Learning low-dimensional models of microscopes

Valentin Debarnot [*], Paul Escande [†], Thomas Mangeat [‡], Pierre Weiss [§]

Abstract

We propose original, accurate and computationally efficient procedures to calibrate fluorescence microscopes from micro-beads images. The designed algorithms present many singularities. First, they allow to estimate space-varying blurs, which is a critical feature for large fields of views. Second, we propose a novel approach for calibration: instead of describing an optical system through a single operator, we suggest to vary the imaging conditions (temperature, focus, active elements) to get indirect images of its different states. Our algorithms then allow to represent the microscope responses as a low-dimensional convex set of operators. This novel approach is shown to significantly improve the estimation on a wide-field microscope. It is deemed as an essential step towards the effective resolution of blind inverse problems. We illustrate the potential of the approach by designing an original procedure for blind image deblurring of point sources and show a massive improvement compared to commercial software.

1 Introduction

Many recent breakthroughs in optics pertain to the field of computational microscopy: computers play a critical role to generate images. This evolution allowed to observe objects with unprecedented contrasts, temporal/spatial resolutions or gave access to new quantitative features. To name a few examples, let us mention Single Molecule Localization Microscopy (SMLM), Structured Illumination Microscopy (SIM), Total Internal Reflection Fluorescence microscopy (TIRF) or Stimulated Emission Depletion (STED) microscopy.

A common prerequisite for these techniques is the design of an accurate mathematical model of the optical system. This step is critical since the generation of images usually relies on an explicit or implicit inversion of this model. The advent of these microscopes therefore makes it more and more important to finely characterize their transfer function.

The dominant model by far in the literature are space invariant systems: the point spread function (PSF) is identical wherever its position in space. While simplifying the theoretical, numerical and experimental aspects, this assumption is however often unrealistic. For instance, when dealing with 3D microscopy, there is clear theoretical and experimental evidence that the variations of the PSF in the z-axis need to be accounted for [1, 2]. Similarly, the invariance in the (x, y) plane becomes questionable for large fields of view (see for instance Fig. 3a). Neglecting this aspect can have dramatic consequences. For instance, it was shown in [3,4] that this approximation can severely damage the reconstruction of images in single molecule imaging, with localization errors of more than 20% for a displacement of less than 200nm. The effects would be even more stringent for large fields of views which are a current challenge with the improved quality of SCMOS detectors. These model mismatches can significantly downgrade the performance of all computational microscopy systems and it is hence critical to finely estimate the optical response of the system.

Existing works A well spread approach to describe the response of an imaging system consists in using Fourier optics [5]. This theory provides a nice description of the system through the pupil function of the objective. In this domain, it is possible to derive mathematical models of space variant PSFs [4, 6, 7] and to infer the parameters of these models (e.g. Zernike coefficients) from experimental data. There are however two limitations to these approaches. First, they are often based on parameters such as the numerical aperture, the wavelengths, or many other physical quantities. The more parameters, the more precise the model, but the harder it becomes to finely characterize them experimentally. In addition, some active components such as micro-mirrors introduce additional perturbations which cannot be easily modeled or inferred. The second problem comes from numerical considerations: the dependency between the model and its parameters is non-linear, which inevitably leads to non-convex estimation procedures, leading to local minima and additional inaccuracies.

Instead of relying on a physical model it is possible to directly estimate the optical response from data. We will follow this approach in this paper. Imaging fluorescent micro-beads in a cover-slide in 2D or a cylinder of agarose in 3D gives a partial idea of the system by providing an access to a few sampled and noisy scattered PSFs. This information can be used to estimate a space invariant system by averaging multiple micro-beads [8]. When the images are aliased it is even possible to obtain a super-resolved estimation [9, 10]. It becomes more delicate to estimate a space variant system. A few researchers - especially in the field of astrophysics [1] - have addressed this issue. The general framework is

---
[*]IMT, CNRS and Université de Toulouse, France
[†]I2M, CNRS and Université de Aix-Marseille
[‡]CBI, CNRS and Université de Toulouse
[§]IMT, CNRS and Université de Toulouse, France

---
[1]In astrophysics, the PSFs variations may be due to weak

the following: a parameterized PSF model is designed either from physics equations or from the data itself. The observed PSFs are then interpolated to cover the whole field of view. In [11, 12], the authors propose to decompose the PSFs over a low dimensional basis and to interpolate the coefficients using thin-plate splines. A subset of the authors proved that this method was minimax optimal in [13] and we will propose a refined version in this paper. It is also possible to use more advanced interpolation methods, using e.g. optimal transport [10], but this method would not scale computationally to the large field of views considered in here.

Our contribution While there now exists a solid theoretical and algorithmic framework to estimate space varying optical responses of optical systems, these methods are often tested on synthetic data that do not reproduce all the complexity of real microscopy images. For instance, the estimation of a PSF requires a very careful treatment of the background and of the noise. Its interpolation requires specific care to avoid obtaining unrealistic results far from the observed responses. Our first objective is to provide *precise estimation algorithms adapted to real data emanating from fluorescence microscopy*. The second and arguably most original contribution is a *new way to calibrate an optical system by learning all its possible states*. Instead of trying to estimate a single operator to describe the microscope, we propose to learn a whole family of possible states by varying the experimental conditions, following the recent theoretical work [14]. We finally show how the proposed methodology allows to *design new algorithms for blind image deblurring*, which significantly outperform more traditional models proposed in the literature.

2 Operator estimation

2.1 Notation

In all the paper, bold fonts refer to vectors, matrices or vectorial functions while regular fonts refer to scalar numbers or functions. The i-th value of a vector $\boldsymbol{x}$ is denoted either x_i or $\boldsymbol{x}[i]$. The ℓ^p norm of a vector $\boldsymbol{x}$ is denoted $\|\boldsymbol{x}\|_p$. The value of a function f is $f(x)$ and its ℓ^p-norm is denoted $\|f\|_p$. The delta Dirac function at a position $\boldsymbol{x} \in \mathbb{R}^d$ is denoted $\delta_{\boldsymbol{x}}$. In all the paper, the integers I, J, K, M and N refer to a number of components ($I \hateq$ size of PSF basis, $J \hateq$ size of space variations basis, $K \hateq$ number of observed microbeads images, $M \hateq$ number of observed microbeads, $N \hateq$ number of pixels of an image).

2.2 Preliminaries

A space varying blurring operator $H : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ can be seen as a linear integral operator mapping an image u to its degraded version Hu through

gravitational lensing and reveal distant massive galaxies.

the formula

$$Hu(\boldsymbol{x}) = \int_{\mathbb{R}^d} L(\boldsymbol{x}, \boldsymbol{y}) u(\boldsymbol{y}) \, d\boldsymbol{y}. \tag{1}$$

The function $L(\cdot, \cdot)$ is called *kernel* of the operator. It describes the impulse response of the system at every location $\boldsymbol{z} \in \mathbb{R}^d$ of the image domain since:

$$(H\delta_{\boldsymbol{z}})(\cdot) = L(\cdot, \boldsymbol{z}). \tag{2}$$

The PSF $S(\cdot, \boldsymbol{z})$ of the system at $\boldsymbol{z}$ is defined as the impulse response centered at 0, i.e.

$$S(\cdot, \boldsymbol{z}) = L(\cdot - \boldsymbol{z}, \boldsymbol{z}). \tag{3}$$

The function S is called the space varying impulse response of the system. This work is based on two important assumptions.

Assumption 1 (PSF approximation). *Every PSF in the field of view is well approximated by its projection over a low-dimensional orthogonal basis $(h_i)_{1 \leq i \leq I}$, i.e.*

$$S(\cdot, \boldsymbol{z}) \simeq \sum_{i=1}^{I} \langle S(\cdot, \boldsymbol{z}), h_i \rangle h_i. \tag{4}$$

This assumption is valid both from a theoretical and an empirical viewpoint. It is indeed well known that any smooth function can be well approximated by its projection on a low dimensional subspace. Typical bases include splines or low frequency Fourier atoms [15]. In practice, we can also construct the family (h_i) by computing the principal component analysis of a family of sampled PSFs. The numerical experiments performed in the paper reveal that for our imaging systems, as little as $I = 5$ elements are enough to capture all possible PSFs accurately. In addition, restricting the PSFs to live on a low dimensional subspace is an efficient method to denoise them as will be illustrated in the numerical section.

Assumption 2 (PSF variations). *Letting*

$$\alpha_i(\boldsymbol{x}) := \langle S(\cdot, \boldsymbol{x}), h_i \rangle \tag{5}$$

denote the i-th coefficient of the PSF at $\boldsymbol{x} \in \mathbb{R}^d$, we assume that α_i varies slowly in space.

The spatial variations of the PSFs are related to field dependent aberrations. For instance, field curvatures produce phase delays which are not shift-invariant. A sufficient condition for having smooth variations of the PSFs is that the phase delays vary slowly in space. This hypothesis has been verified experimentally using arrays of micro-beads for instance, see e.g. [3, 4].

Under Assumptions 1 and 2, a space varying operator is completely characterized by the pair $(\alpha_i, h_i)_{1 \leq i \leq I}$. Of interest, this representation of the operator also leads to fast numerical computations using a structure called *product-convolution*. This decomposition has been developed and improved for the last two decades [1, 16–20]. Its precise approximation rates have been studied in [21] and we refer to the previous references for more insight on these structures. A key property of this decomposition is the following.

Proposition 1 (Product-convolution [18,19,21]). *Assume that a blurring operator H has a space varying impulse response S defined by the tensor:*

$$S(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{I} \alpha_i(\boldsymbol{x}) h_i(\boldsymbol{y}),$$

then for any image u, we have

$$Hu = \sum_{i=1}^{I} h_i \star (\alpha_i u),$$

where the symbol $\star$ stands for the convolution operator.

Hence, the numerical complexity of computing a space-varying operator is just I times the one of a convolution, which can be achieved efficiently using Fast Fourier Transforms for instance.

2.3 Estimating a single operator

Under Assumptions 1 and 2, the problem of estimating the operator reduces to recovering the low dimensional bases (h_i) and (α_i), or at least their discretized counterparts $(\boldsymbol{h}_i)$ and $(\boldsymbol{\alpha}_i)$. In this paragraph, we describe the general principle of the estimation of a blurring operator from a single image of fluorescent micro-beads. The general process is described in Fig. 1.

The first step consists in extracting the most relevant PSFs in the form of small patches (see Appendix A). Then the background is estimated and removed on each patch independently to avoid biases in the PSFs estimation. We propose an original and lightweight procedure in Appendix B. A principal component analysis is then performed to estimate the basis $(\boldsymbol{h}_i)$ (see Appendix A). Each PSF is projected on this basis and the resulting coefficients are interpolated spatially to provide an estimate of the functions (α_i), which can then be discretized as $(\boldsymbol{\alpha}_i)$ (see Appendix C). All those steps are subtle and need to be performed carefully to obtain precise estimates. The technical details reported in the appendices are therefore of great importance.

As an output of the algorithm, the pair $(\boldsymbol{h}_i, \boldsymbol{\alpha}_i)_{1 \leq i \leq I}$ provides a complete description of the operator, since we know an approximation of the PSFs at each image location. The integer I is a user provided parameter.

2.4 Estimating a subspace of operators

Motivation A microscope produces different transfer functions depending on physical parameters that can be hard to control. Typical examples include temperature variations, focal screws, small tilts of optical elements, surface flatness of cover-slides, slight variations of a spatial light modulator rest state,... In those conditions, capturing a single operator (as proposed in the previous section) to describe the microscope might lead to model mismatches and reconstruction errors. In this section, we propose an alternative approach where we aim at learning a *family of plausible operators* that capture all the possible states of a microscope.

The principle and the mathematical foundations behind this approach (statistical properties and fields of application) were recently established by a subset of the authors in [14]. We refer the interested reader to this paper for more details. We provide a simplified description below.

Principle The first requirement to apply this technique is to image stacks of fluorescent micro-beads (in 2D or 3D) under multiple conditions. This process can be automatized when using advanced optical tables with motorized stages and thermostatic chambers. An alternative is to probe only the "extreme" conditions (e.g. highest and lowest plausible temperatures and tilts). After this experimental process is achieved, we have access to a set of images $(\boldsymbol{u}_k)_{1 \leq k \leq K}$. The idea of our estimation procedure is to apply the following procedure:

1. For each image $\boldsymbol{u}_k$, extract the most relevant PSF patches (see Appendix A) and remove the background (see Appendix B).

2. Apply a principal component analysis to the set of patches over multiple images and keep I principal components.

3. Apply a z-score test to discard the patches that are likely outliers (e.g. multiple PSFs in a patch).

4. Reapply a principal component analysis to better estimate the principal components.

5. For each image $\boldsymbol{u}_k$ and each coefficient i, interpolate the coefficient maps $\alpha_{i,k}$ (see Section C).

6. Apply a randomized principal component analysis [22] to the whole set of sampled interpolation maps $(\boldsymbol{\alpha}_{i,k})_{1 \leq i \leq I, 1 \leq k \leq K}$. It is often necessary to apply a randomized SVD here since the interpolation maps $\boldsymbol{\alpha}_{i,k}$ are typically large images.

7. Keep the J largest principal components $(\boldsymbol{a}_j)_{1 \leq j \leq J}$.

8. Project each interpolation map $\boldsymbol{\alpha}_{i,k}$ onto the basis $(\boldsymbol{a}_j)_{1 \leq j \leq J}$, to obtain the matrices $\boldsymbol{\Gamma}_k \in \mathbb{R}^{I \times J}$ defined by
$$\boldsymbol{\Gamma}_k[i,j] = \langle \boldsymbol{\alpha}_{i,k}, \boldsymbol{a}_j \rangle.$$

The output of this process is two orthogonal bases $(\boldsymbol{h}_i)_{1 \leq i \leq I}$ (which describe the PSFs compactly) and $(\boldsymbol{a}_j)_{1 \leq j \leq J}$ (which describe the PSFs variations compactly) as well as a set of matrices $(\boldsymbol{\Gamma}_k)_{1 \leq k \leq K}$ in $\mathbb{R}^{I \times J}$ (which describe the operators associated to each image $\boldsymbol{u}_k$). The operator $\boldsymbol{H}_k$ associated to the k-th input image $\boldsymbol{u}_k$ is then defined for all $\boldsymbol{u}$ by:

$$\boldsymbol{H}_k \boldsymbol{u} = \sum_{1 \leq i \leq I} \sum_{1 \leq j \leq J} \boldsymbol{\Gamma}_k[i,j] \boldsymbol{h}_i \star (\boldsymbol{a}_j \odot \boldsymbol{u}).$$
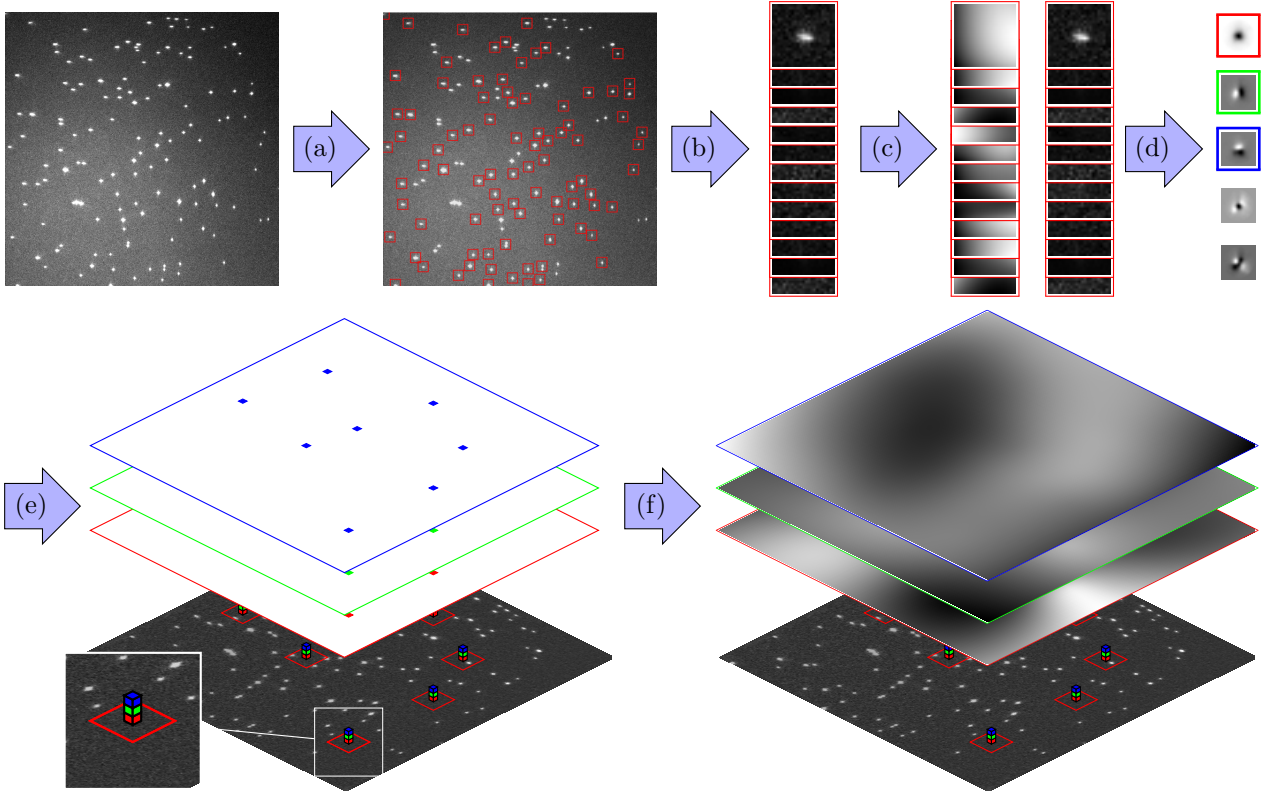
Figure 1: Structure of the algorithm for single operator estimation. (a) Background removal procedure. (b) Selection of well isolated PSFs. (c) Extraction of relevant PSF patches. (d) Principal Component Analysis of the PSFs to find a low dimensional basis. (e) Projection of each selected PSF on the low-dimensional basis. (f) Interpolation of the PSFs coefficients using radial basis functions and correction to ensure admissbile PSFs.

Reducing the family of admissible operators The subspace of operators $\mathcal{H}$ that compactly describes the possible operators is defined by

$$\mathcal{H} \stackrel{\text{def.}}{=} \text{span}(\boldsymbol{L}_{i,j}, 1 \le i \le I, 1 \le j \le J)$$

where $\boldsymbol{L}_{i,j}\boldsymbol{u} \stackrel{\text{def.}}{=} \boldsymbol{h}_i \star (\boldsymbol{a}_j \odot \boldsymbol{u})$ is a simple product-convolution operator. The dimension of $\mathcal{H}$ is $I \times J$.

However, all operators in the subspace $\mathcal{H}$ are not plausible and it is possible to further restrict the family of operators as follows. Assuming that all the extreme conditions have been explored, we can construct the *convex hull* of the coefficients $\boldsymbol{\Gamma}_k$:

$$\mathcal{C} \stackrel{\text{def.}}{=} \text{conv}(\boldsymbol{\Gamma}_k, 1 \le k \le K)$$
$$\stackrel{\text{def.}}{=} \left\{ \sum_{k=1}^{K} \lambda_k \boldsymbol{\Gamma}_k, \lambda_k \ge 0, \sum_{k=1}^{K} \lambda_k = 1 \right\}.$$

The quality of the estimate $\mathcal{C}$ with respect to the number of observations K was studied in [14].

2.5 Implementation details

Normalizing the operators Of importance, let us mention that the procedures described previously suffers from a well known identifiability issue. Since the micro-beads intensity is usually unknown, the operator can be estimated only up to multiplicative constant. To address this problem, it is possible to replace the matrices $\boldsymbol{\Gamma}_k$ by a normalized version

$$\bar{\boldsymbol{\Gamma}}_k = \boldsymbol{\Gamma}_k / \sum_{i,j} \boldsymbol{\Gamma}_k[i,j]$$

and to replace the convex hull by the conic hull

$$\mathcal{C} = \text{cone}(\boldsymbol{\Gamma}_k, 1 \le k \le K) \stackrel{\text{def.}}{=} \left\{ \sum_{k=1}^{K} \lambda_k \boldsymbol{\Gamma}_k, \lambda_k \ge 0 \right\}.$$

Normalizing the PSFs It may also happen that the micro-beads are not perfectly identical and have different fluorescence levels. In that case it is important to normalize the PSF patches (after background removal) by imposing that they sum to 1. By doing so, the operators will be estimated without accounting for the variations of intensity that they may induce due to non uniform illumination. This effect can still be captured by assuming that the loss of intensity is proportional to the background. It then suffices to multiply the normalized patches by the background estimate.

We also normalize the ℓ^2-norm of the PSFs prior to compute principal component analysis, in order to give the same importance to every PSF.

Selecting the subspace sizes The subspace sizes I and J are the two values that a user needs to provide in order to estimate the subspace. If the number I (related to the subspace of PSFs) is too small, then

the PSFs will be badly reproduced, while a value that is too large will result noisy operators (the so-called over-fitting in machine learning). Similarly, the number J captures the variations of the PSFs and has to be chosen with caution. Finally, we would like I and J to be as small as possible to reduce the computing times: the cost of applying a product-convolution operator is directly proportional to I.

The simplest way to choose I is to test different values on a subset of representative PSFs and keep the lowest value that leads to a visually decent reconstruction of the PSFs. The same can be done with J. In practice, we observed that the values $I = 5$ and $J = 5$ faithfully reproduce the operators from a perceptual point of view.

Another possibility is to apply recent results in statistics [23] that provide a simple and optimal way (under a Gaussian noise assumption) to choose I and J. The rule consists in keeping the principal components associated to a singular value larger that $2.858 \cdot \sigma_{med}$, where σ_{med} is the median of the set of singular values. This procedure requires computing the set of all singular values to evaluate the median. In practice, this is possible only for the PSF patches which are low dimensional.

2D vs 3D PSF models   All the proposed algorithms are implemented in 2D, but their extension to 3D is straightforward. From a practical point of view, the estimation of a 3D operator requires to image uniformly scattered microbeads in a medium such as an agarose gel. We do not report experimental results for this problem in the paper.

3   Results

In this section, we test the proposed algorithms against 2 different data-sets: the first one is simulated while the other comes from a wide-field microscope. We start with the estimation of a single operator and of a subspace of operators. We conclude the paper with an application to the deblurring of sparse objects.

3.1   Data-sets

3.1.1   Simulation

We generate several realistic product-convolution operators by designing a collection of admissible PSFs and space variations. We use the following simple model to express the PSF h at position z

$$h(z) = \left| \mathcal{F} \left(\rho \exp \left(2i\pi (\Phi + f(z)) \right) \right) \right|^2, \qquad (6)$$

where ρ is the indicator function of a disk of radius NA/λ, $\mathcal{F}$ denotes the Fourier transform, $f(z)$ models the variation of the phase along the z axis, and

$$\Phi = \sum_i c_i Z_i,$$

with Z_i the i-th Zernike polynomial. Parameters NA and λ denote respectively the numerical aperture and

the emission wavelength. This model allows to reproduce common effects encountered in PSF engineering such as astigmatism. We display an example of simulated astigmatic PSF at different z positions in Fig. 2a. Each image is then produced by taking random values of the different parameters c_i, NA, and λ.

The spatial variations are modeled by z-variations of the sample. This phenomenon occurs naturally when facing spherical aberrations of non perfectly flat coverslips. We use random polynomials of low degree to generate various operators. An additive background is generated with a smooth Gaussian random process, and the final image is degraded with Poisson noise, see Fig. 2 for a simulation example.
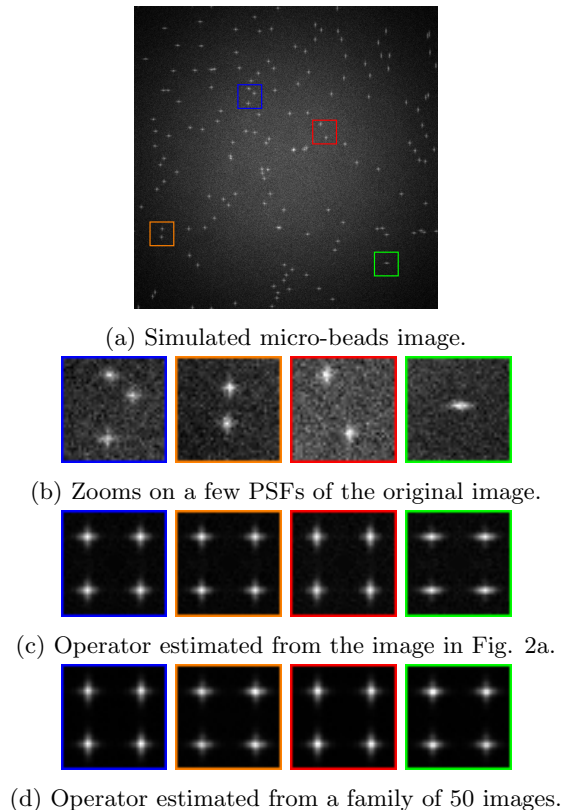

(a) Simulated micro-beads image.

(b) Zooms on a few PSFs of the original image.

(c) Operator estimated from the image in Fig. 2a.

(d) Operator estimated from a family of 50 images.

Figure 2: Simulation experiment: synthesized operators are applied to randomly scattered Dirac masses (512×512 pixels).

3.1.2   Real data

In all experiments, we used a perfectly plane monolayer of micro-beads of a 100nm diameter. There is no refractive index mismatch between the cover-slide and the immersion oil, allowing to avoid spherical aberrations. The data-set is composed of micro-beads of 100nm of diameter acquired with a standard wide-field microscope. The camera is a CMOS with pixel size of 43nm and produces images of 2304×2304 pixels. We collected 18 stacks of fluorescent micro-beads, each one is 8µm thick and is composed of 21 z-stacks. We keep only the 5 central slices since the beads are too degraded when far away from the focal plane. This amounts to a total of 90 images and more than 9700

2D PSFs. We display one image in Fig. 3a.

(a) 2304 × 2304 image of fluorescent micro-beads

(b) Zooms on a few PSFs of the original image.

(c) Estimated operator from the image in Fig. 3a.

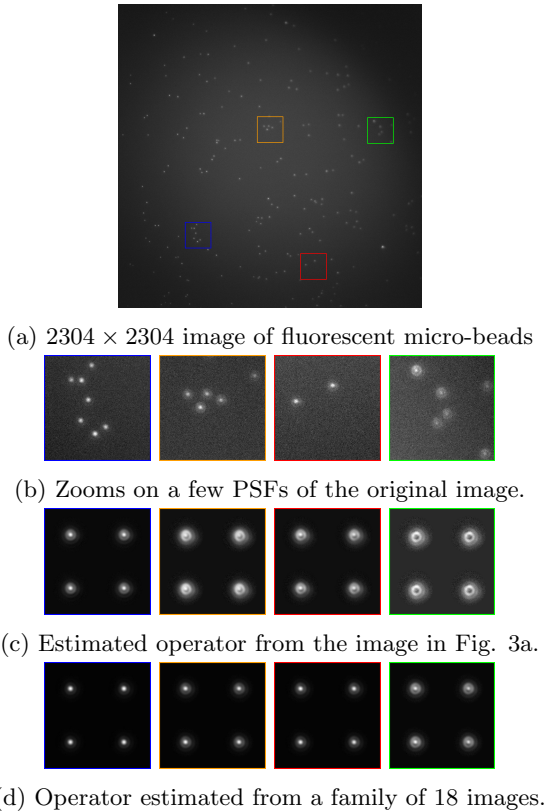(d) Operator estimated from a family of 18 images.

Figure 3: Image of micro-beads taken with a wide field microscope and estimation results. The contrasts have been stretched for a better visualization.

3.2   Estimating operators

In this section we illustrate some features of the proposed methods by estimating a single operator and a subspace of operators from the image of micro-beads generated by the previously described microscopes. Each experiment is performed on a workstation equipped with Intel Xeon E5 and a GPU card Nvidia Tesla K20c from 2012 (2019 technologies are expected to be 4 times faster).

3.2.1   Simulation

We apply the proposed estimation procedures both for a single operator and for a subset of operators. The computing time for a single operator is 15 seconds when estimating 3 principal elements for the PSFs and 3 principal components to describe the coefficients variations. To estimate the subspace of simulated operators, we used 50 different micro-beads images. The computing times increased to 500 seconds (i.e. 10 second per image). The results are displayed in Fig. 4. Of importance, notice that the results obtained with the subspace of operators are based on micro-beads images generated with *different* operators and in particular different PSFs. Despite this higher variety of possible shapes, the method is able to automatically infer the

common patterns and to achieve better denoising and estimation performance thanks to the redundancy.

The estimated operators can be visualized by applying them to a Dirac comb, see Fig. 2c and 2d. To compare the quality of reconstruction, we simply evaluate a rescaled version ℓ^2 distance between the resulting images and the true one. It is compulsory to rescale the distance, since there is an inherent ambiguity between the micro-beads intensity and the microscope gain.

Estimating a subspace of operators rather than a single operator improves the quality of the reconstruction allowing to go from 50% to 15% of relative distance between the estimation and the ground truth. While this figure is not per se impressive, the PSFs family is no longer corrupted with noise, and the coefficients maps family seems smoother. It is likely that other metrics would better reflect this fact.

3.2.2   Wide-field microscopy

We estimate a single operator and a whole subspace of product-convolution operators based on the 2304×2304 images from the data-set from wide-field microscopy. We set $I = 5$ and $J = 5$. Estimating a single operator takes about 150 seconds using 120 PSFs, while estimating the whole subspace takes about 3 hours (i.e 2 minutes per operator) using 9700 PSFs. These computing times are remarkable given the computer features and that the complete dataset weighs about 5Gb.

The PSFs and space variations bases are displayed in Fig. 5 and the estimated operators are displayed in Fig. 3. Similarly to the previous section, we observe that the basis and operators obtained using a large set of images is significantly less noisy. While the principal components beyond 3 contain a significant amount of noise for the single image, the 5-th component of the subspace approach still seems to contain useful geometrical features. The improvement of the coefficients maps is harder to evaluate since the corresponding convolution kernels have changed. Overall, learning the subspace of operators led to a significantly improved reconstruction of the operator with no visible residual noise remaining in Fig. 5.

3.3   Blind deblurring

Image deblurring is a technique than can lead to significant improvements of image resolution and quality. In most acquisitions, this technique is however neglected since it requires strong skills both in optics, image processing and computer science. In particular, the prior calibration of a microscope is critical: model mismatches can lead to dramatic performance losses and oftentimes lead biologists to prefer using the raw images. In this paragraph, we show that the proposed methodology of learning a whole family of operators to describe the microscope allows to avoid a precise calibration before each experiment and therefore significantly eases the application of a deblurring algorithm.

The key observation is that identifying an operator from a single degraded image becomes rather easy

(a) Single image estimation.
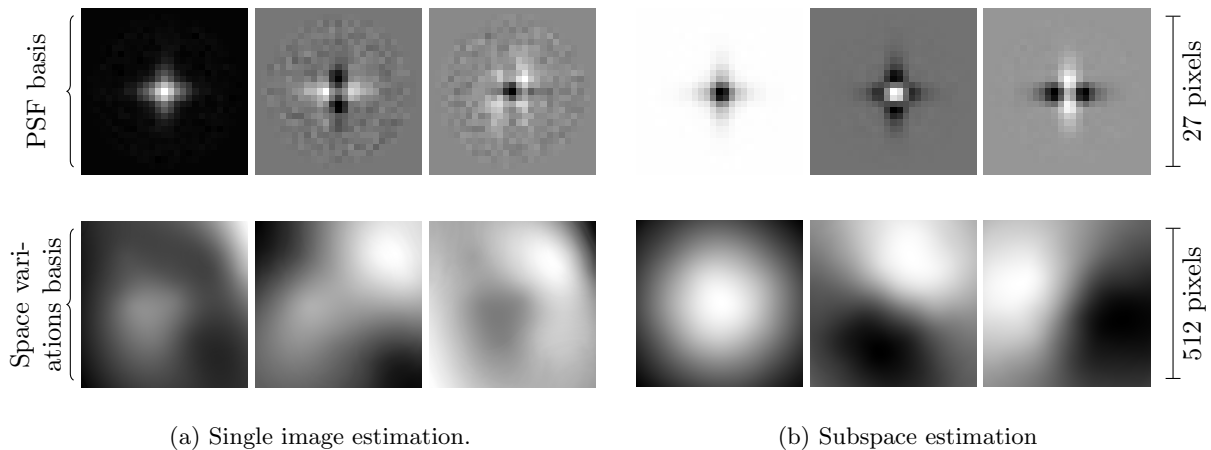
(b) Subspace estimation

Figure 4: Estimation on simulated operators. We compare the estimation with a single operator (left) and with a set of 50 randomly sampled operators (right). Top: the PSF basis $(h_i)_{1 \leq i \leq 3}$. Bottom: the space variations basis $(a_j)_{1 \leq j \leq 3}$. Observe that the PSF basis is significantly less noisy and that the space variations are significantly smoother when estimating over the set of 50 images.

when the operator depends linearly on a small number of parameters. In particular, if we know beforehand that the degraded image contains a few point sources, we show that a simple constrained least squares problem allows to recover the operator in Appendix D. We then design an original non-blind deblurring algorithm with a known operator for sparse + smooth images. The overall algorithm is called BSS for Blind Sparse + Smooth deblurring.

To assess the proposed methodology, we test the proposed algorithm on a real image of fluorescent micro-beads aligned along filament like structures. We perform this experiment on an image obtained with the same wide-field microscope as the one used to collect data-set from the real data. The blurry image is acquired at a distance 300nm from the focal plane Fig. 6a and 6b, (i). For comparison, the same sample at the focal plane is displayed in Fig. 6b, (ii).

We propose a comparison with the software Huygens Professional version 19.04 (Scientific Volume Imaging, The Netherlands). It allows to perform a patch-wise deblurring of the full image. Here, we used a 3x3 patch decomposition. On each patch, the PSF is chosen theoretically based on the specifications of the microscope. The reconstruction is displayed in Fig. 6b, (iii).

While the Huygens correctly removes the background, it fails to really increase the resolution: the micro-beads still consist of large spots which cannot be easily separated. The output of our blind deblurring algorithm in Fig. 6b, (iv) on its side is really convincing. It rather faithfully reproduces the image obtained at the focal plane in Fig. 6b, (ii) with an even better resolution. Observe that this is a really challenging setting: the input image has a significant amount of noise and while the PSFs on the left part are rather small, their diameter is about 40 pixels on the right of the field of view.

4   Conclusion

We proposed a set of fine algorithms to learn a set of product-convolution representations of optical responses in fluorescence microscopy. One of the main originality is to estimate a subspace of operators to capture the whole diversity of possible space-varying blurs of a given microscope. This is in sharp contrast with existing approaches which simply characterize the microscope by a single PSF, or - at best - by a single unstationary operator.

An important outcome of this work is that it strongly improves the identifiability in blind-inverse problems such as blind deblurring or blind super resolution. These arguably constitute two of the most challenging issues in computational imaging. For instance, recent theoretical progresses based on lifting techniques [24] require the prior knowledge of a low dimensional subspace. We have proposed an original blind deblurring approach coined BSS to efficiently solve this problem when imaging point sources with a smooth background.

Future works will consist in extending the existing codes to 3D images, providing an open-source toolbox together with realistic responses of microscopes. This will enable the optics and signal processing communities to test their algorithms against realistic operators.

We expect the proposed work to have far reaching applications ranging from the metrology of imaging systems to new advanced microscopy methods such as supercritical angle localization microscopy, metal enhanced fluorescence, polarization microscopy. All these applications require highly accurate models which are currently unavailable for large fields of view.

Acknowledgments
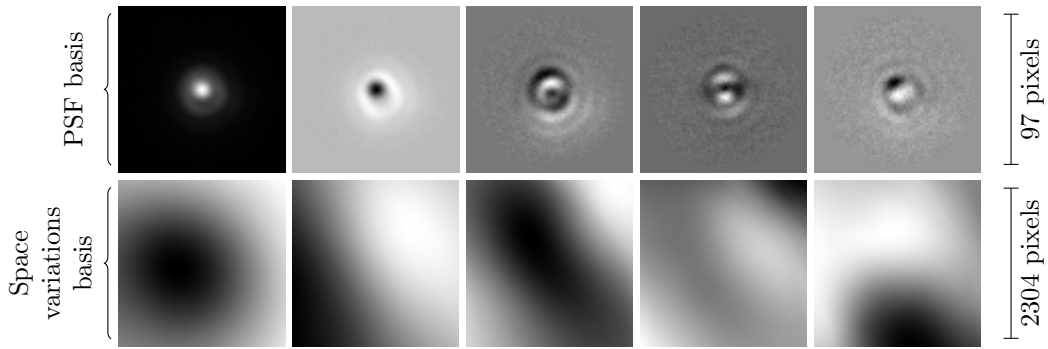
(a) Operator estimation with a single image containing 94 micro-beads, see Fig. 3a. Top: the PSF basis $(h_i)_{1 \leq i \leq 5}$. Bottom: the space variations basis $(a_j)_{1 \leq j \leq 5}$.
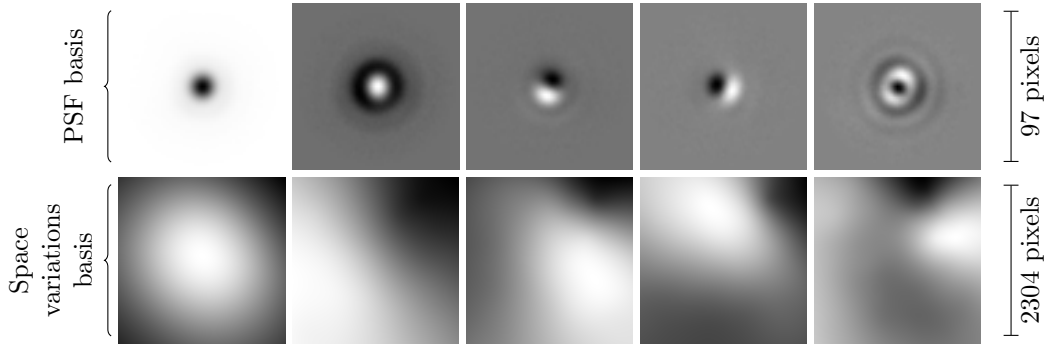

(b) Subspace estimation with 90 images and 9700 PSFs.

Figure 5: Learning the PSF and space variations bases for a standard wide-field microscope.

References

[1] C. Preza and J.-A. Conchello, "Depth-variant maximum-likelihood restoration for three-dimensional fluorescence microscopy," *JOSA A*, vol. 21, no. 9, pp. 1593–1601, 2004.

[2] J. W. Shaevitz and D. A. Fletcher, "Enhanced three-dimensional deconvolution microscopy using a measured depth-varying point-spread function," *JOSA A*, vol. 24, no. 9, pp. 2622–2627, 2007.

[3] G. Zheng, X. Ou, R. Horstmeyer, and C. Yang, "Characterization of spatially varying aberrations for wide field-of-view microscopy," *Optics express*, vol. 21, no. 13, pp. 15 131–15 143, 2013.

[4] A. von Diezmann, M. Y. Lee, M. D. Lew, and W. Moerner, "Correcting field-dependent aberrations with nanoscale accuracy in three-dimensional single-molecule localization microscopy," *Optica*, vol. 2, no. 11, pp. 985–993, 2015.

[5] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[6] A. Aristov, B. Lelandais, E. Rensen, and C. Zimmer, "Zola-3d allows flexible 3d localization microscopy over an adjustable axial range," *Nature communications*, vol. 9, no. 1, p. 2409, 2018.

[7] T. Yan, C. J. Richardson, M. Zhang, and A. Gahlmann, "Computational correction of spatially variant optical aberrations in 3d single-molecule localization microscopy," *Opt. Express*, vol. 27, no. 9, pp. 12 582–12 599, Apr 2019.

[8] Y. Li, M. Mund, P. Hoess, J. Deschamps, U. Matti, B. Nijmeijer, V. J. Sabinina, J. Ellenberg, I. Schoen, and J. Ries, "Real-time 3d single-molecule localization using experimental point spread functions," *Nature methods*, vol. 15, no. 5, p. 367, 2018.

[9] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1187–1193, 2001.

[10] F. N. Mboula, J.-L. Starck, S. Ronayette, K. Okumura, and J. Amiaux, "Super-resolution method using sparse regularization for point-spread function recovery," *Astronomy & Astrophysics*, vol. 575, p. A86, 2015.

[11] M. Gentile, F. Courbin, and G. Meylan, "Interpolating point spread function anisotropy," *Astronomy & Astrophysics*, vol. 549, p. A1, 2013.
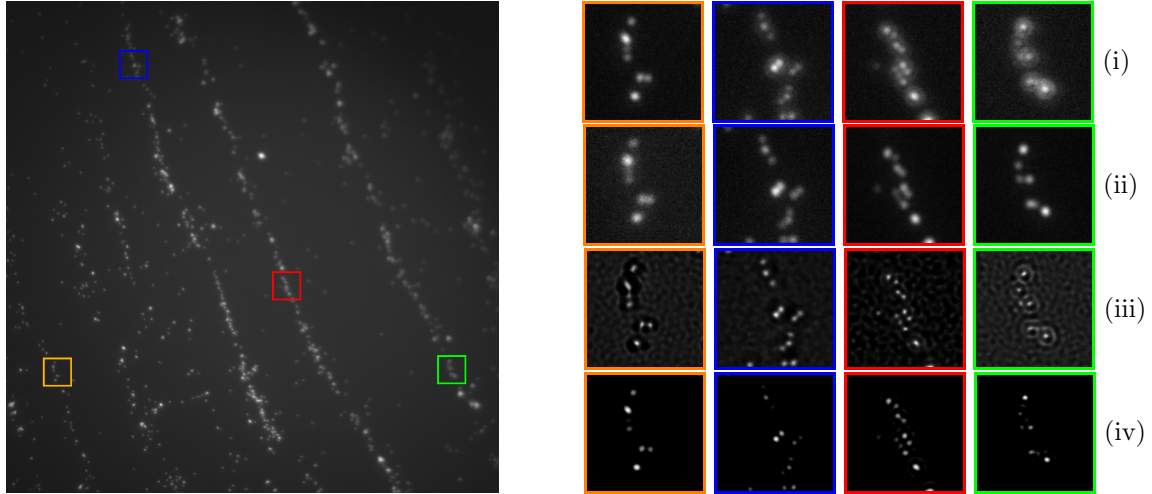
(a) Large crop of the original image (2048 × 2048 pixels) taken at 300nm from the focal plane. The contrasts have been stretched for a better visualization.

(b) (i) original out of focus. (ii) original at focal plane. (iii) deblurred with Huygens software. (iv) BSS deblurring.

Figure 6: Dense flow of micro-beads taken with a wide-field microscope. The Huygens deblurring software based on a theoretical PSF model is compared to the proposed BSS deblurring algorithm based on the subspace estimation.

[12] C. Chang, P. Marshall, J. Jernigan, J. Peterson, S. Kahn, S. Gull, Y. AlSayyad, Z. Ahmad, J. Bankert, D. Bard *et al.*, "Atmospheric point spread function interpolation for weak lensing in short exposure imaging data," *Monthly Notices of the Royal Astronomical Society*, vol. 427, no. 3, pp. 2572–2587, 2012.

[13] J. Bigot, P. Escande, and P. Weiss, "Estimation of linear operators from scattered impulse responses," *Applied and Computational Harmonic Analysis*, vol. 47, no. 3, pp. 730 – 758, 2019.

[14] V. Debarnot, P. Escande, and P. Weiss, "A scalable estimator of sets of integral operators," *Inverse Problems*, 2019.

[15] A. Pinkus, *N-widths in Approximation Theory*. Springer Science & Business Media, 1985, vol. 7.

[16] N. Patwary and C. Preza, "Image restoration for three-dimensional fluorescence microscopy using an orthonormal basis for efficient representation of depth-variant point-spread functions," *Biomedical optics express*, vol. 6, no. 10, pp. 3826–3841, 2015.

[17] M. Arigovindan, J. Shaevitz, J. McGowan, J. W. Sedat, and D. A. Agard, "A parallel product-convolution approach for representing depth varying point spread functions in 3d widefield microscopy based on principal component analysis," *Opt. Express*, vol. 18, no. 7, pp. 6461–6476, Mar 2010.

[18] L. Denis, E. Thiébaut, F. Soulez, J.-M. Becker, and R. Mourya, "Fast approximations of shift-variant blur," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 253–278, 2015.

[19] R. C. Flicker and F. J. Rigaut, "Anisoplanatic deconvolution of adaptive optics images," *J. Opt. Soc. Am. A*, vol. 22, no. 3, pp. 504–513, Mar 2005.

[20] J. G. Nagy and D. P. O'Leary, "Restoring images degraded by spatially variant blur," *SIAM Journal on Scientific Computing*, vol. 19, no. 4, pp. 1063–1082, 1998.

[21] P. Escande and P. Weiss, "Approximation of integral operators using product-convolution expansions," *Journal of Mathematical Imaging and Vision*, vol. 58, no. 3, pp. 333–348, 2017.

[22] N. Halko, P.-G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM review*, vol. 53, no. 2, pp. 217–288, 2011.

[23] M. Gavish and D. L. Donoho, "The optimal hard threshold for singular values is $4/\sqrt{3}$," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 5040–5053, 2014.

[24] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1711–1732, 2013.

[25] M. D. Buhmann, *Radial basis functions: theory and implementations*. Cambridge university press, 2003, vol. 12.

[26] Y. Nesterov, *Lectures on convex optimization*. Springer, 2018, vol. 137.

[27] A. Kumar, V. Sindhwani, and P. Kambadur, "Fast conical hull algorithms for near-separable non-negative matrix factorization," in *International*

Conference on Machine Learning, 2013, pp. 231–239.

A   PSFs selection and processing

Rough PSF detection   The rough detection of PSFs on an image $\boldsymbol{u}$ is quite easy: we simply identify the maxima of the Laplacian of a blurred version of $\boldsymbol{u}$ and keep those above a user-provided threshold.

Rejection of bad patches   The analytical expression of a PSF yields functions that are not compactly supported. However, they decay quite fast at infinity (an Airy pattern decays as $1/|\boldsymbol{x}|^3$) and the information brought by the image is dominated by noise far away from the PSF center. In what follows, we let r denote an upper-bound on the maximal radius of a PSF. To avoid patches containing more than one PSF or PSFs too close to the image boundaries, we select the subset of PSFs with a center being at least r pixels from the image boundary and $2r$ pixels from the other centers. At the end of this procedure, we obtain a set of patches containing isolated PSFs.

Shift and re-sampling   The maxima of the discrete Laplacian are localized on the pixel grid and might not correspond to the PSF center. In addition, there is no reason for the micro-beads to be perfectly centered on a pixel. To obtain a better PSF centering, we find the maxima of correlation with a continuous Gaussian function. This allows us to re-interpolate the patches around this center using a bi-cubic interpolation. The output is a sequence of patches $(\boldsymbol{p}_m)_{1\leq j\leq M}$.

Compact supports   To avoid learning noise, we select two radii $r_{ext} \geq r$ beyond which the PSF image is dominated by noise. In practice, we set $r_{ext} = 1.2r$. We then apply a multiplier to each patch that smoothly decays to 0 from r to r_{ext} (a bump function). The analytical expression of the radial multiplier function $m(t)$ is given by

$$
m(t) = \begin{cases} 1 & \text{if } t < r, \\ 0 & \text{if } t > r_{ext}, \\ \exp\left(\frac{-1}{1-\left(\frac{t-r}{r_{ext}-r}\right)^2}+1\right) & \text{otherwise.} \end{cases}
$$

Principal component analysis   Once all selected PSFs have been treated, we perform a principal component analysis to obtain an optimal representation basis. Depending on the number of sampled PSFs, we can use a standard singular value decomposition or a randomized one [22].

Finer rejection of false detections   Finally, it may happen that the initial selection of patches did not screen out all unwanted PSFs (e.g. too degraded or multiple ones on a patch). To get a finer selection and avoid biases, we project all the PSFs on the learned basis $(h_i)_{1\leq i\leq I}$. We then remove the coefficients in $\mathbb{R}^I$ obtained by the projection which are likely outliers. We simply use a z-score test: we evaluate the empirical mean μ_i and variance σ_i of each coefficient and discard the coefficients that deviate by more than $3\sigma_i$ from their mean. Once the unlikely patches have been removed we recompute a PCA on the remaining patches to get a better estimate of $(h_i)_{1\leq i\leq I}$.

B   Background removal

Getting a correct estimate of the background is a critical step since not accounting for it would strongly bias the PSFs estimates. The background is due to the auto-fluorescence of the sample and may vary spatially due to a non homogeneous illumination. In general this variation is smooth, allowing for an efficient estimation described hereafter.

We assume that the PSF is dominated by noise in a domain ω made of pixels outside a disk of radius r from the patch center. In this region, the image signal is therefore constituted of background and noise only. Assuming that the background is a smooth component, we can try to fit a low degree polynomial to these pixels. Letting $\boldsymbol{p}_j$ denote the j-th patch, this amounts to solving the following quadratic problem:

$$
\inf_{\boldsymbol{\lambda}\in\mathbb{R}^P} \frac{1}{2}\|\boldsymbol{p}_j - \boldsymbol{M}\boldsymbol{\lambda}\|_{l^2(\omega)}^2, \tag{7}
$$

where $\boldsymbol{M} = [\boldsymbol{m}_1,\ldots,\boldsymbol{m}_p] \in \mathbb{R}^{|\omega|\times P}$ is a matrix containing the sampled low degree monomials $\boldsymbol{m}_p$ and $\boldsymbol{\lambda}$ represents the coefficients of the polynomial. This problem boils down to a low dimensional linear system which can be solved with a linear conjugate algorithm. Letting $\boldsymbol{\lambda}^\star$ denote the solution of this problem, the estimated background is simply $\boldsymbol{M}\boldsymbol{\lambda}^\star$. In practice, we simply polynomials of order 2. In 2D, this yields the value $P = 6$ for the monomials 1, x, y, xy, x^2, y^2.

C   Estimating space variations

Thin-plate approximation   Once a basis $(\boldsymbol{h}_i)_{1\leq i\leq I}$ is computed (see section A), it is possible to project each noisy patch on this basis to get a low dimensional representation of the selected PSFs. This provides an estimate $\beta_{i,m} = \langle \boldsymbol{p}_m, \boldsymbol{h}_i\rangle$ of the values $\alpha_i(\boldsymbol{z}_m)$. In order to estimate the space variations, we can use surface fitting techniques on the set $(\boldsymbol{z}_m, \beta_{i,m})_{1\leq m\leq M}$ to get an approximation of the functions α_i.

There exists numerous surface fitting techniques. Following the numerical experiments led in [11], it seems that the use of radial basis function [25] is significantly more efficient than other approaches in the context of astronomy. We therefore resort to this technique.

Radial basis functions approximation can be interpreted as a variational problem in the framework of Reproducible Kernel Hilbert Spaces. In this context,

the estimators $\hat{\alpha}_i$ of α_i can be expressed as

$$\hat{\alpha}_i = \operatorname*{argmin}_{\alpha \in H^2(\mathbb{R}^2)} \frac{1}{2} \sum_{m=1}^{M} w_m |\alpha(\boldsymbol{z}_m) - \beta_{i,m}|^2 + \frac{\eta}{2} |\alpha|_{H^2}^2, \quad (8)$$

where $|\alpha|_{H^2} \overset{\text{def.}}{=} \langle \Delta u, \Delta u \rangle_{L^2(\mathbb{R}^d)}$ and where $\eta > 0$ is a parameter that allows to trade off the proximity to the samples $\beta_{i,m}$ for the smoothness of the surface. In order to balance the importance of each PSFs in the approximation, the weights w_m are chosen equal to the area of the Voronoï cell associated to each location $\boldsymbol{z}_m$.

The solution of (8) is known to be a thin-plate spline [15] and can be computed by solving a $(M+3) \times (M+3)$ linear system.

In what follows, we will let $\boldsymbol{\alpha}_i$ or $\hat{\boldsymbol{\alpha}}_i$ denote a version of α_i sampled on a Euclidean grid.

Enforcing realistic PSFs There is no reason for the thin-plate approximation method to generate realistic PSFs everywhere in the field of view. Indeed, the coefficients are interpolated *independently* of each other while there exists strong dependencies between them. In practice we observed that the previous method was not good at extrapolating the PSFs outside of the convex hull of the sampled PSFs. Important features like positivity for instance might be lost far away from the sampled PSFs. To avoid this effect, we propose an original framework below.

The procedure for estimating the PSF basis $(\boldsymbol{h}_i)$ also yields the projected coefficients $(\beta_{i,m})_{1 \le i \le I, 1 \le m \le M}$. We propose to define the set of *admissible* PSFs coefficients as

$$\mathcal{B} \overset{\text{def.}}{=} \operatorname{cone}(\boldsymbol{\beta}_{.,m}, 1 \le m \le M) \subset \mathbb{R}^I.$$

This roughly amounts to say that admissible PSFs correspond to the conic hull of the already observed and denoised PSFs. Taking the conic hull seems natural: if a PSF is in the set, all its scaled versions by a non-negative factor also belong to the set. Let $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_I) \in \mathbb{R}^{I \times N}$ denote a - possibly unfeasible - estimate of interpolation map. We can generate a feasible one $\hat{\boldsymbol{\alpha}}$ by projection: $\hat{\boldsymbol{\alpha}}[n] \overset{\text{def.}}{=} \Pi_{\mathcal{B}}(\boldsymbol{\alpha}[n])$ for all $1 \le n \le N$.

The projection algorithm Let $\boldsymbol{B} = [\boldsymbol{\beta}_{.,1}, \dots, \boldsymbol{\beta}_{.,M}] \in \mathbb{R}^{I \times M}$ denote the matrix of observed coefficients. Projecting the coefficients $\boldsymbol{\alpha} \in \mathbb{R}^I$ of a PSF onto $\mathcal{B}$ amounts to solving the following convex variational problem:

$$\inf_{\boldsymbol{\lambda} \in \mathbb{R}^M, \boldsymbol{\lambda} \ge 0} \frac{1}{2} \| \boldsymbol{B}\boldsymbol{\lambda} - \boldsymbol{\alpha} \|_2^2.$$

It can be solved using an accelerated projected gradient descent [26].

Unfortunately, if M is very large, applying matrix-vector products with B for every pixel $n \le N$ becomes untractable and we need to simplify the cone $\mathcal{C}$. Following [27], we propose to select a small subset of the columns of $\boldsymbol{B}$ using a simple greedy algorithm. We

start with a matrix $\hat{\boldsymbol{B}}$ containing a single vector equal to the average of the columns of $\boldsymbol{B}$. We then update it by iteratively adding the column in $\boldsymbol{B}$ which maximizes the angle with the current conic hull of the columns in $\hat{\boldsymbol{B}}$. We stop when the angle is below a given threshold. In our experiments with $M = 14000$ PSFs and $I = 5$, we could obtain a very good approximation of the hull with only 20 components instead of 14000, making the projection algorithm relevant even for very large images.

D BSS-deblurring

Here, we propose a method called BSS-deblurring, where BSS stands for Blind Sparse+Smooth. Given a blurred image $\boldsymbol{u}_0$, the method provides an estimate of the associated operator and a deblurred image. It consists of two separate steps: first the operator is estimated using isolated micro-beads in the image. This estimate is then used inside an original variational problem.

Identifying the operator We assume that the user is able to select P patches out of $\boldsymbol{u}_0$ supported on $(\omega_p)_{1 \le p \le P}$ that contain isolated micro-beads. The patches $\boldsymbol{p}_p$ are assumed to be well centered and without background, which can be achieved using the previously proposed methods. We also assume that each micro-bead can have a different amplitude lying in the interval $[a, b]$ with $0 < a < b$.

Our aim is to estimate a discrete operator $\boldsymbol{H} : \mathbb{R}^N \to \mathbb{R}^N$ of the form:

$$\boldsymbol{H}\boldsymbol{u} = \sum_{1 \le i \le I} \sum_{1 \le j \le J} \boldsymbol{\Gamma}[i,j] \boldsymbol{h}_i \star (\boldsymbol{a}_j \odot \boldsymbol{u})$$

where the pairs of orthogonal bases $((\boldsymbol{h}_i)_i, (\boldsymbol{a}_j)_j)$ are known and the coefficients $(\boldsymbol{\Gamma}[i,j])_{i,j}$ are unknown.

By projecting the patch $\boldsymbol{p}_p$ onto the PSFs basis $(\boldsymbol{h}_i)_i$, we obtain the coefficient $c_{p,i} = \langle \boldsymbol{p}_p, \boldsymbol{h}_i \rangle$ which is a noisy estimate of the interpolation map $\boldsymbol{\alpha}_i$ at the position $\boldsymbol{z}_p$:

$$\boldsymbol{H}\delta_{\boldsymbol{z}_p} = \sum_{1 \le i \le I} \sum_{1 \le j \le J} \boldsymbol{\Gamma}[i,j] \boldsymbol{\alpha}_i(\boldsymbol{z}_p) \boldsymbol{h}_i \approx \sum_{1 \le i \le I} c_{p,i} \boldsymbol{h}_i.$$

We propose to identify $\boldsymbol{\Gamma}$ by solving the following bilinear inverse problem:

$$\operatorname*{argmin}_{\boldsymbol{\Gamma} \in \mathcal{C}, \boldsymbol{g} \in [a,b]^P} \frac{1}{2} \sum_{1 \le p \le P} \left\| \sum_{1 \le i \le I} \sum_{1 \le j \le J} g_p \boldsymbol{\Gamma}_{i,j} \boldsymbol{\alpha}_j - c_{p,i} \right\|_2^2, \tag{9}$$

where g_p is the unknown amplitude of the bead at position z_p and $\mathcal{C}$ is the conic hull of the sampled operators.

We solve this problem using an alternating minimization algorithm: we first solve the problem w.r.t. $\boldsymbol{g}$ with fixed $\boldsymbol{\Gamma}$ and then solve the problem w.r.t. $\boldsymbol{\Gamma}$ with fixed $\boldsymbol{g}$. The individual minimization problems are convex and can be solved with accelerated projected gradient descents.

If the amplitudes $\boldsymbol{g}$ were known, the problem (9) would boil down to a constrained least square problem

of size $I \times J$. Since each patch $\boldsymbol{p}_p$ yields I coefficients, the condition $P \geq J$ should be enough to ensure the identification. It is remarkable that such a low value (typically 5) is enough to identify the operator! Higher values of P would however make the method more robust to noise.

To validate the proposed approach, we randomly select an operator in the conic hull $\mathcal{C}$, and apply it to a grid of 25 Dirac masses (for a field of view of 2304×2304 pixels). A significant amount of noise is added to the image and the true locations of the beads are randomly perturbed (with Gaussian random variable of variance 0.5). We then estimate the operator and display the result in Figure 7.

Figure 8: Identification of a real operator with 38 micro-beads. Top: 3 patches used for the estimation. Bottom: estimated PSFs.

manually so as to obtain a visually pleasant result. This problem can be solved efficiently using accelerated proximal gradient descent algorithm [26].
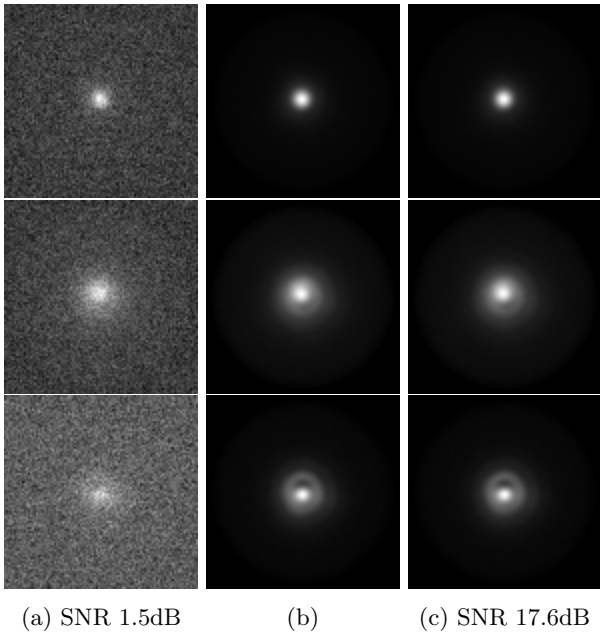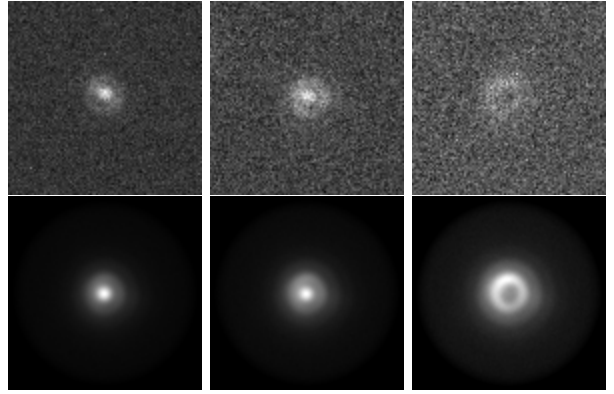

(a) SNR 1.5dB   (b)   (c) SNR 17.6dB

Figure 7: Identification of an operator in simulated images. a) Noisy crops used for identification. b) Impulse responses of the true operator at some locations. c) Estimated operator. Notice that the method is able to denoise the PSFs very efficiently.

We finally apply the method to the image used in the blind-deblurring example in Figure 6a. We display the reconstructed operator in Figure 8.

Sparse+Smooth-deblurring Let $\boldsymbol{u}_0 \in \mathbb{R}^N$ denote a blurry image and $\boldsymbol{H} : \mathbb{R}^N \to \mathbb{R}^N$ a discrete integral operator. We aim at deblurring an image composed of a sparse component $\boldsymbol{u}_1$ (e.g. scattered micro-beads), and a smooth component $\boldsymbol{u}_2$ (e.g. auto-fluorescent background). In order to recover $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, we propose to solve the following original variational problem:

$$\inf_{\boldsymbol{u}_1 \in \mathbb{R}_+^N, \boldsymbol{u}_2 \in \mathbb{R}_+^N} \frac{1}{2}\|H(\boldsymbol{u}_1+\boldsymbol{u}_2)-y\|_2^2+\gamma_1\|\boldsymbol{u}_1\|_1+\gamma_2\|\Delta\boldsymbol{u}_2\|_2^2.$$

The term $\frac{1}{2}\|H(\boldsymbol{u}_1 + \boldsymbol{u}_2) - y\|_2^2$ is the data fitting term, the term $\gamma_1\|\boldsymbol{u}_1\|_1$ promotes the sparsity of the $\boldsymbol{u}_1$ component and the term $\gamma_2\|\Delta\boldsymbol{u}_2\|_2^2$ promotes smoothness of $\boldsymbol{u}_2$. The non-negative parameters γ_1 and γ_2 allow to balance each term and have been tuned