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

Guillaume Giroussens, Simon Labouesse, Marc Allain, Thomas Mangeat, Lorry Mazzella, Loïc Le Goff, Anne Sentenac, Jérôme Idier

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

**December 11, 2024** - Guillaume Giroussens, Simon Labouesse, Marc Allain, Thomas Mangeat, Lorry Mazzella, Loïc le Goff, Anne Sentenac, and Jérôme Idier

***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 from random speckle 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 that prevents 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 provide a 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 standard 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 $\rho$ and an excitation function $E$:

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

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 $2{\rm NA}/\lambda$ where NA is the numerical aperture of the microscope objective and $\lambda$ the wavelength of the fluorescence light. When accounting for the third dimension, $\mathcal{D}_{\rm PSF}$ becomes akin to a solid torus with pointy edges. It keeps the same radius, with maximal height about ${\rm NA}^2/(2\lambda)$ and a missing cone as we move toward the zero transverse frequency [1, Sec. 5.4].

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

G. Giroussens, M. Allain, A. Sentenac, L. Mazzella and L. le Goff 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).

When the illumination $E$ 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 microscope image are limited to $\mathcal{D}_{\rm PSF}$. Super-resolution microscopy [2] refers to any technique that can estimate frequency components of the sample beyond $\mathcal{D}_{\rm PSF}$. The issue of super-resolution imaging can thus be stated as: how can the frequency components of the sample beyond $\mathcal{D}_{\rm PSF}$ be recovered from images that are limited to $\mathcal{D}_{\rm PSF}$?

The main answers belong to the *instrumental super-resolution* category [3]. Compared to standard wide-field imaging, these methods imply a modification of the acquisition strategy, 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, rather than as a standalone possibility [4].

Instrumental super-resolution fluorescence microscopy techniques make use of multiple diffraction-limited images of the sample under different inhomogeneous excitations of the fluorescence. The easiest way to generate various excitation is by using series of known, inhomogeneous illumination pattern, like a rotated or translated periodic light grid in Structured Illumination Microscopy (SIM), or a scanned focused beam for Image Scanning Microscopy (ISM). The illumination pattern being also diffraction limited, at most a two-fold improvement is achieved by solving a linear system.

On the other hand, considering the stochastic response of the fluorophores, one can model the excitation function $E$ as a sparse function, located precisely on the active fluorophores. In this case, there is no theoretical upper bound to the achieved resolution, but data processing becomes generally more complex, involving additional *a priori* information on the sparsity of the sample. For instance, Single Molecule Localization Methodes (SMLM) generally relies on a sparse hypothesis. Other methods, such as Super Resolution Optical Fluctuation (SOFI) microscopy, rely on image statistics, but the use of higher orders compromises the quantitative property of the technique, as the signal from the brightest emitters is over-amplified.

In between these two approaches, the Random Illumination Microscopy (RIM) technique consists in illuminating the sample with *unknown* speckle patterns, eliminating the constraint

of precisely knowing or estimating them, as it is the case in blind-SIM or speckle SIM techniques [5], [6]. RIM remains a quantitative method by retrieving the frequency components of the sample from the statistics of the recorded images. RIM being a linear structured illumination method, it has at most a two-fold resolution improvement, which can be achieved in practice [7], [8].

In the existing RIM method [9], the fluorescence density estimate is computed *iteratively* to minimize the distance between the empirical (experimental) variance and the expected (asymptotic) variance model (which is a quadratic function of $\rho$). This variance-matching solver (RIM-VAR) is robust, but its slow convergence impairs the recovery of the highest spatial frequencies that should be accessible with RIM. This issue is particularly detrimental for large-scale problems because RIM-VAR does not even converge within a realistic computational time. As a result, RIM-VAR has only been implemented in a simplified two-dimensional configuration where only the thin slice of the sample located at the focal plane of the microscope is accounted for in the reconstruction.

The implementation of a full three-dimensional (3D) RIM approach requires a faster and computationally more efficient solver. More generally, such a solver is pivotal as a building-block for a high-throughput RIM method, hence for the future dissemination of RIM.

In this paper, we present two novel RIM reconstruction methods. The first one, RIM-CF, is a fast non-iterative linear deconvolution method of the standard-deviation image, which provides a super-resolved estimate of the sample. This estimate can be used as the initial guess for RIM-STD, an iterative standard deviation matching algorithm that converges much faster than RIM-VAR. RIM-STD speeds up the reconstruction process, typically by two orders of magnitude.

The article is organized as follows. Section II presents the minimal theoretical background for understanding RIM, as well as a summary of the variance matching procedure RIM-VAR introduced in [9]. Section III introduces RIM-CF, an approximate *closed-form*, non-iterative estimator of the sample. Then, Section IV proposes an iterative inversion scheme based on the standard deviation of the images. Finally, Section V provides comparisons between RIM-CF, RIM-STD and RIM-VAR on 1D, 2D and 3D synthetic data, as well as on 3D experimental images.

## II. Principles of Random Illumination Microscopy

This first section introduces the notations, assumptions and basic principles of RIM and summarizes the main features of RIM-VAR, the variance matching procedure presented in [9] and used in all existing RIM implementations.

### A. Modeling and theoretical resolution bound for RIM

Let us consider $M$ images $(z_1, \ldots, z_M)$ of the sample $\rho$, 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 $\varepsilon_m$, supposed to be additive, so that the true acquisition model for any image is given by

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

This convolution model is general enough to encompass both 2D and 3D acquisition cases[1]. For the sake of simplicity, we shall assume that the following (standard) properties hold for the random quantities in (2):

- $E$ and $\varepsilon$ are mutually independent and second-order stationary;
- The auto-correlation functions of $E$ and $\varepsilon$, denoted hereafter $\gamma_E$ and $\gamma_\varepsilon$, are real-valued and known *a priori* (see for instance [10, Chap. 4] for a justification).

From Eq. 2, we deduce that each acquisition $z_m$ is sensitive to the sample frequencies within the domain[2]

$$\mathcal{D}_{\mathrm{SR}} = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{spec}}. \tag{3}$$

With a sufficient number of known speckle illuminations, using simple arguments based on linear (Fourier) analysis, we can prove that all the frequencies of the sample within $\mathcal{D}_{\mathrm{SR}}$ are identifiable. However, because RIM works with *unknown* illuminations, the situation is theoretically (*what frequencies can be retrieved?*) and practically (*how can we retrieve them?*) more complex [7], [11], [8]. The natural way to address these questions is through the image statistics, and more specifically through the image variance:

$$v_z(\boldsymbol{r}; \rho) = v_s(\boldsymbol{r}; \rho) + v_\varepsilon \tag{4}$$

with $v_s$ the contribution of the speckle excitation to the variance, and $v_\varepsilon$ the contribution of the instrumental (*e.g.*, camera) noise. Furthermore, $v_s$ can be written as

$$v_s(\boldsymbol{r}; \rho) = \iint t(\boldsymbol{r} - \boldsymbol{x}, \boldsymbol{r} - \boldsymbol{x}')\rho(\boldsymbol{x})\rho(\boldsymbol{x}')\, d\boldsymbol{x}\, d\boldsymbol{x}' \tag{5}$$

with the following kernel

$$t(\boldsymbol{x}, \boldsymbol{x}') = h(\boldsymbol{x})h(\boldsymbol{x}')\gamma_E(\boldsymbol{x} - \boldsymbol{x}'). \tag{6}$$

Clearly, the pivotal quantity $v_s$ is a *quadratic* (nonlinear) mapping of the unknown sample $\rho$. This nonlinear relationship does not allow us to directly identify the frequency components of $\rho$. However, it has been proven that, when $\gamma_E = h$, the frequency components of the sample are in bijection with the frequency components of the variance [8], and can be retrieved without bias through numerical iterative optimization [11]. In this situation, RIM can be essentially considered as a linear system, fully characterized by a convolution kernel that is the inverse Fourier transform of the indicator function over $\mathcal{D}_{\mathrm{SR}}$; in the noise-free case, the retrieved sample and the true sample are indeed identical within the super-resolved domain $\mathcal{D}_{\mathrm{SR}}$. Nevertheless, the computation of $v_s$ is a challenge: (5) is a convolution over 4 or 6 spatial dimensions (in 2D or 3D, respectively), making its numerical evaluation cumbersome. To decrease the computational burden, one can note that $t$ is a *positive definite* operator that can be decomposed as,

$$t(\boldsymbol{x}, \boldsymbol{x}') = \textstyle\sum_{n \geq 1} \lambda_n \psi_n(\boldsymbol{x})\psi_n(\boldsymbol{x}') \tag{7}$$

[1] In the latter case, the acquisition of a 3D data stack $z_m$ (*i.e.,* the scan of the fluorescent signal along the optical axis) must be performed with a static 3D speckle illumination setup. See Sec. V-D for more details.

[2] $\ominus$ stands for the Minkowski difference : $A \ominus B := \{a - b \,|\, a \in A,\, b \in B\}$, *i.e.*, the cross-correlation of the two domains.

where $\{\psi_n \,|\, n \geq 1\}$ is a countable family of orthogonal functions with $\int |\psi_n(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} = 1$, and $\{\lambda_n\}_{n\geq 1}$ are real numbers ordered by decreasing magnitude, with $\lim_{n\to\infty} \lambda_n = 0$, see *e.g.,* [12] and Appendix B-A, Eq. (27). Eq. (7) is a decomposition into a sum of separable functions. Inserting it into (5) provides an equivalent formulation for the variance *via* a sum of squared convolution integrals [8]

$$v_s(\cdot;\rho) = \sum_{n\geq 1} (u_n \otimes \rho)^2 \tag{8}$$

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 (8) to the first $K$ leading terms, which is equivalent to using a low rank approximation of the original kernel. The quality of $t_K$, the approximation of $t$ using the first $K$ eigenvectors, can be assessed by computing the energy ratio of the operators:

$$\tau_K = \frac{\|t_K\|_2^2}{\|t\|_2^2} = \frac{\sum_{k=1}^{K} \lambda_k^2}{\sum_{k=1}^{\infty} \lambda_k^2}. \tag{9}$$

For the 2D epi-fluorescence microscope setting, we checked numerically that $\tau_{10} > 1 - 10^{-6}$. This computational trick allows us to compute the image variance at a cost of $O(KN\log(N))$ arithmetic operations.

### *B. Variance matching algorithm, RIM-VAR [9]*

Since RIM-VAR will serve as a baseline comparison with the new methods proposed in this paper, it is useful to introduce its main features and drawbacks. RIM-VAR consists in minimizing the cost functional

$$Q(\rho\,;\,\mu) = \|\widehat{v}_z(\boldsymbol{r}) - v_z(\boldsymbol{r};\rho)\|_2^2 + \mu\,\|\rho(\boldsymbol{r})\|_2^2\,. \tag{10}$$

In the right hand-side (r.h.s.) of (10), the first term is a discrepancy measure between the variance model $v_s$ computed using (8) and the *empirical* variance $\widehat{v}_s$ computed from the set of speckled images. The second term is a zero-order Tikhonov regularization, with $\mu$ its parameter tuning the precision *vs.* robustness-to-noise trade-off one expects in the retrieved sample [13, Sec. 4], [14, Sec. 5.6]. This criterion is then minimized with an iterative conjugate-gradient approach, for which an exact step-size can be calculated since the restriction of $Q$ to any line is a $4^{th}$-order scalar polynomial.

*1) 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:

$$\widetilde{g} = \frac{\widetilde{h}^*}{|\widetilde{h}|^2 + \eta} \tag{11}$$

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 PSF by enhancing the weight of its high frequencies. Second, it reduces the influence of the noise [14, Chap. 5]: it suppresses the noise frequency components outside the OTF support, and it simultaneously prevents the detrimental amplification of the frequency components of the noise close to the limit of the OTF support. This latter effect relies on the tuning of the parameter $\eta$, which is then mainly determined by the noise level.

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 PSF and noise variance. This pre-processing was applied to all RIM acquisitions as it significantly improved the final resolution in the sample estimate provided by RIM-VAR.

*2) Limit of RIM-VAR:* Even if RIM-VAR is now routinely used to produce super-resolved images of biological samples, see for instance [9], [15]–[17], it was observed in many examples that it suffers 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 only very slowly retrieved [18, Sec. 5], [14, Sec. 6.5], [19]. 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

Although the spectral decomposition (8) was introduced first as a computational trick, it provides key insights into the structure of the second-order statistics of the images and it 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 [20, p.18] ensures that the largest eigenvalue $\lambda_1$ is of multiplicity one, so that the first eigenvector $u_1$ is dominant in the decomposition (8). To estimate the quality of the rank-one approximation, we numerically computed (9) for classical microscopes in epi-fluorescence and found that this indicator is close to $0.9$, *i.e.,* the first eigenvector in the spectral decomposition of $t$ accounts for 90% of its total energy. We thus propose to approximate the variance with the first term of the decomposition,

$$\boldsymbol{v}_s(\cdot;\rho) \;\approx\; (u_1 \otimes \rho)^2\,. \tag{12}$$

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. (12) as

$$\sigma_s(\boldsymbol{r};\rho) \approx\; (u_1 \otimes \rho)\,(\boldsymbol{r}). \tag{13}$$

At this point, it is worth noting that the pre-filtering of the speckled images introduced in II-B1 generally breaks the

positivity assumption for $u_1$, and especially in the case of the Wiener-type pre-filter defined by (11). Such a choice of pre-filter is thus *a priori* inappropriate for the non-iterative method.

Under the approximation (13), the incoherent RIM microscope can be understood as a linear (spatially invariant) system equipped with a PSF 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.

**Proposition 1** *Let $\widetilde{u}_1$ be the Fourier transform of $u_1$ [the eigenvector associated with $\lambda_1$ in the spectral decomposition (7)]. Under standard assumptions (see $\mathbf{A}_1 - \mathbf{A}_7$ in Appendix A), $\widetilde{u}_1$ is a non-negative function whose support is exactly $\mathcal{D}_{\mathrm{SR}}$.*

*Proof:* See Appendix B-B. ∎

Proposition 1 is particularly fruitful for 3D RIM because the bandwidth associated with $u_1$ fills the so-called *missing cone*, which otherwise prevents the *optical sectioning* [1, Sec.5.4].

This result then explains the optical sectioning property of the image standard deviation, as proposed in the Dynamic Speckle Illumination (DSI) approach presented in [21], [22]. With knowledge of the convolution kernel $u_1$, we can leverage on (13) to further improve the sample estimate over DSI. More specifically, we define our *closed-form* RIM estimator, RIM-CF, as

$$\widehat{\rho} \,=\, \mathcal{F}^{-1}\left(\frac{\widetilde{u}_1}{|\widetilde{u}_1|^2+\mu}\mathcal{F}(\widehat{\sigma}_s)\right), \tag{14}$$

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 (4)

$$\widehat{\sigma}_s(\boldsymbol{r}) = \left(\max\left\{0,\, \widehat{v}_z(\boldsymbol{r}) - v_\varepsilon\right\}\right)^{\frac{1}{2}} \tag{15}$$

with $\widehat{v}_z$ the empirical variance built from the set of microscope images $z_m, m = 1,\ldots,M$; 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 $v_\varepsilon$ is subtracted (if $v_\varepsilon$ is unknown, one may use the technique proposed in [11, Sec. IV]). The (Tikhonov) regularization parameter $\mu > 0$ is adjusted to get the expected regularity using RIM-CF [14, Sec. 5.6].

Before concluding this section, some important remarks can be made about the expected performance of the method. Let us recall first that RIM-CF is an approximate solution for the quadratic integral equation (5). Whereas the RIM-STD solver (see next section) computes an exact solution that is unbiased for the frequency components of the sample inside $\mathcal{D}_{\mathrm{SR}}$, such a property is lost for RIM-CF. An important consequence is that the resolution in RIM-CF is spatially variable. Interestingly, we also found that this spatial resolution is locally driven by the spatial density of fluorescent sources. This effect is illustrated by considering first a single pointwise fluorescent emitter. In this simple case, it is easy to derive from (5) that

$$v_s(\boldsymbol{r}) = I_0^2 h^2(\boldsymbol{r}) \quad \Longrightarrow \quad \sigma_s(\boldsymbol{r}) = I_0 h(\boldsymbol{r}). \tag{16}$$

with $I_0$ the average speckle intensity. Because the standard-deviation is proportional to $h$ in this case, no super-resolution is expected from RIM-CF (since it relies on a *linear* deconvolution of $\sigma_s$). Now, let us consider a pair of fluorescent emitters. The variance now reads

$$\begin{aligned} v_s(\boldsymbol{r}) \,=\, & \gamma_E(\boldsymbol{r}_0) h\left(\boldsymbol{r}-\tfrac{\boldsymbol{r}_0}{2}\right) h\left(\boldsymbol{r}+\tfrac{\boldsymbol{r}_0}{2}\right) + \cdots \\ & I_0^2\left[h^2\left(\boldsymbol{r}-\tfrac{\boldsymbol{r}_0}{2}\right) + h^2\left(\boldsymbol{r}+\tfrac{\boldsymbol{r}_0}{2}\right)\right] \end{aligned} \tag{17}$$

with $\boldsymbol{r}_0$ the separating distance. When the two emitters are *well separated*, the PSFs located in $\frac{\boldsymbol{r}_0}{2}$ and $-\frac{\boldsymbol{r}_0}{2}$ have almost disjoint supports, so the standard-deviation becomes proportional to $h\left(\boldsymbol{r}-\frac{\boldsymbol{r}_0}{2}\right)+h\left(\boldsymbol{r}+\frac{\boldsymbol{r}_0}{2}\right)$. The encountered situation is that of two isolated emitters, for which RIM-CF cannot provide any super-resolution. Things change if the emitters *get closer*, so that the first term in the right-hand side of (17) cannot be neglected and the two squared PSF in the second term interact. These interactions are significant when $\boldsymbol{r}_0$ is comparable to the resolution limit of the microscope. In this case, a resolution enhancement arises, both in the standard deviation and in RIM-CF, see Fig. 7 and 8 in the supplemental. The same effect can be observed with a random set of $N$ separated emitters: RIM-CF is found more resolved where the density of emitters is locally the highest. In other words, the resolving power of RIM-CF is locally variable, and it is higher where it matters the most. The supplementary section gives a detailed presentation of this effect.

In summary, RIM-CF is expected to provide more resolution with higher density samples, and no superresolution with isolated emitters. Obviously, even in a favorable scenario, RIM-CF remains based on the approximation (12) and modeling errors deteriorate the resolution gain, *e.g.,* RIM-CF cannot reach the theoretical resolution limit one can expect from RIM in a noise-free setting. An iterative solver, based on a more accurate model is required to reach this limit.

## 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 (7). 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 [23], [24], 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

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

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 $\mathbf{Diag}(\cdot)$ be a diagonal matrix built from a vector, and $\mathbf{diag}(\cdot)$ be 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 [13, Chap. 4]. Once discretized, the observation model (2) reads

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

where the unknown fluorescence map $\boldsymbol{\rho}$, the $m$th 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 reads

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

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

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

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 (4) reads

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

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

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

where $\|\cdot\|_2$ is the usual Euclidean 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, we provide a low-rank approximation of $\boldsymbol{v}_s(\boldsymbol{\rho})$ directly from the spectral decomposition of a matrix operator that is analogue to the kernel $t(\cdot,\cdot)$ given in (6). 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 [18, Sec. 5.2]. In addition, the NGC iterations can be appropriately "scaled" with a preconditioning matrix to accelerate the convergence [18, 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 preconditioner using a Majorization-Minimization (MM) construction, following the lines of [25, 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

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

with

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

where $\boldsymbol{g}_k$ is the gradient of (23) evaluated at the current estimate $\boldsymbol{\rho}_k$, $\alpha_k$ 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 conjugation formula (PR$^+$) [18, p.122] for the conjugation factor $\beta_k$:

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

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

Since this NCG scheme only fulfils 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

[3]From the structure of model (5), 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 an 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 (5) is insensitive to any frequency component of $\rho$ 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).

estimator RIM-CF proposed in the previous section is a natural initial point.

The memory resource required to build the update (24) is clearly seen from (26): we only need to store and perform convolutions with $\boldsymbol{s}$ such that $\mathbf{S} = \mathbf{BCCB}(\boldsymbol{s})$. This is done in $O(N \log N)$ operations *via* the Fast Fourier Transform (FFT).

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 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 $\mathcal{D}_{\text{spec}} = \mathcal{D}_{\text{PSF}}$ so that the sample frequencies should, in theory, be recovered in $\mathcal{D}_{\text{SR}} = \mathcal{D}_{\text{PSF}} \ominus \mathcal{D}_{\text{PSF}}$, (which corresponds to the Fourier support of $h^2$).

### A. 1D simulation

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 $\mathcal{D}_{\text{PSF}} = [-\nu_{\text{PSF}}, \nu_{\text{PSF}}]$ where $\nu_{\text{PSF}} = 2/\lambda$ with $\lambda$ 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. Furthermore, it is clear that the convergence speed of RIM-VAR is too low to produce the full resolution gain ($\mathcal{D}_{\text{SR}} = [-2\nu_{\text{PSF}}, 2\nu_{\text{PSF}}]$), even after thousands of iterations. In comparison, RIM-STD reaches the theoretical super-resolution bound 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.

Finally, we recall after our analysis in Sec. III that well-separated point-wise florescent emitters are somewhat pathological cases for our non-iterative method RIM-CF. With a single emitter, actually, RIM-CF cannot provide a better estimate than a deconvoluted wide-field. For this reason, the RIM-CF estimate is not shown in Fig. 1.

### B. 2D simulation

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 $256 \times 256$ 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). RIM-VAR reaches the same resolution level as RIM-STD but with a significantly larger computation time (about 450 seconds). The comparison with the deconvolution of the wide-field image shows unambiguously the resolution gain brought by the iterative and non-iterative versions.

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

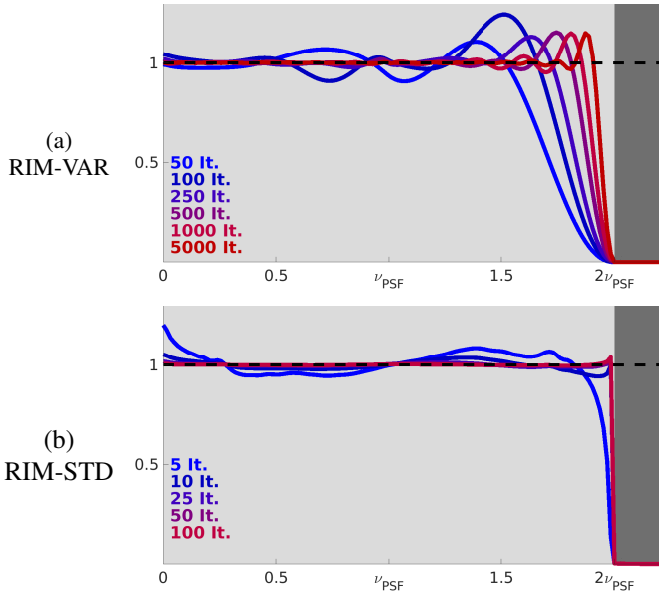


Fig. 1. (a)-(b) Resolution gain of RIM-VAR and RIM-STD (with pre-filtering) as a function of the iteration number, on a toy 1D problem involving a single pointwise emitter, with $\mathcal{D}_{\text{PSF}} = \mathcal{D}_{\text{spec}} = [-\nu_{\text{PSF}}, \nu_{\text{PSF}}]$, so $\mathcal{D}_{\text{SR}}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

In this ideal noise-free 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 (11), even though the positivity of $u_1$ is not granted in that case (hence the derivation of (13) from (12)). The absence of visible artefacts in this non-ideal case can be explained by the fact that $u_1$ is nearly non-negative.

Additional validation based on the processing of 2D biological data is available in appendix F.

### C. 3D Simulation

The controlled computational complexity of RIM-CF and RIM-STD being particularly attractive for the 3D setting, our last simulation is performed with a synthetic 3D sample made of microtubules [26]. Fig. 3 displays RIM-CF, RIM-STD and RIM-VAR reconstructions in both transverse and axial directions. The results are produced either from the asymptotic variance image, or from an empirical one using 1000 speckled images. Fig. 3 displays RIM-CF, RIM-STD and RIM-VAR reconstructions in both transverse and axial directions. The results are produced either from the asymptotic variance image, or from an empirical one using 1000 speckled images. In this 3D context, as predicted by the analysis in [7], RIM should provide optical sectioning since the accessible sample Fourier domain $\mathcal{D}_{\text{SR}} = \mathcal{D}_{\text{PSF}} \ominus \mathcal{D}_{\text{PSF}}$ 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

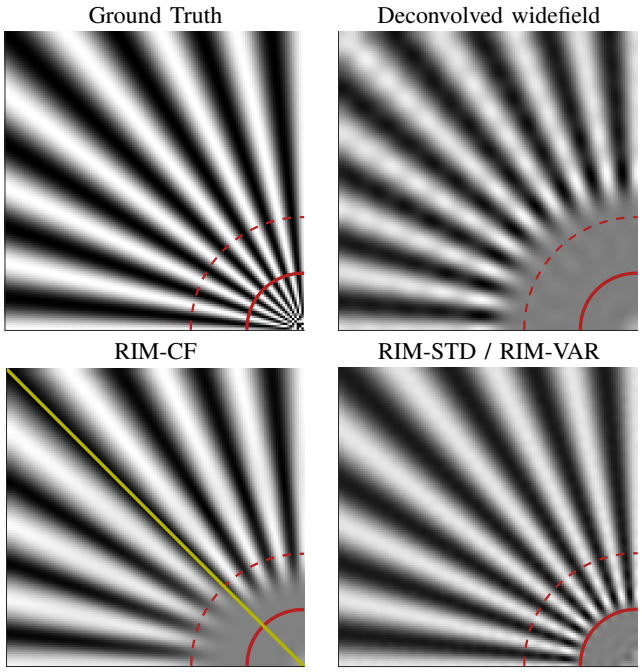


Fig. 2. Proof of the super-resolution induced by RIM in a 2D asymptotical noiseless case. The dashed and solid circular arcs indicate the resolution level achievable with a cutoff frequency $\nu_{\mathrm{PSF}}$ (RMSE=$1.06 \cdot 10^{-1}$) and $2\nu_{\mathrm{PSF}}$ (RMSE reference), respectively. For the RIM-CF reconstruction, the lower left part is without data pre-filtering (RMSE=$1.21 \cdot 10^{-1}$) and the upper-right part with (RMSE=$8.21 \cdot 10^{-2}$). RIM-VAR yields the same reconstruction as RIM-STD (RMSE=$1.41 \cdot 10^{-2}$), albeit with an increased reconstruction time (3.5s against 450s)

approaches start from a constant initial guess, showing that RIM-STD dramatically outperforms RIM-VAR in terms 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 still improves the resolution of the sample estimation as compared to the wide-field deconvolution or the raw standard deviation image, Fig. 3(c,d).

A more realistic simulation was also performed with the reconstruction of a fluorescence map from an empirical variance statistics built from 1,000 noisy speckled images. Each microscope image was plagued by an additive Gaussian white noise with zero mean and in order to simulate Poisson fluctuations, its variance was chosen equal to the expected signal in a given pixel. The integrated SNR on a single image was set to 30 dB. For this more realistic situation, RIM-STD exhibits a loss of resolution compared to the asymptotic regime, while RIM-CF is minimally affected, see Fig. 3(f,g). Actually, and contrary to the asymptotic regime (see Fig. 2 and Fig. 3(b,e)), RIM-CF and RIM-STD now reach comparable results in terms of quality, see Fig. 3(f,g).

A possible explanation is that with, the frequency components mostly impacted by the noise in the reconstruction are the one close to the limit of the super-resolved domain $\mathcal{D}_{\mathrm{SR}}$. The same frequency components happen to be poorly approximated with the rank-one approximation and RIM-CF would be unable to retrieve them even in a noiseless setting, see Fig. 2(c,d). As such, the addition of noise (in a reasonable quantity) primarily affects frequency components that are inaccessible to RIM-CF, so that RIM-CF reconstructions are less subject to deterioration compared to the one of RIM-STD and RIM-VAR.

In order to deepen our analysis on this phenomenon, we performed a mono-dimensional Monte-Carlo simulation of RIM-CF and RIM-STD, which is presented as supplementary material. The same conclusion can be drawn out from it: RIM-STD is clearly superior to RIM-CF in ideal conditions, but the difference in quality between the two methods gradually decreases under noisy, non-asymptotic configurations.

Finally, we recall that we used a constant (spatially uniform) initialization to run the iterative solvers. Such an initial guess is totally agnostic with respect to the sample to retrieve, thus setting a suited ground to assess if the frequency components that *should* be retrieved are *actually* retrieved during the course of the iterations. This is an important aspect of our comparison because RIM-VAR fails to converge properly with this initialization, at least when we address the full 3D problem, *i.e.,* in a frontal way instead of building a sub-optimal "slice-by-slice" solution as in [9]. This situation is clearly shown in Fig. 3(c,h). Here, RIM-VAR was started with a constant initial guess and run for the computation time required by RIM-STD to converge. The reconstruction achieved by RIM-VAR in this case produces almost no super-resolution[5]. This pathological behavior of RIM-VAR is precisely what motivated us to propose RIM-STD and RIM-CF.

### D. 3D Experimental Validation

Finally, we propose to reconstruct a real 3D fluorescent sample to further illustrate that RIM-CF and RIM-STD are efficient means to produce super-resolution in the three dimensions. We first note that several acquisition strategies are possible in a 3D RIM setup. Nevertheless, in order to keep an imaging modality that is consistent with the imaging model (2), the position of speckle relative to the sample must remain constant during the 3D acquisition process[6]. For this purpose, we used a custom remote focusing microscope using an electrically tuned lens in the optical setup (see supplemental and Fig. 10 for additional information). This setup allows us to remotely change the position of the object focal plane [27].

The test sample consists of a set of isolated fluorescent beads, 100 nm in diameter, hovering above a dense fluorescent layer (which is a continuous deposit of the same 100 nm fluorescent beads).

The continuous layer was produced first, and set on a glass slide. A drop of optical glue (of refractive index about 1.5) embedding individual beads was then spilled above this layer.

[5]We note that RIM-VAR in [9] was initialized with the empirical-variance image. This initialization provides the solver with an estimate for the frequency components of the sample lying beyond the cutoff frequency enforced by the OTF, *i.e.*, this initial guess provides some super-resolution which is further refined by RIM-VAR.

[6]The 3D acquisition strategy used in [9], translating the sample in the speckle pattern, would produce a model mismatch in the 3D case.

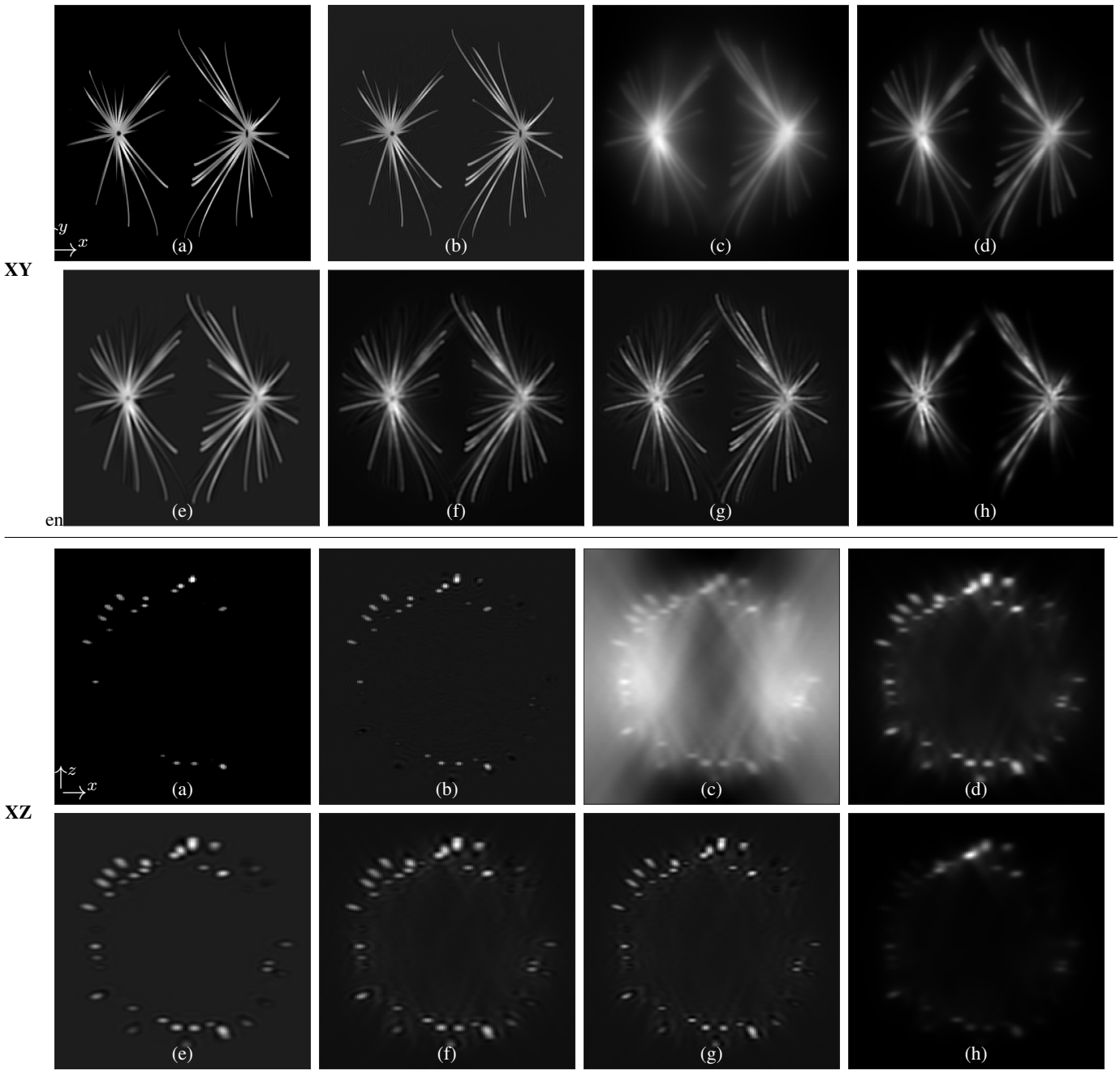


Fig. 3. RIM simulation with an object representing microtubules (*cf.* [26]). The top and bottom halves represent XY and XZ cuts, respectively. (a) Ground truth. (b) Ground-truth bandwidth limited image, matching the reconstruction of the asymptotic, standard deviation with RIM-STD. (c) Average of $M = 1000$ speckled microscope images (RMSE=1.14).(d) Empirical standard deviation of the same $M = 1000$ speckled images. (RMSE=0.66) (e) Non-iterative estimation RIM-CF, with the asymptotic standard deviation (case corresponding to $M = \infty$, RMSE=0.32). (f,g,h) Reconstruction with RIM-CF (RMSE $=$ 0.42), RIM-STD (RMSE $=$ 0.32) and RIM-VAR (RMSE=0.68) 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 $\mathcal{D}_{\mathrm{SR}} = \mathcal{D}_{\mathrm{PSF}} \ominus \mathcal{D}_{\mathrm{PSF}}$.

Large non-fluorescent beads ($5\,\mu$m) were also put above the layer to act as 'spacers' when the mixture is compressed with another glass slide.

With this sample, 200 3D stacks of images under speckle illuminations were then acquired. For each stack, 35 axial positions have been recorded with an axial step of 150 nm. Then, the 3D empirical variance was used as an input for RIM-CF, RIM-VAR and RIM-STD.

Several reconstructions produced for this sample are shown in Fig. 4. In particular, we see that the (deconvolved) widefield and the standard deviation strongly differ in their ability to produce optical sectioning (a situation that is actually consistent with Proposition 1). With RIM-CF, the quality of the sample estimate can be further improved from the standard deviation, at a minimal additional computing cost. Ultimately, both iterative procedures RIM-VAR and RIM-STD provide an even larger improvement, while the latter achieves faster convergence, reducing computation times from minutes to

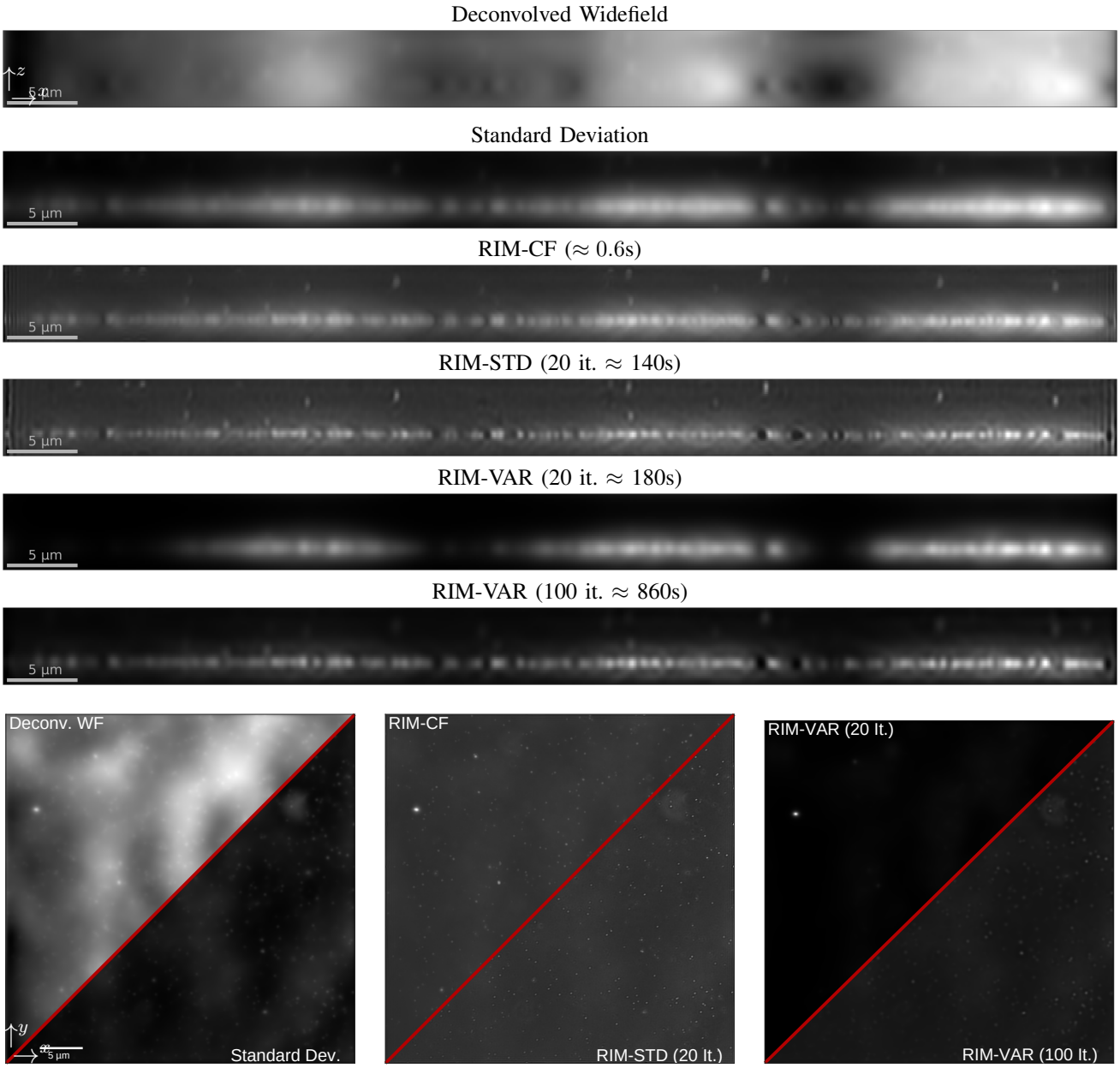


Fig. 4. 3D reconstruction on experimental data. The sample consists of fluorescent beads above a uniform fluorescent layer. The total field of view is roughly $40\mu$m $\times$ $40\mu$m $\times$ $6\mu$m and the data stack is $1024 \times 1024 \times 70$ pixels. All images have their contrast level calculated on the whole value range of the considered region of interest.

seconds and hours to minutes.

The presence of Gibbs oscillations is also noticeable in the RIM-VAR and RIM-STD reconstructions.

The phenomenon is inherent with the fact that RIM cannot recover any frequency component from the sample outside $\mathcal{D}_{\mathrm{SR}}$, but the oscillations are eventually producing negative values in the reconstructed sample. A positivity constraint added in the RIM-STD solver could prevent this effect, but this would also introduce some bias in the retrieved spectrum. Finally, we also note the presence of oscillations close to the lateral border of the sample. These oscillations are induced by our implementation *via* BCCB matrices of the convolution operator, see Sec. IV-A. These border artifacts could be circumvented by implementing more realistic boundary assumptions [28].

## VI. Conclusion

In this paper, we have shown that the standard deviation of microscope images obtained under random speckle illumination is essentially similar to conventional wide-field images but with a "super-resolved OTF", covering a frequency domain twice as large as that of the original microscope.

This result leads to RIM-CF, a reconstruction strategy based on a specific linearization of the RIM reconstruction problem. RIM-CF is a fast, FFT-based inversion of the standard deviation that provides a super-resolved estimate of the sample. However, because it is based on an approximation, this

estimator cannot achieve the maximal theoretical resolution gain for RIM.

To further improve the resolution gain, we also developed an iterative estimator based on a standard deviation matching procedure. We have shown that this novel algorithm, RIM-STD, is able to provide a resolution gain that is close to the one expected in ideal conditions, with less than one hundred iterates. This fast convergence is due to the quasi-linear behaviour of the image standard deviation with respect to the sample, and the use of an appropriate preconditioner. 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. In particular, RIM-CF and RIM-STD are 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)

- $\mathbf{A}_1$) $\rho$ 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 $\varepsilon$ are mutually independent and second-order stationary.
- $\mathbf{A}_6$) The auto-correlation functions of $E$ and $\varepsilon$, denoted hereafter $\gamma_E$ and $\gamma_\varepsilon$, are real-valued positive functions known *a priori* (see for instance [10, Chap. 4] for a justification).
- $\mathbf{A}_7$) The support $\mathcal{D}_{\mathrm{spec}}$ of the Fourier transform of $\gamma_E$ (*i.e.*, of the spectral energy density) is bounded and such that $\mathcal{D}_{\mathrm{spec}} \subseteq \mathcal{D}_{\mathrm{PSF}}$.

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 [7, Sec. II], *i.e.*, the main identification results of RIM [7, Prop. 3] and [8, 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 (6):

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

Lemma 1 below shows that $T$ has many good properties. In particular, it is a compact self-adjoint operator, so the spectral theorem [29, 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 $\psi_n$ is an orthogonal basis of eigenvectors for the operator $T$ and $\lambda_n$ 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** *[20, 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 (6).

**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 $\gamma_E$ 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 (6). 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.*, $\gamma_E$ 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.,* [20, Chap.6]) do not apply to the $L_2(\mathbb{R})$-space considered here.

### *B. Proof of Proposition 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

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

Inserting (6) 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 $\gamma_E$ 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 (*i.e.*, cross-correlation) $\mathcal{D}_{\widetilde{g}_{\boldsymbol{\nu}}} \ominus \mathcal{D}_{\rm PSF}$, 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

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

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 $\lambda_1$ 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 (3), $\exists \boldsymbol{\xi}_1 \in \mathcal{D}_{\rm PSF}, \boldsymbol{\xi}_2 \in \mathcal{D}_{\rm spec}$ such that $\boldsymbol{\nu}_0 = \boldsymbol{\xi}_1 - \boldsymbol{\xi}_2 \in \mathcal{D}_{\rm SR}$, or equivalently

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

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 $\gamma_E$ 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

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

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 (6) as a definite non-negative matrix

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

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 (4) reads $\boldsymbol{v}_z = (v_{z;1}, \dots, v_{z;N})^{\mathrm{t}}$ with $v_{z;n} := v_z(\boldsymbol{r}_n)$ given by

$$v_{z;n} = \boldsymbol{\rho}^{\mathrm{t}}\mathbf{P}_n^{\mathrm{t}}\mathbf{T}\mathbf{P}_n\boldsymbol{\rho} + v_\varepsilon = \boldsymbol{\rho}^{\mathrm{t}}\mathbf{T}_n\boldsymbol{\rho} + v_\varepsilon \tag{35}$$

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

$$\mathbf{T} = \sum_{k=1}^{K} \lambda_k \boldsymbol{\psi}_k \boldsymbol{\psi}_k^{\mathrm{t}} = \sum_{k=1}^{K} \boldsymbol{u}_k \boldsymbol{u}_k^{\mathrm{t}} \tag{36}$$

where $K \leq N$ is the rank of $\mathbf{T}$, $\lambda_1, \dots, \lambda_K > 0$, $\mathbf{T}\boldsymbol{\psi}_k = \lambda_k \boldsymbol{\psi}_k$, 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

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

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 (8). Criterion $J$ given in (23) then reads

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

with $\sigma_n(\boldsymbol{\rho}) = (\boldsymbol{\rho}^{\mathrm{t}}\mathbf{T}_n\boldsymbol{\rho} + v_\varepsilon)^{\frac{1}{2}}$ 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

$$\begin{aligned}\frac{\partial J}{\partial \boldsymbol{\rho}} &= \sum_{n=1}^{N} \frac{\partial \sigma_n}{\partial \boldsymbol{\rho}}(\boldsymbol{\rho})\Big(\sigma_n(\boldsymbol{\rho}) - \hat{\sigma}_n\Big) + \mu\boldsymbol{\rho} \\ &= \sum_{n=1}^{N}\left(\Big(1 - \frac{\widehat{\sigma}_n}{\sigma_n(\boldsymbol{\rho})}\Big)\mathbf{T}_n\boldsymbol{\rho}\right) + \mu\boldsymbol{\rho}. \qquad (39) \\ &= \sum_{k=1}^{K}\left[\mathbf{U}_k^{\mathrm{t}}\Big(\mathbf{U}_k\boldsymbol{\rho} \odot \frac{\boldsymbol{\sigma}(\boldsymbol{\rho}) - \widehat{\boldsymbol{\sigma}}}{\boldsymbol{\sigma}(\boldsymbol{\rho})}\Big)\right] + \mu\boldsymbol{\rho}, \qquad (40)\end{aligned}$$

The division in this expression is done entrywise. We recall that thanks to the contribution of the electronic noise, it is guaranteed 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:

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

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 preconditioner, 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 $\alpha_k$ 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 [18]. 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:

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

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.* [23], [30]–[32]). 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 [33], [34]. 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 [7], [11].

## REFERENCES


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

[2] 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.

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

[4] 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.

[5] E. Mudry, K. Belkebir, J. Girard, J. Savatier, E. Le Moal, C. Nicoletti, M. Allain, and A. Sentenac, "Structured illumination microscopy using unknown speckle patterns", *Nat. Photonics*, vol. 6, no. 5, pp. 312–315, 2012.

[6] L.-H. Yeh, S. Chowdhury, N. A. Repina, and L. Waller, "Speckle-structured illumination for 3d phase and fluorescence computational microscopy", *Biomed. Opt. Express*, vol. 10, no. 7, pp. 3635–3653, Jul 2019.

[7] 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.

[8] 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.

[9] 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.

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

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

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

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

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

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

[16] 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.

[17] 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.

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

[19] 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.

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

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

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] W. J. Shain, N. A. Vickers, A. Negash, T. Bifano, A. Sentenac, and J. Mertz, "Dual fluorescence-absorption deconvolution applied to extended-depth-of-field microscopy", *Opt. Lett.*, vol. 42, no. 20, pp. 4183–4186, Oct 2017.

[28] P. Hansen, J. Nagy, and D. O'Leary, *Deblurring Images*, SIAM, 2006.

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

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

[31] 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.

[32] 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.

[33] 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.

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

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

# Supplemental

## APPENDIX E
## PARTICULAR CASE OF POINT-WISE SOURCES

The study of respectively a single source and a pair of source is of particular interest since it can be done analytically.

### A. Case of a single emitter

First, let us focus on the case of a single fluorescence source, *i.e.*, such that the sample reads $\rho(\boldsymbol{r}) = \delta(\boldsymbol{r})$. In this case, one can directly compute the asymptotic variance of the image by using (5). This yields, after taking the standard deviation:

$$v_s(\boldsymbol{r}) = \gamma_E(0)h^2(\boldsymbol{r}) \quad \Longrightarrow \quad \sigma(\boldsymbol{r}) = \sqrt{\gamma_E(0)}h(\boldsymbol{r}) \tag{49}$$

Two remarks can be made about this equality.

The first concerns the lack of super-resolution capabilities possible with RIM-CF, as in this particular case, the standard deviation contains no frequencies outside $\mathcal{D}_{\mathrm{PSF}}$. In this case, a linear deconvolution, for instance the RIM-CF estimation (14), will not be able to retrieve any useful information beyond $\mathcal{D}_{\mathrm{PSF}}$. Even worse, in the case of RIM-CF, the spectrum of $u_1$ does not match that of $h$ on $\mathcal{D}_{\mathrm{PSF}}$, and RIM-CF has a slightly worse resolution than a deconvolved widefield. This lack of super-resolution capacity from the RIM-CF approximation is displayed on Fig. 5. In a noisy configuration, RIM-CF would perform even worse, because of the unnecessary enhancement of noise frequencies existing beyond $\mathcal{D}_{\mathrm{PSF}}$.

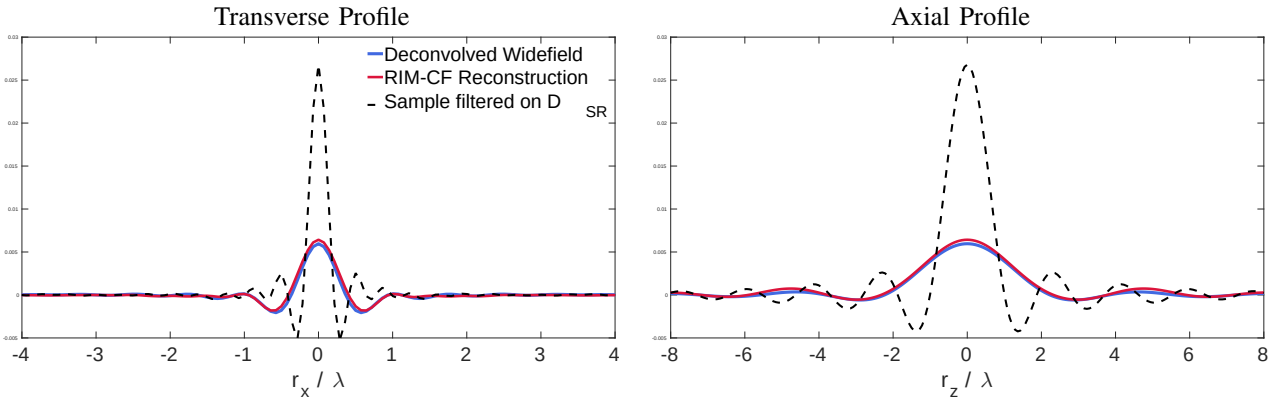


Fig. 5. Illustration of the lack of super-resolution of RIM-CF compared to a widefield deconvolution in the case of a single isolated source

Even in the case of a single emitter, it should be noted that the iterative estimators (RIM-STD and RIM-VAR), manage to retrieve the sample with improved resolution. This can be explained considering the results of [11], stating the variance of object is entirely defined by the frequencies of the object on $\mathcal{D}_{\mathrm{SR}}$, and that no two different positive objects yield the same variance.

The second remark deals with the meaning of the approximation (13), in this particular case:

$$u_1 \approx h. \tag{50}$$

It is somewhat surprising, since the PSF is approximated by a function with a larger Fourier support. In this regard, considering the full rank variance expression (51) with only a single source yields another remarkable equality:

$$h^2 = \sum_{n\geq 1} u_n^2 \quad \Rightarrow \quad h = \left(\sum_{n\geq 1} u_n^2\right)^{\frac{1}{2}} \tag{51}$$

It is known from Proposition 1 that the first term of this sum $u_1$, has a Fourier support $\mathcal{D}_{\mathrm{SR}} \supset \mathcal{D}_{\mathrm{PSF}}$. As such, the inclusion of all the other terms must cancel out all frequencies beyond $\mathcal{D}_{\mathrm{PSF}}$. This phenomenon is illustrated on Fig. 6, showing the convergence of (51) to $h$ as an increasing number of eigenvectors are considered.

### B. Case of multiple emitters

The previous discussion may let us think that if an object mainly consists of single emitters, the RIM-CF method will not be able to produce an improvement in resolution. However, such an analysis supposes a linear relationship between the standard deviation and the sample, which is not the case. In fact the resolution capability of RIM-CF highly depends on the local density of the sample (see Fig 8), performing better when the sample density is moderately high.

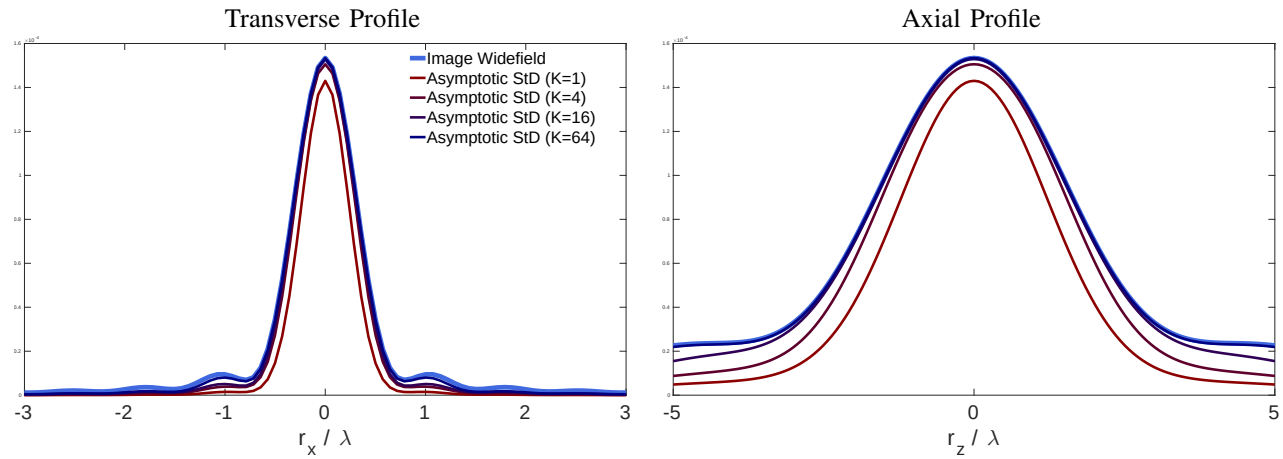


Fig. 6. Convergence of square root of the sum of the squared eigenvectors to $h$.

To understand this phenomenon, let us consider the simplest case of a pair of emitters separated by a distance $\boldsymbol{r}_0$, so that $\rho(\boldsymbol{r}) = \delta(\boldsymbol{r} - \frac{\boldsymbol{r}_0}{2}) + \delta(\boldsymbol{r} - \frac{\boldsymbol{r}_0}{2})$. In that case the variance can once again be computed exactly, leading to:

$$v_s(\boldsymbol{r}) = \gamma_E(\boldsymbol{r}_0) \times h\left(\boldsymbol{r} - \frac{\boldsymbol{r}_0}{2}\right) h\left(\boldsymbol{r} + \frac{\boldsymbol{r}_0}{2}\right) + I_0^2\left[h^2\left(\boldsymbol{r} - \frac{\boldsymbol{r}_0}{2}\right) + h^2\left(\boldsymbol{r} + \frac{\boldsymbol{r}_0}{2}\right)\right] \tag{52}$$

We note the presence of a correlation term $\gamma_E(\boldsymbol{r}_0)h(\boldsymbol{r} - \frac{\boldsymbol{r}_0}{2})h(\boldsymbol{r} + \frac{\boldsymbol{r}_0}{2})$. When the sources are far apart, they do not interact with each other, and can thus be seen as single emitters, and no super-resolution is possible from RIM-CF (see Fig. 7). This is further justified by the fact, that when the correlation term is negligible, the standard deviation of images is well approximated by:

$$\sigma_s(\boldsymbol{r}) \approx h \otimes \left[\delta(\boldsymbol{r} - \frac{\boldsymbol{r}_0}{2}) + \delta^2(\boldsymbol{r} - \frac{\boldsymbol{r}_0}{2})\right], \tag{53}$$

and the standard deviation is once again simply the convolution of the sample with the PSF.

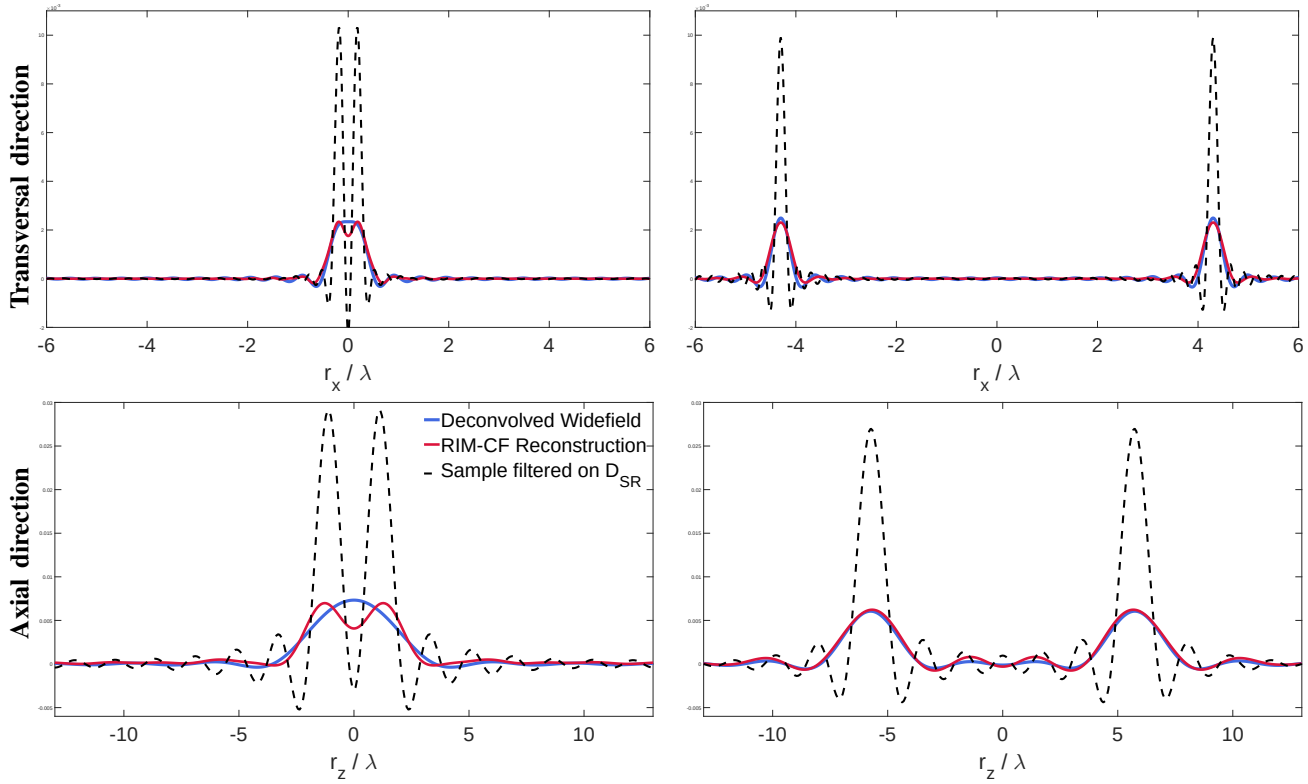


Fig. 7. Evolution of the transverse (above) and axial (below) resolution as a function of the separation distance between a pair of point-wise emitters.

On the other hand, when the distance between the beads is close to the diffraction limit, and the correlation term does not simplify anymore. In this case, the presence of a high frequency interaction enables an improved resolution from RIM-CF.

When there are more than two beads, several correlation terms appear in the variance, leading to an even more complex analysis of the situation. In this case, the resolution gain of RIM-CF depends on the local density of the sample, and is hopefully higher in denser regions. To illustrate this, we consider a 2D problem, where the sample consists of $N = 50$ fluorophores randomly spread within a disk of radius $r$, so that the sample density is given by $d = \frac{N}{\pi r^2}$. Fig. 8 compares a deconvolved widefield image with both RIM-CF and RIM-STD reconstructions for various density levels. We can observe here that RIM-STD leads to the same resolution gains, whatever the sample density.

## APPENDIX F
## PROCESSING OF 2D EXPERIMENTAL DATA

To supplement the simulation and experiments proposed in this article, we propose an additional processing of biological data, based on the data presented in [9]-Fig.2A (Vimentin network). The reconstructions obtained with the new estimators are compared with the original reconstruction, Fig. 9.

This processing of experimental biological data summarizes all the points already presented in the article, namely:

- The standard deviation of the images already presents a significative improvement over the widefield image, but its deconvolution using a suitable filter (RIM-CF) provides additional resolution gains.
- Both iterative methods achieve the same final resolution level, but the new accelerated method RIM-STD converges six times faster than RIM-VAR.
- The convergence speed gain achieved with RIM-STD over RIM-VAR is reduced on experimental data, where a full two-fold improvement is impossible due to deteriorated imaging conditions. In that case, RIM-CF, yields a final resolution very close to the one achieved with iterative estimators, as its already imperfect model is less impacted by images perturbations.

In this reconstruction, it appears that RIM-STD creates a reconstruction with a higher background noise level, which may reduce the interpretation of lower intensity regions. This phenomenon that doesn't appear in simulation with standard noise hypothesis may need further investigations. We suspect that it originates from a sharper inversion of hot and cold pixels of the camera, that are not taken into account in the proposed version.

## APPENDIX G
## 3D EXPERIMENTAL INFORMATION

The microscope used for the 3D acquisition is a custom-built one, made of optomechanical elements presented in Fig. 10. fO is a 100x objectify (Nikon, N100X-PFO, nominal NA=1.3, effective NA=1.21), fT, fR1, fR2 are Thorlabs TTL165-A, AC254-125-A-ML, AC254-125-A-ML. The exact position of the ETL and its associated offset lens (Optotune EL-10-30-Ci Series) is finely adjusted to optimize the detection PSF. We found that placing specifically the ETL in the Fourier plane and not the offset lens or the mid-point between the offset lens and the ETL was optimal. The lenses fO, fT , fR1, fR2 are arranged to ensure telecentry of the imaging path. In the excitation branch of the set-up, light from a 488 nm diode laser (Oxxius, LBX-488-200-CSB), impinging on a diffuser, is relayed to the pupil of the objective by fR3 and fR4 (Thorlabs LA1986-A-ML and LA1131-A-ML). Successive speckles are triggered through rotation of the diffuser with a rotary stage (2-axis controller 8SMC4-USB-B9-2). The illumination beam is combined with the imaging path with a dichroic miror (semrock, Di02-R488-25x36). Fluorescence is further filtered with a bandpass filter (Semrock, FF01-525/45-25) and imaged with a camera (Prime BSI Express Scientific CMOS Camera). A Matlab interface controls all instruments of the setup.

## APPENDIX H
## 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

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

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 [35, Chap. 6]. The cutoff frequency of the OTF is set to $f_c = 40$, so that iterative methods using noise-free (*i.e.*, asymptotic) 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)[7]. 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

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

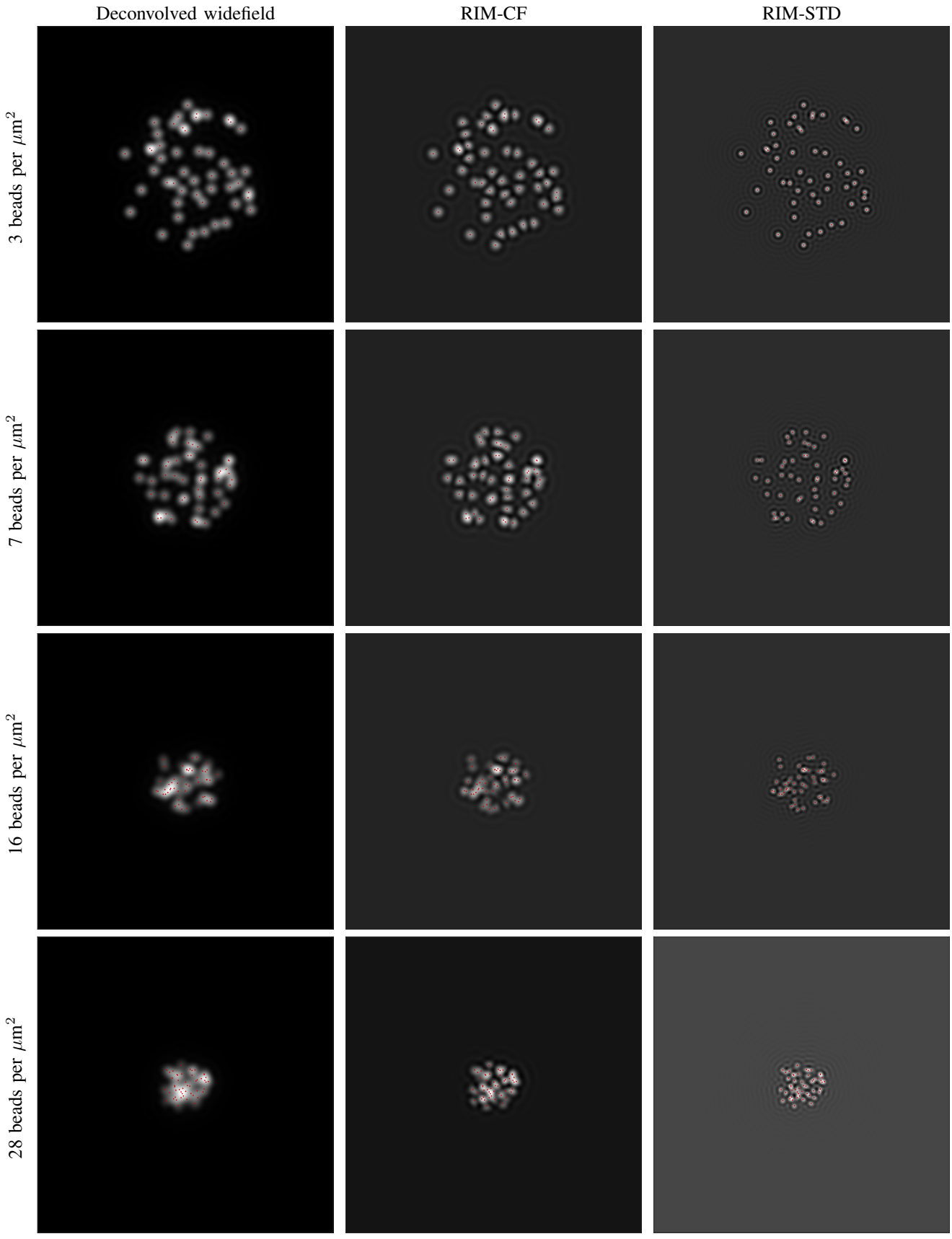


Fig. 8. Comparison of RIM-CF and RIM-STD to the widefield image depending on the fluorophore density. The numerical aperture is set to 1, and the illumination and collection wavelenght are set to 545nm.

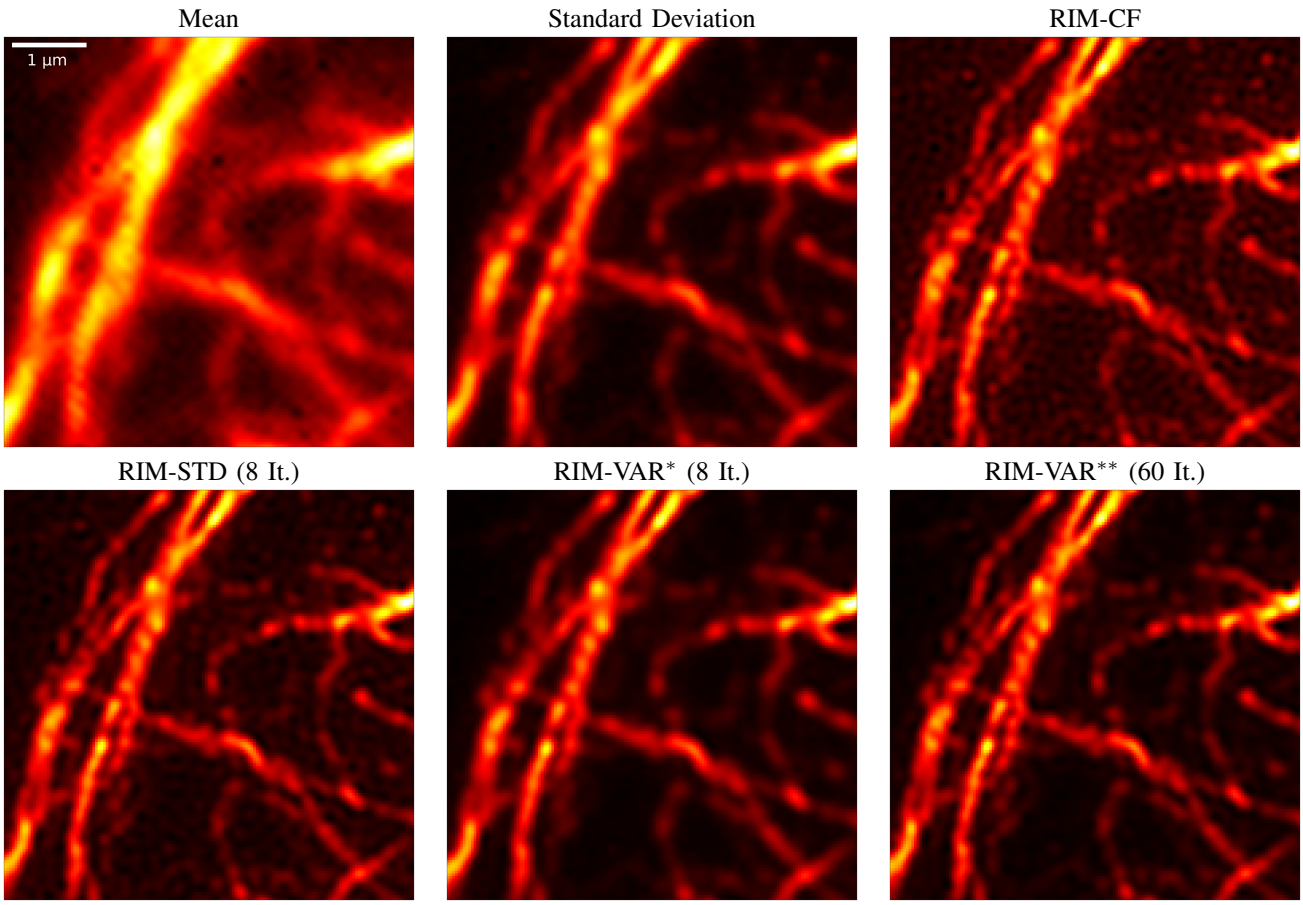


Fig. 9. Reconstruction of the different estimators on the same dataset used in [Mangeat2021]-Fig.2A (Vimentin network). *Same computation time as RIM-STD. **Computation until convergence (60 It.). All images are displayed in min/max scale and contrast has not been adjusted.

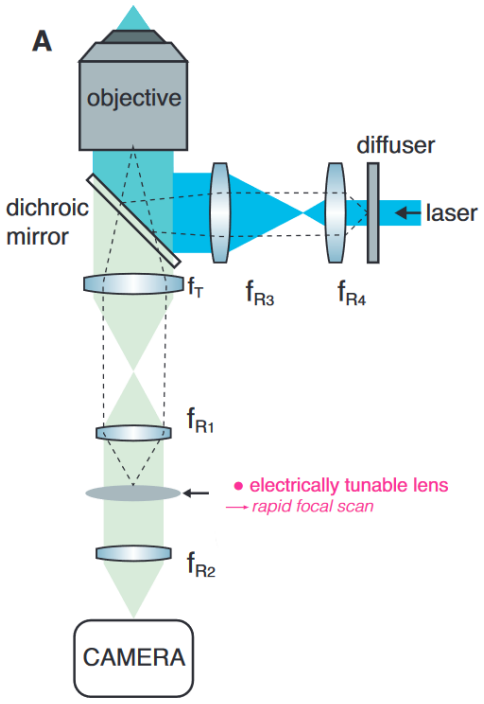


Fig. 10. Remote focusing setup used for the acquisition of the data presented in Fig. 4

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 [10]); 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. 11 shows some intermediate quantities generated in one simulation, namely one acquisition $\boldsymbol{z}_m$, as well as the statistics derived[8] from the $M = 500$ observations in the current dataset.

The results of the MC simulation are given for several values of the regularization parameter $\mu$ in Figs. 12 and 13, respectively for RIM-CF and RIM-STD, with the truncation level K = 10. Let us recall that a given tuning of $\mu$ achieves a reconstruction quality within a bias *vs.* variance tradeoff [13, Chap. 4]. We can then identify three distinct regimes in these results. When $\mu$ 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 $\mu$ 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 $\mu$ 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.

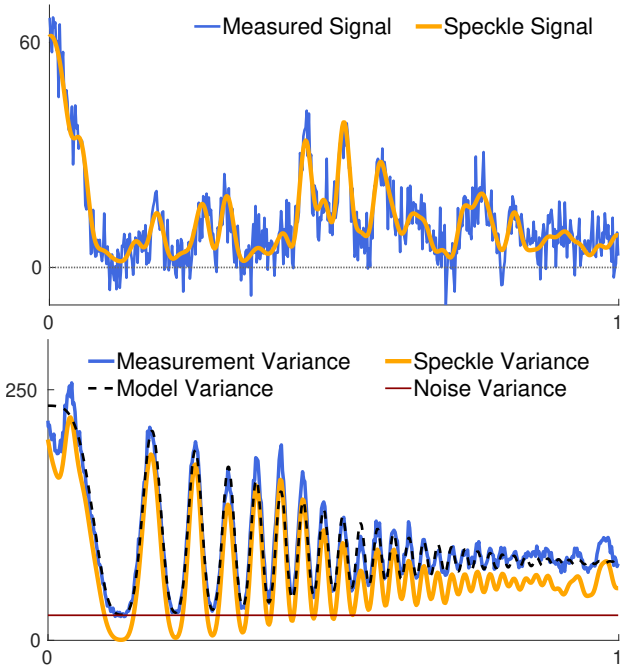


Fig. 11. Illustration of the simulation environment. The top image represents a speckled measurement before and after the addition of the noise. The bottom image represents both the empirical and model variances used within RIM-VAR, in the case where 500 speckles have been generated. It is noticeable 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.

Similarly, Figs. 14 and 15 represent the proposed estimators applied after the data pre-filtering procedure proposed in Sec. II-B1. Raw images have then been convolved with the filter $g$ defined by (11), with $\eta = 10^{-5}$. As a result, the triangular

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

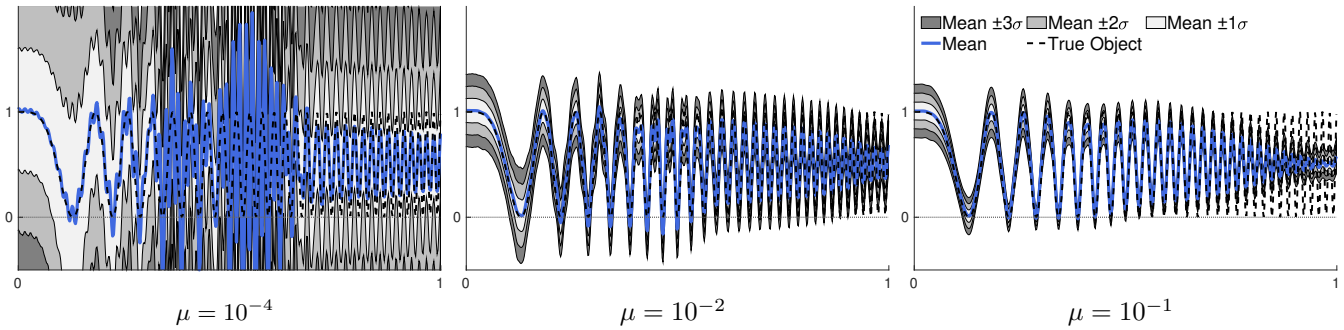


Fig. 12. Statistical result of Monte-Carlo reconstruction with RIM-CF.

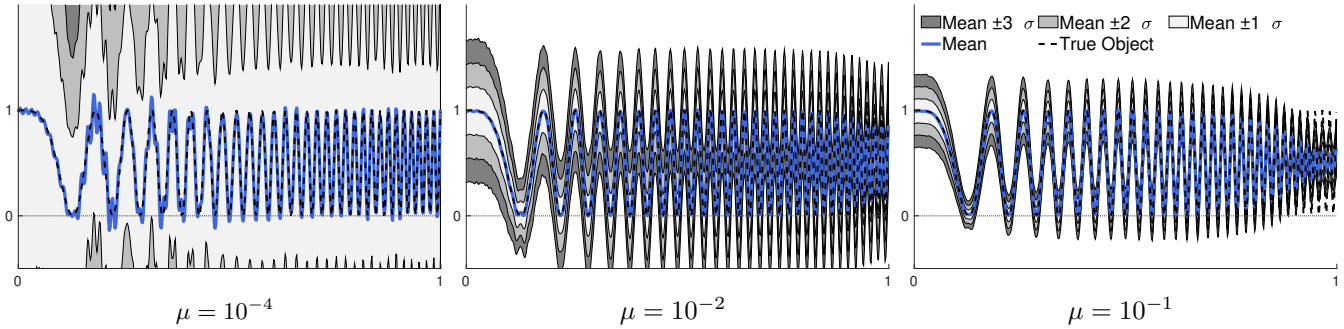


Fig. 13. Statistical result of Monte-Carlo reconstruction with RIM-STD

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. 14) with pre-whitened data provides an overall improvement of the solution.

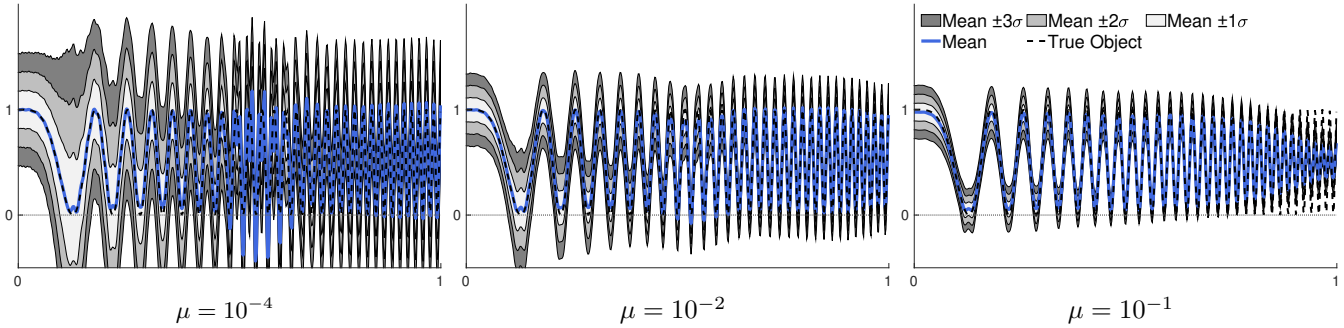


Fig. 14. Statistical result of Monte-Carlo reconstruction with RIM-CF, after data pre-filtering

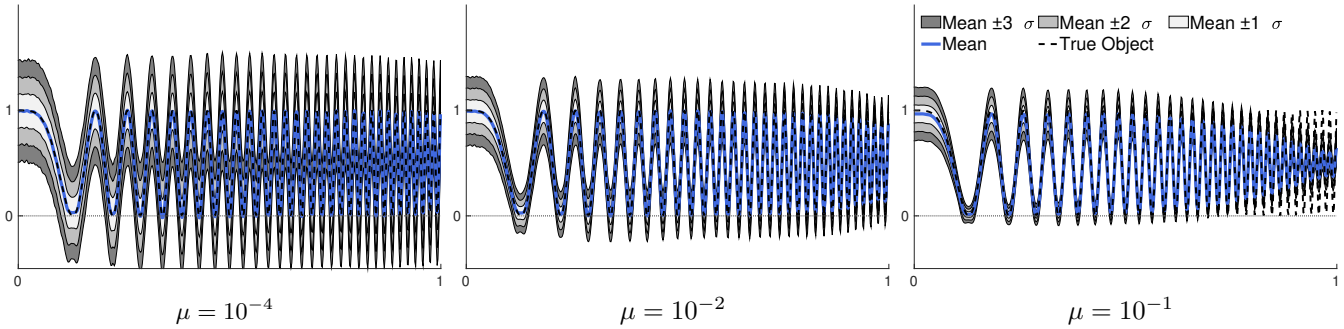


Fig. 15. Statistical result of Monte-Carlo reconstruction with RIM-STD, after data pre-filtering