Fast super-resolved reconstructions in fluorescence random illumination microscopy (RIM)

Guillaume Giroussens, Simon Labouesse, Marc Allain, Thomas Mangeat, Anne Sentenac, Jérôme Idier

Fast super-resolved reconstructions in fluorescence random illumination microscopy (RIM)

June 21, 2023 - Guillaume Giroussens, Simon Labouesse, Marc Allain, Thomas Mangeat, Anne Sentenac, and Jérôme Idier, *Member, IEEE*

***Abstract*—Random Illumination Microscopy (RIM) is a recent super-resolved fluorescence imaging technique in which the sample is recovered iteratively by matching the empirical variance of low-resolution images obtained under random speckled illuminations with the expected variance model. RIM was shown theoretically to achieve a two-fold resolution gain and its performances have proven very robust to deteriorated imaging conditions. However, the reconstruction algorithm suffers from a slow convergence which can prevent the method from being used to its full potential. Here, we show that a simple, non-iterative, linear deconvolution of the empirical standard-deviation image using an appropriate kernel can be sufficient to obtain a satisfactory super-resolved reconstruction of the sample. This first estimate can be further improved with a new accelerated iterative strategy which convergence speed is about two orders of magnitude better than that of variance matching.**

***Index Terms*—Multi-illumination imaging, High-resolution, Cutoff frequency, Second-order statistics, Optical microscopy**

I. Introduction

In fluorescence microscopy, the light intensity recorded by the camera, y, can be modeled as the convolution of a point spread function (PSF) h with the product of the sample fluorescence density ρ with an excitation function E (which depends on the illumination and on the quantum nature of the fluorophores):

$$y = h \otimes (\rho E), \tag{1}$$

where $\otimes$ stands for the convolution operator, either in two or three spatial dimensions. The free-space light propagation from the sample to the camera prevents the wavefield high frequencies from reaching the detector [1, Sec. 3.3.4]. As a result, h has necessarily a bounded Fourier support (denoted by $\mathcal{D}_{\rm PSF}$ in the following). In a two-dimensional imaging configuration (when the sample is assumed to be infinitively thin along the optical axis), $\mathcal{D}_{\rm PSF}$ is a disk of radius $2\mathrm{NA}/\lambda$ where NA is the numerical aperture of the microscope objective and λ the wavelength of the fluorescent light. When accounting for the third dimension, $\mathcal{D}_{\rm PSF}$ becomes a solid torus of the same radius, with maximal height about $\mathrm{NA}^2/2\lambda$ and a missing cone as we move toward the zero transverse frequency [1, Sec. 5.4].

When the excitation E is spatially uniform, as in standard wide-field microscopy, it is clear from (1) that the spatial frequencies of the sample that can be recovered from the image are limited to $\mathcal{D}_{\rm PSF}$. The issue of super-resolution imaging can thus be stated in this way: how can we recover frequencies of the sample beyond $\mathcal{D}_{\rm PSF}$ from images that are frequency limited to $\mathcal{D}_{\rm PSF}$?

The main answers belongs to the *instrumental super-resolution* category [2]. Compared to standard wide-field imaging, they imply non-trivial modifications of the acquisition process, so that the acquired data contain information beyond $\mathcal{D}_{\rm PSF}$. In this paper, we put aside other possibilities that would rely on *a priori* or learned models of the sample fluorescence density to extrapolate the missing spatial frequencies of the fluorophore density from a wide-field image. Indeed, the resort to learned models in super-resolution fluorescence microscopy is a growing tendency, but it is generally considered as an add-on to instrumental super-resolution, in order to simplify or to improve the post-processing step, rather than as a stand-alone possibility [3].

All instrumental super-resolution fluorescence microscopy techniques make use of multiple diffraction-limited images of the sample under different inhomogeneous excitations of the fluorescence and processing (numerically or analogically) the data to form a super-resolved reconstruction of the sample. The basic idea is that the modulation of the fluorescence by the non-uniform excitations brings into $\mathcal{D}_{\rm PSF}$ some of the frequency components of the sample that were outside $\mathcal{D}_{\rm PSF}$. Super-resolved methods in fluorescence microscopy fall essentially in two categories depending on either the fluorescence excitation is assumed known or unknown [4]. The first ones, such as Stimulated Emission Depletion (STED) confocal microscopy, Image Scanning Microscopy (ISM) or Structured Illumination Microscopy (SIM) yield super-resolved images that are linearly linked to the sample with a well defined resolution gain. Yet, they require the knowledge and thus the stringent control of the excitations which may be experimentally challenging. The second ones, such as Super-resolution Optical Fluctuation Imaging (SOFI) or Single Molecule Localization Methods (SMLM) are simpler to implement, as the excitations are minimally controlled, but their reconstructions, based on a non-linear manipulation of the data and often using *a priori* information on the sample, are usually more difficult to interpret. A notable exception in this category is the recent Random Illumination Microscopy (RIM) which, by exciting

This work was supported by the Agence Nationale de la Recherche under grant ANR-20-CE45-0024

G. Giroussens, M. Allain and A. Sentenac are with Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, Marseille, France (e-mail: guillaume.giroussens@fresnel.fr, marc.allain@fresnel.fr, anne.sentenac@fresnel.fr).

S. Labouesse and T. Mangeat are with Toulouse Université, CNRS, Centre de Biologie Intégrative, Toulouse, France (e-mail: simon.labouesse@univ-tlse3.fr, thomas.mangeat@univ-tlse3.fr).

J. Idier is with Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France (e-mail: jerome.idier@ls2n.fr).

the sample with random speckled light, is able to provide super-resolved images that are linearly linked to the sample with a well defined resolution gain, comparable to that of SIM [5], [6].

In practice, RIM consists in recording multiple images under different realizations of speckled illumination patterns obtained, for example, by passing a laser beam through a rotating diffuser. In the existing version of RIM, the super-resolved image is formed from the variance of these low-resolution speckled images [7]. More precisely, the fluorescence density ρ is estimated iteratively so as to minimize the distance between the empirical (experimental) variance and the expected (asymptotic) variance model. This variance model is a quadratic functional of ρ and involves the speckle auto-correlation. If this auto-correlation function corresponds to the point-spread function of the microscope, we established that a bijection exists between the variance image and the sample frequencies in the super-resolved domain $\mathcal{D}_{\mathrm{SR}} = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$ [6]. Indeed, a two-fold resolution gain was observed on calibrated samples and spectacular super-resolved images were obtained on fixed and live biological samples [8].

However, while robust and successful, RIM variance-matching reconstruction scheme suffers from slow convergence and many iterations are necessary to recover the sample highest spatial frequencies in $\mathcal{D}_{\mathrm{SR}}$. This issue appears as a major drawback especially for three-dimensional imaging. In this paper, we propose to facilitate and extend RIM applicability. First, we recall the main features of RIM variance-matching technique. Then, we derive a direct and simple inversion technique that is able to provide, under some approximations, a rapid super-resolved estimate of the sample. Last, we ameliorate this reconstruction with a novel iterative estimator based on the standard deviation of the images. We show that the standard-deviation matching algorithm converges much faster than the variance matching technique.

II. Principles of Random Illumination Microscopy

In this first section, we introduce the notations, assumptions and basic principles of RIM and recall the main features of the variance matching procedure that is used presently to build the sample super-resolved image.

A. Modeling and theoretical resolution bound for RIM

Let us consider M images $(z_1, \ldots, z_M)$ of the sample ρ, acquired following model (1) using M distinct random (typically, fully developed speckled) illuminations E_m. Each observation z_m is also plagued by some noise ε_m, supposed to be additive, so that the true acquisition model for any image is given by

$$z_m = h \otimes (\rho E_m) + \varepsilon_m. \tag{2}$$

For the sake of simplicity, we shall assume that the following (standard) properties hold for the random quantities in (2):

- E and ε are mutually independent and second-order stationary.
- The auto-correlation functions of E and ε, denoted hereafter γ_E and γ_ε, are real-valued and known *a priori* (see for instance [7, Chap. 4] for a justification).

Then, the covariance function of the observations (2) reads

$$\gamma_z(\boldsymbol{r}, \boldsymbol{r}'; \rho) \;=\; \gamma_s(\boldsymbol{r}, \boldsymbol{r}'; \rho) + \gamma_\varepsilon(\boldsymbol{r} - \boldsymbol{r}'). \tag{3}$$

The covariance induced by the modulation with random illuminations *alone* is given by

$$\gamma_s(\boldsymbol{r}, \boldsymbol{r}'; \rho) = \iint h(\boldsymbol{r} - \boldsymbol{x})\rho(\boldsymbol{x})\gamma_E(\boldsymbol{x} - \boldsymbol{x}')\rho(\boldsymbol{x}')h(\boldsymbol{r}' - \boldsymbol{x}') \; d\boldsymbol{x}\, d\boldsymbol{x}' \tag{4}$$

which is a *quadratic* (nonlinear) mapping of the unknown sample ρ. This nonlinear relationship prevents *a priori* a direct identification of the frequency components of ρ. However, it has been proved that, provided the support of the Fourier transform of γ_E (*i.e.*, of the energy spectral density of E) belongs to $\mathcal{D}_{\mathrm{spec}} \subseteq \mathcal{D}_{\mathrm{PSF}}$, any spatial frequency components of the sample inside[1]

$$\mathcal{D}'_{\mathrm{SR}} := \mathcal{D}_{\mathrm{spec}} \ominus \mathcal{D}_{\mathrm{spec}} \tag{5}$$

can be unambiguously defined from the data covariance matrix [5, Proposition 3].

More recently, it has been shown further that the full covariance was not required: the sample frequencies in the domain (5) can also unambiguously be defined from the data variance $\gamma_s(\boldsymbol{r}, \boldsymbol{r}; \rho)$ *alone* [6].

In the present implementations of RIM, the super-resolved image of the sample is obtained iteratively using RIM-VAR, a variance-matching approach[2]. Let us first recall its main features.

B. A variance matching algorithm, RIM-VAR

RIM-VAR consists in minimizing (using a conjugate gradient technique) the cost functional

$$Q(\rho\,;\,\mu) = \|\widehat{v}_s(\boldsymbol{r}) - v_s(\boldsymbol{r}; \rho)\|_2^2 + \mu \, \|\rho(\boldsymbol{r})\|_2^2 \tag{6}$$

where $\|\cdot\|_2$ is the usual norm for the L_2 (square integrable functions) space. The variance model

$$v_z(\boldsymbol{r}; \rho) \equiv \gamma_z(\boldsymbol{r}, \boldsymbol{r}; \rho) = v_s(\boldsymbol{r}; \rho) + v_\varepsilon \tag{7}$$

with $v_\varepsilon \equiv \gamma_\varepsilon(\mathbf{0})$ is derived from (3). In the right hand-side (r.h.s.) of (6), the first term is a discrepancy measure between the variance model v_s and the *empirical* variance $\widehat{v}_s$ computed from the set of speckled images, the second term is a zero-order Tikhonov regularization, with μ its parameter tuning the precision *vs*. robustness-to-noise tradeoff one expects in the retrieved sample [9, Sec. 4], [10, Sec. 5.6].

[1]The notation $\ominus$ (resp. $\oplus$) stands for the Minkowski difference (resp. sum) between two sets:

$$A \ominus B = \{a - b \,|\, a \in A,\, b \in B\}.$$

[2]The RIM-VAR reconstruction algorithm is accessible from the repository `https://github.com/teamRIM/tutoRIM` under the name AlgoRIM.

1) Fast computation for the variance model: A first difficulty in the implementation of RIM-VAR is the computation of the model output v_s for the different estimates of ρ. Indeed, $v_s(\boldsymbol{r};\rho)$ involves a quadruple (in 2D) or sextuple (in 3D) integral that is not numerically tractable as it stands,

$$v_s(\boldsymbol{r};\rho) = \iint t(\boldsymbol{r}-\boldsymbol{x},\boldsymbol{r}-\boldsymbol{x}')\rho(\boldsymbol{x})\rho(\boldsymbol{x}')\,d\boldsymbol{x}\,d\boldsymbol{x}' \quad (8)$$

with the following kernel

$$t(\boldsymbol{x},\boldsymbol{x}') = h(\boldsymbol{x})h(\boldsymbol{x}')\gamma_E(\boldsymbol{x}-\boldsymbol{x}'). \quad (9)$$

This major issue was solved by observing that t is a *positive definite* operator that can be decomposed as,

$$t(\boldsymbol{x},\boldsymbol{x}') = \textstyle\sum_{n\geq 1}\lambda_n\psi_n(\boldsymbol{x})\psi_n(\boldsymbol{x}') \quad (10)$$

where $\{\psi_n \,|\, n\geq 1\}$ is a countable family of orthogonal functions with $\int|\psi_n(\boldsymbol{x})|^2\,d\boldsymbol{x}=1$, and $\{\lambda_n\}_{n\geq 1}$ are real numbers ordered by decreasing magnitude, with $\lim_{n\to\infty}\lambda_n = 0$, see *e.g.,* [11] and Appendix B-A, Eq. (27). The decomposition (10) can be inserted into (8) to provide an equivalent formulation for the variance *via* as a sum of squared convolution integrals [6]

$$v_s(\cdot;\rho) = \sum_{n\geq 1}\left(u_n\otimes\rho\right)^2 \quad (11)$$

with $u_n = \sqrt{\lambda_n}\psi_n$. Although this reformulation does not reduce *per se* the computational burden, an approximation of the variance v_s can be obtained by truncating the sum (11) to the first K leading terms. We found that $K\approx 10$ was generally sufficient to provide an accurate enough variance estimation [8], [12] with a dramatically reduced computation time.

2) A preprocessing of the raw speckled images: Another important feature of RIM-VAR lays in the preprocessing of the data. We applied to each speckled image a Wiener-type pre-filter defined in the Fourier domain as:

$$\widetilde{g} = \frac{\widetilde{h}^*}{|\widetilde{h}|^2+\eta} \quad (12)$$

where " $\widetilde{\cdot}$ " denotes the Fourier transform and $\eta>0$ is a regularization parameter. Note that the support of $\widetilde{g}$ is $\mathcal{D}_{\mathrm{PSF}}$. This preprocessing amounts to replacing the microscope point spread function h by $h\otimes g$ and the noise variance, $\gamma_\varepsilon(\mathbf{0})$ by $(g\otimes\gamma_\varepsilon\otimes g)(\mathbf{0})$. This pre-filtering has three major interests. First it narrows the image point spread function by enhancing the weight of its high frequencies. Second, it reduces the influence of noise by discarding all the noise Fourier components laying outside $\mathcal{D}_{\mathrm{PSF}}$. Third, the spectral decomposition of the pre-filtered variance model appeared to converge more rapidly than that of the standard variance.

Applying RIM-VAR to the pre-filtered microscope images is straightforward with the above changes on the microscope point spread function and noise variance. This preprocessing was applied systematically to all RIM applications and significantly improved the final resolution of the sample reconstruction.

3) Limit of RIM-VAR: RIM-VAR is now routinely used to produce super-resolved images of biological samples, see for instance [8], [13]–[15]. However, it was observed in many examples that RIM-VAR suffered from slow convergence. Indeed, the number of iterations (and associated computation time) required to reach convergence is so high that an early stopping of the iterations is usually needed. This results in a resolution gain that is often not optimal since the high frequencies of the sample (that should be retrieved with RIM) have not reached convergence yet. Note that such a convergence issue often happens with *unscaled* gradient-based iterations, which tend to restore the low-frequency components of the object first, whereas the high-frequency components are much slowly retrieved [16, Sec. 5], [10, Sec. 6.5], [17]. This effect is detrimental to the performances of RIM, especially in 3D imaging.

In the following, we present two novel inversion schemes which significantly accelerate the reconstruction process.

III. A Non-iterative Estimator, RIM-CF

The spectral decomposition (11) of the variance paves the way towards a simple non-iterative estimator of the sample. As we underlined in the previous section, the eigenvalues $\{\lambda_n\}$ are expected to decrease as n grows. The Perron-Frobenius theorem [18, p.18] ensures that the largest eigenvalue λ_1 is of multiplicity one so that the first eigenvector u_1 is dominant in the decomposition (11). We have observed numerically that, for a classical imaging configuration (epi-fluorescence, NA close to one), $\|u_1\|_2^2 \,/\, \|t\|_2^2 \approx 0.9$ *i.e.*, the contribution of the first eigenvector accounts for nearly 90% of the energy of the kernel (9). We thus propose to approximate the variance with the first term of the decomposition,

$$\boldsymbol{v}_s(\cdot;\rho) \;\approx\; \left(u_1\otimes\rho\right)^2. \quad (13)$$

To take the square root of this equation and obtain a linearization of the RIM problem, we prove in Appendix B-A that u_1 is a real non-negative function, so we can safely rewrite Eq. (13) as

$$\sigma_s(\boldsymbol{r};\rho) \approx \left(u_1\otimes\rho\right)(\boldsymbol{r}). \quad (14)$$

At this point, it is worth noting that the pre-filtering of the speckled images breaks the positivity assumption for u_1, unless the pre-filter g is itself a non-negative function, which is not the case of the Wiener-type pre-filter defined by (12). Such a choice of pre-filter is thus *a priori* inappropriate for the non-iterative method, but not redhibitory in practice.

Under the approximation (14), the incoherent RIM microscope can be understood as a linear (spatially invariant) system equipped with a point spread function given by u_1. The frequency components of the sample might then be retrieved within the support of the Fourier transform of u_1. The following proposition ensures that this frequency support extends beyond the diffraction limit.

Property 1 *Let $\widetilde{u}_1$ be the Fourier transform of u_1 [the eigenvector associated with λ_1 in the spectral decomposition* (10)*].*

Under the assumptions given in Appendix A, $\widetilde{u}_1$ is a non-negative function whose support is given by

$$\mathcal{D}_{\mathrm{SR}} = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{spec}}. \tag{15}$$

Proof: See Appendix B-B. ∎

Our mathematical analysis shows that, in first approximation, and provided $\mathcal{D}_{\mathrm{spec}} = \mathcal{D}_{\mathrm{PSF}}$, the standard deviation of speckled images can be modeled as the convolution of the sample with a super-resolved point-spread function whose Fourier support is that of h^2, (*i.e.*, without missing cone in 3D). This result explains the optical sectioning property of the image standard deviation as proposed in the Dynamic Speckle Illumination (DSI) approach presented in [19], [20].

Here, thanks to our knowledge of the convolution kernel u_1, we can leverage on (14) to further improve the sample estimate over DSI. More specifically, we define our *closed-form* RIM estimator, RIM-CF, as

$$\widehat{\rho} \,=\, \mathcal{F}^{-1}\left(\frac{\widetilde{u}_1}{|\widetilde{u}_1|^2+\mu}\mathcal{F}(\widehat{\sigma}_s)\right), \tag{16}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ stand, respectively, for the Fourier transform and its inverse. RIM-CF is basically a non-iterative regularized deconvolution (performed in the Fourier domain) of the *empirical* standard-deviation derived from (7)

$$\widehat{\sigma}_s(\boldsymbol{r}) = \left(\max\left\{0,\, \widehat{v}_z(\boldsymbol{r}) - v_\varepsilon\right\}\right)^{\frac{1}{2}} \tag{17}$$

with $\widehat{v}_z$ the empirical variance built from the set of microscope images $\{z_m\}_m$; in this latter relation, the $\max$ function is required because the fluctuations in the empirical statistic $\widehat{v}_z(\boldsymbol{r})$ may lead to negative values when v_ε is subtracted (if v_ε is unknown, one may use the technique proposed in [12, Sec. IV]). The (Tikhonov) regularization parameter $\mu > 0$ is adjusted to get the expected regularity using RIM-CF [10, Sec. 5.6].

IV. An Accelerated Iterative Estimator, RIM-STD

The RIM-CF estimator presented in the previous section is based on an approximate modeling of the standard deviation of the speckled images. To further improve the reconstruction, it may be interesting to get back to a more accurate model of the asymptotic standard deviation, *i.e.*, many terms in the spectral decomposition of the kernel t given in (10). In this case, a direct inversion is not possible. Yet, bearing in mind the quasi-linear behavior of the standard-deviation with respect to the sample, an iterative reconstruction based on a standard-deviation matching procedure is expected to converge quickly and by all means faster than the variance-matching approach. Note that replacing the variance by the standard deviation is comparable to adopting an *amplitude-based* criterion rather than an *intensity-based* one in a phase retrieval (PR) problem, as recommended in [21], [22], for instance. Indeed, the RIM reconstruction problem shares some common features with PR. The similarities and differences between both problems are detailed in Appendix D.

The standard-deviation matching algorithm, RIM-STD, iteratively estimates the sample so as to minimize a discretized form of the cost functional

$$J(\rho\,;\,\mu) = \frac{1}{2}\|\widehat{\sigma}_z(\boldsymbol{r}) - \sigma_z(\boldsymbol{r}\,;\,\rho)\|_2^2 + \frac{\mu}{2}\,\|\rho(\boldsymbol{r})\|_2^2\,. \tag{18}$$

The minimization of J is obtained using a preconditioned (or, scaled) conjugate-gradient technique, as detailed below. A key point of the algorithm lays in the preconditioning which allows a spectacular improvement of the convergence speed while requiring minimal computation time.

A. Numerical implementation of RIM-STD

The practical implementation of RIM-STD starts with an explicit discretization of the continuous problem at hand, see Appendix C-A for details. Let $\mathbf{Diag}(\cdot)$ a diagonal matrix built from a vector, and $\mathbf{diag}(\cdot)$ a column vector extracted from the main diagonal of a matrix. We also use the notation $\mathbf{BCCB}(\cdot)$ to define a Block Circulant with Circulant Blocks (BCCB) matrix with an input vector as its first row [9, Chap. 4]. Once discretized, the observation model (2) reads

$$\boldsymbol{z}_m = \mathbf{H}\,\mathbf{Diag}(\boldsymbol{E}_m)\boldsymbol{\rho} + \boldsymbol{\varepsilon}_m \tag{19}$$

where the unknown fluorescence map $\boldsymbol{\rho}$, the mth microscope image $\boldsymbol{z}_m$ and the random quantities $\boldsymbol{E}_m$ and $\boldsymbol{\varepsilon}_m$ are all N-dimensional vectors. For the sake of simplicity, we adopt a circular convolution model $\mathbf{H} = \mathbf{BCCB}(\boldsymbol{h})$.

The $N \times N$ covariance matrix corresponding to (3) reads

$$\boldsymbol{\Gamma}_z(\boldsymbol{\rho}) = \boldsymbol{\Gamma}_s(\boldsymbol{\rho}) + v_\varepsilon\mathbf{I} \tag{20}$$

where the first term in the r.h.s. is the covariance matrix associated with the random illumination

$$\boldsymbol{\Gamma}_s(\boldsymbol{\rho}) = \mathbf{H}\,\mathbf{Diag}(\boldsymbol{\rho})\boldsymbol{\Gamma}_E\mathbf{Diag}(\boldsymbol{\rho})\mathbf{H} \tag{21}$$

with $\boldsymbol{\Gamma}_E = \mathbf{BCCB}(\boldsymbol{\gamma}_E)$ the covariance matrix associated with the second-order stationary random vectors $\boldsymbol{E}_m$. The second term in the r.h.s. is the identity matrix $\mathbf{I}$ scaled by a scalar variance $v_\varepsilon \geq 0$ since it is the covariance of the assumed white CCD readout noise. These covariance matrices play the role of a model in our fitting strategy, hence the explicit dependency on the (unknown) sample $\boldsymbol{\rho}$ in our notations. Finally, the discretized version of the variance equation (7) reads

$$\boldsymbol{v}_z(\boldsymbol{\rho}) = \boldsymbol{v}_s(\boldsymbol{\rho}) + v_\varepsilon \quad \text{with} \quad \boldsymbol{v}_s = \mathbf{diag}\left(\boldsymbol{\Gamma}_s\right). \tag{22}$$

The computation of the (super-resolved) solution then relies on the iterative (and local) minimization of a criterion derived from (18)

$$J(\boldsymbol{\rho}\,;\,\mu) = \frac{1}{2}\|\widehat{\boldsymbol{\sigma}}_z - \boldsymbol{\sigma}_z(\boldsymbol{\rho})\|_2^2 + \frac{\mu}{2}\,\|\boldsymbol{\rho}\|_2^2 \tag{23}$$

where $\|\cdot\|_2$ is the usual Euclidian norm in $\mathbb{R}^N$, $\widehat{\boldsymbol{\sigma}}_z$ is the empirical standard deviation of the stack of M microscope images and $\boldsymbol{\sigma}_z(\boldsymbol{\rho})$ is the (pointwise) square-root of the expected variance vector given in (22). To derive the latter quantity, we stress that the actual computation of $\boldsymbol{\Gamma}_s$ is not needed. Instead, and in a strategy similar to the one developed in Sec. II-B1, we provide a low-rank approximation of $\boldsymbol{v}_s(\boldsymbol{\rho})$ directly from

the spectral decomposition of a matrix operator that is analog to the kernel $t(\cdot,\cdot)$ given in (9). See Appendix C-A for details.

In order to minimize (23), we adopt a nonlinear conjugate-gradient (NGC) method, which is an effective, reference tool to tackle large-scale nonlinear least-square problems [16, Sec. 5.2]. In addition, the NGC iterations can be appropriately "scaled" with a preconditioning matrix to accelerate the convergence [16, p.118]. In the framework of RIM-STD, we propose to rely on a preconditioning matrix $\boldsymbol{S}$ that is BCCB and non-negative definite (NND), *cf.* Appendix C where such a choice is justified with respect to the Hessian matrix. Indeed, we would obtain the same preconditionner using a Majorization-Minimization (MM) construction, following the lines of [23, Sec. III.B], once adapted to RIM according to Appendix D.

Starting from a given initial guess $\boldsymbol{\rho}_0 \in \mathbb{R}^N$, the updated sample estimate $\boldsymbol{\rho}_k$ is given by

$$\boldsymbol{\rho}_{k+1} \;=\; \boldsymbol{\rho}_k + \alpha_k \boldsymbol{d}_k \qquad k = 1, 2, \ldots \tag{24}$$

with

$$\boldsymbol{d}_k = \begin{cases} -(\mathbf{S} + \mu \mathbf{I}_N)^{-1} \boldsymbol{g}_k & \text{if } k = 0, \\ -(\mathbf{S} + \mu \mathbf{I}_N)^{-1} \boldsymbol{g}_k + \beta_k \boldsymbol{d}_{k-1} & \text{otherwise} \end{cases} \tag{25}$$

where $\boldsymbol{g}_k$ is the gradient of (23) evaluated at the current estimate $\boldsymbol{\rho}_k$, α_k is the current step-length and $\boldsymbol{S}$ is the BCCB preconditioner that is invariant over the course of iterations. We adopt the Polak-Ribière conjugaison formula (PR$^+$) [16, p.122] for the conjugaison factor β_k:

$$\beta_k = \max\left(\frac{(\boldsymbol{g}_k - \boldsymbol{g}_{k-1})^{\mathrm{t}} \mathbf{S} \boldsymbol{g}_k}{\boldsymbol{g}_{k-1}^{\mathrm{t}} \mathbf{S} \boldsymbol{g}_{k-1}}, 0 \right). \tag{26}$$

The expression of the gradient $\boldsymbol{g}_k$, of the precondtioner $\boldsymbol{S}$ and the derivation of the step-length α_k, are given in Appendix C.

Since this NCG scheme only fulfills a *local* convergence property, the initialization step must be carefully considered. For instance, with a sample that consists in a pair of pointwise emitters, we extensively tested randomly chosen initial points without any sign constraint. Some of them converged to distinct critical points, which were presumably local minimizers. Whereas most of these local solutions were not very different from the global one, some of them were indeed spurious solutions associated with relatively high criterion values. In the same situation, all randomly chosen, non-negative initial points converged to the global minimizer. Whereas we have no formal proof to support that spurious local minima can be avoided with any specific initialization[3], non-negative initial points were found effective in all the practical cases we tested so far, be they simulated or real. Specifically, the non-iterative estimator RIM-CF proposed in the previous section is a natural initial point.

The memory ressource required to build the update (24) is clearly seen from (26): we need to store and manipulate the BCCB matrix $\boldsymbol{S}$ and two (N-dimensional) successive gradients. Matrix-vector products involving $\boldsymbol{S}$ (like the ones in (25) and (26)) are performed in $O(N \log N)$ operations *via* the Fast Fourier Transform (FFT), and the storage of $\boldsymbol{S}$ boils down to an N-dimensional vector precomputed beforehand.

With the low-rank approximation described above, RIM-STD benefits from a reasonably low computational burden per iteration. Actually, this algorithm has a similar computational cost per iteration than RIM-VAR, but it requires far fewer iterations to reach convergence as seen in the next section.

V. Results in Simulation

In this section, we compare the three different reconstruction strategies RIM-VAR, RIM-CF and RIM-STD on synthetic data. Hereafter, we only consider imaging configurations in which $\mathcal{D}_{\mathrm{spec}} = \mathcal{D}_{\mathrm{PSF}}$ so that the sample frequencies should, in theory, be recovered in $\mathcal{D}_{\mathrm{SR}} = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$, (which corresponds to the Fourier support of h^2).

We first illustrate the convergence issue of RIM-VAR and RIM-STD on a toy, one-dimensional (1D) problem where the sample is a Dirac and $\mathcal{D}_{\mathrm{PSF}} = [-\nu_{\mathrm{PSF}}, \nu_{\mathrm{PSF}}]$ where $\nu_{\mathrm{PSF}} = 2/\lambda$ with λ the wavelength of the fluorescent light. We display the recovered Fourier spectrum of the sample for different number of the iterations of RIM-VAR or RIM-STD. It is clear from Fig. 1 that the convergence speed of RIM-VAR is too low to produce the full resolution gain ($\mathcal{D}_{\mathrm{SR}} = [-2\nu_{\mathrm{PSF}}, 2\nu_{\mathrm{PSF}}]$), even after thousands of iterations. In comparison, RIM-STD reaches the resolution limit in less than 100 iterations. Generally, the computation time of RIM-STD to reach convergence is two orders of magnitude lower than that of RIM-VAR.

Then, we analyse the resolution gain of the novel reconstruction schemes RIM-CF and RIM-STD. Fig. 2 displays the 2D reconstructions of a star-like pattern performed from asymptotical (noise-free) standard-deviation images using RIM-CF and RIM-STD. For a 256×256 image, on a standard desktop computer[4], RIM-CF is obtained in about 0.01 seconds, while RIM-STD about 3.5 seconds (100 iterations to reach convergence). The comparison with the deconvolution of the wide-field image shows unambiguously the resolution gain brought by the iterative and non-iterative versions. However, in this ideal noiseless configuration, the error model in RIM-CF takes its toll and RIM-CF is significantly less resolved than RIM-STD. Interestingly, we note that the reconstruction of RIM-CF is better when the microscope images are pre-filtered according to (12), even though the positivity of u_1 is not granted in that case (hence the derivation of (14) from (13)). The absence of visible artefacts in this non-ideal case can be explained by the fact that u_1 is nearly non-negative.

Last, we consider a synthetic 3D sample made of micro-tubules [24] and we compare RIM-CF, RIM-STD and RIM-VAR transverse and axial reconstructions produced from the

[3]From the structure of model (8), we derive that criterion (23) is symmetric about $\rho = 0$, which is a local maximizer (hence not an appropriate start for the algorithm). It follows that if $\widehat{\rho}$ is a local minimizer, then $-\widehat{\rho}$ is. When the algorithm is started with unsigned random initial-guesses, this structural symmetry is probably at the origin of the spurious minima that are found by the iteration. Finally, we recall that model (8) is insensitive to any frequency component of ρ outside the frequency domain $\mathcal{D}_{\mathrm{SR}} := \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{spec}}$. Thus, the gradient of the STD-fitting is also insensitive to those frequency components and any frequency component outside $\mathcal{D}_{\mathrm{SR}}$ in the initial-guess will be ultimately suppressed by the algorithm provided that $\mu > 0$ (which is required in practice).

[4]4 CPUs, Intel(R) Xeon(R) CPU E5-1607 v3 @ 3.10GHz, 16 GB of RAM

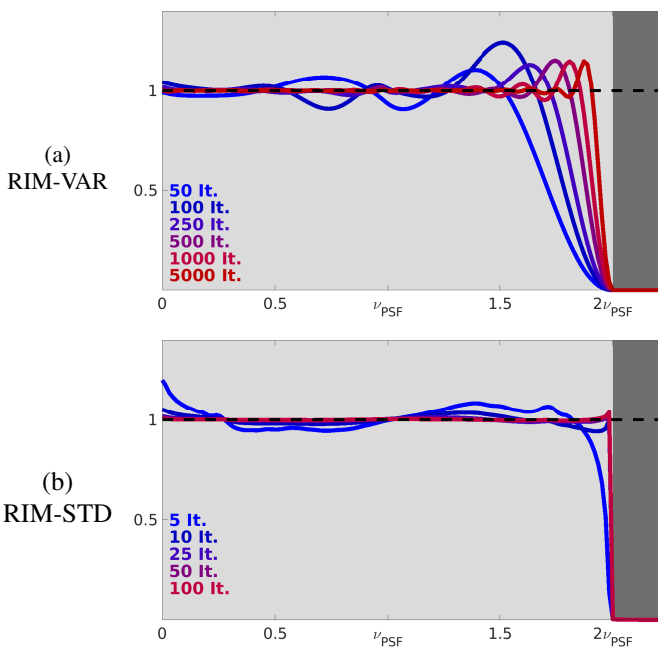

Fig. 1. (a)-(b) Resolution gain of RIM-VAR and RIM-STD (with pre-filtering) as a function of the iteration number, on a toy 1D problem involving a single pointwise emitter, with $\mathcal{D}_{\mathrm{PSF}} = \mathcal{D}_{\mathrm{spec}} = [-\nu_{\mathrm{PSF}}, \nu_{\mathrm{PSF}}]$, so $\mathcal{D}_{\mathrm{SR}} = [-2\nu_{\mathrm{PSF}}, 2\nu_{\mathrm{PSF}}]$. The function $\tilde{\rho}$ being symmetric, we only represent its positive frequency components. The portion in dark gray is beyond $2\boldsymbol{\nu}_{\mathrm{PSF}}$, *i.e.*, out of reach for RIM.

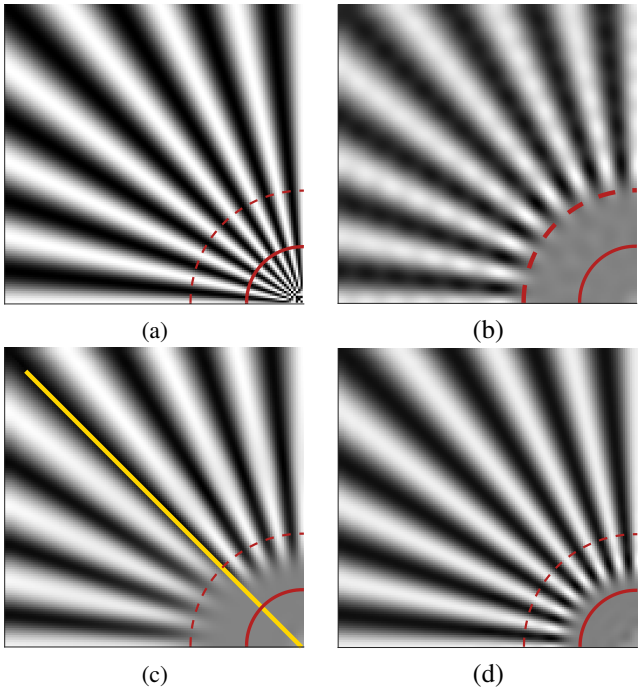

Fig. 2. Proof of the super-resolution induced by RIM in a 2D asymptotical noiseless case. (a) True object. (b) Wide-field deconvolution. Obtained by taking E equal to a constant in (1). (c) Reconstruction using the non-iterative estimator RIM-CF, with (lower) and without (upper) data pre-filtering. (d) Reconstruction using RIM-STD The dashed and solid lines are indicators of the resolution level achievable with a cutoff frequency ν_{PSF} and $2\nu_{\mathrm{PSF}}$ respectively.

asymptotic variance image and from an empirical (using 1000 speckled images) variance image. Note that in this 3D context, as predicted by the analysis in [5], RIM should provide optical sectioning since the accessible sample Fourier domain $\mathcal{D}_{\mathrm{SR}} = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$ does not exhibit any *missing cone* [1, Sec. 5.4]. For this data stack $512 \times 256 \times 128$, the computation time for RIM-CF is around 0.6 seconds, while for both RIM-STD and RIM-VAR, the computation time is 25 seconds, corresponding to the time it took for RIM-STD to reach convergence. Both these iterative approaches start from a constant initial guess showing that RIM-STD dramatically outperforms RIM-VAR in term of convergence speed. Different observations can be drawn from Fig. 3.

On asymptotic data, RIM-STD achieves RIM theoretical resolution bounds Fig. 3(b) both in the transverse and axial directions. RIM-CF Fig. 3(e) is not as efficient but it improves signifciantly the sample estimation as compared to the wide-field deconvolution or the raw standard-deviation image, Fig. 3(c,d).

In more realistic conditions, using 1000 noisy speckled images to compute the empirical variance, RIM-STD exhibits a loss of resolution compared to the asymptotic regime while RIM-CF is minimally affected. As a result, RIM-STD reconstruction is now very similar to RIM-CF, see Fig. 3(f,g). In this case, the model error due to the truncation of (11) in RIM-CF may be negligible compared to the model error consisting in estimating the variance with an asymptotic expression while the actual data is an empirical variance formed with a finite number of illuminations. In practice, the data are also deteriorated with Poisson or electronic noise and RIM-STD and RIM-CF are going to behave differently under these various sources of error. A theoretical analysis being out of reach, a more pragmatic approach is to investigate how these methods behave *via* Monte-Carlo simulations. Such an analysis is reported in Appendix E, from which we conclude that, while RIM-STD is clearly superior to RIM-CF in ideal conditions, RIM-STD resolution and overall statistical behavior is only slightly better than that of RIM-CF in noisy, non asymptotic configurations.

VI. Conclusion

The main contribution of this paper has been to show that, with a slight model error, the standard deviation of microscope images obtained under random speckled illuminations can be written as the sample convolved with a super-resolved point spread function. This result leads to a very simple and fast RIM reconstruction scheme, RIM-CF, which proved efficient in various imaging configurations. To further improve the resolution gain, we also developed an iterative estimator based on a standard-deviation matching procedure, avoiding the model error. We have shown that this novel algorithm, RIM-STD, was able to provide resolution gains similar to the expected theoretical ones with less than one hundred iterates. This fast convergence is due to the quasi-linear behavior of the image standard deviation with respect to the sample and the use of an appropriate preconditionner. We believe that the speed and the performances of RIM-STD and RIM-CF are major assets for extending the applicability of RIM and its ease of use. The two-fold resolution improvement and

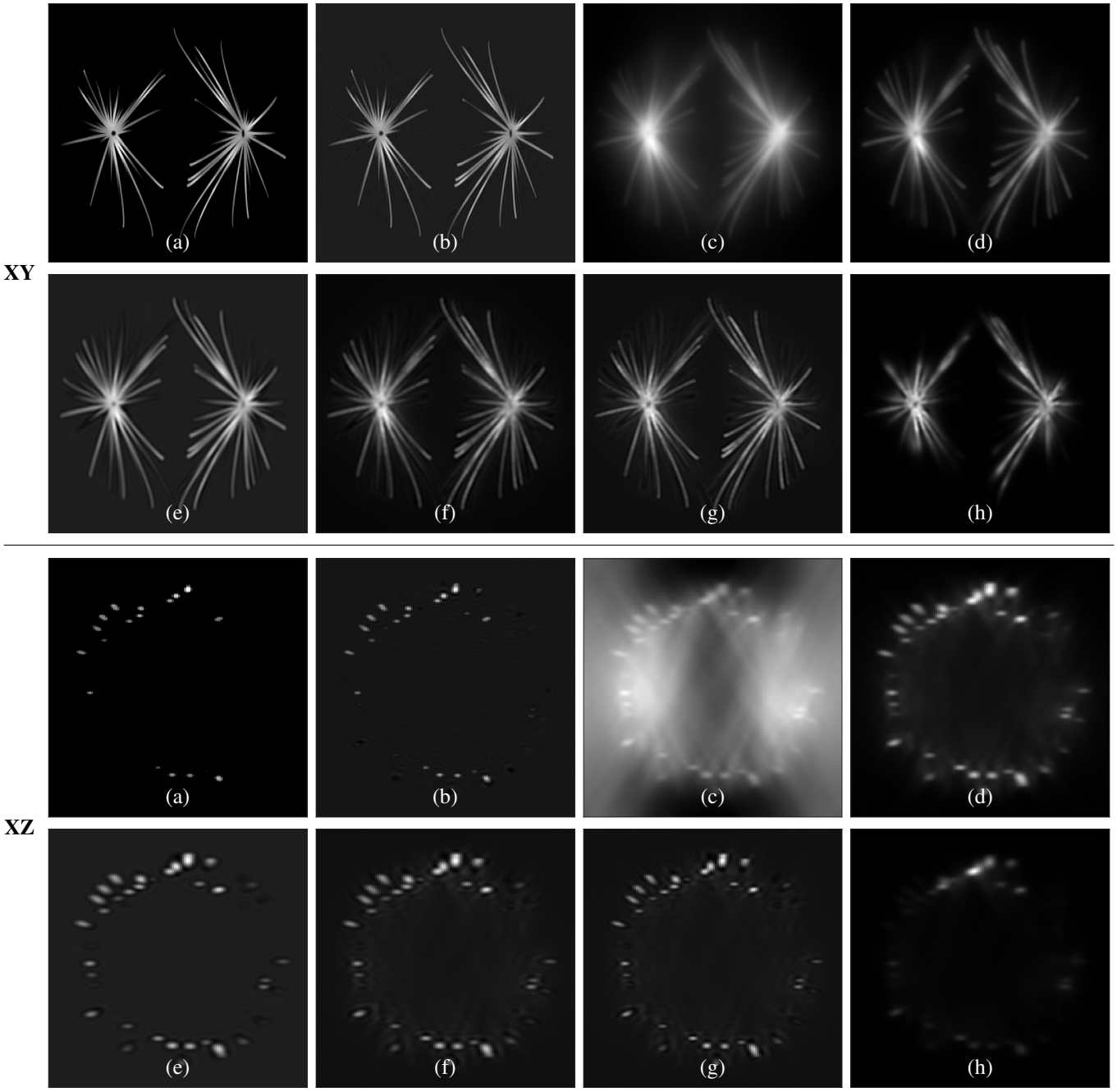

Fig. 3. RIM simulation with an object representing microtubules (*cf.* [24]). The top and bottom halves represent XY and XZ cuts, respectively. (a) Ground truth. (b) Ground-truth bandwidth limited image, matching the inversion of the asymptotical, standard deviation with RIM-STD. (c) Average of $M = 1000$ speckled microscope images. (d) Empirical standard deviation of the same $M = 1000$ speckled images. (e) Non iterative estimation RIM-CF, with the asymptotic standard deviation (case corresponding to $M = \infty$). (f)-(h) Reconstruction with RIM-CF, RIM-STD and RIM-VAR respectively, from the empirical standard deviation image shown in (d). In all simulations, we set $\mathcal{D}_{\mathrm{spec}} = \mathcal{D}_{\mathrm{PSF}}$, so the identifiability domain is $\mathcal{D}'_{\mathrm{SR}} = \mathcal{D}_{\mathrm{SR}} = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$.

optical sectioning of the reconstructions, together with the low computational burden, make RIM-CF and RIM-STD the methods of choice for three-dimensional imaging.

Appendix A
Adopted Assumptions

The scope of this paper is restricted to incoherent imaging, which implies the following standard assumptions concerning the quantities in the observation model (2)

- $\mathbf{A}_1$) ρ is both integrable and square-integrable and takes finite, real non-negative values over $\mathbb{R}^d$;
- $\mathbf{A}_2$) h is a (real-valued) non-negative, symmetric function that is non-zero almost everywhere;
- $\mathbf{A}_3$) h has a finite energy, which implies the existence of its Fourier transform $\widetilde{h}$.
- $\mathbf{A}_4$) $\widetilde{h}$ is a non-negative function with a bounded support $\mathcal{D}_{\mathrm{PSF}} := \{\boldsymbol{x} \in \mathbb{R}^d \mid \widetilde{h}(\boldsymbol{x}) \neq 0\}$.
- $\mathbf{A}_5$) E and ε are mutually independent and second-order stationary.

$\mathbf{A}_6$) The auto-correlation functions of E and ε, denoted hereafter γ_E and γ_ε, are real-valued positive functions known *a priori* (see for instance [7, Chap. 4] for a justification).

$\mathbf{A}_7$) The support $\mathcal{D}_{\rm spec}$ of the Fourier transform of γ_E (*i.e.*, of the spectral energy density) is bounded and such that $\mathcal{D}_{\rm spec} \subseteq \mathcal{D}_{\rm PSF}$.

We note that assumptions $\mathbf{A}_1$-$\mathbf{A}_2$ are specific to the case of incoherent imaging, and are then slightly more restrictive than the ones given in [5, Sec. II], *i.e.*, the main identification results of RIM [5, Prop. 3] and [6, Th. 2] hold here.

APPENDIX B
PROPERTIES OF EIGENVECTOR u_1

A. Positivity

In order to prove that u_1 is positive, let us first define an integral operator associated with kernel t introduced in (9):

$$T : \Phi \mapsto \int t(\boldsymbol{x}, \boldsymbol{x}')\Phi(\boldsymbol{x}')\, d\boldsymbol{x}. \tag{27}$$

Lemma 1 below shows that T has many good properties. In particular, it is a compact self-adjoint operator, so the spectral theorem [25, Chap. 3] applies. Hence, T can be decomposed on the orthonormal basis of its eigenvectors *via* a decomposition of its kernel t, *i.e.*, there exists a countable family $\{\psi_n \,|\, n \geq 1\}$ such that

$$t(\boldsymbol{x}, \boldsymbol{x}') = \sum_{n \geq 1} \lambda_n \psi_n(\boldsymbol{x}) \psi_n(\boldsymbol{x}') \tag{28}$$

where ψ_n is an orthogonal basis of eigenvectors for the operator T and λ_n are the associated eigenvalues. Since the operator is self-adjoint, $\{\lambda_n\}_{n\geq 1}$ are real-valued. The eigenvalues, ordered by decreasing magnitude, are such that the limit of the sequence is zero as n grows to infinity.

The Krein-Rutman theorem can be used in order to prove the positivity of $u_1 := \sqrt{\lambda_1}\psi_1$. For the sake of completeness, this theorem is stated below with our notations.

Theorem 1 *[18, Th. 19.2]. Let X be a Banach space and $K \subseteq X$ a convex cone so that $K \ominus K$ is dense in X (K is a total cone). Let $T : X \mapsto X$ be a positive (i.e., $T(K) \subseteq K$) compact operator with a positive spectral radius $r(T)$. Then $r(T)$ is an eigenvalue of T and the associated eigenvector lies within $K \setminus \{0\}$.*

In what follows, we check that the assumptions of the Krein-Rutman theorem hold for the integral operator (27). Let us first review the key properties of kernel t defined in (9).

Lemma 1 *$t : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ is symmetric, i.e., $t(\boldsymbol{x}, \boldsymbol{x}') = t(\boldsymbol{x}', \boldsymbol{x})$, and it is strictly positive almost everywhere. Furthermore, we have $\iint |t(\boldsymbol{x}', \boldsymbol{x}')|^2\, d\boldsymbol{x}\, d\boldsymbol{x}' < \infty$.*

Proof: With the assumptions on h and γ_E in mind (see $\mathbf{A}_2$ and $\mathbf{A}_6$ in Appendix A, the first part of the lemma are direct consequences of the definition of the kernel given in (9). Finally, E being a second-order stationary process ($\mathbf{A}_5$), its auto-correlation is such that $|\gamma_E(\boldsymbol{x})| \leq \gamma_E(\mathbf{0}) < \infty$, *i.e.*, γ_E is bounded. The assumption $h \in L_2(\mathbb{R}^d)$ then leads to $t \in L_2(\mathbb{R}^d \times \mathbb{R}^d)$, which completes the proof. ∎

We are now in position to prove that Theorem 1 applies. In particular,

1) Let $X = L_2(\mathbb{R}^d)$ be the set of square integrable functions over $\mathbb{R}^d$. X is an Hilbert space, and thus is also a Banach space. Let K be the set of non-negative functions in X. K is a total cone since $K \ominus K = X$ as any function is the difference of its positive and negative parts: $f = f^+ - f^-$ with $f^+ = \max(0, f)$ and $f^- = -\min(0, f)$, both being non-negative functions.
2) As a consequence of Lemma 1, T is an endomorphism of $L_2(\mathbb{R}^d)$.
3) Since t is a symmetric kernel, T is self-adjoint and can be decomposed on a countable basis of eigenvectors.
4) Consequently, the spectral radius of T is by definition $r(T) := \sup\{|\lambda_n| \mid n > 1\}$. Without loss of generality, the eigenvalues can be ordered by decreasing magnitude, so that $r(T) = |\lambda_1| \geq 0$. Furthermore, for Hilbert-Schmidt operators, the energy of the kernel is finite and given by $\|t\|_2^2 = \sum_{n\geq 1} \lambda_n^2$. In particular, $|\lambda_1| > 0$ when $\|t\|_2 > 0$, which is the case according to Lemma 1.
5) Finally, t being a non-negative integral kernel, T is a positive operator. Jointly with the previous point, this ensures that $\lambda_1 > 0$. In addition, since T is an endomorphism, any L_2 function applied to T is also in an L_2, which implies $T(K) \subset K$.

The conditions in Theorem 1 being met, the first eigenvector $u_1 := \sqrt{\lambda_1}\psi_1$ is in $K \setminus \{0\}$, *i.e.*, it is a non-negative function.

Let us stress that we were not able to show the desirable property that the spectral radius is associated to a single eigenvector. In a finite-dimensional setting (*i.e.*, when t is a matrix), such a result holds as a consequence of the Perron-Frobenius theorem. In the infinite-dimensional setting, a stronger version of the Krein-Rutman theorem would be necessary, but the existing ones (*e.g.,* [18, Chap.6]) do not apply to the $L_2(\mathbb{R})$-space considered here.

B. Proof of Propositiion 1

Here, we prove that the Fourier support of u_1 is exactly $\mathcal{D}_{\rm SR}$. We first note that $\widetilde{u}_n := \mathcal{F}(u_n)$ are the eigenvectors of an integral operator, with a kernel given by

$$\widetilde{t}(\boldsymbol{\nu}, \boldsymbol{\nu}') = \iint t(\boldsymbol{x}, \boldsymbol{x}') e^{-2i\pi(\boldsymbol{\nu}\cdot\boldsymbol{x} - \boldsymbol{\nu}'\cdot\boldsymbol{x}')}\, d\boldsymbol{x}\, d\boldsymbol{x}'. \tag{29}$$

Inserting (9) into (29) leads to

$$\widetilde{t}(\boldsymbol{\nu}, \boldsymbol{\nu}') = \int \widetilde{h}(\boldsymbol{\xi} - \boldsymbol{\nu})\widetilde{h}(\boldsymbol{\xi} - \boldsymbol{\nu}')\widetilde{\gamma}_E(\boldsymbol{\xi})\, d\boldsymbol{\xi}. \tag{30}$$

When $\widetilde{h}$ and $\widetilde{\gamma}_E$ are positive (as it is the case here, see Appendix A), $\widetilde{t}$ is also positive and the non-negativity of $\widetilde{u}_1$ follows from the Krein-Rutman theorem with arguments similar to the ones used in Appendix B-A. We can now show the following result.

Lemma 2 *For all $\boldsymbol{\nu}$, the support of $\widetilde{f}_{\boldsymbol{\nu}} : \boldsymbol{\nu}' \mapsto \widetilde{t}(\boldsymbol{\nu}, \boldsymbol{\nu}')$ is*

$$\mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}}} = \big(\mathcal{D}_{\rm spec} \cap (\{\boldsymbol{\nu}\} \oplus \mathcal{D}_{\rm PSF})\big) \ominus \mathcal{D}_{\rm PSF}.$$

Proof: According to (30), $\widetilde{f}_{\boldsymbol{\nu}}$ is the correlation between $\widetilde{h}$ and $\widetilde{g}_{\boldsymbol{\nu}} : \boldsymbol{\xi} \mapsto \gamma_E(\boldsymbol{\xi})\widetilde{h}(\boldsymbol{\xi} - \boldsymbol{\nu})$, given $\boldsymbol{\nu} \in \mathbb{R}^d$. The support of $\widetilde{g}_{\boldsymbol{\nu}}$ is the intersection of the supports of γ_E and $\widetilde{h}(\cdot - \boldsymbol{\nu})$:

$$\mathcal{D}_{\widetilde{g}_{\boldsymbol{\nu}}} = \mathcal{D}_{\rm spec} \cap (\{\boldsymbol{\nu}\} \oplus \mathcal{D}_{\rm PSF}).$$

Then, as $\widetilde{f}_{\boldsymbol{\nu}}$ is the correlation of two non-negative functions $\widetilde{h}$ and $\widetilde{g}_{\boldsymbol{\nu}}$, its support is the Minkowski difference $\mathcal{D}_{\widetilde{g}_{\boldsymbol{\nu}}} \ominus \mathcal{D}_{\rm PSF}$, hence the result. ∎

The following corollary is a direct consequence of Lemma 2.

Corollary 1 *We have*

(i) $\forall \boldsymbol{\nu} \notin \mathcal{D}_{\rm SR}$, $\mathcal{D}_{\rm spec} \cap (\{\boldsymbol{\nu}\} \oplus \mathcal{D}_{\rm PSF}) = \varnothing$, *so* $\mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}}} = \varnothing$.

(ii) $\mathcal{D}_{\rm spec} \subseteq \mathcal{D}_{\rm PSF}$ *(see assumption* $\mathbf{A}_7$ *in App. A)* $\implies \mathcal{D}_{\rm spec} \cap (\{\mathbf{0}\} \oplus \mathcal{D}_{\rm PSF}) = \mathcal{D}_{\rm spec}$, *so* $\mathcal{D}_{\widetilde{f}_{\mathbf{0}}} = \mathcal{D}_{\rm SR}$.

We can now proceed to the proof of Proposition 1. First, we have

$$\forall \boldsymbol{\nu},\ \big[\widetilde{T}\widetilde{u}_1\big](\boldsymbol{\nu}) = \lambda_1 \widetilde{u}_1(\boldsymbol{\nu}) = \int \widetilde{t}(\boldsymbol{\nu}, \boldsymbol{\nu}')\widetilde{u}_1(\boldsymbol{\nu}')\, d\boldsymbol{\nu}'. \quad (31)$$

as a direct implication that $\widetilde{u}_1$ is an eigenvector of integral operator $\widetilde{T}$. Then, we consider first the case $\boldsymbol{\nu} \notin \mathcal{D}_{\rm SR}$: we already noticed from Corollary 1(i) that the kernel of this integral is uniformly zero, as its support is empty. Since λ_1 is not zero, (31) leads to $\forall \boldsymbol{\nu} \notin \mathcal{D}_{\rm SR}, \widetilde{u}_1(\boldsymbol{\nu}) = 0$. In particular, this means that the support of $\widetilde{u}_1$ (denoted $\mathcal{D}_{\widetilde{u}_1}$ in the sequel) is such that $\mathcal{D}_{\widetilde{u}_1} \subseteq \mathcal{D}_{\rm SR}$.

Now that we established that $\mathcal{D}_{\widetilde{u}_1}$ is smaller than $\mathcal{D}_{\rm SR}$, we aim at showing that $\widetilde{u}_1$ cannot vanish inside $\mathcal{D}_{\rm SR}$. Let $\boldsymbol{\nu}_0 \in \mathcal{D}_{\rm SR}$ be such that $\widetilde{u}_1(\boldsymbol{\nu}_0) = 0$. Then, we have from (31)

$$\lambda_1 \widetilde{u}_1(\boldsymbol{\nu}_0) = \int \widetilde{t}(\boldsymbol{\nu}_0, \boldsymbol{\nu}')\widetilde{u}_1(\boldsymbol{\nu}')\, d\boldsymbol{\nu}' = 0.$$

As both $\widetilde{t}$ and $\widetilde{u}_1$ are positive functions, this implies that $\widetilde{u}_1$ must vanish wherever the function $\widetilde{f}_{\boldsymbol{\nu}_0} : \boldsymbol{\nu}' \mapsto \widetilde{t}(\boldsymbol{\nu}_0, \boldsymbol{\nu}')$ is not zero. Following Lemma 2, this means that $\widetilde{u}_1$ vanishes on the domain $\mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}_0}}$. We now prove the following result.

Lemma 3 $\forall \boldsymbol{\nu}_0 \in \mathcal{D}_{\rm SR}$, *we have* $\mathbf{0} \in \mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}_0}}$.

Proof: By definition of $\mathcal{D}_{\rm SR}$ given by (15), $\exists\, \boldsymbol{\xi}_1 \in \mathcal{D}_{\rm PSF}, \boldsymbol{\xi}_2 \in \mathcal{D}_{\rm spec}$ such that $\boldsymbol{\nu}_0 = \boldsymbol{\xi}_1 - \boldsymbol{\xi}_2 \in \mathcal{D}_{\rm SR}$, or equivalently

$$\mathbf{0} = (\boldsymbol{\nu}_0 - \boldsymbol{\xi}_1) + \boldsymbol{\xi}_2. \quad (32)$$

We then deduce that

1) With h real (assumption $\mathbf{A}_2$), $\mathcal{D}_{\rm PSF}$ is symmetric and $-\boldsymbol{\xi}_1 \in \mathcal{D}_{\rm PSF}$, so $(\boldsymbol{\nu}_0 - \boldsymbol{\xi}_1) \in \{\boldsymbol{\nu}_0\} \oplus \mathcal{D}_{\rm PSF}$.
2) With γ_E real (assumption $\mathbf{A}_6$), $\mathcal{D}_{\rm spec}$ is symmetric and $-\boldsymbol{\xi}_2 \in \mathcal{D}_{\rm spec}$. As a consequence, $(\boldsymbol{\nu}_0 - \boldsymbol{\xi}_1) = -\boldsymbol{\xi}_2 \in \mathcal{D}_{\rm spec}$.
3) $\mathbf{A}_7 \implies \mathcal{D}_{\rm spec} \subseteq \mathcal{D}_{\rm PSF}$, so $-\boldsymbol{\xi}_2 \in \mathcal{D}_{\rm PSF}$.

We have $(\boldsymbol{\nu}_0 - \boldsymbol{\xi}_1) \in \mathcal{D}_{\rm spec} \cap (\{\boldsymbol{\nu}_0\} \oplus \mathcal{D}_{\rm PSF})$ according to 1) and 2), and $(-\boldsymbol{\xi}_2) \in \mathcal{D}_{\rm PSF}$ according to 3). Since (32) is equivalent to

$$\mathbf{0} = (\boldsymbol{\nu}_0 - \boldsymbol{\xi}_1) - (-\boldsymbol{\xi}_2), \quad (33)$$

by definition of $\mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}_0}}$, (33) is equivalent to $\mathbf{0} \in \mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}_0}}$. ∎

If $\widetilde{u}_1$ vanishes over the domain $\mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}_0}}$, we have in particular $\widetilde{u}_1(\mathbf{0}) = 0$ since $\mathbf{0} \in \mathcal{D}_{\widetilde{f}_{\boldsymbol{\nu}_0}}$ from the lemma above. We can now use the same derivation with $\boldsymbol{\nu}_0 = \mathbf{0}$. Then, Corollary 1(ii) yields that $\widetilde{u}_1$ vanishes over $\mathcal{D}_{\rm SR}$, which contradicts the fact that $\widetilde{u}_1$ is the first eigenvector of a nonzero integral operator. Thus, there is no frequency $\boldsymbol{\nu}_0 \in \mathcal{D}_{\rm SR}$ such that $\widetilde{u}_1(\boldsymbol{\nu}_0) = 0$, and $\mathcal{D}_{\rm SR} \subseteq \mathcal{D}_{\widetilde{u}_1}$.

Since $\mathcal{D}_{\widetilde{u}_1} \subseteq \mathcal{D}_{\rm SR}$ and $\mathcal{D}_{\rm SR} \subseteq \mathcal{D}_{\widetilde{u}_1}$, we have $\mathcal{D}_{\widetilde{u}_1} = \mathcal{D}_{\rm SR}$.

APPENDIX C
NUMERICAL IMPLEMENTATION OF RIM-STD

A. Discretization of the problem

Let us discretize a d-dimensional space variable $\boldsymbol{r} \in \mathbb{R}^d$ on a regular grid $\mathcal{G}$. This grid consists in N elements (*i.e.*, segments, pixels or voxels) indexed by their spatial coordinate vector $\boldsymbol{r}_n$, $n = 0, \ldots, N-1$. For any band-limited function $f : \mathbb{R}^d \to \mathbb{R}$, the N element vector $\boldsymbol{f} := \big(f(\boldsymbol{r}_1), \ldots, f(\boldsymbol{r}_N)\big)^{\rm t}$ defines a lossless sampling of f as long as each discretization step meets the Nyquist criterion. In a similar way, any kernel $t : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ can be discretized to provide $N \times N$ matrices. In particular, it is easy to discretize kernel (9) as a definite non-negative matrix

$$\mathbf{T} = \mathbf{Diag}(\boldsymbol{h})\, \boldsymbol{\Gamma}_E\, \mathbf{Diag}(\boldsymbol{h}) \quad (34)$$

with $\boldsymbol{h}$ the discrete version of the PSF h and $\boldsymbol{\Gamma}_E = \mathbf{BCCB}(\boldsymbol{\gamma}_E)$, a BCCB matrix defined by the auto-correlation of the speckle. After introducing shifting matrices $\mathbf{P}_n$, performing circular shifts so that pixel n is now at position 1, the discretized version of the variance (7) reads $\boldsymbol{v}_z = (v_{z;1}, \ldots, v_{z;N})^{\rm t}$ with $v_{z;n} := v_z(\boldsymbol{r}_n)$ given by

$$v_{z;n} = \boldsymbol{\rho}^{\rm t}\mathbf{P}_n^{\rm t}\mathbf{T}\mathbf{P}_n\boldsymbol{\rho} + v_\varepsilon = \boldsymbol{\rho}^{\rm t}\mathbf{T}_n\boldsymbol{\rho} + v_\varepsilon \quad (35)$$

with $\mathbf{T}_n := \mathbf{P}_n^{\rm t}\mathbf{T}\mathbf{P}_n$. For the sake of notational convenience, we will drop the subscript z hereafter. Matrix $\mathbf{T}$ can be decomposed on its basis of eigenvectors $\{\psi_1, \ldots, \psi_N\}$ as

$$\mathbf{T} = \sum_{k=1}^{K} \lambda_k \boldsymbol{\psi}_k \boldsymbol{\psi}_k^{\rm t} = \sum_{k=1}^{K} \boldsymbol{u}_k \boldsymbol{u}_k^{\rm t} \quad (36)$$

where $K \leq N$ is the rank of $\mathbf{T}$, $\lambda_1, \ldots, \lambda_K > 0$, $\mathbf{T}\boldsymbol{\psi}_k = \lambda_k \boldsymbol{\psi}_k$, and $\boldsymbol{u}_k = \sqrt{\lambda_k}\boldsymbol{\psi}_k$. Using the shift properties of the eigenvectors of $\mathbf{T}_n$, the variance vector can be given the following expression

$$\boldsymbol{v} = \sum_{k=1}^{K} (\mathbf{U}_k\boldsymbol{\rho}) \odot (\mathbf{U}_k\boldsymbol{\rho}) + v_\varepsilon \quad (37)$$

with $\mathbf{U}_k = \mathbf{BCCB}(\boldsymbol{u}_k)$ and $\odot$ the Hadamard (*e.g.,* entry-wise) product. This equation is the discretized counterpart of (11). Criterion J given in (23) then reads

$$J(\boldsymbol{\rho}) = \frac{1}{2} \sum_{n=1}^{N} \Big(\sigma_n(\boldsymbol{\rho}) - \widehat{\sigma}_n\Big)^2 + \frac{\mu}{2} \|\boldsymbol{\rho}\|_2^2 \quad (38)$$

with $\sigma_n(\boldsymbol{\rho}) = (\boldsymbol{\rho}^{\rm t}\mathbf{T}_n\boldsymbol{\rho} + v_\varepsilon)^{\frac{1}{2}}$ and $\widehat{\sigma}_n := \sqrt{\widehat{v}_n}$. We recall that in practice, a reduced-rank approximation of $\mathbf{T}$ is usually performed with ten to twenty eigenvectors, resulting in a fast yet accurate evaluation of (35).

B. Computation of the gradient

We have

$$\frac{\partial J}{\partial \boldsymbol{\rho}} = \sum_{n=1}^{N} \frac{\partial \sigma_n}{\partial \boldsymbol{\rho}}(\boldsymbol{\rho})\Big(\sigma_n(\boldsymbol{\rho}) - \hat{\sigma}_n\Big) + \mu\boldsymbol{\rho}$$

$$= \sum_{n=1}^{N} \left(\Big(1 - \frac{\widehat{\sigma}_n}{\sigma_n(\boldsymbol{\rho})}\Big)\mathbf{T}_n\boldsymbol{\rho} \right) + \mu\boldsymbol{\rho}. \tag{39}$$

$$= \sum_{k=1}^{K} \left[\mathbf{U}_k^{\mathrm{t}}\mathbf{U}_k\boldsymbol{\rho} \odot \frac{\boldsymbol{\sigma}(\boldsymbol{\rho}) - \widehat{\boldsymbol{\sigma}}}{\boldsymbol{\sigma}(\boldsymbol{\rho})} \right] + \mu\boldsymbol{\rho}, \tag{40}$$

The division in this expression is done entrywise. We recall that thanks to the contribution of the electronic noise, it is guarantee that $\boldsymbol{\sigma}(\boldsymbol{\rho})$ only has non zero entries, so the gradient is well defined. Furthermore all operations involved in this computation can be done entry-wise in either the direct or the Fourier space, so the computation can be done efficiently.

C. Specification of a preconditioner

Differentiating expression (39) yields the Hessian matrix:

$$\mathbf{H}(\boldsymbol{\rho}) = \sum_{n=1}^{N} \left(\frac{\sigma_n(\boldsymbol{\rho}) - \widehat{\sigma}_n}{\sigma_n(\boldsymbol{\rho})}\mathbf{T}_n - \frac{\widehat{\sigma}_n \mathbf{T}_n\boldsymbol{\rho}\boldsymbol{\rho}^{\mathrm{t}}\mathbf{T}_n}{\sigma_n(\boldsymbol{\rho})^3} \right) + \mu\mathbf{I}_N$$

$$= \mathbf{S} + \mu\mathbf{I}_N - \sum_{n}^{N} \frac{\widehat{\sigma}_n}{\sigma_n(\boldsymbol{\rho})}\Big(\frac{1}{v_n(\boldsymbol{\rho})}\mathbf{T}_n\boldsymbol{\rho}\boldsymbol{\rho}^{\mathrm{t}}\mathbf{T}_n + \mathbf{T}_n\Big)$$

with $\mathbf{S} := \sum_{n=1}^{N}\mathbf{T}_n$. For our problem, we stress that $\mathbf{H}(\boldsymbol{\rho})$ may not be NND. Moreover, its size prevents *a priori* its direct use in any Newton or quasi-Newton scheme. Matrix $\mathbf{S}$ is nevertheless NND (as a sum of NND matrices). Moreover, we have $\mathbf{S} = \mathbf{BCCB}(\boldsymbol{s})$ with

$$\widetilde{\boldsymbol{s}} = \sum_{k=1}^{K} |\widetilde{\boldsymbol{u}}_k|^2. \tag{41}$$

As a consequence, matrix $\mathbf{S}+\mu\mathbf{I}_N$ is a positive definite BCCB matrix, and thus a natural candidate to be a preconditionner, yielding a scaling of the gradient at an $\mathcal{O}(N\log(N))$ complexity.

D. Stepsize computation

We propose to use a line-search backtracking strategy to find a suitable step α_k for iteration (24). In order to grant convergence of the PCG iteration, we ensure that the stepsize meets the standard Armijo-Goldstein condition through backtracking [16]. The initial step of the backtracking subroutine is given by a 1D Newton step along the current descent direction

$$\bar{\alpha}_k = \frac{W_k'(0)}{W_k''(0)} \tag{42}$$

with $W_k(\alpha) = J(\boldsymbol{\rho}_k + \alpha\boldsymbol{d}_k)$. Such an initial step being often accepted by the Armijo rule, the average number of backtracking iterations is limited over the whole minimization process. Furthermore, this step can be given a closed-form expression. This computation is straighforward, and yields the following expression for the initial step:

$$\bar{\alpha} = \frac{\sum_{n=1}^{N} B_n\left(1 - \frac{\widehat{\sigma}_n}{\sigma_n}\right) + \mu\boldsymbol{d}^{\mathrm{t}}\boldsymbol{\rho}}{\sum_{n=1}^{N}\left(A_n - \frac{B_n^2}{\sigma_n}\right)\left(1 - \frac{\widehat{\sigma}_n}{\sigma_n}\right) + \frac{B_n^2}{\sigma_n} + \mu||\boldsymbol{d}||^2}. \tag{43}$$

With

$$A_n = \boldsymbol{d}^{\mathrm{t}}\mathbf{T}_n\boldsymbol{d} \tag{44}$$

$$B_n = \boldsymbol{d}^{\mathrm{t}}\mathbf{T}_n\boldsymbol{\rho} \tag{45}$$

Using the decomposition of $\mathbf{T}_0$, the whole set of A_n and B_n can be computed with an $\mathcal{O}(N\log N)$ complexity, leaving of the overall complexity of the algorithm unchanged.

Appendix D
Parallels Between RIM and Phase Retrieval

A common expression for phase retrieval (PR) problems found in the literature is

$$y_n^2 = |\boldsymbol{a}_n^{\dagger}\boldsymbol{x}|^2 \tag{46}$$

in the noiseless version (*cf.* [21], [26]–[28]). If we restrict our attention to real-valued problems, $\boldsymbol{a}_n, \boldsymbol{x} \in \mathbb{R}^N$ and we have

$$y_n^2 = \boldsymbol{x}^{\mathrm{t}}\mathbf{A}_n\boldsymbol{x}, \tag{47}$$

$\mathbf{A}_n = \boldsymbol{a}_n\boldsymbol{a}_n^{\mathrm{t}}$ being a rank-one NND matrix. For RIM, the noiseless variance model reads

$$\sigma_n^2 = \boldsymbol{\rho}^{\mathrm{t}}\mathbf{T}_n\boldsymbol{\rho} \tag{48}$$

where $\mathbf{T}_n$ is an NND matrix whose rank is usually $K \ll N$. Clearly, this relation retains the quadratic structure of the PR problem with the rank-one condition removed. Indeed, RIM falls within the category of *Generalized Phase Retrieval* (GPR) problems, according to [29], [30]. since RIM-CF makes use of a rank-one approximation for $\mathbf{T}_n$, it is formally identical to a real-valued, non-negative (and thus trivially solved) instance of a PR problem.

A standard question in PR and GPR problems is to ensure that the measurement elements $\mathbf{A}_n$ are sufficiently diverse to ensure that the solution is unique (up to a factor of modulus one). Recent contributions focus on random measurement operators to provide such a diversity with a high probability. However, this cannot be applied to RIM, since in the latter case, the available data correspond to Fourier measurements at the output of an optical system. As a consequence, identifiability results are rather to be derived in a specific way, as explored in [5], [12].

Appendix E
Characterization via Monte Carlo Simulations

Let us resort to Monte-Carlo simulations (MC) to investigate the behavior of our iterative and non-iterative RIM estimators. To prevent an inevitable explosion of the total simulation time, we use a 1D "chirp" function (*i.e.*, the 1D equivalent of the spoke-pattern shown in Fig. 2) that reads

$$\rho(r) = \frac{1}{2} + \frac{1}{2}\cos\left(2\pi f_0 r^2\right)$$

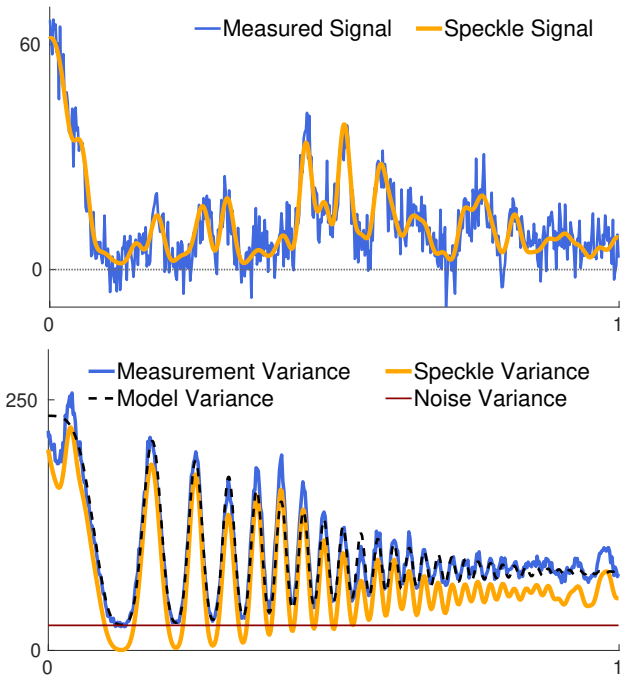

Fig. 4. Illustration of the simulation environment. The top image represents a speckled measurement before and after the addition of the noise. The bottom image represents both the empirical and model variances used within RIM-VAR, in the case where 500 speckles have been generated. It is noticable that the empirical variance is roughly the speckled variance shifted by a fixed amount (here, roughly 25), corresponding precisely to the variance of the additive noise pledging the data. With 500 speckles, there remains significant differences between the empirical variance and the model variance.

where $f_0 = 30$ controls the frequency domain spanned by the chirp. A total of N = 1024 points evenly sampled over the domain $r \in [-1, 1]$ are computed, resulting in a symmetric numerical object $\boldsymbol{\rho} \in \mathbb{R}_+^N$ (the symmetry also helps to mitigate circular boundary effects that may arise from replacing FT by FFT in the reconstruction methods).

We also consider a 1D imager model whose OTF $\widetilde{\boldsymbol{h}} = \boldsymbol{W}^\dagger \boldsymbol{h}$ is a discretized triangle function —a rather standard assumption, see [31, Chap. 6]. The cutoff frequency of the OTF is set to $f_c = 40$, so that iterative methods using noise-free (*i.e.*, asymptotical) statistics should retrieve the chirp perfectly.

In order to get statistically meaningful results, we consider 2.500 reconstructions of the same object from independent datasets, all generated from the (discretized) RIM observation model given in (19)[5]. Each reconstruction first requires the generation of a series of $M = 500$ (1D) microscope observations $\{\boldsymbol{z}_m\}_{m=1}^M$ obtained from an identical number of random speckle illuminations $\{\boldsymbol{E}_m\}_{m=1}^M$. The autocorrelation function of the illuminations is set to $\boldsymbol{\gamma}_E = E_0^2 \times \boldsymbol{h}$, with $E_0 \in \mathbb{R}_+$ the expected value of the illuminations and $\boldsymbol{h}$ the PSF of the microscope (a setting consistent with a standard fluorescence microscope working in epi-illumination, see for instance [7]); the expected illumination is set to $E_0 = \sqrt{20}$ for the whole experiment. Finally, each image $\boldsymbol{z}_m$ is plagued with additive Gaussian noise, with a variance of $\gamma_\varepsilon = 25$. Fig. 4 shows some intermediate quantities generated in one simulation, namely one acquisition $\boldsymbol{z}_m$, as well as the statistics derived[6] from the $M = 500$ observations in the current dataset.

The results of the MC simulation are given for several values of the regularization parameter μ in Figs. 5 and 6, respectively for RIM-CF and RIM-STD, with the truncature level K = 10. Let us recall that a given tuning of μ achieves a reconstruction quality within a bias *vs.* variance tradeoff [9, Chap. 4]. We can then identify three distinct regims in these results. When μ is too low (*e.g.*, $\mu < 10^{-2}$ here), the retrieved solutions are subject to a large amplification of the various sources of noise; in this situation, we note that the iterative estimates remain almost free of bias. When μ is too large (*e.g.*, $\mu \geq 10^0$ here), over-regularization is killing the noise amplification at the expense of a severe loss in the maximal resolution. For both estimators, the reconstruction variance is then small, but the bias is large. Finally, when the value of μ is intermediate, the fluctuation and the bias in the estimates are kept "under control". In such a case, the iterative reconstruction RIM-STD shows almost no bias and achieves a lower variance than RIM-CF.

Similarly, Figs. 7 and 8 represent the proposed estimators applied after the data pre-filtering procedure proposed in Sec. II-B2. Raw images have then been convolved with the filter g defined by (12), with $\eta = 10^{-5}$. As a result, the triangular OTF has been nearly inverted on its domain, so the equivalent OTF $h \otimes g$ is a top-hat function with an unchanged cutoff frequency. In this case, the iterative algorithm RIM-STD yields comparable results, both with and without pre-filtering. This is not surprising since the main goal of pre-filtering is to increase the convergence speed of the algorithm by increasing the sensitivity to higher frequency components. Finally, we note that performing RIM-CF (see Figs. 5 and 7) with pre-whitened data provides an overall improvement of the solution.

REFERENCES

[1] J. Mertz, *Introduction to Optical Microscopy*, Cambridge University Press, 2nd edition, 2019.

[2] J. Lindberg, "Mathematical concepts of optical superresolution", *Journal of Optics*, vol. 14, no. 8, pp. 083001, July 2012.

[3] K. Prakash, B. Diederich, R. Heintzmann, and L. Schermelleh, "Super-resolution microscopy: a brief history and new avenues", *Phil. Trans. R. Soc. A.*, vol. 380, pp. 20210110, Feb. 2022.

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

[5] J. Idier, S. Labouesse, M. Allain, P. Liu, S. Bourguignon, and A. Sentenac, "On the superresolution capacity of imagers using unknown speckle illuminations", *IEEE Transactions on Computational Imaging*, vol. 4, no. 1, pp. 87–98, 2018.

[6] S. Labouesse, J. Idier, A. Sentenac, M. Allain, and T. Mangeat, "Proof of the resolution-doubling of random illumination microscopy using the variance of the speckled images", in *29th EUSIPCO*, 2021, pp. 1159–1162.

[7] J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*, Coberts & Co, 2007.

[8] T. Mangeat, S. Labouesse, M. Allain, A. Negash, E. Martin, A. Guénolé, R. Poincloux, C. Estibal, A. Bouissou, S. Cantaloube, E. Vega, T. Li, C. Rouvière, S. Allart, D. Keller, V. Debarnot, X. B. Wang, G. Michaux, M. Pinot, R. Le Borgne, S. Tournier, M. Suzanne, J. Idier, and A. Sentenac, "Super-resolved live-cell imaging using random illumination microscopy", *Cell Rep. Meth.*, vol. 1, no. 1, pp. 100009, 2021.

[5] In the 1D case, the **BCCB** matrices are simply circulant matrices.

[6] Because the object is symmetric, the data and reconstructions of the chirp shown hereafter are given over half the simulated domain $r \in [0, 1]$

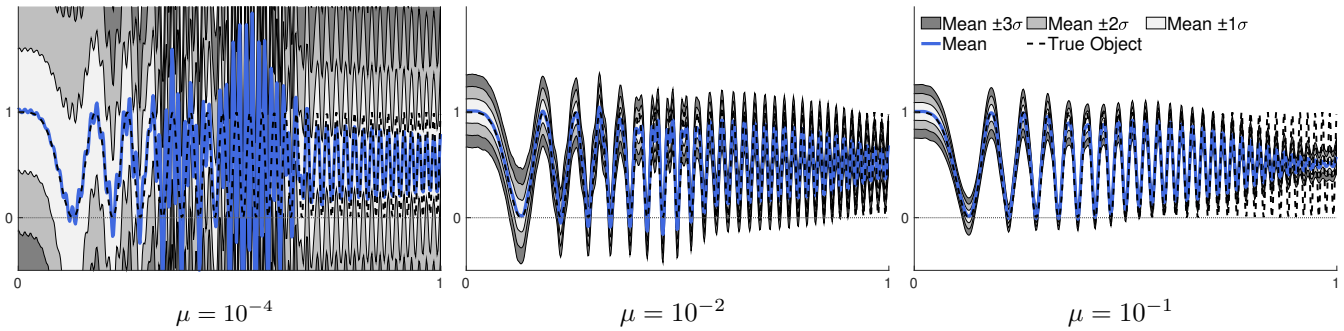

Fig. 5. Statistical result of Monte-Carlo inversion with RIM-CF.

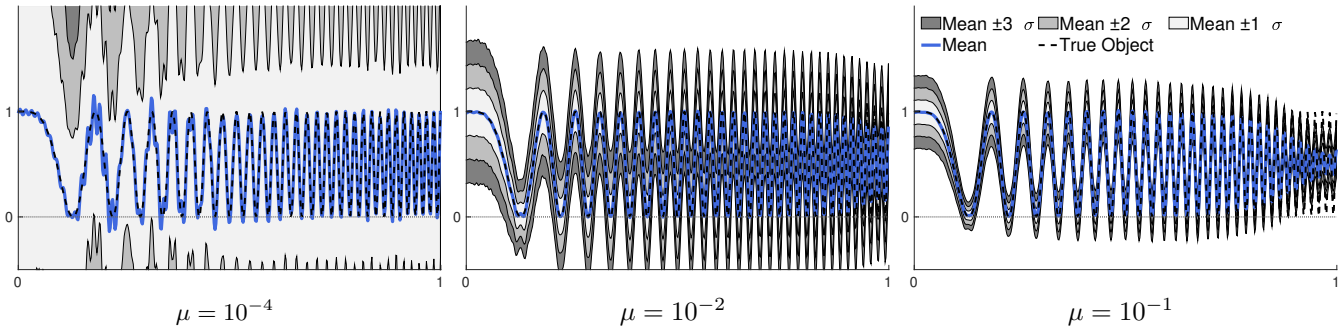

Fig. 6. Statistical result of Monte-Carlo inversion with RIM-STD

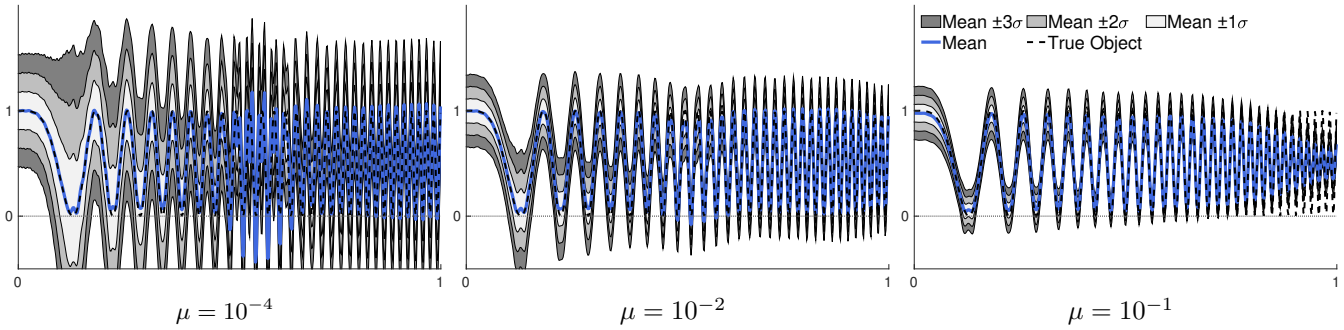

Fig. 7. Statistical result of Monte-Carlo inversion with RIM-CF, after data pre-filtering

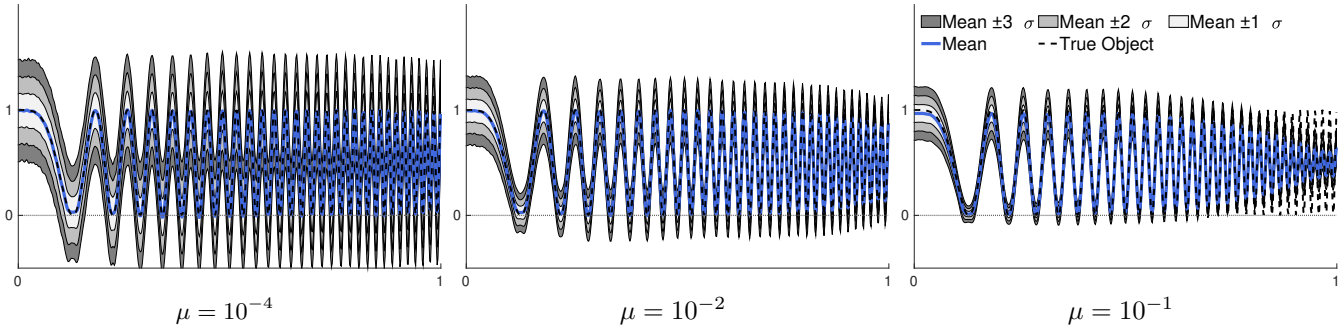

Fig. 8. Statistical result of Monte-Carlo inversion with RIM-STD, after data pre-filtering

[9] J. Idier, Ed., *Bayesian Approach to Inverse Problems*, ISTE Ltd and John Wiley & Sons Inc, Apr. 2008.

[10] M. Bertero and P. Boccacci, *Introduction to inverse problems in imaging*, Institute of Physics Publishing, 1997.

[11] E. Schmidt, "Zur theorie der linearen und nichtlinearen integralgleichungen, i", *Mathematische Annalen*, vol. 63, pp. 433–476, 1907.

[12] S. Labouesse, J. Idier, A. Sentenac, T. Mangeat, and M. Allain, "Random illumination microscopy from variance images", in *28th EUSIPCO*, 2021, pp. 785–789.

[13] N. Schweizer, L. Haren, I. Dutto, R. Viais, C. Lacasa, A. Merdes, and J. L uders, "Sub-centrosomal mapping identifies augmin-γ turc as part of a centriole-stabilizing scaffold", *Nat. Commun.*, vol. 12, 2021.

[14] M. Jasnin, J. Hervy, S. Balor, A. Bouissou, A. Proag, R. Voituriez, J. Schneider, T. Mangeat, I. Maridonneau-Parini, W. Baumeister, S. Dmitrieff, and R. Poincloux, "Elasticity of podosome actin networks produces nanonewton protrusive forces", *Nat. Commun.*, vol. 13, 2022.

[15] M. Portes, T. Mangeat, N. Escallier, O. Dufrancais, B. Raynaud-Messina, C. Thibault, I. Maridonneau-Parini, C. Vérollet, and R. Poincloux, "Nanoscale architecture and coordination of actin cores within the sealing zone of human osteoclasts", *eLife*, vol. 11, pp. e75610, jun 2022.

[16] S. Wright and J. Nocedal, *Numerical Optimization*, Springer, 2nd edition, 1999.

[17] K. Sauer and C. Bouman, "A local update strategy for iterative reconstruction from projections", *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 534–548, 1993.

[18] K. Deimling, *Nonlinear Functional Analysis*, Springer, 1985.

[19] C. Ventalon and J. Mertz, "Quasi-confocal fluorescence sectioning with dynamic speckle illumination", *Opt. Lett.*, vol. 30, no. 24, pp. 3350–3352, Dec 2005.

[20] 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.

[21] H. Zhang, Y. Zhou, Y. Liang, and Y. Chi, "A nonconvex approach for phase retrieval: Reshaped wirtinger flow and incremental algorithms", *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 5164–5198, Jan. 2017.

[22] S. Kandel, S. Maddali, Y. Nashed, S. Hruszkewycz, C. Jacobsen, and M. Allain, "Efficient ptychographic phase retrieval via a matrix-free Levenberg-Marquardt algorithm", *Opt. Express*, vol. 29, no. 15, pp. 23019, July 2021.

[23] T. Qiu, P. Babu, and D. P. Palomar, "PRIME: Phase retrieval via majorization-minimization", *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5174–5186, oct 2016.

[24] D. Sage, L. Donati, F. Soulez, D. Fortun, G. Schmit, A. Seitz, R. Guiet, C. Vonesch, and M. Unser, "Deconvolutionlab2: An open-source software for deconvolution microscopy", *Methods*, vol. 115, pp. 28–41, 2017.

[25] I. Gohberg and S. Goldberg, *Basic Operator Theory*, Birkhäuser, 1st edition, 2001.

[26] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase", *Appl. Comput. Harmon. Anal.*, vol. 20, no. 3, pp. 345–356, 2006.

[27] Y. Shechtman, Y. Eldar, O. Cohen, H. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging", *IEEE Sig. Proc. Mag.*, vol. 32, no. 3, pp. 87–109, 2015.

[28] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms", *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[29] Y. Wang and Z. Xu, "Generalized phase retrieval: Measurement number, matrix recovery and beyond", *Appl. Comput. Harmon. Anal.*, vol. 47, no. 2, pp. 423–446, 2019.

[30] M. Huang, Y. Rong, Y. Wang, and Z. Xu, "Almost everywhere generalized phase retrieval", *Appl. Comput. Harmon. Anal.*, vol. 50, pp. 16–33, 2021.

[31] J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 2nd edition, 1996.