric mounting, see for instance [21], [22]. Moreover, if the counting rate is high enough, an additive fluctuation model over the real and the imaginary part is relevant.

with

$$y_m(\boldsymbol{r}) = \int h(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}')E_m(\boldsymbol{r}')\,d\boldsymbol{r}', \tag{3}$$

and where E_m and ε_m are random quantities: E_m is the mth random illumination and ε_m stands for electronic noise. Furthermore, the following assumptions will be made concerning the observation model (2):

i) The PSF h is both integrable and square-integrable (*i.e.*, $\int |h(\boldsymbol{x})|^p \; d\boldsymbol{x} < \infty$ for $p = 1, 2$). Moreover, its Fourier transform[2] $\widetilde{h}$ takes finite values and vanishes outside a bounded set $\mathcal{D}_{\mathrm{PSF}} = \{\boldsymbol{u} \,|\, \widetilde{h}(\boldsymbol{u}) \neq 0\}$. These assumptions are met when the measurement $\boldsymbol{z}_m$ is obtained in the far-field domain, which is the case of most imaging systems [4].

ii) The sample ρ is integrable and takes finite, nonnegative values over $\mathbb{R}^d$. Moreover, it approaches zero at large distance from the origin.

iii) The data grid $\mathcal{G}$ is sufficiently large to make the influence of finite data extent negligible. As a consequence, we will identify $\mathcal{G}$ with $\mathbb{Z}^d$ in the sequel. This is indeed a legitimate simplification given Assumptions (i) and (ii) since we can show that $h \otimes (\rho E)$ tends towards zero at large distance from the origin, provided that the illumination pattern E is bounded.

iv) $\mathcal{G}$ is fine enough to sample the PSF h with no discretization error. According to Parzen's multidimensional extension of Shannon theorem [25], such a condition is met as soon as $\mathcal{D}_{\mathrm{PSF}}$ belongs to the baseband $\mathcal{B} = [-1/2, 1/2]^d$.

v) The noise and illuminations are second-order stationary, mutually decorrelated random processes. This is a standard hypothesis which is verified for most imagers [12, Sec. 4.4]. Moreover, a direct extension would be possible to cases where the statistical mean of the illuminations is spatially varying. Without loss of generality, we will also assume that the noise is zero-mean.

vi) The first two moments of the illuminations and of the noise are known. This assumption is at the core of our approach. It is expected to be less difficult to satisfy than the knowledge of the illuminations patterns.

In this work, we restrict the analysis of the data by considering only second-order statistics, *i.e.*, the statistical mean and covariance of the data. More precisely, our aim is to determine the spatial frequency domain over which the sample spectrum can be identified from these statistics. Such a restriction is legitimate for several reasons.

On the one hand, the empirical mean and covariance are easily accessible statistical quantities, that can provide reliable second-order information from a practically acceptable number of illuminations.

On the other hand, the statistical mean and covariance are exhaustive statistics if the data are Gaussian, whether it is real-valued or complex circular. For instance, the latter assumption is suited to coherent imaging techniques such as tomographic diffraction microscopy. In other situations, such as optical fluorescence microscopy, the speckle illumination, and hence the data, are non-Gaussian. The statistical mean and covariance do not summarize all the information about the sample available in the measurements in such situations, but our results still provide a *lower bound* on the information retrievable from the complete data statistics.

With the goal of characterizing the SR potential of second-order methods, we wish to assign each component of the spatial Fourier transform $\widetilde{\rho}(\boldsymbol{u})$ of the imaged sample to one of the three classes:

1) *Non-identifiable* spectral components are those for which the second-order data statistics bring no information.
2) *Partially identifiable* components are those for which the second-order data statistics bring some information, but for which some ambiguity remains.
3) *Identifiable* components are those which are uniquely determined given the second-order data statistics.

Obviously, the support of each class in the Fourier domain may depend on the frequency support $\mathcal{D}_{\mathrm{PSF}}$ and on the covariance structure of the speckle illumination.

III. First and Second-Order Statistics of the Data

The statistical mean and covariance of the data are now derived. In what follows, $\mathcal{E}\{\cdot\}$ and * denote the statistical expectation operator and complex conjugation, respectively. According to assumptions (v) and (vi), let $\mathcal{E}\{E\} = E_0$ and $\gamma_E(\boldsymbol{r}) = \mathcal{E}\{E(\boldsymbol{x})E^*(\boldsymbol{x}-\boldsymbol{r})\} - |E_0|^2$ denote the mean and covariance function of the speckle, and let $\gamma_\varepsilon(\boldsymbol{r}) = \mathcal{E}\{\varepsilon(\boldsymbol{x})\varepsilon^*(\boldsymbol{x}-\boldsymbol{r})\}$ denote the covariance function of the noise.

A. First-Order Information Content

From the observation model (2)-(3) and from the assumption of centered noise, we deduce the statistical mean:

$$\mu_z(\boldsymbol{r}) = \mathcal{E}\{z_m(\boldsymbol{r})\} = E_0 \int h(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}')\,d\boldsymbol{r}', \quad \boldsymbol{r} \in \mathbb{Z}^d. \tag{4}$$

The continuous-space Fourier transform of μ_z reads

$$\widetilde{\mu}_z(\boldsymbol{u}) = E_0\,\widetilde{h}(\boldsymbol{u})\widetilde{\rho}(\boldsymbol{u}) \tag{5}$$

for all $\boldsymbol{u} \in \mathbb{R}^d$. Function $\widetilde{\mu}_z$ has a support limited to $\mathcal{D}_{\mathrm{PSF}}$, so, according to assumption (iv), the sampling of μ_z on $\mathbb{Z}^d$ is lossless. A straightforward deduction from expression (5) is that any spectral component of ρ belonging to the support $\mathcal{D}_{\mathrm{PSF}}$ is identifiable, provided that $E_0 \neq 0$. In particular, if $E_m(\boldsymbol{r})$ is a complex circular Gaussian process, then $E_0 = 0$ [12] and $\mu_z(\boldsymbol{r})$ brings no information about the unknown sample. This conclusion leads to the following property.

Property 1: The frequency component $\widetilde{\rho}(\boldsymbol{u})$ is identifiable from μ_z if and only if $\boldsymbol{u} \in \mathcal{D}_1$ with

$$\mathcal{D}_1 = \begin{cases} \mathcal{D}_{\mathrm{PSF}} & \text{if } E_0 \neq 0, \\ \varnothing & \text{otherwise.} \end{cases}$$

[2]Hereafter, the tilde sign $\widetilde{\ }$ denotes the d-dimensional continuous-space Fourier transform.

In any case, the first-order moment does not convey any information on the spectral components outside $\mathcal{D}_{\rm PSF}$, *i.e.*, it brings no SR capacity.

B. Second-Order Information Content

Now let us focus on the data covariance function that reads

$$\gamma_z(\boldsymbol{r},\boldsymbol{r}') = \mathcal{E}\left\{z_m(\boldsymbol{r}) z_m^*(\boldsymbol{r}')\right\} - \mu_z(\boldsymbol{r})\,\mu_z^*(\boldsymbol{r}')$$

with $\boldsymbol{r}$, $\boldsymbol{r}' \in \mathbb{Z}^d$. If (2) and (3) hold, then we immediately get

$$\gamma_z(\boldsymbol{r},\boldsymbol{r}') = \gamma_y(\boldsymbol{r},\boldsymbol{r}') + \gamma_\varepsilon(\boldsymbol{r}-\boldsymbol{r}') \tag{6}$$

with

$$\gamma_y(\boldsymbol{r},\boldsymbol{r}') = \iint \rho(\boldsymbol{x})\rho(\boldsymbol{x}')\,h(\boldsymbol{r}-\boldsymbol{x})h^*(\boldsymbol{r}'-\boldsymbol{x}')\,\gamma_E(\boldsymbol{x}-\boldsymbol{x}')\,d\boldsymbol{x}\,d\boldsymbol{x}'. \tag{7}$$

The noise covariance function γ_ε is known according to assumption (vi), but it conveys no information about the sample. The knowledge of γ_z is thus equivalent to that of γ_y, the latter term being the only potential source of information about the spectral components outside $\mathcal{D}_{\rm PSF}$. Let us examine the Fourier content of γ_y, first neglecting its discrete character. Using the continuous-space Fourier transform of (7), we obtain

$$\widetilde{\gamma}_y(\boldsymbol{u},\boldsymbol{u}') = \widetilde{h}(\boldsymbol{u})\,\widetilde{h}^*(-\boldsymbol{u}') \times \iint e^{-2i\pi(\boldsymbol{u}\cdot\boldsymbol{x}+\boldsymbol{u}'\cdot\boldsymbol{x}')}\rho(\boldsymbol{x})\rho(\boldsymbol{x}')\gamma_E(\boldsymbol{x}-\boldsymbol{x}')\,d\boldsymbol{x}\,d\boldsymbol{x}' \tag{8}$$

for all $\boldsymbol{u}$, $\boldsymbol{u} \in' \mathbb{R}^d$, where $\boldsymbol{u}\cdot\boldsymbol{r}$ denotes the usual scalar product in $\mathbb{R}^d$. Given that

$$\gamma_E(\boldsymbol{x}-\boldsymbol{x}') = \int e^{2i\pi(\boldsymbol{x}-\boldsymbol{x}')\cdot\boldsymbol{u}''}\widetilde{\gamma}_E(\boldsymbol{u}'')\,d\boldsymbol{u}'',$$

it is easy to express $\widetilde{\gamma}_y(\boldsymbol{u},\boldsymbol{u}')$ as follows:

$$\widetilde{\gamma}_y(\boldsymbol{u},\boldsymbol{u}') = \widetilde{h}(\boldsymbol{u})\,\widetilde{h}^*(-\boldsymbol{u}')\,\widetilde{g}(\boldsymbol{u},\boldsymbol{u}'), \tag{9}$$

with

$$\widetilde{g}(\boldsymbol{u},\boldsymbol{u}') = \int \widetilde{\rho}(\boldsymbol{u}-\boldsymbol{u}'')\,\widetilde{\rho}(\boldsymbol{u}'+\boldsymbol{u}'')\,\widetilde{\gamma}_E(\boldsymbol{u}'')\,d\boldsymbol{u}''. \tag{10}$$

According to (9) and to assumption (iv), $\widetilde{\gamma}_y$ has a support limited to $\mathcal{B}\times\mathcal{B}$. Hence, $\widetilde{\gamma}_y(\boldsymbol{u},\boldsymbol{u}')$ identifies with the discrete-space Fourier transform of γ_y for all $\boldsymbol{u}$, $\boldsymbol{u}' \in \mathcal{B}$. We conclude that the available information on the sample ρ from the discrete data covariance is contained in (and limited to) $\widetilde{g}(\boldsymbol{u},-\boldsymbol{u}')$, for all $\boldsymbol{u}$, $\boldsymbol{u}' \in \mathcal{D}_{\rm PSF}$.

IV. Super-Resolution Capacity of Second-Order Methods

According to expressions (7) or (9)-(10), the spectral density $\widetilde{\gamma}_E$ clearly plays a central role in identifying the spectral components of the sample. However, a difficulty in analyzing the SR capacity of second-order methods comes from the fact that the data covariance is not a linear but a quadratic functional of the unknown sample ρ. As a consequence, no general theory can be applied to solve equations (7) or (9)-(10) for ρ. However, two cases lend themselves to a deeper analysis. The first one corresponds to "sufficiently" correlated speckles, in the sense that the frequency support of the covariance function $\widetilde{\gamma}_E$ is contained in the frequency support of the PSF. At the opposite, the case of uncorrelated speckles can also be treated. These two cases are examined in the next two subsections, whereas handling the intermediate case remains an open issue. In the sequel, we make use of the Minkowski difference between two sets

$$A \ominus B = \{\boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}\in A, \boldsymbol{y}\in B\}$$

to define the frequency domains over which the identification (or partial identification) of the frequency components of the sample is possible.

A. Case of "Sufficiently" Correlated Speckle

Let us assume that the unknown speckle illuminations are spatially correlated, and that the frequency support of its covariance function is $\mathcal{D}_{\rm spec} = \{\boldsymbol{u} \,|\, \widetilde{\gamma}_E(\boldsymbol{u}) \neq 0\}$. According to expression (9), $\widetilde{\gamma}_y(\boldsymbol{u},\boldsymbol{u}')$ vanishes when either $\boldsymbol{u}$ or $-\boldsymbol{u}'$ is outside $\mathcal{D}_{\rm PSF}$. On the other hand, according to (10), $\widetilde{g}(\boldsymbol{u},\boldsymbol{u}')$ conveys no information on the frequency components $\widetilde{\rho}(\boldsymbol{v})$ such that either $\boldsymbol{v}\pm\boldsymbol{u}$ or $\boldsymbol{v}\pm\boldsymbol{u}'$ falls outside $\mathcal{D}_{\rm spec}$. Then, the following property holds.

Property 2: Any spectral component $\widetilde{\rho}(\boldsymbol{u})$ such that $\boldsymbol{u} \notin \mathcal{D}_1 \cup \mathcal{D}_2'$ with $\mathcal{D}_2' = \mathcal{D}_{\rm PSF} \ominus \mathcal{D}_{\rm spec}$ is non-identifiable from the mean $\mu_z(\boldsymbol{r})$ and the covariance function $\gamma_y(\boldsymbol{r},\boldsymbol{r}')$.

Remark 1: If each speckle pattern was known, the set of identifiable frequency components would be $\mathcal{D}_1 \cup \mathcal{D}_2'$ for a sufficiently large number of speckles. Moreover, the components outside $\mathcal{D}_1 \cup \mathcal{D}_2'$ would remain non-identifiable. In the same way, if the complete data statistics were available (the speckle patterns being unknown), the components outside $\mathcal{D}_1 \cup \mathcal{D}_2'$ would also be non-identifiable, since the latter situation is not more favorable than the former. We thus conclude that frequency components outside $\mathcal{D}_1 \cup \mathcal{D}_2'$ cannot be retrieved from standard (*i.e.*, non Bayesian) statistical information, even including higher moments.

Property 2 is of negative nature. Fortunately, a positive partial converse can be established in the important situation where the frequency support of the illuminations $\mathcal{D}_{\rm spec}$ is not larger than that of the PSF. The following non trivial property holds. Its proof is reported in Appendix B.

Property 3: Provided that γ_E is such that $\mathcal{D}_{\rm spec} \subseteq \mathcal{D}_{\rm PSF}$, any spectral component $\widetilde{\rho}(\boldsymbol{u})$ is identifiable from the mean $\mu_z(\boldsymbol{r})$ and the covariance function $\gamma_y(\boldsymbol{r},\boldsymbol{r}')$ if $\boldsymbol{u} \in \mathcal{D}_1 \cup \mathcal{D}_2''$ with $\mathcal{D}_2'' = \mathcal{D}_{\rm spec} \ominus \mathcal{D}_{\rm spec}$.

Remark 2: An alternative definition of $\mathcal{D}_2''$ is obtained *via*

$$\mathcal{D}_2'' = \{\boldsymbol{u} \,|\, (\widetilde{\gamma}_E \star \widetilde{\gamma}_E)(\boldsymbol{u}) > 0\} \tag{11}$$

where $\star$ is the (deterministic) cross-correlation[3] operator.

[3]The cross-correlation between two square-integrable functions f_1 and f_2 is defined by

$$(f_1 \star f_2)(\boldsymbol{x}) = \int f_1^*(\boldsymbol{x}')f_2(\boldsymbol{x}'+\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}'.$$

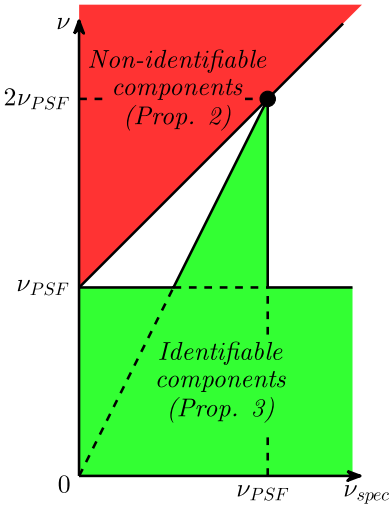

Fig. 1. Illustration of Props. 2 and 3 when $\mathcal{D}_{\rm PSF}$ and $\mathcal{D}_{\rm spec}$ are centered disks of respective radii $\nu_{\rm PSF}$ and $\nu_{\rm spec}$ (and $E_0 \neq 0$). The cutoff frequency $\nu_{\rm PSF}$ is fixed, while the speckle maximal frequency $\nu_{\rm spec}$ varies along the X-axis. The range of identifiable and of non-identifiable frequency components are represented along the Y-axis.

Let us consider a two-dimensional (2D) incoherent (*e.g.*, fluorescence) microscopy problem as an illustrative example for Properties 1, 2 and 3. In this case, $\mathcal{D}_{\rm PSF}$ and $\mathcal{D}_{\rm spec}$ are centered disks of respective radii $\nu_{\rm PSF}$ and $\nu_{\rm spec}$ and $\mathcal{D}_2''$ is a centered disk of radius $2\nu_{\rm spec}$. As a consequence, if $\nu_{\rm PSF}/2 < \nu_{\rm spec} < \nu_{\rm PSF}$, we get $\mathcal{D}_1 \subset \mathcal{D}_2''$ and a SR factor of $2\nu_{\rm spec}/\nu_{\rm PSF}$.

Fig. 1 gives a graphical illustration of the situation. Let us remark that the status of the frequency components outside the colored areas remains an open question. Our conjecture is that they are only partially identifiable from the second-order data statistics.

In the same conditions, according to Remark 1, the SR factor would be equal to $1 + \nu_{\rm spec}/\nu_{\rm PSF}$ if each speckle pattern was known. On Fig. 1, such a limit corresponds to the boundary line of the set of non-identifiable components. In a similar way, classical (harmonic) SIM would yield an SR factor equal to $1 + \nu_{\rm harm}/\nu_{\rm PSF}$ if many known harmonic illuminations were used, at frequencies spread around a centered circle of radius $\nu_{\rm harm}$.

The important case $\mathcal{D}_{\rm spec} = \mathcal{D}_{\rm PSF}$ is encountered in practice when illuminations and observations are performed *via* the same components (same antenna array for emission and detection, or same microscope objective for illumination and collection). The latter two properties then allow us to reach a tight conclusion in this context: second-order data statistics are sufficient to identify all the frequency components of the sample within $\mathcal{D}_2'' \equiv \mathcal{D}_2' \equiv \mathcal{D}_2 = \mathcal{D}_{\rm PSF} \ominus \mathcal{D}_{\rm PSF}$, and bring no information outside (such a situation corresponds to the black dot in Fig. 1). In other words, they should permit to recover the sample with a resolution equivalent to that of $|h|^2$, akin to classical SIM in fluorescence microscopy.

Let us also add a few comments about our main result for three-dimensional (3D) problems:

1) For *coherent* imaging systems, $\widetilde{h}$ is typically a hollow spherical cap, as depicted in Fig. 2(a), and thus a single incoming excitation (plane wave) E cannot provide any 3D information about the permittivity contrast ρ [4]. In contrast, the same setup using coherent, but random excitations such that $\mathcal{D}_{\rm spec} = \mathcal{D}_{\rm PSF}$, is able to retrieve $\widetilde{\rho}$ within a domain $\mathcal{D}_2''$ that is a centered solid torus, hence providing 3D information[4] about the permittivity contrast.
2) In *incoherent* imaging, $\widetilde{h}$ is typically a solid torus, as depicted in Fig. 2(b), which provides very poor sectioning capability along the axial direction z; this is the so-called "missing-cone" problem in wide-field incoherent microscopy [26]. In this case, speckle intensity illuminations such that $\mathcal{D}_{\rm spec} = \mathcal{D}_{\rm PSF}$ give access to a frequency domain $\mathcal{D}_2''$ that provides an extended lateral and axial resolution without any missing-cone, see Fig. 2(c). This

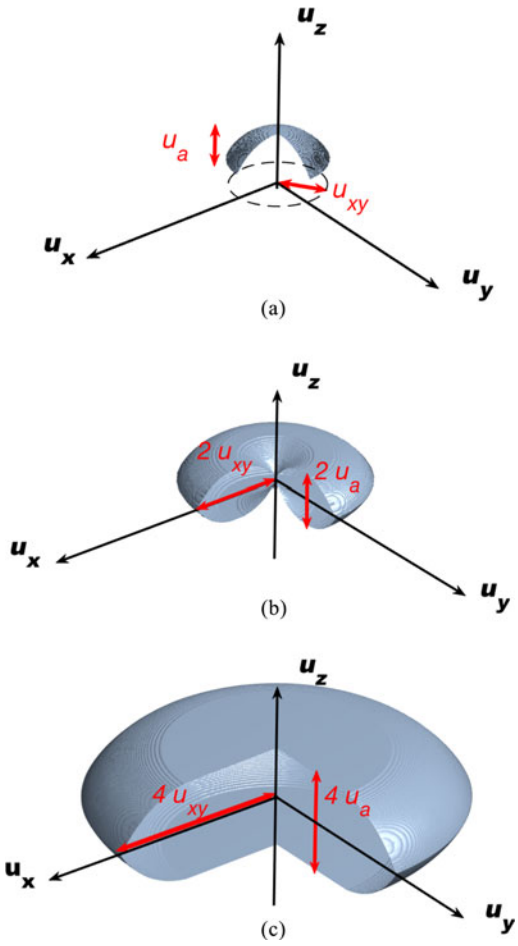

Fig. 2. **Coherent imaging system**: (a) $\mathcal{D}_{\rm PSF}$ is a surface in the 3D Fourier domain with an isotrope *lateral* cutoff frequency u_{xy} and an *axial* frequency extension u_a; (b) Assuming random excitations such that $\mathcal{D}_{\rm spec} = \mathcal{D}_{\rm PSF}$, the frequency components of the permittivity contrast ρ should be further identified over a domain $\mathcal{D}_2''$ that is a torus with a lateral (resp. axial) frequency extension of $2u_{xy}$ (resp. $2u_a$). **Incoherent imaging system**: (b) $\mathcal{D}_{\rm PSF}$ is a solid torus exhibiting a "missing cone" along u_z; (c) Assuming random illuminations such that $\mathcal{D}_{\rm spec} = \mathcal{D}_{\rm PSF}$, ρ should be further identified over an extended frequency domain $\mathcal{D}_2''$ providing a lateral (resp. axial) resolution of $4u_{xy}$ (resp. $4u_a$) without any "missing-cone" along u_z.

[4]The same identification domain can be obtained from a set of plane waves with various incoming angles, *i.e.,* with the additional difficulty and slowness of controlling the angles of illuminations, see [4], [22] for details.

domain is actually equivalent to that of a perfect confocal microscope with an infinitively small pinhole [27], but it is obtained with no transverse scanning and no loss of photons.

As a final note, we stress that Property 3 deals with the identifiability of the frequency components of the sample, and it does not predict the reachable estimation precision in the realistic situation of a limited set of noisy data and of a limited field of view. Nevertheless, the numerical reconstruction example proposed in Section V suggests that most of the frequency components within $\mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$ can be reliably retrieved in practice.

B. Case of Uncorrelated Speckle

Let us now assume that $\gamma_E(\boldsymbol{r}) = \gamma_E(\boldsymbol{0})\delta(\boldsymbol{r})$, where δ is a Dirac distribution. This assumption can be considered valid when the speckle correlation typical size is much smaller than that of the point spread function. Then (7) becomes

$$\gamma_y(\boldsymbol{r}, \boldsymbol{r}') = \gamma_E(\boldsymbol{0}) \int \rho^2(\boldsymbol{x})\, h(\boldsymbol{r} - \boldsymbol{x}) h^*(\boldsymbol{r}' - \boldsymbol{x})\, d\boldsymbol{x}. \quad (12)$$

In the Fourier domain, $\widetilde{\gamma}_E(\boldsymbol{u}) = \gamma_E(\boldsymbol{0})$, so (9) read

$$\widetilde{\gamma}_y(\boldsymbol{u}, \boldsymbol{u}') = \gamma_E(\boldsymbol{0}) \widetilde{h}(\boldsymbol{u})\, \widetilde{h}^*(-\boldsymbol{u}')\, \widetilde{\rho^2}(\boldsymbol{u} + \boldsymbol{u}'). \quad (13)$$

The latter relation is important since it yields that ρ^2 is accessible over all frequencies $\boldsymbol{u} + \boldsymbol{u}'$ such that both $\boldsymbol{u}$ and $-\boldsymbol{u}'$ belong to $\mathcal{D}_{\mathrm{PSF}}$, *i.e.*, over the set $\mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$. As a conclusion, the following property holds.

Property 4: The frequency component $\widetilde{\rho^2}(\boldsymbol{u})$ is identifiable from the covariance function $\gamma_y(\boldsymbol{r}, \boldsymbol{r}')$ if and only if $\boldsymbol{u} \in \mathcal{D}_2 = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$.

A remarkable fact is that Property 4 still holds if only the variance $v(\boldsymbol{r}) = \gamma_y(\boldsymbol{r}, \boldsymbol{r})$ is considered instead of the full covariance function $\gamma_y(\boldsymbol{r}, \boldsymbol{r}')$, provided that the data grid fulfills a more stringent condition than assumption (iv). The starting point is the following relation obtained from (12):

$$v(\boldsymbol{r}) = \gamma_E(\boldsymbol{0})\, (\rho^2 \otimes |h|^2)(\boldsymbol{r}), \quad (14)$$

The continuous-space Fourier transform of v reads

$$\widetilde{v}(\boldsymbol{u}) = \gamma_E(\boldsymbol{0})\, \widetilde{\rho^2}(\boldsymbol{u}) \times (\widetilde{h} \star \widetilde{h})(\boldsymbol{u}).$$

Since the support of $\widetilde{h} \star \widetilde{h}$ spans the domain $\mathcal{D}_2$, the discrete-space Fourier transform of v identifies with $\widetilde{v}$ provided that $\mathcal{D}_2 \subset \mathcal{B}$. This assumption is stronger than (ii). Typically, it means that the data must be acquired at least at twice the Nyquist rate. Alternatively, the acquisition rate could be unchanged, but then the data should be interpolated on a twice finer grid to yield the variance (14). Obviously, interpolation will not bring any new information. It will simply allow us to preserve the SR information contained in the sampled variance function v, and more precisely to avoid aliasing on v. On the other hand, interpolation will also apply to the noise component, so that the corresponding statistics should be modified accordingly.

Property 5: Provided that $\mathcal{D}_2 \subset \mathcal{B}$, the frequency component $\widetilde{\rho^2}(\boldsymbol{u})$ is identifiable from the variance function $\gamma_y(\boldsymbol{r}, \boldsymbol{r})$ if and only if $\boldsymbol{u} \in \mathcal{D}_2$.

A statement somewhat similar to Property 5 was already made in [15], assuming a circular aperture and a single incoherent point source (for ρ). Whereas the authors of [15] assert that uncorrelated speckle illumination has the ability to improve the resolution of the imaging setup beyond the diffraction limit, it should be stressed that if $\widetilde{\rho^2} = \widetilde{\rho} \otimes \widetilde{\rho}$ can be retrieved on $\mathcal{D}_2$, this does not mean that $\widetilde{\rho}$ can be retrieved on the same domain, nor in any other domain. In practice, additional constraints (*e.g.*, the positivity of the sample assumed in this paper) can be further considered [19], [23, Sec. 6.6.4], but with no formal guarantee about the super-resolution property obtained on ρ, to our best knowledge.

V. Numerical Implementation for 2D Speckle SIM

The goal of the present section is to give a practical illustration of Property 3, which is the main theoretical result of Section IV. For this purpose, we consider a 2D speckle illumination fluorescence microscopy problem. In the standard assumption of a perfect circular lens, h is the so-called Airy pattern [23, Sec. 4.4.2], and the optical transfer function (OTF) $\widetilde{h}$ defines a support $\mathcal{D}_{\mathrm{PSF}} = \{\boldsymbol{u}, \|\boldsymbol{u}\|_2 < 2\mathrm{NA}/\lambda\}$ with NA the numerical aperture of the microscope and λ the emission/excitation wavelength. We further assume that the illumination of the sample and the collection of the emitted light is performed through the same optical device. Ignoring the Stokes-shift,[5] we consider hereafter that $\gamma_E = E_0^2\, h$. According to Property 3, a SR effect approaching a factor two is expected from the empirical second-order statistics of a set of M collected images, for asymptotically large values of M. The goal here is to show empirically that this SR effect does happen in realistic conditions, and in particular for moderately large values of M.

A. Discretized Model for 2D Speckle SIM

For the sake of computer implementation, (3) must be replaced by its discretized counterpart

$$\boldsymbol{z}_m = \boldsymbol{H}\boldsymbol{R}\boldsymbol{E}_m + \boldsymbol{\varepsilon}_m, \quad (15)$$

where $\boldsymbol{H}$ is a symmetric convolution matrix, and $\boldsymbol{R} = \mathrm{Diag}(\boldsymbol{\rho})$, so that $\boldsymbol{R}\boldsymbol{E}_m$ corresponds to the product between the vectorized sample $\boldsymbol{\rho}$, and the vectorized mth illumination pattern $\boldsymbol{E}_m$. The mean vector and the covariance matrix of the acquisition $\boldsymbol{z}_m$ now read

$$\boldsymbol{\mu}_z = E_0 \boldsymbol{H}\boldsymbol{\rho}, \qquad \boldsymbol{\Gamma}_z = \boldsymbol{H}\boldsymbol{R}\boldsymbol{\Gamma}_E\boldsymbol{R}\boldsymbol{H} + \boldsymbol{\Gamma}_\varepsilon, \quad (16)$$

where $\boldsymbol{\Gamma}_E$ and $\boldsymbol{\Gamma}_\varepsilon$ are the covariance matrix of the speckle patterns and of the additive noise, respectively. For any finite number of illuminations M, the empirical mean $\hat{\boldsymbol{\mu}}_z$ and covariance $\hat{\boldsymbol{\Gamma}}_z$ are defined as

$$\widehat{\boldsymbol{\mu}}_z = \frac{1}{M}\sum_{m=1}^{M} \boldsymbol{z}_m, \qquad \widehat{\boldsymbol{\Gamma}}_z = \frac{1}{M}\sum_{m=1}^{M} \boldsymbol{z}_m \boldsymbol{z}_m^\dagger - \widehat{\boldsymbol{\mu}}_z \widehat{\boldsymbol{\mu}}_z^\dagger, \quad (17)$$

[5]The Stokes-shift [28] implies that the support $\widetilde{h}$ is slightly smaller than the support of $\widetilde{\gamma_E}$. This difference between supports is small (about 10%) and we expect that it will have a negligible impact on the SR effect that should reach twice the cutoff frequency of the OTF.

where the symbol † stands for the transpose conjugate operator. As M grows, $\hat{\boldsymbol{\mu}}_z$ and $\hat{\boldsymbol{\Gamma}}_z$ respectively converge toward $\boldsymbol{\mu}_z$ and $\boldsymbol{\Gamma}_z$, so that spatial frequency components of the sample within $\mathcal{D}_2''$ become identifiable, according to Prop. 3. With a view to propose a computationally effective strategy to retrieve the latter components, we first directly formulate the main elements of Prop. 3 and of its proof in a finite dimensional (*i.e.*, discretized) framework.

B. Matrix Transposition of Property 3

In the discrete framework of model (15), the matrix formulation of Property 3 mostly relies on the one-to-one mapping between the asymptotically available covariance $\boldsymbol{\Gamma}_y = \boldsymbol{H}\boldsymbol{R}\boldsymbol{\Gamma}_E\boldsymbol{R}\boldsymbol{H}$ and the matrix $\boldsymbol{S} = \boldsymbol{\Gamma}_E^{1/2}\boldsymbol{R}\boldsymbol{\Gamma}_E^{1/2}$, provided that $\operatorname{Ker}\boldsymbol{H} \subseteq \operatorname{Ker}\boldsymbol{\Gamma}_E$. The latter condition is the discrete-space counterpart of the condition $\mathcal{D}_{\rm spec} \subseteq \mathcal{D}_{\rm PSF}$ that allows the identification result stated in Prop. 3. ($\operatorname{Ker}\boldsymbol{M}$ denotes the set of vectors $\boldsymbol{v}$ such that $\boldsymbol{M}\boldsymbol{v}$ is the null vector). Then, we can show that $\boldsymbol{S}$ is the unique Hermitian positive semi-definite[6] square-root of

$$\boldsymbol{F} = \boldsymbol{\Gamma}_E^{1/2}\boldsymbol{H}^{+}\boldsymbol{\Gamma}_y\boldsymbol{H}^{+}\boldsymbol{\Gamma}_E^{1/2},$$

where $^+$ denotes the generalized inverse [29, ch. 5]. Indeed, matrices $\boldsymbol{F}$ and $\boldsymbol{S}$ respectively correspond to kernels F and f introduced in the proof of Property 3 (see Appendix B). Whereas $\boldsymbol{\Gamma}_y$ quadratically depends on $\boldsymbol{\rho}$, $\boldsymbol{S}$ exhibits a linear dependency with respect to ρ, paving the way to an identifiability analysis *via* a standard eigenvalue decomposition.

C. Numerical Estimation Strategy

The reconstruction principle from the second-order data statistics amounts to finding $\boldsymbol{\rho}$ that makes the mean vector $\boldsymbol{\mu}_z$ and the covariance matrix $\boldsymbol{\Gamma}_z$ in (16) best match with the empirical quantities $\hat{\boldsymbol{\mu}}_z$ and $\hat{\boldsymbol{\Gamma}}_z$ defined by (17). Given the previous subsection, a simple idea to recover the identifiable components of $\widetilde{\boldsymbol{\rho}}$ would be to compute an approximation $\widehat{\boldsymbol{F}}$ of matrix $\boldsymbol{F}$ from the empirical data statistics:

$$\widehat{\boldsymbol{F}} = \boldsymbol{\Gamma}_E^{1/2}\boldsymbol{H}^{+}\widehat{\boldsymbol{\Gamma}}_y\boldsymbol{H}^{+}\boldsymbol{\Gamma}_E^{1/2},$$

where $\widehat{\boldsymbol{\Gamma}}_y = \widehat{\boldsymbol{\Gamma}}_z - \boldsymbol{\Gamma}_\varepsilon$, with a view to extract a positive semi-definite square-root matrix $\widehat{\boldsymbol{S}}$. However, neither $\widehat{\boldsymbol{\Gamma}}_y$ nor $\widehat{\boldsymbol{F}}$ are guaranteed to be positive semi-definite, so the existence of $\widehat{\boldsymbol{S}}$ is not granted.

A preferable procedure consists in introducing an appropriate dissimilarity measure between the empirical and the theoretical second-order statistics of the data, and to minimize the dissimilarity to obtain an estimated sample $\widehat{\boldsymbol{\rho}}$. One possible choice of dissimilarity measure is the Kullback-Leibler divergence (KLD) $D(\boldsymbol{\rho}) = D_{\rm KL}(\mathcal{N}(\widehat{\boldsymbol{\mu}}_z, \widehat{\boldsymbol{\Gamma}}_z)\|\mathcal{N}(\boldsymbol{\mu}_z, \boldsymbol{\Gamma}_z))$, where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ is the normal distribution of mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Gamma}$. According to [30, Sec. 9.1], an explicit expression of $D(\boldsymbol{\rho})$ is:

$$D(\boldsymbol{\rho}) = \frac{1}{2}\operatorname{Tr}\big(\boldsymbol{\Gamma}_z^{-1}\widehat{\boldsymbol{\Gamma}}_z\big) + \frac{1}{2}(\boldsymbol{\mu}_z - \widehat{\boldsymbol{\mu}}_z)^{\rm t}\boldsymbol{\Gamma}_z^{-1}(\boldsymbol{\mu}_z - \widehat{\boldsymbol{\mu}}_z) + \frac{1}{2}\log\frac{|\boldsymbol{\Gamma}_z|}{|\widehat{\boldsymbol{\Gamma}}_z|} - \frac{N}{2} \tag{18}$$

where $|\cdot|$ and $\operatorname{Tr}(\cdot)$ are the determinant and the trace of a square matrix, respectively. Let us mention that D is proportional to the log-likelihood of the data under the assumption that the latter follow the normal distribution $\mathcal{N}(\boldsymbol{\mu}_z, \boldsymbol{\Gamma}_z)$ [31]. However, the minimizer of D is an unregularized solution, which is unstable with respect to the random fluctuations in the dataset. Therefore, a penalization term must be added to D. In the sequel, we choose a quadratic penalization term to stabilize the solution, so that the SR effect remains purely driven by the data term. The criterion to minimize is then

$$J(\boldsymbol{\rho}) = D(\boldsymbol{\rho}) + \frac{\beta}{2}\|\boldsymbol{\rho}\|_2^2, \tag{19}$$

with $\beta \geq 0$ and $\|\cdot\|_2$ is the usual Euclidian norm. From a computational perspective, a closed-form minimizer cannot be found, so the minimization problem must be solved iteratively. Indeed, it is a so-called *structured covariance* type problem, for which the Expectation-Maximization (EM) algorithm can be implemented [31]–[33]. However, our tests indicate that the EM algorithm converges very slowly in the speckle SIM context. For this reason, we rather rely on a nonlinear conjugate gradient method, which turns out to produce more efficient iterations. It relies on the expression of the gradient of the penalized KLD (19) with respect to $\boldsymbol{\rho}$ (see Appendix C for a derivation):

$$\nabla J(\boldsymbol{\rho}) = -\big(\big[\boldsymbol{\Omega}^{\rm t}\big(\boldsymbol{\Delta}_\Gamma + \boldsymbol{\delta}_\mu\boldsymbol{\delta}_\mu^{\rm t}\big)\boldsymbol{\Omega}\big] \circ \boldsymbol{\Gamma}_E\big)\,\boldsymbol{\rho} - E_0\boldsymbol{\Omega}^{\rm t}\boldsymbol{\delta}_\mu + \beta\boldsymbol{\rho}, \tag{20}$$

displayed as a column vector, with $\boldsymbol{\Omega} = \boldsymbol{\Gamma}_z^{-1}\boldsymbol{H}$, $\boldsymbol{\delta}_\mu = \widehat{\boldsymbol{\mu}}_z - \boldsymbol{\mu}_z$, $\boldsymbol{\Delta}_\Gamma = \widehat{\boldsymbol{\Gamma}}_z - \boldsymbol{\Gamma}_z$, and $\circ$ stands for the Hadamard (componentwise) product. Let us stress that each computation of the gradient needs the construction and the inversion of an $N \times N$ matrix (for an N-pixel size problem), which represents a prohibitive computing cost for realistic imaging problems. The design of less costly iterations for large-size problems is out of the scope of the present paper, but we are currently working on this crucial issue.

D. Numerical Illustration for 2D Speckle SIM

Numerical simulations are now considered to support that a significant SR effect can be obtained in speckle fluorescence SIM, even with a moderately large number of illumination patterns $\boldsymbol{E}_m$. The ground truth $\boldsymbol{\rho}^\star$ consists in the 2D 'star-like' fluorescence pattern depicted in Fig. 3(a). The convolution matrix $\boldsymbol{H}$ modeling the microscope is built from the discretized OTF associated with a circular aperture [23, Eq. (6)–(32)]; the numerical aperture NA is set to 1.49 and the emission/excitation wavelength λ is arbitrary set to 1. For this configuration, the resolution limit of standard wide-field imaging is clearly visible in Fig. 3(c). According to (15), a set of $M \in \{100, 1000\}$ speckle patterns are simulated to produce M low-resolution microscope

[6] A Hermitian matrix is positive semi-definite if and only if all of its eigenvalues are nonnegative.

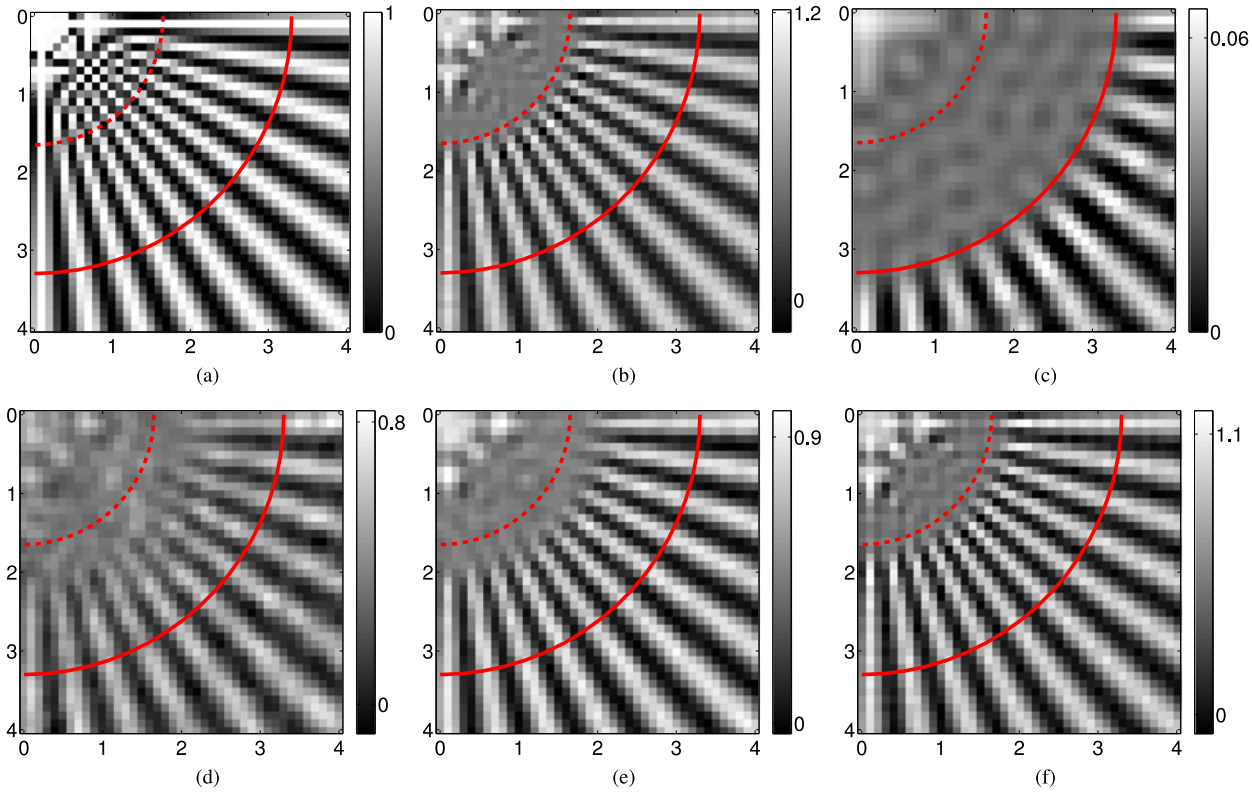

Fig. 3. (a) Lower quarter of the (80 × 80 pixels) ground-truth fluorescence pattern considered in [13]. (b) Filtered ground-truth retaining only the spatial frequency lower than twice the OTF limit. (c) The deconvolution of the wide-field (constant illumination) microscope acquisition. (d,e) Estimator of $\boldsymbol{\rho}$ obtained from the minimization of the penalized KLD (19) with $M = 100$ (d) and $M = 1000$ (e) speckle patterns; the regularization parameter is set to $\beta = \beta_0/M$ with $\beta_0 = 100$. (f) Estimator of $\boldsymbol{\rho}$ obtained from the minimization the KLD (18) with the asymptotic statistics $\widehat{\boldsymbol{\mu}} = \boldsymbol{\mu}^\star$ and $\widehat{\boldsymbol{\Gamma}} = \boldsymbol{\Gamma}^\star$. The distance units along the horizontal and vertical axes are given in wavelength λ. The image sampling step for all simulations is set to λ/20. The dashed (resp. solid) lines corresponds to the spatial frequencies transmitted by the OTF support (resp. twice the OTF limit).

images $\{\boldsymbol{z}_m\}_{m=1}^M$. The covariance matrix $\boldsymbol{\Gamma}_E$ is set to $E_0^2\,\boldsymbol{H}$ (we assume $\gamma_E = E_0^2\,h$) and each acquisition $\boldsymbol{z}_m$ is corrupted with an independent and identically distributed Gaussian noise such that the signal-to-noise ratio in each frame is set to 40 dB. From the dataset $\{\boldsymbol{z}_m\}_{m=1}^M$, the statistics $\widehat{\boldsymbol{\mu}}_z$ and $\widehat{\boldsymbol{\Gamma}}_z$ (17) are built. The case of an infinite illumination number ($M = \infty$) is also adressed by considering the expected (*i.e.*, asymptotical) statistics $\widehat{\boldsymbol{\mu}}_z = \boldsymbol{\mu}_z^\star$ and $\widehat{\boldsymbol{\Gamma}}_z = \boldsymbol{\Gamma}_z^\star$, where $\boldsymbol{\mu}_z^\star$ and $\boldsymbol{\Gamma}_z^\star$ are obtained from (16) by setting $\boldsymbol{\rho} = \boldsymbol{\rho}^\star$. In all cases, we proceed to the iterative minimization of the penalized KLD (19) to estimate the sample, using the deconvolved wide-field image of Fig. 3(c) as an initial point. For the (noise-free) asymptotic statistics, the regularization parameter is set to $\beta = 0$ and, as expected, the reconstruction exhibits the doubled resolution predicted by Prop. 3, see Fig. 3(f) compared to Fig. 3(b), (c). With 100 and 1000 illuminations, the SR factor is lower, but the reconstructions shown on Figs. 3(d), (e) are still much more resolved than the wide-field image of Fig. 3(c). Moreover, the SR factor progressively grows with the illumination number M, the result at $M = 1000$ being close to the asymptotic regime.

VI. Conclusion and Perspectives

We have mathematically demonstrated that the mean and the covariance function of low resolution images obtained with unknown, random illuminations permit to recover a super-resolved image of the sample, provided the first two statistical moments of the illuminations are fully characterized. Since this condition is expected to be less stringent to meet than the knowledge of each illumination, we believe that this result can be interesting in many practical situations.

In fluorescence microscopy, if the speckle is generated through the same objective as the one used to collect the light, its covariance function is almost identical to the microscope PSF and Proposition 3 is also expected to apply. We believe that this is a particularly important result. Indeed, it shows that speckle microscopy has the potential to generate a super-resolved image corresponding to the PSF $h_{\text{ext}} = |h|^2$. In other words, the SR would be equivalent to that of a perfect confocal microscope with infinitively small pinhole [27], but it would be obtained with no transverse scanning and with no loss of photons.

For coherent imaging system, the consequence could be even more spectacular. In holographic systems, the Fourier support of the PSF h is generally a cap of sphere. As a result, the three-dimensional information on the sample is lost if only one illumination is used. This is clearly observed in tomographic diffraction microscopy where the reconstruction of a target from its unique 2D hologram obtained under a monochromatic plane wave illumination is significantly deteriorated along one axis [34]. On the other hand, by processing 2D images obtained

under different speckle illuminations, one should be able to reconstruct the target in three dimensions with a PSF comparable to that obtained in tomographic diffraction imaging [4], but without the difficulty and slowness of controlling the angles of illuminations.

In photoacoustic imaging using speckle illuminations, the autocorrelation length of the random optical intensity can drop to a few hundreds of nanometers while the acoustic PSF h has a typical width of tens of microns. Hence, from the acoustic point of view, the illumination can be seen as an uncorrelated random process [18], which is the case studied in SubSection IV-B. Using optical speckle in a photoacoustic experiment would allow to retrieve the square of the optical absorption density with a resolution corresponding to the PSF $|h|^2$.

Finally, it is important to stress the limits of the present analysis. First, the case of complex-valued samples (*i.e.*, with both dielectric and absorptive components) remains to be investigated, since it could have important implications in electromagnetic tomography. The present study can be easily adapted to the case of pure absorptive (imaginary) samples, but an extension to the more general case is not so direct. Second, our theoretical results are of asymptotic nature, in that they only predict the SR capacity of the imagers with an arbitrarily large number of illuminations. In particular, Proposition 3 does not provide the *sensitivity* of the retrievable sample frequency components. The simulation results shown in Section V-D nonetheless suggest that these frequency components can be retrieved with only a few hundreds of illuminations. Third, our results do not take into account the potential impact of advanced regularization in the inversion schemes (for instance, exploiting a sparsity prior [14] could yield an additional increase of resolution). Fourth, there exist many imaging configurations where the second-order statistics do not entirely characterize the probability distribution of the data (*e.g.*, when the speckle illuminations are positive intensities). In such cases, our identifiability results only provide a lower bound on the super-resolution factor that could be reached from the complete data statistics, since accounting more precisely for the speckle statistics could still ameliorate the resolution. According to Remark 1, no amelioration can be expected when the support of the speckle covariance identifies with that of the PSF, but the question remains open in other cases. For instance, for uncorrelated speckles, one can write the following extension of Eq. (14),

$$\mathrm{Cum}_y^n(\boldsymbol{r},\,\ldots,\,\boldsymbol{r}) = \mathrm{Cum}_E^n(\mathbf{0},\,\ldots,\,\mathbf{0})\,(\rho^n \otimes |h|^n)(\boldsymbol{r}), \quad (21)$$

where Cum^n denotes the nth circular cumulant of a given random process [35]. Equation (21) indicates that data higher-order statistics yield additional information on higher spatial frequencies of the sample. Such a property is reminiscent of the principle of SOFI [36]. The computational issue also remains broadly open, both in terms of memory requirements (to store the empirical data covariance matrix) and of computing time. The iterative scheme proposed in SubSection V-D is clearly limited to small-sized images. A challenge will be to accelerate the reconstruction process while preserving the SR capacity of speckle-based imaging, as characterized in this paper.

As a final remark, let us stress that controlled and random illuminations lead in our opinion to distinct "resolution *vs.* cost *vs.* versatility" trade-offs for the setup. In particular, when an accurate control of the illumination can be obtained within the sample volume, random illuminations may not be the best option to maximize the resolution for a given photon budget. On the contrary, random illuminations should achieve a better trade-off when the illumination cannot be controlled, or if one aims at designing *versatile* and *cheap* setups. The cautious evaluation of these trade-offs is a clear perspective of this work.

APPENDIX A
CASE OF POISSON DATA

For an incoherent imaging setup (*e.g.*, optical fluorescence microscopy), the intensity measurement relies on counting discrete particles and the model (2) can be replaced by

$$z_m(\boldsymbol{r}) = p_m(\boldsymbol{r}) + \varepsilon_m(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{Z}^d \quad (22)$$

where $p_m(\boldsymbol{r})$ is a Poisson random variable with mean

$$\mathcal{E}\left\{p_m(\boldsymbol{r}) \,|\, E_m\right\} = \int y_m(\boldsymbol{\zeta})\,\Pi(\boldsymbol{\zeta}-\boldsymbol{r})\,d\boldsymbol{\zeta}, \quad \boldsymbol{r} \in \mathbb{Z}^d,$$

where Π is the indicator function of a centered detector pixel. Assuming further that the pixel size is "small" with respect to the spatial variation in y_m, the expected counting rate for all $\boldsymbol{r} \in \mathbb{Z}^d$ is approximated by

$$\mathcal{E}\left\{p_m(\boldsymbol{r}) \,|\, E_m\right\} = a\, y_m(\boldsymbol{r})$$

where a is the area of a single detector pixel. We also assume that the Poisson outcomes are jointly statistically independent. The expression for $\mu_z(\boldsymbol{r}) = \mathcal{E}\left\{z_m(\boldsymbol{r})\right\} = \mathcal{E}\left\{p_m(\boldsymbol{r})\right\}$ then reads, according to the law of iterated expectations:

$$\mathcal{E}\left\{p_m(\boldsymbol{r})\right\} = \mathcal{E}\left\{\mathcal{E}\left\{p_m(\boldsymbol{r}) \,|\, E_m\right\}\right\} = a\,\mathcal{E}\left\{y_m(\boldsymbol{r})\right\}$$

where $\mathcal{E}\left\{y_m(\boldsymbol{r})\right\} = E_0(h \otimes \rho)(\boldsymbol{r})$. Concerning the data covariance function, we have for $\boldsymbol{r}$, $\boldsymbol{r}' \in \mathbb{Z}^d$:

$$\gamma_z(\boldsymbol{r},\boldsymbol{r}') = \gamma_p(\boldsymbol{r},\boldsymbol{r}') + \gamma_\varepsilon(\boldsymbol{r}-\boldsymbol{r}')$$

with:

$$\gamma_p(\boldsymbol{r},\boldsymbol{r}') = \mathcal{E}\left\{p_m(\boldsymbol{r})p_m(\boldsymbol{r}')\right\} - \mathcal{E}\left\{p_m(\boldsymbol{r})\right\}\,\mathcal{E}\left\{p_m(\boldsymbol{r}')\right\}. \quad (23)$$

According to the law of iterated expectations,

$$\mathcal{E}\left\{p_m(\boldsymbol{r})p_m(\boldsymbol{r}')\right\} = \mathcal{E}\left\{\mathcal{E}\left\{p_m(\boldsymbol{r})p_m(\boldsymbol{r}') \,|\, E_m\right\}\right\}.$$

For $\boldsymbol{r}' \neq \boldsymbol{r}$, since $p_m(\boldsymbol{r})$ and $p_m(\boldsymbol{r}')$ are decorrelated Poisson variables given E_m, we get

$$\mathcal{E}\left\{p_m(\boldsymbol{r})p_m(\boldsymbol{r}')\right\} = a^2\,\mathcal{E}\left\{y_m(\boldsymbol{r})y_m(\boldsymbol{r}')\right\}$$

while for $\boldsymbol{r}' = \boldsymbol{r}$,

$$\mathcal{E}\left\{p_m(\boldsymbol{r})^2\right\} = a^2\mathcal{E}\left\{y_m(\boldsymbol{r})^2\right\} + a\mathcal{E}\left\{y_m(\boldsymbol{r})\right\},$$

since a Poisson variable is of equal mean and variance. Therefore, (6) must be replaced by

$$\gamma_z(\boldsymbol{r},\boldsymbol{r}') = a^2\,\gamma_y(\boldsymbol{r},\boldsymbol{r}') + \mu_z(\boldsymbol{r})\delta_{\mathrm{K}}(\boldsymbol{r}-\boldsymbol{r}') + \gamma_\varepsilon(\boldsymbol{r}-\boldsymbol{r}') \quad (24)$$

where γ_y is given in (7) and $\delta_{\mathrm{K}}(\boldsymbol{r}) = 1$ if $\boldsymbol{r} = \boldsymbol{0}$ and zero otherwise. The data covariance hence only differs from (7) when $\boldsymbol{r} = \boldsymbol{r}'$. In this case, we note that the additional term $\mu_z(\boldsymbol{r})$ is proportional to E_0, whereas the variance $\gamma_y(\boldsymbol{r}, \boldsymbol{r})$ varies as $\gamma_E(\boldsymbol{r})$, which is usually proportional to E_0^2 for intensity speckle patterns [12]. Therefore, accounting for the Poisson statistics of the data in may be useful in the low counting-rate regime only.

Appendix B
Proof of Property 3

Let q denote the impulse response of the filter defined in the Fourier domain by

$$\widetilde{q}(\boldsymbol{u}) = \widetilde{\gamma}_E^{1/2}(\boldsymbol{u}) \quad \text{if } \boldsymbol{u} \in \mathcal{D}_{\mathrm{spec}}, \quad 0 \text{ otherwise.}$$

Akin to γ_E, q is positive semi-definite, and hence it is a Hermitian symmetric function. We have then $\widetilde{\gamma}_E = \widetilde{q}^2$, and hence

$$\gamma_E = q \otimes q. \tag{25}$$

Let us also define the following kernels:

$$f(\boldsymbol{r}, \boldsymbol{r}') = \int q(\boldsymbol{r} - \boldsymbol{x}) q^*(\boldsymbol{r}' - \boldsymbol{x}) \rho(\boldsymbol{x})\, d\boldsymbol{x}, \tag{26}$$

$$F(\boldsymbol{r}, \boldsymbol{r}') = \int f(\boldsymbol{r}, \boldsymbol{r}'') f^*(\boldsymbol{r}', \boldsymbol{r}'')\, d\boldsymbol{r}''. \tag{27}$$

and the induced integral operators K_f and K_F:

$$K_f \phi(\boldsymbol{r}) = \int f(\boldsymbol{r}, \boldsymbol{r}') \phi(\boldsymbol{r}')\, d\boldsymbol{r}',$$

$$K_F \phi(\boldsymbol{r}) = \int F(\boldsymbol{r}, \boldsymbol{r}') \phi(\boldsymbol{r}')\, d\boldsymbol{r}'.$$

According to the Cauchy-Schwarz inequality,

$$|f(\boldsymbol{r}, \boldsymbol{r}')|^2 \leq \int |q(\boldsymbol{r} - \boldsymbol{x})|^2 \rho(\boldsymbol{x})\, d\boldsymbol{x} \int |q(\boldsymbol{r}' - \boldsymbol{x})|^2 \rho(\boldsymbol{x})\, d\boldsymbol{x}.$$

As a consequence,

$$\iint |f(\boldsymbol{r}, \boldsymbol{r}')|^2 \, d\boldsymbol{r}\, d\boldsymbol{r}' \leq \left(\int |q(\boldsymbol{r})|^2 \, d\boldsymbol{r} \int \rho(\boldsymbol{x})\, d\boldsymbol{x} \right)^2,$$

where ρ is integrable, according to assumption (*ii*), and

$$\int |q(\boldsymbol{r})|^2 \, d\boldsymbol{r} = \int \widetilde{\gamma}_E(\boldsymbol{u})\, d\boldsymbol{u} = \gamma_E(0) < \infty.$$

Therefore, we have

$$\iint |f(\boldsymbol{r}, \boldsymbol{r}')|^2 \, d\boldsymbol{r}\, d\boldsymbol{r}' < \infty,$$

i.e., $f \in L^2(\mathbb{R}^d \times \mathbb{R}^d; \boldsymbol{C})$, and consequently, K_f is a Hilbert-Schmidt integral operator [37, Proposition 3.4.16]. On the other hand, the integral operator K_F is the square of K_f, in the sense that $K_F \phi = K_f K_f \phi$ for any ϕ. Thus, K_F is also a Hilbert-Schmidt operator.

Now let us go to the heart of the proof, which is threefold. The first step allows us to show that kernel F is uniquely defined from γ_y. In a second step, we establish that f is uniquely defined from F given (27). At this point, we conclude that the knowledge of γ_y implies that of f, which is a linear functional of ρ (whereas the dependency of γ_y in ρ is quadratic). The last step consists in a Fourier analysis of f, in order to determine which spectral components of ρ are identifiable from the knowledge of f.

Step 1) Given (26) and (25), we have the following alternate expression for (27):

$$F(\boldsymbol{r}, \boldsymbol{r}') = \iint \rho(\boldsymbol{x}) \rho(\boldsymbol{x}')\, q(\boldsymbol{r} - \boldsymbol{x}) q^*(\boldsymbol{r}' - \boldsymbol{x}')\, \gamma_E(\boldsymbol{x} - \boldsymbol{x}')\, d\boldsymbol{x}\, d\boldsymbol{x}'. \tag{28}$$

Comparing the latter expression to (7), it is clear that $F = \gamma_y$ in the case $q = h$, *i.e.*, when the speckle covariance is $h \otimes h$. More generally, using a double Fourier transform on (28), in the same way as we obtained (9) from (7), we get

$$\widetilde{F}(\boldsymbol{u}, \boldsymbol{u}') = \widetilde{q}(\boldsymbol{u}) \widetilde{q}(-\boldsymbol{u}') \widetilde{g}(\boldsymbol{u}, \boldsymbol{u}')$$

$$= \frac{\widetilde{q}(\boldsymbol{u}) \widetilde{q}(-\boldsymbol{u}')}{\widetilde{h}(\boldsymbol{u}) \widetilde{h}^*(-\boldsymbol{u}')} \widetilde{\gamma}_y(\boldsymbol{u}, \boldsymbol{u}') \quad \text{if } \boldsymbol{u}, \boldsymbol{u}' \in \mathcal{D}_{\mathrm{spec}}, \tag{29}$$

$$= 0 \quad \text{otherwise.} \tag{30}$$

Let us remark that $\widetilde{h}(\boldsymbol{u}) \neq 0$ if $\boldsymbol{u} \in \mathcal{D}_{\mathrm{spec}}$ because we have assumed $\mathcal{D}_{\mathrm{spec}} \subseteq \mathcal{D}_{\mathrm{PSF}}$.

Step 2) Kernel f is obviously symmetric. Moreover, it is positive semi-definite, since for any square integrable function ϕ,

$$\iint f(\boldsymbol{r}, \boldsymbol{r}') \phi(\boldsymbol{r}) \phi^*(\boldsymbol{r}')\, d\boldsymbol{r}\, d\boldsymbol{r}' = \int |q \otimes \phi|^2(\boldsymbol{x}) \rho(\boldsymbol{x})\, d\boldsymbol{x} \geq 0.$$

It is easy to check that kernel F is also positive semi-definite. Moreover K_F is bounded, since it is a Hilbert-Schmidt operator. Being bounded and positive semi-definite, K_F admits a unique square root [37, Proposition 3.2.11]. In other words, K_f is uniquely defined given K_F, and equivalently, given the kernel F, there exists a unique kernel f that fulfills (27).

Finally, the knowledge of γ_y uniquely determines F through (29)–(30), which in turn determines f.

Step 3) In the Fourier domain, Eq. (26) reads

$$\widetilde{f}(\boldsymbol{u}, \boldsymbol{u}') = \widetilde{q}(\boldsymbol{u})\, \widetilde{q}(-\boldsymbol{u}')\, \widetilde{\rho}(\boldsymbol{u} + \boldsymbol{u}'). \tag{31}$$

The latter identity shows that $\widetilde{\rho}(\boldsymbol{u}' + \boldsymbol{u}'')$ is identifiable for all $(\boldsymbol{u}', \boldsymbol{u}'')$ such that $\boldsymbol{u}'$ and $-\boldsymbol{u}''$ belong to $\mathcal{D}_{\mathrm{PSF}}$. We thus conclude that the frequency components $\widetilde{\rho}(\boldsymbol{u})$ are identifiable from kernel f, and thus from the data covariance γ_y, for all $\boldsymbol{u} \in \mathcal{D}_{\mathrm{spec}} \ominus \mathcal{D}_{\mathrm{spec}}$.

Appendix C
Gradient of the Kullback-Leibler Divergence

We first note that (18) also reads

$$D(\boldsymbol{\rho}) = \frac{1}{2} \log |\boldsymbol{\Gamma}_z| + \frac{1}{2M} \mathrm{Tr}\left(\boldsymbol{\Gamma}_z^{-1} \boldsymbol{V} \boldsymbol{V}^t \right) + C \tag{32}$$

where C is an irrelevant constant term, and

$$\boldsymbol{V} = (\boldsymbol{v}_1 | \cdots | \boldsymbol{v}_M) \quad \text{with} \quad \boldsymbol{v}_m = \boldsymbol{z}_m - \boldsymbol{\mu}_z. \tag{33}$$

The following identities [38, Sec. 2] will be useful for the derivation of the gradient of (32):

$$\begin{aligned}\nabla_\theta \log|\boldsymbol{A}| &= \mathrm{Tr}\left(\boldsymbol{A}^{-1}(\nabla_\theta \boldsymbol{A})\right)\\ \nabla_\theta(\boldsymbol{A}^{-1}) &= -\boldsymbol{A}^{-1}(\nabla_\theta \boldsymbol{A})\boldsymbol{A}^{-1}\\ \nabla_\theta(\boldsymbol{AB}) &= (\nabla_\theta \boldsymbol{A})\boldsymbol{B} + \boldsymbol{A}(\nabla_\theta \boldsymbol{B})\\ \nabla_\theta \mathrm{Tr}(\boldsymbol{A}) &= \mathrm{Tr}\,(\nabla_\theta \boldsymbol{A})\\ \nabla_\theta(\boldsymbol{A}^{\mathrm{t}}) &= (\nabla_\theta \boldsymbol{A})^{\mathrm{t}}\end{aligned} \tag{34}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices that depend on a real scalar parameter θ. From these relations, we get

$$\begin{aligned}\partial_n D(\boldsymbol{\rho}) = &\frac{1}{2}\mathrm{Tr}\left(\boldsymbol{\Gamma}_z^{-1}(\partial_n \boldsymbol{\Gamma}_z)\right)\\ &+ \frac{1}{2M}\mathrm{Tr}\left((\partial_n \boldsymbol{\Gamma}_z^{-1})\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}} + \boldsymbol{\Gamma}_z^{-1}\partial_n(\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}})\right)\end{aligned} \tag{35}$$

where $\partial_n = \nabla_{\rho_n}$. The gradient of (32) is then defined by

$$\nabla D(\boldsymbol{\rho}) = \mathbf{vect}\left\{\partial_n D(\boldsymbol{\rho})\right\} \tag{36}$$

where $\mathbf{vect}\left\{v_n\right\} = (v_1|\cdots|v_N)^{\mathrm{t}}$. According to (35) and (34), the expressions of $\partial_n \boldsymbol{\Gamma}_z$ and $\partial_n(\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}})$ are required. Let $\boldsymbol{e}_n$ be the nth canonical vector, $\boldsymbol{h}_n$ the nth column of $\boldsymbol{H}$ and $\mathbf{1} = (1\cdots 1)^{\mathrm{t}}$. We get from (16)

$$\partial_n \boldsymbol{\Gamma}_z = \boldsymbol{H}\boldsymbol{R}\boldsymbol{\Gamma}_E \boldsymbol{e}_n \boldsymbol{h}_n^{\mathrm{t}} + (\boldsymbol{H}\boldsymbol{R}\boldsymbol{\Gamma}_E \boldsymbol{e}_n \boldsymbol{h}_n^{\mathrm{t}})^{\mathrm{t}} \tag{37}$$

and from (33):

$$\partial_n(\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}}) = -E_0\left(\boldsymbol{V}\mathbf{1}\boldsymbol{h}_n^{\mathrm{t}} + (\boldsymbol{V}\mathbf{1}\boldsymbol{h}_n^{\mathrm{t}})^{\mathrm{t}}\right). \tag{38}$$

The derivative of the three terms in (35) can now be obtained. On the one hand, elementary manipulations involving the trace operator allow to deduce

$$\mathrm{Tr}\left(\boldsymbol{\Gamma}_z^{-1}(\partial_n \boldsymbol{\Gamma}_z)\right) = 2\,\boldsymbol{e}_n^{\mathrm{t}}\boldsymbol{W}\boldsymbol{h}_n \tag{39}$$

from (37), with $\boldsymbol{W} = \boldsymbol{\Gamma}_E\boldsymbol{R}\boldsymbol{H}^{\mathrm{t}}\boldsymbol{\Gamma}_z^{-1}$. On the other hand, we have from (34) and (37):

$$\mathrm{Tr}\left((\partial_n \boldsymbol{\Gamma}_z^{-1})\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}}\right) = -2\,\boldsymbol{e}_n^{\mathrm{t}}(\boldsymbol{W}\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}}\boldsymbol{\Gamma}_z^{-1})\boldsymbol{h}_n \tag{40}$$

and from (38):

$$\mathrm{Tr}\left(\boldsymbol{\Gamma}_z^{-1}\partial_n(\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}})\right) = -2\,E_0\,\boldsymbol{h}_n^{\mathrm{t}}\boldsymbol{\Gamma}_z^{-1}\boldsymbol{V}\mathbf{1}. \tag{41}$$

According to (35) and (36), we need to vectorize the relations (39), (40) and (41) to obtain the full gradient of (32). In particular, according to the identity

$$((\boldsymbol{A}\,\mathrm{Diag}(\boldsymbol{\rho})\boldsymbol{B}^t)\circ\boldsymbol{I})\mathbf{1} = (\boldsymbol{A}\circ\boldsymbol{B})\boldsymbol{\rho}, \tag{42}$$

we deduce from (39) that

$$\begin{aligned}\mathbf{vect}\left\{\mathrm{Tr}\left(\boldsymbol{\Gamma}_z^{-1}(\partial_n \boldsymbol{\Gamma}_z)\right)\right\} &= 2\left((\boldsymbol{W}\boldsymbol{H})\circ\boldsymbol{I}\right)\mathbf{1}\\ &= 2\big((\boldsymbol{H}^{\mathrm{t}}\boldsymbol{\Gamma}_z^{-1}\boldsymbol{H})\circ\boldsymbol{\Gamma}_E\big)\boldsymbol{\rho}.\end{aligned} \tag{43}$$

Similarly, we obtain after a few manipulations

$$\frac{1}{M}\mathbf{vect}\left\{\mathrm{Tr}\left((\partial_n \boldsymbol{\Gamma}_z^{-1})\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}}\right)\right\} = -\frac{2}{M}\big((\boldsymbol{\Omega}^{\mathrm{t}}\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}}\boldsymbol{\Omega})\circ\boldsymbol{\Gamma}_E\big)\boldsymbol{\rho} \tag{44}$$

and

$$\frac{1}{M}\mathbf{vect}\left\{\mathrm{Tr}\left(\boldsymbol{\Gamma}_z^{-1}\partial_n(\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}})\right)\right\} = -\frac{2}{M}E_0\boldsymbol{\Omega}^t\boldsymbol{V}\mathbf{1} \tag{45}$$

where $\boldsymbol{\Omega} = \boldsymbol{\Gamma}_z^{-1}\boldsymbol{H}$. As a result, the gradient of (35) reads

$$\begin{aligned}&\nabla D(\boldsymbol{\rho}) =\\ &-\left(\left(\boldsymbol{\Omega}^{\mathrm{t}}\left(\frac{1}{M}\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}} - \boldsymbol{\Gamma}_z\right)\boldsymbol{\Omega}\right)\circ\boldsymbol{\Gamma}_E\right)\boldsymbol{\rho} - \frac{1}{M}E_0\boldsymbol{\Omega}^t\boldsymbol{V}\mathbf{1}.\end{aligned} \tag{46}$$

Finally, the following relations hold:

$$\begin{aligned}\boldsymbol{V} &= (\boldsymbol{z}_1 - \widehat{\boldsymbol{\mu}}|\cdots|\boldsymbol{z}_M - \widehat{\boldsymbol{\mu}}) + \boldsymbol{\delta}_\mu\mathbf{1}^{\mathrm{t}},\\ \frac{1}{M}\boldsymbol{V}\boldsymbol{V}^{\mathrm{t}} &= \widehat{\boldsymbol{\Gamma}} + \boldsymbol{\delta}_\mu\boldsymbol{\delta}_\mu^{\mathrm{t}},\end{aligned}$$

which allow us to obtain the gradient expression (20), given that $\nabla J(\boldsymbol{\rho}) = \nabla D(\boldsymbol{\rho}) + \beta\boldsymbol{\rho}$.

ACKNOWLEDGMENT

The authors are grateful to the Associate Editor and to the anonymous reviewers for their valuable comments that helped in improving the quality of this article.

REFERENCES

[1] P. Van den Berg and J. Fokkema, "Removal of undesired wavefields related to the casing of a microwave scanner," *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 1, pp. 187–192, Jan. 2003.

[2] D. C. Munson, J. O'Brien, and K. W. Jenkins, "A tomographic formulation of spotlight mode synthetic aperture radar," *Proc. IEEE*, vol. 71, no. 8, pp. 917–925, Aug. 1983.

[3] M. G. Gustafsson *et al.*, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," *Biophys. J.*, vol. 94, no. 12, pp. 4957–4970, 2008.

[4] O. Haeberlé, K. Belkebir, H. Giovaninni, and A. Sentenac, "Tomographic diffractive microscopy: Basics, techniques and perspectives," *J. Modern Opt.*, vol. 57, no. 9, pp. 686–699, 2010.

[5] A. J. Devaney, "A filtered backpropagation algorithm for diffraction tomography," *Ultrason. Imag.*, vol. 4, pp. 336–350, 1982.

[6] R. Heintzmann and C. G. Cremer, "Laterally modulated excitation microscopy: Improvement of resolution by using a diffraction grating," *Proc. SPIE*, vol. 3568, pp. 185–196, 1999.

[7] M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, "Doubling the lateral resolution of wide-field fluorescence microscopy using structured illumination," *Proc. SPIE*, vol. 3919, pp. 141–150, 2000.

[8] A. Jost, E. Tolstik, P. Feldmann, K. Wicker, A. Sentenac, and R. Heintzmann, "Optical sectioning and high resolution in single-slice structured illumination microscopy by thick slice Blind-SIM reconstruction," *PLoS ONE*, vol. 10, no. 7, pp. 1–10, Jul. 2015.

[9] R. Ayuk *et al.*, "Structured illumination fluorescence microscopy with distorted excitations using a filtered blind-SIM algorithm," *Opt. Lett.*, vol. 38, no. 22, pp. 4723–4726, Nov. 2013.

[10] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer, "Probe retrieval in ptychographic coherent diffractive imaging," *Ultramicroscopy*, vol. 109, no. 4, pp. 338–343, 2009.

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

[12] J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*. Greenwood Village, CO, USA: Roberts & Company, 2007.

[13] E. Mudry *et al.*, "Structured illumination microscopy using unknown speckle patterns," *Nature Photon.*, vol. 6, no. 5, pp. 312–315, 2012.

[14] J. Min *et al.*, "Fluorescent microscopy beyond diffraction limits using speckle illumination and joint support recovery," *Sci. Rep.*, vol. 3, 2013, Art. no. 2075.

[15] J.-E. Oh, Y.-W. Cho, G. Scarcelli, and Y.-H. Kim, "Sub-Rayleigh imaging via speckle illumination," *Opt. Lett.*, vol. 38, no. 5, pp. 682–684, Mar. 2013.
[16] A. Negash *et al.*, "Improving the axial and lateral resolution of three-dimensional fluorescence microscopy using random speckle illuminations," *J. Opt. Soc. Amer.*, vol. 33, no. 6, pp. 1089–1094, Jun. 2016.
[17] S. Labouesse *et al.*, "Joint reconstruction strategy for structured illumination microscopy with unknown illuminations," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 1–14, May 2017.
[18] J. Gateau, T. Chaigne, O. Katz, S. Gigan, and E. Bossi, "Improving visibility in photoacoustic imaging using dynamic speckle illumination," *Opt. Lett.*, vol. 38, no. 23, pp. 5188–5191, 2013.
[19] T. Chaigne *et al.*, "Super-resolution photoacoustic fluctuation imaging with multiple speckle illumination," *Optica*, vol. 3, no. 1, pp. 54–57, Jan. 2016.
[20] J. W. Goodman *Statistical Optics*. New York, NY, USA: John Wiley, 1985.
[21] E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrast imaging," *Opt. Lett.*, vol. 4, pp. 291–293, 1999.
[22] C. Godavarthi *et al.*, "Superresolution with full-polarized tomographic diffractive microscopy," *J. Opt. Soc. Amer. A*, vol. 32, pp. 287–292, 2015.
[23] J. W. Goodman, *Introduction to Fourier Optics* (Physical and Quantum Electronics Series), 2nd ed. San Francisco, CA, USA: McGraw-Hill, 1996.
[24] L. Jofre, M. Hawley, A. Broquetas, E. de los Reyes, M. Ferrando, and A. Elias-Fuste, "Medical imaging with a microwave tomographic scanner," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 3, pp. 303–312, Mar. 1990.
[25] E. Parzen, "A simple proof and some extensions of the sampling theorem," Stanford University, Stanford, CA, USA, Tech. Rep.7, 1956.
[26] F. Macias-Garza, A. Bovik, K. Diller, S. Aggarwal, and J. Aggarwal, "The missing cone problem and low-pass distortion in optical serial sectioning microscopy," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Apr. 1988, pp. 890–893.
[27] S. Kimura and C. Munakata, "Dependence of 3-D optical transfer functions on the pinhole radius in a fluorescent confocal optical microscope," *Appl. Opt.*, vol. 29, no. 20, pp. 3007–3011, Jul. 1990.
[28] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 3rd ed. New York, NY, USA: Springer, 2006.
[29] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.
[30] S. Kullback, *Information Theory and Statistics*. New York, NY, USA: Wiley, 1959.
[31] M. I. Miller and D. L. Snyder, "The role of likelihood and entropy in incomplete-data problems: Applications to estimating point-process intensities and Toeplitz constrained covariances," *Proc. IEEE*, vol. 75, no. 7, pp. 892–906, Jul. 1987.
[32] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statistical Soc. B*, vol. 39, pp. 1–38, 1977.
[33] A. D. Lanterman, "Statistical radar imaging of diffuse and specular targets using an expectation-maximization algorithm," *Proc. SPIE*, vol. 4053, pp. 20–31, 2000.
[34] B. Simon, M. Debailleul, V. Georges, V. Lauer, and O. Haeberlé, "Tomographic diffractive microscopy of transparent samples," *Eur. Phys. J., Appl. Phys.*, vol. 44, no. 1, pp. 29–35, 2008.
[35] P. Comon and L. De Lathauwer, "Algebraic identification of under-determined mixtures," in *Handbook of Blind Source Separation*, P. Comon and C. Jutten, Eds. Oxford, U.K.: Academic, 2010, ch. 9, pp. 325–365.
[36] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)," *Proc. Nat. Acad. Sci.*, vol. 106, no. 52, pp. 22287–22292, 2009.
[37] G. K. Pedersen, *Analysis Now*. New York, NY, USA: Springer-Verlag, 1995.
[38] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Technical University of Denmark, Lyngby, Denmark, Tech. Rep., Nov. 2012.

Authors' photographs and biographies not available at the time of publication.