Multispectral image reconstruction of faint circumstellar environments from high contrast angular spectral differential Imaging (ASDI) data

Olivier Flasseur, Samuel Thé, Loïc Denis, Éric Thiébaut, Maud Langlois

Multispectral Image Reconstruction of Faint Circumstellar Environments from High Contrast Angular Spectral Differential Imaging (ASDI) Data

Olivier Flasseur[a], Samuel Thé[b], Loïc Denis[c], Éric Thiébaut[b], and Maud Langlois[b]

[a]Laboratoire d'Études Spatiales et d'Instrumentation en Astrophysique, Observatoire de Paris, Univ. PSL, Sorbonne Univ., Univ. Paris Diderot, France
[b]Centre de Recherche Astrophysique de Lyon, CNRS, Univ. de Lyon, Univ. Claude Bernard Lyon 1, ENS de Lyon, France
[c]Univ. Lyon, UJM-Saint-Etienne, CNRS, Institut d Optique Graduate School, Laboratoire Hubert Curien, Saint-Étienne, France

ABSTRACT

We recently proposed `REXPACO`,[1] an algorithm for imaging circumstellar environments from high-contrast angular differential imaging (ADI) data. In the context of high-contrast imaging where the signal of interest is largely dominated by a nuisance term due to the stellar light leakages and the noise, our algorithm amounts to jointly estimating the object of interest and the statistics (mean and covariance matrix) of the nuisance component. In this contribution, we first extend the `REXPACO` algorithm by refining the statistical model of the nuisance component it embeds. Capitalizing on the improved robustness of this new method named `robust REXPACO`, we then show how it can be modified to deal with angular plus spectral differential imaging (ASDI) datasets. We apply our methods on several ADI and ASDI datasets from the IRDIS and IFS imagers of the VLT/SPHERE instrument and we show that the proposed algorithms significantly reduce the typical artifacts produced by state-of-the-art algorithms. By also taking into account the instrumental point spread function (PSF), our algorithms yield a deblurred estimate of the object of interest without the artifacts observed with other methods.

Keywords: High-contrast imaging, angular and spectral differential imaging, circumstellar environment, reconstruction algorithm, regularized inverse problems, statistical method

1. INTRODUCTION

High-contrast imaging is an observational method used to study the close-environment of stars.[2–4] The very high-contrast achieved in direct imaging (typically $\geq 10^5$ in infrared) allows for major discoveries in the studies of circumstellar environment made of gas and/or dust. In particular, *protoplanetary disks* are widely studied since exoplanets can form inside the disks by accretion. In that context, cutting-edge ground-based facilities like VLT/SPHERE[5] Gemini/GPI,[6] Magellan/MagAO[7] or SUBARU/SCExAO[8] have revealed an impressive variety of protoplanetary or transition disks, see[9–13] for case-study examples. The morphological structure of the disks is of primary interest since the process of exoplanet formation and disk evolution leaves an imprint on the distribution of material at different locations in a protoplanetary disk,[14–16] resulting in a variety of substructures like rings, gaps, spirals, vortices, clumps or arms. The PDS 70 system is one of the flagship examples illustrating the power of direct imaging to analyze interactions between exoplanets and protoplanetary disks. PDS 70 is to date the only system hosting two exoplanets in formation inside a protoplanetary disk that was detected by direct imaging.[17–20] Whatever the targeted star, the observation strategy and the design of the data processing method are two keystones of direct imaging to reach a high-contrast and to achieve the best extraction of relevant astrophysical information.

Further author information: olivier.flasseur@obspm.fr
O. Flasseur was previously with *Centre de Recherche Astrophysique de Lyon, CNRS, Univ. de Lyon, Univ. Claude Bernard Lyon 1, ENS de Lyon, France*, where most of this work was performed.

In this paper, we focus on observations acquired with angular (plus spectral) differential imaging (A(S)DI[21,22]). ADI consists in tracking the observed target over time following a particular strategy: the telescope derotator is tuned to maintain the telescope pupil stable while the field-of-view rotates. Consequently, in the resulting 3-D datasets (2-D + time), the off-axis objects of interest (i.e., point-like sources plus circumstellar disk surrounding the star) follow an apparent motion along a predictable circular trajectory while the telescope pupil remains static. ADI can be combined with spectral differential imaging (SDI[22]) that consists in recording simultaneously multiple images of the same astrophysical scene in several spectral channels with an integral field spectrograph (IFS). Within A(S)DI observations, speckles resulting from residual aberrations are thus strongly correlated from one exposure to the other. Images can be combined in a post-processing step to cancel out most of the speckles while part of the signal from the off-axis objects is preserved.

There are very few post-processing methods[1,23] specifically designed for the reconstruction of extended features in the circumstellar environment from ADI observations. In addition, to the best of our knowledge, none of them are able to perform a joint multi-spectral processing of ASDI observations (i.e., each spectral channel is processed independently in a sub-optimal fashion). Among the existing reconstruction methods from ADI observations, we recently proposed `REXPACO`[1] that builds a statistical model, self-calibrated on the data, that accounts for the spatial covariances of the nuisance component (i.e., speckles and noise) at the scale of small patches of a few tens of pixels. This statistical modeling is combined with a modeling of the instrumental effects. The intensity distribution of the sought object (i.e., disk and exoplanets) is estimated by formalizing the reconstruction task as a regularized inverse-problem. The method is fully unsupervised, the unknown (hyper)-parameters are estimated in a data-driven fashion, and allows to restore images simultaneously (i) deconvolved from the instrumental blur, (ii) significantly less impacted by artifacts than state-of-the-arts methods of the fields, and (iii) that allow an unmixing between point-like sources and spatially extended features.

While `REXPACO` demonstrates state-of-the-art performance,[1] there is still room for improvements. In this contribution, we follow two complementary paths in this direction. First, in Sect. 2 we refine the statistical model of `REXPACO` by extending its multi-variate Gaussian mixture model by a *scaled* mixture of Gaussian. We name this new algorithm "`robust REXPACO`" since it is able to (partly) capture the temporal variability of the observations induced by the evolution of the observing conditions during the acquisitions. Second, in Sect. 3 we develop the cornerstones of an extension* of `robust REXPACO` specifically designed for disk reconstruction from ASDI datasets. Finally, Sect. 4 gives comparative reconstruction results obtained on several datasets from the IRDIS (ADI datasets) and IFS (ASDI datasets) imagers of the VLT/SPHERE instrument. The two approaches significantly reduce the typical artifacts produced by state-of-the-art algorithms and yield a deblurred estimate of the object of interest. In addition, `robust REXPACO` is able to retrieve finer structures than `REXPACO`, especially at short angular separations.

2. RECONSTRUCTION ALGORITHM FROM ADI DATASETS

2.1 Modeling of the measured intensity

An ADI dataset is formed by N-pixels images recorded at different times $t \in [\![1;T]\!]$. For convenience, we represent an ADI dataset by a vector $\boldsymbol{r} \in \mathbb{R}^{NT}$ formed by the concatenation of its pixel values. The resulting vector $\boldsymbol{r}$ of measured intensities is modeled by the additive contribution of two components:

$$\boldsymbol{r} = \mathbf{M}\,\boldsymbol{x} + \boldsymbol{f}\,, \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}_+^M$ is the intensity distribution of the light coming from the sought objects (i.e., disk and point-like sources) that is transformed by a linear operator $\mathbf{M} : \mathbb{R}^M \to \mathbb{R}^{NT}$ modeling the instrumental effects that govern the sequence of acquisitions, and $\boldsymbol{f} \in \mathbb{R}^{NT}$ is the nuisance component (i.e., speckles plus noise) corrupting the object component. The spatial extent of $\boldsymbol{x}$ can be larger than the spatial extent of $\boldsymbol{r}$ (i.e., $M \geq N$) since the apparent rotation induced by differential imaging (ADI or ASDI) allows to reconstruct part of the object component lying within the sensor field-of-view at less than T exposures. Besides, the intensity distribution $\boldsymbol{x}$ is assumed to be time-invariant, i.e. (i) we neglect the proper motion of the off-axis objects around their host

*This algorithm performing a joint estimation of the nuisance component and of the sought objects from ASDI datasets will be described in more details in a paper (Flasseur *et al.* in prep.) currently in preparation.

star, and (ii) we neglect the temporal variability of the intensity of the objects of interest. Assumptions (i) and (ii) are fulfilled since such effects are negligible at the time scale of the few hours of observations.

2.2 Modeling of the instrumental effects

2.2.1 Direct model

The contribution $\mathbf{M}_t\,\boldsymbol{x}$ of the off-axis objects $\boldsymbol{x}$ in the frame $\boldsymbol{r}_t \in \mathbb{R}^N$ at time t can be decomposed by a cascade of linear operations:

$$\mathbf{M}_t = \mathbf{TBAR}_t \quad \text{such that} \quad \mathbf{M} = \begin{pmatrix} \mathbf{M}_1 : \mathbb{R}^M \to \mathbb{R}^N \\ \vdots \\ \mathbf{M}_t : \mathbb{R}^M \to \mathbb{R}^N \\ \vdots \\ \mathbf{M}_T : \mathbb{R}^M \to \mathbb{R}^N \end{pmatrix}, \tag{2}$$

where $\mathbf{M}_t$ accounts for the following instrumental effects:

- a *rotation* $\mathbf{R}_t$ describing the apparent motion of the off-axis objects induced by the rotation of the field-of-view between time t and a reference time t_{ref} (e.g., $t_{\text{ref}} = t_1$). This operator is implemented as a (sparse) interpolation matrix.
- an *attenuation* $\mathbf{A}$ describing the decrease in transmitted intensity induced by the coronagraph in the vicinity of the host star. This operator is implemented as a diagonal matrix: $\mathbf{A} = \text{diag}(\boldsymbol{a})$ with $\boldsymbol{a} \in [0;1]^N$ is the vectorization of a 2-D coronagraphic transmission map. In practice, we assume that the transmission map has a radial profile (i.e., all pixels belonging at a given separation are supposed to be impacted by the same transmission factor), we neglect deformations (beyond an overall attenuation) of the off-axis PSF induced by the coronagraph near the star, and the attenuation is supposed to be time-invariant. Assumption (i) is in agreement with measurements and simulations performed for the VLT/SPHERE instrument,[5] assumption (ii) is reasonable given that signals from the off-axis objects are largely dominated by the starlight near the star, and assumption (iii) is reasonable given that the observations by cutting-edge instruments are quite stable. Besides, the operator $\mathbf{A}$ can be generalized to $\mathbf{A}_t$ to account for time-dependent effects such as a decentering of the coronagraph during the observations.
- a *blur* $\mathbf{B}$ describing the instrumental blurring. This operator is implemented as a bi-dimensional discrete convolution by the off-axis PSF. The blurring effect is supposed to be isotropic as well as spatially and temporally independent. We checked with numerical simulations that these assumptions are reasonable. The blur operator $\mathbf{B}$ can be generalized easily to a time-dependent operator $\mathbf{B}_t$ (if such a measurement or estimate is available) to account for the evolution of the observing conditions during the sequence of acquisition.
- a *truncation* $\mathbf{T}$ performing a selection of pixels (corresponding to the actual data within the field-of-view) from the larger M-pixels image received as output from the blurring operator $\mathbf{H}$.

2.2.2 Implementation details

Considering implementation constraints, we aim to obtain a computationally efficient formulation for the global operator $\mathbf{M}$ and its adjoint $\mathbf{M}^\top$ since these are evaluated repeatedly in the iterative reconstruction framework we describe in Sect. 2.4. In particular, Eq. (2) involves T convolutions by $\mathbf{B}$ of the *rotated plus attenuated* contributions $\{\mathbf{AR}_t\,\boldsymbol{x}\}_{t=1:T}$ of the off-axis objects $\boldsymbol{x}$. To cope with this issue, we reformulate the direct model by performing a re-ordering of the operators $\mathbf{T}$, $\mathbf{B}$, $\mathbf{A}$, and $\mathbf{R}_t$ involved in Eq. (2) and we define the direct model

$\mathbf{M}_t$ at time t by:

$$\mathbf{M}_t = \underbrace{\mathbf{T}\mathbf{A}\mathbf{R}_t}_{\mathbf{Q}_t}\mathbf{B} \quad \text{such that} \quad \mathbf{M} = \underbrace{\begin{pmatrix} \mathbf{Q}_1 : \mathbb{R}^M \to \mathbb{R}^N \\ \vdots \\ \mathbf{Q}_t : \mathbb{R}^M \to \mathbb{R}^N \\ \vdots \\ \mathbf{Q}_T : \mathbb{R}^M \to \mathbb{R}^N \end{pmatrix}}_{\mathbf{Q}} \mathbf{B}\,, \tag{3}$$

where the set of T operators $\{\mathbf{Q}_t\}_{t=1:T}$ is implemented by sparse matrices and performs sequentially time-dependent rotations, and time-independent attenuations and truncations. Permutations involved between Eqs. (2) and (3) are valid under the anisotropic and time-invariant assumptions made for the off-axis PSF and assuming a radial profile of the transmission map. Such a reformulation of the direct model $\mathbf{M}$ allows to perform a single convolution of the off-axis objects $\boldsymbol{x}$ with the operator $\mathbf{B}$ that leads to an improvement by about one to two orders of magnitude in terms of algorithmic complexity. Such a gain is especially critical to reconstruct the spectral flux distribution of off-axis objects from ASDI datasets, see Sect. 3.

2.3 Modeling of the nuisance component

2.3.1 Statistical model

In ground-based direct imaging at high-contrast, difficulties in modeling the nuisance component are triple:

- First, the nuisance component varies drastically across sequences of observations (e.g., performed during different nights), as it is highly dependent on the observing conditions, on the physical properties of the star, and on the instrument settings. For those reasons, we build a *dataset-dependent model*, whose parameters are estimated solely from the dataset of interest $\boldsymbol{r}$.
- Second, even for a given dataset $\boldsymbol{r}$, the nuisance component exhibits a complex evolution: it displays strong fluctuations between temporal frames $\{\boldsymbol{r}_t\}_{t=1:T}$, especially near the star where the stellar leakages dominate. This limits the possibility to describe accurately such an evolution with a end-to-end physics-based model of the residual aberrations uncorrected by the optical system. So, we opt for a *data-driven model*, whose parameters are self-calibrated on the observations. Among the wide literature of the field, machine learning and deep learning approaches showed to be very efficient to build complex models from large training sets. However, specifying these methods to our problem remains a challenge given the lack of training data and the strong non-stationarity of the nuisance component. For that reason, we opt for a *statistical model* that amounts to describe the spatially non-stationary correlations and the temporal fluctuations of the nuisance component.
- Third, the number of measurements available to perform the estimation of the model parameters is very limited, so that a trade-off should be reached between the complexity of the model and its fidelity with respect to the observations. For that reason, we opt for *a patch-based model* that amounts to describe locally (i.e., at the scale of a small window of a few tens of pixels) the non-stationary fluctuations and spatial correlations of the observations.

Given those requirements, we model the distribution of the nuisance component at a pixel location n of the field-of-view by a multi-variate Gaussian whose parameters are estimated locally from the vectorized patch $\mathbf{P}_n\,\boldsymbol{f} = \boldsymbol{f}_n \in \mathbb{R}^{KT}$, where $\mathbf{P}_n$ is a linear operator performing the extraction and the pixel-wise vectorization of the collection $\{\boldsymbol{f}_{n,t}\}_{t=1:T} \in \mathbb{R}^{K\times T}$ of the T K-pixels patches centered on pixel n. In that context, and assuming statistical independence between patches, the distribution of the nuisance component $\boldsymbol{f}$ writes:

$$\mathrm{p}_F(\boldsymbol{f}) = \prod_{n\in\mathbb{P}} \mathrm{p}_{F_n}(\boldsymbol{f}_n) \quad \text{with} \quad \mathrm{p}_{F_n}(\boldsymbol{f}_n) \propto \det{}^{-\frac{1}{2}}(\mathbf{C}_n) \exp\left(-\frac{1}{2}\left(\boldsymbol{f}_n - \boldsymbol{m}_n\right)^\top \mathbf{C}_n^{-1}\left(\boldsymbol{f}_n - \boldsymbol{m}_n\right)\right), \tag{4}$$

with $\mathbb{P}$ the set of locations where the statistics of the nuisance component are evaluated. The cardinal of $\mathbb{P}$ depends solely on the patch shape and on the patch stride selected to pave the whole field-of-view. We consider

non-overlapping square patches (i.e., $\mathrm{card}(\mathbb{P}) = \lfloor N/K^2 \rceil$). Each distribution p_{F_n} is parameterized by a mean $\boldsymbol{m}_n \in \mathbb{R}^{KT}$ and a covariance matrix $\mathbf{C}_n \in \mathbb{R}^{KT \times KT}$ that should be estimated locally, around pixel n, from the observations $\boldsymbol{r}_n$. Given the high number of degrees of freedom involved in such a model (the number of parameters being larger than the number of measurements by several orders of magnitudes), we chose to enforce a specific structure for the mean $\boldsymbol{m}_n$ and for the covariance matrix $\mathbf{C}_n$. Under that constraints, we simplify the spatio-temporal mean by keeping only its spatial variability:

$$\boldsymbol{m}_n = \boldsymbol{m}_n^{\mathrm{spat}} \otimes \mathbf{1}_T \,, \tag{5}$$

which is the result of the Knonecker product (denoted by the operator $\otimes$) between the temporal mean vector $\boldsymbol{m}^{\mathrm{spat}} \in \mathbb{R}^K$ of the collection of patches $\{\boldsymbol{f}_{n,t}\}_{t=1:T}$ and the unit vector $\mathbf{1}_T \in 1^T$. Similarly, we also simplify the spatio-temporal covariance matrix $\mathbf{C}_n$ by neglecting the temporal covariances:

$$\mathbf{C}_n = \mathbf{I}_T \otimes \mathbf{C}_n^{\mathrm{spat}} \,, \tag{6}$$

which is the result of the Kronecker product between the spatial covariance matrix $\mathbf{C}_n^{\mathrm{spat}} \in \mathbb{R}^{K \times K}$ of the collection of patches $\{\boldsymbol{f}_{n,t}\}_{t=1:T}$ and the identity matrix $\mathbf{I}_T$ of size $T \times T$. Equations (5) and (6) amount to model each spatial patch $\boldsymbol{f}_{n,t}$ as a random variable following a multi-variable Gaussian $\mathcal{N}(\boldsymbol{m}_n^{\mathrm{spat}}, \mathbf{C}_n^{\mathrm{spat}})$.

Besides, in our previous work[24] on the `PACO` algorithm for exoplanet detection, we investigate a refinement of that formulation for the statistical model of the nuisance component by replacing the multi-variate Gaussian assumption with a multi-variate Gaussian scale mixture (GSM) model.[25,26] The former includes additional parameters $\boldsymbol{\sigma}_n^2 = \{\sigma_{n,t}^2\}_{t=1:T}$ scaling temporally patches of the collection $\{\boldsymbol{f}_{n,t}\}_{t=1:T}$ according to their relative degree of fluctuations. Such a model showed to be very effective to identify and to neutralize, in a data-driven fashion, patches displaying larger fluctuations than others. We showed empirically that the robustness (e.g., against the presence of bad pixels or strong stellar leakages induced by a sudden degradation of the adaptive-optics correction) of the statistics of the nuisance component is significantly improved. As a consequence, the fidelity of the model with respect to the observations is also improved. Concerning the direct outputs of the algorithm, we showed that the stationarity of the detection map, the control of the false alarm rate, and the detection sensitivity are improved. In this present contribution, we propose to combine the GSM model we investigated in the context of point-source detection[24] with the reconstruction framework we developed for the reconstruction of extended features.[1] This amount to modeling each spatial patch $\boldsymbol{f}_{n,t}$ as a random variable following a scaled multi-variable Gaussian $\mathcal{N}(\boldsymbol{m}_n^{\mathrm{spat}}, \sigma_{n,t}^2\, \mathbf{C}_n^{\mathrm{spat}})$, i.e. Eq. (6) is replaced by:

$$\mathbf{C}_n = \boldsymbol{\Lambda}_n \otimes \mathbf{C}_n^{\mathrm{spat}} \,, \tag{7}$$

where $\boldsymbol{\Lambda}_n \in \mathbb{R}^{T \times T}$ is a diagonal matrix such that $[\boldsymbol{\Lambda}_n]_{tt} = \sigma_{n,t}^2$.

2.3.2 Estimation of the model parameters

Under the statistical model defined in Sect. 2.3.1, the maximum-likelihood estimators $\{\widehat{\boldsymbol{m}}_n^{\mathrm{spat}}, \widehat{\mathbf{S}}_n^{\mathrm{spat}}, \widehat{\boldsymbol{\sigma}}_{n,t}^2\}_{n\in\mathbb{P},t=1:T}$ of the parameters $\{\boldsymbol{m}_n^{\mathrm{spat}}, \mathbf{C}_n^{\mathrm{spat}}, \boldsymbol{\sigma}_{n,t}^2\}_{n\in\mathbb{P},t=1:T}$ describing the nuisance component writes:[24]

$$\begin{cases} \widehat{\boldsymbol{m}}_n^{\mathrm{spat}} = \left(\sum\limits_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2} \right)^{-1} \sum\limits_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2} \boldsymbol{r}_{n,t} \,, \\ \widehat{\mathbf{S}}_n^{\mathrm{spat}} = \sum\limits_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2} \widehat{\boldsymbol{u}}_{n,t}\, \widehat{\boldsymbol{u}}_{n,t}^{\top} \,, \\ \widehat{\sigma}_{n,t}^2 = \frac{1}{K} \widehat{\boldsymbol{u}}_{n,t}^{\top}\, \widehat{\mathbf{S}}_n^{\mathrm{spat}^{-1}} \widehat{\boldsymbol{u}}_{n,t} \,, \end{cases} \tag{8}$$

with the residuals (i.e. after centering):

$$\widehat{\boldsymbol{u}}_{n,t} = \boldsymbol{r}_{n,t} - \widehat{\boldsymbol{m}}_n^{\mathrm{spat}} \,. \tag{9}$$

The estimators $\widehat{\boldsymbol{m}}_n^{\mathrm{spat}}$ (resp., $\widehat{\mathbf{S}}_n^{\mathrm{spat}}$) can be interpreted as the weighted sample mean (resp., as the weighted sample covariance matrix) of the collection of patches $\{\boldsymbol{r}_{n,t}\}_{t=1:T}$, and $\widehat{\sigma}_{n,t}^2$ is the temporal variance of the residuals patches $\{\boldsymbol{u}_{n,t}\}_{t=1:T}$ (i.e., the weights $\widehat{\sigma}_{n,t}^{-2}$ are larger for frames with small fluctuations).

Since the number T of patch samples available to perform the estimation of $\widehat{\mathbf{S}}_n^{\text{spat}}$ is typically lower than the number K of pixels in a patch, the sample covariance $\widehat{\mathbf{S}}_n^{\text{spat}}$ is very noisy and can even be rank deficient. A form of regularization should be enforced to stabilize its estimation and to guarantee its invertibility. As in our previous works,[1,24,27–29] we resort to a *shrinkage* estimator[30,31] formed by a convex combination between the low bias/high variance estimator $\widehat{\mathbf{S}}_n^{\text{spat}}$ and a high bias/low variance estimator $\widehat{\mathbf{F}}_n^{\text{spat}}$:

$$\widehat{\mathbf{C}}_n^{\text{spat}} = (1 - \widehat{\rho}_n^{\text{spat}})\, \widehat{\mathbf{S}}_n^{\text{spat}} + \widehat{\rho}_n^{\text{spat}}\, \widehat{\mathbf{F}}_n^{\text{spat}}\,, \tag{10}$$

where $\widehat{\mathbf{F}}_n^{\text{spat}}$ is a diagonal matrix with only the sample variances:

$$\left[\widehat{\mathbf{F}}_n^{\text{spat}}\right]_{kk'} = \begin{cases} \left[\widehat{\mathbf{S}}_n^{\text{spat}}\right]_{kk'} & \text{if} \quad k = k' \\ 0 & \text{if} \quad k \neq k'\,. \end{cases} \tag{11}$$

By introducing the matrix $\widehat{\mathbf{W}}_n^{\text{spat}} \in \mathbb{R}^{K \times K}$ whose entries are defined as:

$$\left[\widehat{\mathbf{W}}_n^{\text{spat}}\right]_{kk'} = \begin{cases} 1 - \widehat{\rho}_n^{\text{spat}} & \text{if } k \neq k' \\ 1 & \text{if } k = k'\,, \end{cases} \tag{12}$$

Eq. (10) can be rewritten as:

$$\widehat{\mathbf{C}}_n^{\text{spat}} = \widehat{\mathbf{W}}_n^{\text{spat}} \odot \widehat{\mathbf{S}}_n^{\text{spat}}\,, \tag{13}$$

where the operator $\odot$ denotes the Hadamard (i.e., point-wise) product. The hyper-parameter $\widehat{\rho}_n^{\text{spat}}$ plays a key role since it governs a bias-variance trade-off. In our previous work,[24] we derived its closed-form expression, which is an extension of the results of Chen *et al.*[31] in the case of a non-constant valued shrinkage matrix $\widehat{\mathbf{F}}_n^{\text{spat}}$:

$$\widehat{\rho}_n^{\text{spat}} = \frac{\operatorname{tr}\left(\widehat{\mathbf{S}}_n^{\text{spat}^2}\right) + \operatorname{tr}^2\left(\widehat{\mathbf{S}}_n^{\text{spat}}\right) - 2\sum_{k=1}^{K}\left[\widehat{\mathbf{S}}_n^{\text{spat}}\right]_{kk}^2}{(\bar{T}_n^{\text{spat}} + 1)\left(\operatorname{tr}\left(\widehat{\mathbf{S}}_n^{\text{spat}^2}\right) - \sum_{k=1}^{K}\left[\widehat{\mathbf{S}}_n^{\text{spat}}\right]_{kk}^2\right)}\,, \tag{14}$$

where

$$\bar{T}_n^{\text{spat}} = \frac{\left(\sum_{t=1}^{T} 1/\widehat{\sigma}_{n,t}^2\right)^2}{\sum_{t=1}^{T} 1/\widehat{\sigma}_{n,t}^4}\,, \tag{15}$$

is the equivalent average number of patch samples contributing to the estimation of $\widehat{\mathbf{S}}_n^{\text{spat}}$ given the temporal weights $\{\widehat{\sigma}_{n,t}^2\}_{t=1:T}$ estimated at location n. We denote by $\widehat{\boldsymbol{\Omega}} = \{\widehat{\boldsymbol{m}}_n^{\text{spat}}, \widehat{\mathbf{C}}_n^{\text{spat}}, \widehat{\sigma}_{n,t}^2\}_{n\in\mathbb{P},t=1:T}$ the whole set of (shrunk) estimators describing the statistics of the nuisance component.

In addition, the estimators of $\widehat{\boldsymbol{\Omega}}$ are biased by the off-axis objects whose contributions are partly encoded in the mean and in the covariance matrix. As a direct consequence, we have shown[1] that a bias (taking the form of morphological distortions and of intensity under-estimations) can impact the estimation of the intensity distribution $\boldsymbol{x}$. For that reason, performing a joint estimation of $\boldsymbol{x}$ and of $\boldsymbol{\Omega}$ is mandatory (see Sect. 2.4). In that context, we iteratively replace the estimators of $\widehat{\boldsymbol{\Omega}}$ such that (i) they are corrected from the contribution of the off-axis objects $\boldsymbol{x}$ in the observations $\boldsymbol{r}$, (ii) they account for the shrinkage of the covariance matrix:

$$\begin{cases} \widehat{\boldsymbol{m}}_n^{\text{spat}} = \left(\sum_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2}\right)^{-1} \sum_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2} \left(\boldsymbol{r}_{n,t} - [\mathbf{M}_t\, \boldsymbol{x}]_n\right)\,, \\ \widehat{\mathbf{C}}_n^{\text{spat}} = \mathbf{W}_n^{\text{spat}} \odot \sum_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2} \widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})\, \widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})^\top\,, \\ \widehat{\sigma}_{n,t}^2 = \frac{1}{K} \widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})^\top \left(\mathbf{W}_n^{\text{spat}} \odot \widehat{\mathbf{C}}_n^{\text{spat}^{-1}}\right) \widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})\,, \end{cases} \tag{16}$$

with the residuals defined in Eq. (9) updated accordingly:

$$\widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x}) = \boldsymbol{r}_{n,t} - [\mathbf{M}_t\, \boldsymbol{x}]_n - \widehat{\boldsymbol{m}}_n^{\text{spat}}\,. \tag{17}$$

Since estimators defined in Eqs. (16) are inter-dependent, they are estimated iteratively until their convergence.

2.4 Inverse-problem formulation and regularized inversion

We formalize the reconstruction of the intensity distribution $\boldsymbol{x}$ as an inverse-problem in which the minimized objective function $\mathcal{C}$ is the sum of two components; (i) a data-fidelity term $\mathcal{D}$, and (ii) a regularization term $\mathcal{R}$:

$$\widehat{\boldsymbol{x}} \in \underset{\boldsymbol{x}\geq\boldsymbol{0}}{\arg\min} \; \min_{\widehat{\boldsymbol{\Omega}}} \left\{\mathcal{C}(\boldsymbol{r},\mathbf{M},\boldsymbol{x},\boldsymbol{\Omega},\boldsymbol{\mu}) = \mathcal{D}(\boldsymbol{r},\mathbf{M},\boldsymbol{x},\boldsymbol{\Omega}) + \mathcal{R}(\boldsymbol{x},\boldsymbol{\mu})\right\} \tag{18}$$

In that formulation, the statistics $\boldsymbol{\Omega}$ of the nuisance component and the distribution of the sought object $\boldsymbol{x}$ are estimated jointly with a hierarchical approach. Such a strategy is mandatory to avoid biases in $\boldsymbol{x}$ (e.g., morphology distortions, under-estimation of the intensity) induced by the contamination of the nuisance statistics by the off-axis objects (whose contributions are partly encoded in $\boldsymbol{\Omega}$).

Under our model assumptions made in Sects. 2.2 and 2.3, the data-fidelity takes the form of a (modified) co-log-likelihood term penalizing the discrepancy between the observations $\boldsymbol{r}$ and the contribution $\mathbf{M}\,\boldsymbol{x}$ of the off-axis objects in $\boldsymbol{r}$:

$$\mathcal{D}(\boldsymbol{r},\mathbf{M},\boldsymbol{x},\boldsymbol{\Omega},\boldsymbol{\mu}) = \sum_{n\in\mathbb{P}} \left[\sum_{t=1}^{T} \frac{1}{2} \log\det\left(\sigma_{n,t}^2 \mathbf{C}_n^{\text{spat}}\right) + \frac{1}{2}\,\text{tr}\left(\mathbf{C}_n^{\text{spat}\,-1}\left(\mathbf{W}_n^{\text{spat}} \odot \sum_{t=1}^{T} \frac{1}{\sigma_{n,t}^2}\boldsymbol{u}_{n,t}(\boldsymbol{x})\,\boldsymbol{u}_{n,t}(\boldsymbol{x})^{\top}\right)\right)\right], \tag{19}$$

where the shrinkage matrices $\{\mathbf{W}_n^{\text{spat}}\}_{n\in\mathbb{P}}$ are evaluated once with Eq. (12) at the beginning of the optimization procedure and are then considered fixed, i.e. they are supposed to be independent from $\boldsymbol{x}$. Formulation (19) is an extension of our previous work[1] under a GSM model for the nuisance component. That specific form of the data-fidelity term ensures that the shrinkage estimators $\widehat{\boldsymbol{\Omega}}$ are minimizers of $\mathcal{D}$.

A regularization term $\mathcal{R}$ is added to the data-fidelity $\mathcal{D}$ to enforce prior knowledge about the unknown object $\boldsymbol{x}$ and to improve the conditioning of the inversion. As in our previous work,[1] we select a weighted combination between two penalizations. We first penalize the ℓ_1-norm of $\boldsymbol{x}$ in order to promote the sparsity of the solution (i.e., we expect to retrieve a well contrasted disk on an almost null background). The second penalization takes the form of a ℓ_2–ℓ_1 *edge-preserving* regularization[32] promoting the sparsity of the gradient of $\boldsymbol{x}$ (i.e., we expect to retrieve smooth objects with sharp edges). In that context, the regularization term $\mathcal{R}$ writes:

$$\mathcal{R}(\boldsymbol{x},\boldsymbol{\mu}) \underset{\boldsymbol{x}\geq\boldsymbol{0}}{=} \mu_{\ell_1} \sum_{m=1}^{M} x_m + \mu_{\ell_2-\ell_1} \sum_{m=1}^{M} ||\sqrt{||\boldsymbol{\Delta}_m\boldsymbol{x}||_2^2 + \epsilon^2}\,, \tag{20}$$

where operator $\boldsymbol{\Delta}_m$ is the finite differences approximation of the the spatial gradient at location m, $\epsilon = \sqrt{10^{-7}}$ is a minimum-value constant setting the edge-preserving property of the regularization. Regularization $\mathcal{R}$ stays sufficiently generic to be well-suited for various types of disk morphologies (see our previous work[1]); from protoplanetary disks exhibiting a complex intensity distribution to transition and debris disks having generally a simpler morphology. The hyper-parameters $\boldsymbol{\mu} = \{\mu_{\ell_1}, \mu_{\ell_2-\ell_1}\}$ play a key role since they set the relative weight of the two regularizers involved in Eq. (20). Following our previous work,[1] they are optimally estimated in a data-driven fashion by minimizing a quantitative criterion.[33,34]

Once the data-fidelity (Eq. (19)) and the regularization (Eq. (20)) terms are defined, we solve the constrained minimization problem with VMLM-B optimization algorithm.[35] This quasi-Newton method with limited memory implements bound constraints (here, non-negativity of $\boldsymbol{x}$) and requires solely the objective function $\mathcal{C}$ and its gradient.

2.5 Approximation of the uncertainties

In this section, we aim to approximate a confidence map, i.e. to derive an estimator of the standard-deviation $\boldsymbol{\gamma}$ on the reconstructed intensity distribution. Such an estimator is very useful to evaluate the faithfulness of the reconstruction as a function of the location in the field-of-view and to compare the quality of different

observations. Areas where the confidence is low should be analyzed with care since they correspond to areas where the reconstruction is mostly driven by the regularization term $\mathcal{R}$. We resort to the computation of the Cramér-Rao lower bounds (CRLBs) which are good estimates of the covariance of maximum likelihood estimators when the number of samples is large.[36] Computing the CRLBs implies having a parametric model of the data that involves the unknown parameters. Here, we consider a affine model of the data:

$$\mathbf{D} := \left(\mathbf{M} - \frac{1}{T}\overline{\mathbf{M}}\right) , \tag{21}$$

with $\overline{\mathbf{M}} = \mathbf{M}\,\mathbf{1}_T^\top$ the averaging operator. Then, we derive the Fisher's information $\left[\mathrm{I}^{\mathrm{F}}(\alpha)\right]$ about the flux α contained in the data at location n:

$$\left[\mathrm{I}^{\mathrm{F}}(\alpha)\right]_n = \sum_{t=1}^{T} \frac{1}{\sigma_{n,t}^2} \frac{\partial \mathbf{D}(\boldsymbol{x} + \alpha\delta_n)^\top}{\partial \alpha} \mathbf{C}_n^{\mathrm{spat}^{-1}}(\boldsymbol{x}) \frac{\partial \mathbf{D}(\boldsymbol{x} + \alpha\delta_n)}{\partial \alpha} , \tag{22}$$

where δ_n is a discrete dirac at pixel location n in the object space and α the unknown flux at that location. Thus, the product $\alpha\delta_n$ models the signal of a point-like source of flux α. This approximation amounts to neglect the effect of the regularization and to consider the maximum likelihood estimation of the intensity of a point-like source superimposed to the disk. We note that in Eq. (22), the Fisher's information depends on $\boldsymbol{x}$ via the covariance matrices $\{\mathbf{C}_n^{\mathrm{spat}}\}_{n\in\mathbb{P}}$. Following our previous work on the `robust PACO` algorithm, we also derive a *best case* Fisher's information $\mathrm{I}^{\mathrm{best\ F}}(\alpha)$ that corresponds to the amount of information that would have been contained in the data should the observing conditions have been as good as that of the best temporal frame. For a given pixel location n, such a computation amounts to replace each temporal weighting factor $\{\sigma_{n,t}^2\}_{t=1:T}$ by the smallest one:

$$\left[\mathrm{I}^{\mathrm{best\ F}}(\alpha)\right]_n = \frac{T}{\min_t(\{\sigma_{n,t}^2\}_{t=1:T})} \frac{\partial \mathbf{D}(\boldsymbol{x} + \alpha\delta_n)^\top}{\partial \alpha} \mathbf{C}_n^{\mathrm{spat}^{-1}}(\boldsymbol{x}) \frac{\partial \mathbf{D}(\boldsymbol{x} + \alpha\delta_n)}{\partial \alpha} . \tag{23}$$

The estimate γ_n at location n is obtained from the Fisher's information by:

$$\gamma_n = \sqrt{\left[\mathrm{I}^{\mathrm{F}}(\alpha)^{-1}\right]_n} . \tag{24}$$

A similar expression holds for the best case estimate γ_n^{best} by replacing $\mathrm{I}_n^{\mathrm{F}}$ with $\mathrm{I}_n^{\mathrm{F\ best}}$ in the previous expression.

3. RECONSTRUCTION ALGORITHM FROM ASDI DATASETS

Building on the `robust REXPACO` algorithm described in Sect. 2, we describe in this section the cornerstones of its extension to the reconstruction of the circumstellar environments from ASDI datasets[†].

3.1 Modeling of the measured intensity and of instrumental effects

The additive model of the data defined in Eq. (1) stays unchanged except that the observations $\boldsymbol{r}$ and the nuisance component now lies in $\mathbb{R}^{NLT}$, with L is the number of spectral channels ($L = 39$ for VLT/SPHERE-IFS). Similarly, the intensity distribution of the sought object $\boldsymbol{x}$ is now a vector of $\mathbb{R}_+^{ML}$.

The direct model of the instrumental effects defined in Eqs. (3) should be complemented to account for the spectral diversity of the observations, in particular to the linear scaling of speckles with the wavelength. Since we aim at capturing the spatio-spectral correlations of the nuisance component (see Sect. 3.2), the data should be spatially rescaled with a factor $\lambda/\lambda_{\mathrm{ref}}$ (e.g., $\lambda_{\mathrm{ref}} = \lambda_1$) so that each frame at time t and at wavelength λ displays

[†]This algorithm will be described in more details in a paper currently in preparation, Flasseur *et al.*

speckles pattern with superimposed structures. The direct model $\mathbf{M}$ now writes:

$$\mathbf{M}_t = \underbrace{\mathbf{S}\,\mathbf{Q}_t}_{\mathbf{U}_t}\,\mathbf{B} \quad \text{such that} \quad \mathbf{M} = \underbrace{\begin{pmatrix} \mathbf{U}_1 : \mathbb{R}^{ML} \to \mathbb{R}^{NL} \\ \vdots \\ \mathbf{U}_t : \mathbb{R}^{ML} \to \mathbb{R}^{NL} \\ \vdots \\ \mathbf{U}_T : \mathbb{R}^{ML} \to \mathbb{R}^{NL} \end{pmatrix}}_{\mathbf{U}} \mathbf{B}\,, \tag{25}$$

where $\mathbf{S}$ is a time-invariant *scaling* operator compensating for the image scaling applied during a pre-processing step to spatially align the speckles pattern across the spectral axis. The operator $\mathbf{S}$ is implemented as a sparse interpolation matrix.

3.2 Modeling of the nuisance component

We aim to extend the statistical framework developed in Sect. 2.3 to build a spatio-temporo-spectral model of the nuisance component. We enrich the general structures of the mean $\boldsymbol{m}_n$ (Eq. (5)) and of the covariance matrix $\mathbf{C}_n$ (Eq. (7)) as:

$$\begin{cases} \boldsymbol{m}_n = \boldsymbol{m}_n^{\text{spec}} \otimes \mathbf{1}_T \ \in \mathbb{R}^{KLT} \\ \mathbf{C}_n = \boldsymbol{\Lambda}_n \otimes \mathbf{C}_n^{\text{spec}} \otimes \mathbf{C}_n^{\text{spat}} \ \in \mathbb{R}^{KLT \times KLT}\,, \end{cases} \tag{26}$$

where $\boldsymbol{m}_n^{\text{spec}} \in \mathbb{R}^{KL}$ and $\mathbf{C}_n^{\text{spec}} \in \mathbb{R}^{L \times L}$ are respectively the multi-spectral mean and covariance matrix. In that context, the statistics of the nuisance component defined in Eq. (16) are modified to account for the specific form of the model defined in Eq. (26). The expression of the estimator $\widehat{\boldsymbol{m}}_n^{\text{spat}}$ defined in Eqs. (16) stays unchanged and now holds for $\widehat{\boldsymbol{m}}_n^{\text{spec}}$ with $\boldsymbol{r}_{n,t} \in \mathbb{R}^{KL}$ being the results of the vectorization of a spatio-spectral patch extracted around location n. The estimators $\{\widehat{\mathbf{C}}_n^{\text{spat}}, \widehat{\mathbf{C}}_n^{\text{spec}}, \widehat{\sigma}_{n,t}^2\}_{n \in \mathbb{P}, t=1:T}$ write:

$$\begin{cases} \widehat{\mathbf{C}}_n^{\text{spat}} = \widehat{\mathbf{W}}_n^{\text{spat}} \odot \frac{1}{TL} \sum_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2} \text{mat}\left(\widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})\right) \left(\widehat{\mathbf{W}}_n^{\text{spec}} \odot \widehat{\mathbf{C}}_n^{\text{spec}^{-1}}\right) \text{mat}\left(\widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})\right)^\top\,, \\ \widehat{\mathbf{C}}_n^{\text{spec}} = \widehat{\mathbf{W}}_n^{\text{spec}} \odot \frac{1}{TK} \sum_{t=1}^{T} \frac{1}{\widehat{\sigma}_{n,t}^2} \text{mat}\left(\widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})\right)^\top \left(\widehat{\mathbf{W}}_n^{\text{spat}} \odot \widehat{\mathbf{C}}_n^{\text{spat}^{-1}}\right) \text{mat}\left(\widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})\right)\,, \\ \widehat{\sigma}_{n,t}^2 = \frac{1}{KL} \widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})^\top \left(\left(\widehat{\mathbf{W}}_n^{\text{spec}} \odot \widehat{\mathbf{C}}_n^{\text{spec}}\right)\left(\widehat{\mathbf{W}}_n^{\text{spat}} \odot \widehat{\mathbf{C}}_n^{\text{spat}}\right)\right)^{-1} \widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})\,, \end{cases} \tag{27}$$

where the residuals $\widehat{\boldsymbol{u}}_{n,t}(\boldsymbol{x})$ as defined in Eq. (17) now lie in $\mathbb{R}^{KL}$, and $\text{mat}(\boldsymbol{y})$ is an operator performing the reshape of any vector $\boldsymbol{y} \in \mathbb{R}^{AB}$ into a matrix $\mathbf{Y} \in \mathbb{R}^{A \times B}$. In Eqs. (27), the regularization by shrinkage of the sample covariance matrices, and thus the estimation of the weighting matrices $\widehat{\mathbf{W}}_n^{\text{spat}}$ and $\widehat{\mathbf{W}}_n^{\text{spat}}$ are performed in a similar fashion than detailed in Sect. 2.3. Given the estimators defined in Eqs. (27), the objective function is extended to jointly minimize the spatio-temporo-spectral residuals and the constrained optimization problem is also solved with the VMLM-B algorithm.[35]

4. RECONSTRUCTION RESULTS ON DATASETS FROM THE VLT/SPHERE-IRDIS AND VLT/SPHERE-IFS INSTRUMENTS

4.1 Datasets description

To assess the performance of the method proposed in Sect. 2, we consider two datasets recorded with the VLT/SPHERE-IRDIS instrument. These datasets result from the observations of the PDS 70 star (observing dates: 2018-02-24 and 2019-04-14, spectral band: K1 i.e. $\lambda = 2.110\mu$m). The 2018's observations have been obtained under excellent observing conditions and reconstructions from this dataset have been published by several authors[1,17,18,20,23] including our previous work[1] on the `REXPACO` algorithm. The 2019's observations have been obtained under medium to bad observing conditions.

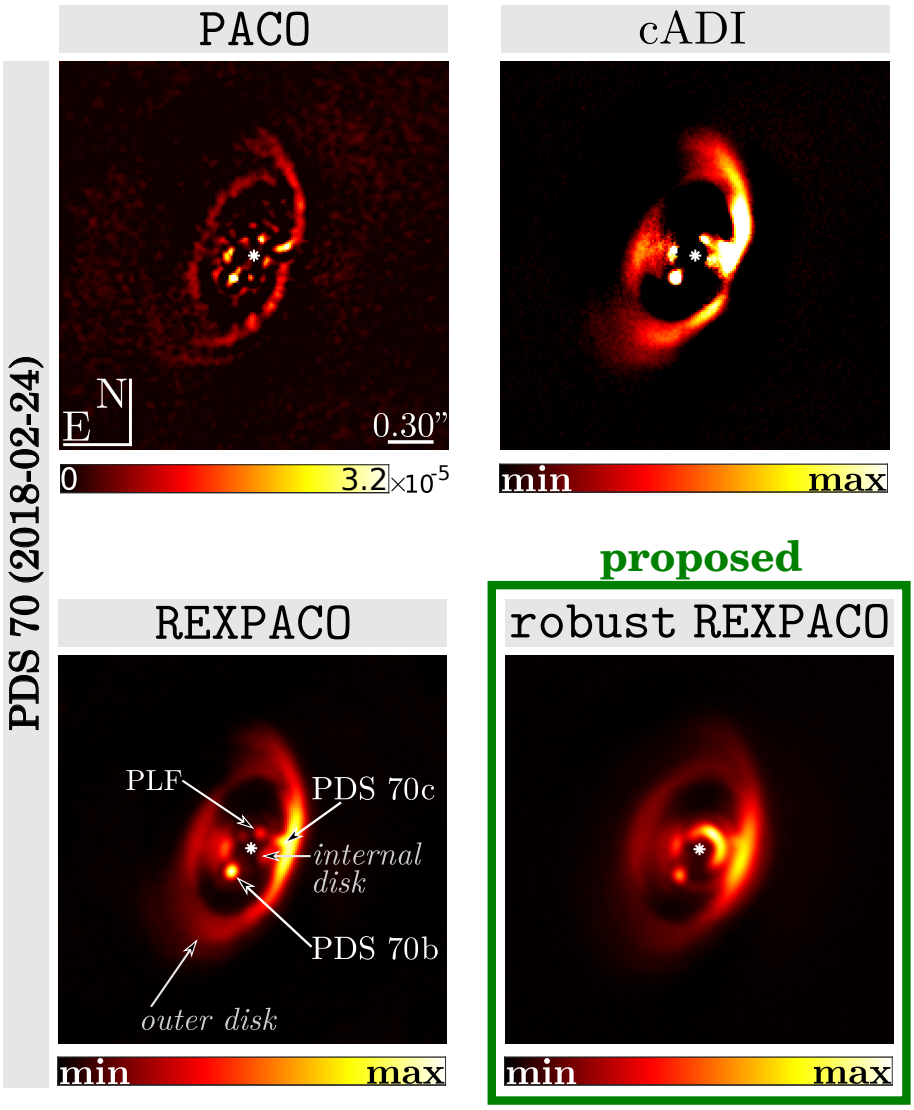

Figure 1. Images of the flux distribution reconstructed by `robust REXPACO` comparatively to the PACO, cADI, and `REXPACO` algorithms. For `REXPACO` and `robust REXPACO`, we display a reblurred version $\mathbf{H}\boldsymbol{x}$ of the deconvolved intensity distribution $\boldsymbol{x}$ to ease comparisons with the other methods that do not produce deblurred images of the flux distribution. Dataset: PDS 70 (2018-02-24, K1 spectral band).

To evaluate qualitatively the gain brought by the method described in Sect. 3, we select two datasets obtained with the VLT/SPHERE-IFS imager from the observations of the HR 4796 star (observing date: 2015-02-02, spectral bands: Y-J i.e. $\lambda \in [0.9; 1.3]\,\mu$m) and of the SAO 206462 star (observing date: 2015-05-15, spectral bands: Y-J-H i.e. $\lambda \in [0.9; 1.6]\,\mu$m).

The four A(S)DI raw observations were pre-reduced with the data reduction handling pipeline[37] of the SPHERE instrument. It performs thermal background subtraction, flat-field correction, anamorphism correction, compensation for spectral transmission, flux normalization, bad pixels identification and interpolation, frame centering, true-North alignment, and frame selection. These operations are complemented by custom routines implemented in the SPHERE data center[38] in particular to refine the spectral calibrations, to attenuate the spectral cross-talk, and to improve the correction of bad pixels. Finally, the SPHERE data center combines the pre-reduced observations and delivers the calibrated ADI and ASDI datasets we consider in this work.

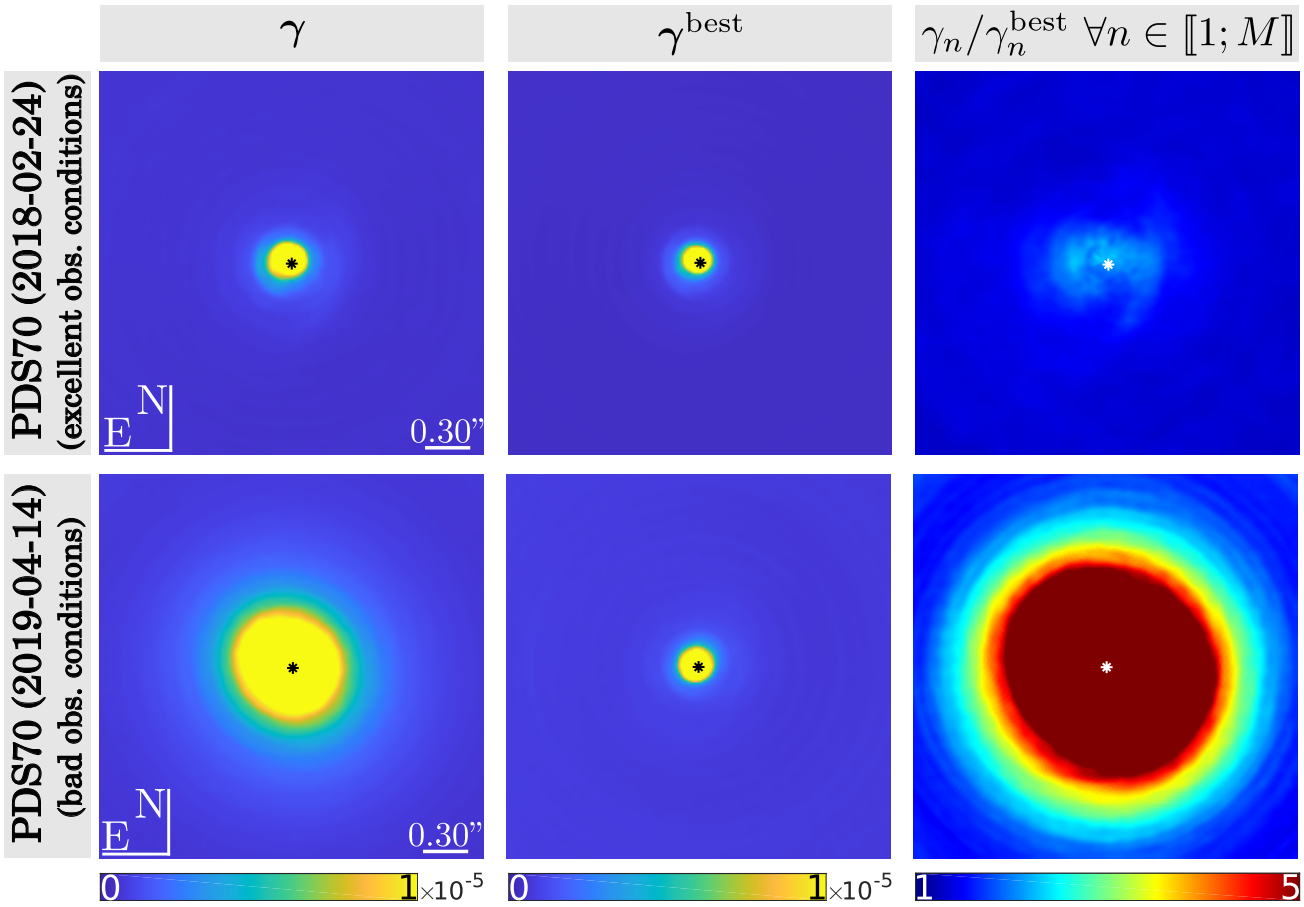

Figure 2. Images of the theoretical accuracy on the flux distribution (minimal standard deviation given by the Cramér-Rao lower bounds). The distribution $\boldsymbol{\gamma}$ is obtained with the actual quality of the datasets while the distribution $\boldsymbol{\gamma}^{\text{best}}$ is obtained by assuming that the observing conditions have been as good as that of the best temporal frame. The third column displays ratio between $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}^{\text{best}}$ (closer than 1 means that the observing conditions are stable during the observations). Datasets: PDS 70 (2018-02-24, K1 spectral band, excellent observing conditions; 2019-04-14, K1 spectral band, medium to bad observing conditions).

4.2 Results on VLT/SPHERE-IRDIS and VLT/SPHERE-IFS datasets

To evaluate the benefits of the `robust REXPACO` algorithm proposed in Sect. 2, we compare qualitatively the intensity distribution $\widehat{\boldsymbol{x}}$ reconstructed with the cADI,[21] `PACO`,[27–29] `REXPACO`[1] and `robust REXPACO`. Figure 1 presents reconstruction results from these four algorithms on the 2018's dataset of PDS 70 described in Sect. 4.1. The results from `PACO`, cADI and `REXPACO` algorithms were already published in our previous work on the `REXPACO` algorithm (see Fig. 11, 4th column of Flasseur *et al.* 2021[1]). We first focus on the comparison between (i) methods specifically designed to the reconstruction of the circumstellar environment (i.e., `REXPACO` and `robust REXPACO`), and (ii) standard processing algorithms of the field like cADI. We observe that `REXPACO` and `robust REXPACO` are able to better reject the nuisance component, especially near the star. They also allow to significantly reduce the typical artifacts induced by cADI like non-physical discontinuities, morphological distortions, and flux attenuation in the structures of the disk. While being grounded on the same statistical model of the nuisance component than `REXPACO`, the state-of-the-art detection algorithm `PACO` is prone to several artifacts in the presence of extended features since it assumes that the pattern to detect takes the form of a point-like source. The advantages of (`robust`) `REXPACO` are brought by the unique joint modeling of the nuisance component and of the instrumental effects embedded in a reconstruction framework. We refer the reader to our previous work[1] for a detailed discussion of these benefits. Focusing on the comparison between `REXPACO` and

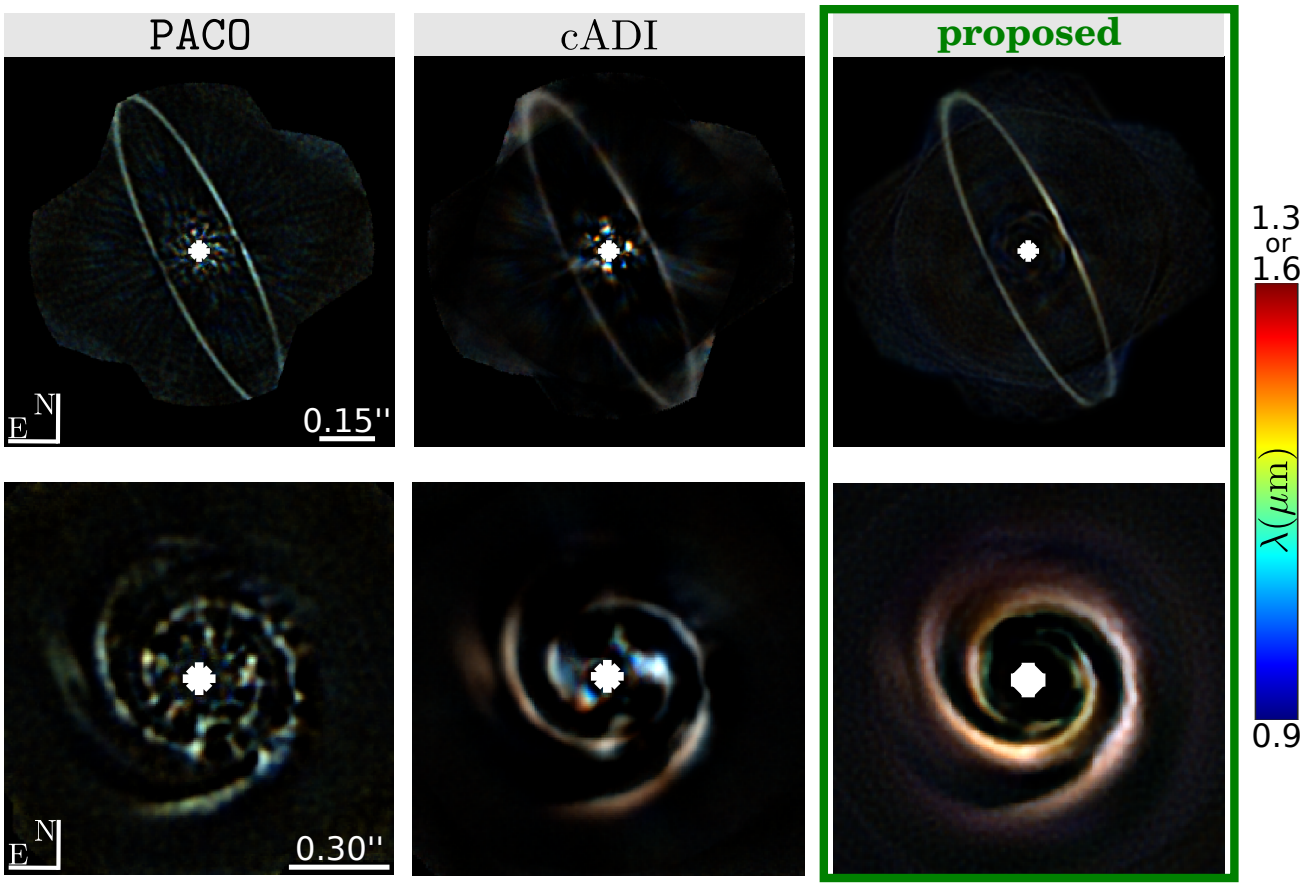

Figure 3. Images of the flux distribution reconstructed by the method proposed in Sect. 3 for ASDI datasets comparatively to the `PACO` and cADI algorithms. The deblurred reconstruction $\widehat{\boldsymbol{x}}$ is displayed for the proposed algorithm. Datasets: HR 4796 (2015-02-02, Y-J spectral bands, medium observing conditions) and SAO 206462 (2015-05-15, Y-J-H spectral bands, medium observing conditions).

`robust REXPACO`, we observe that the data-driven temporal robustness embedded in `robust REXPACO` allows to reconstruct fine structures of the disk (especially at short angular separations) that were rejected within the nuisance component by `REXPACO`. In particular, the internal disk of PSD 70 is well visible in the `robust REXPACO` reconstruction (the discontinuity on the South-East is due to a slight decentering of the coronagraph during the whole sequence of observations). Besides, the `REXPACO` algorithm exhibits a point-like feature (PLF) first pointed out by Mesa *et al.* 2021[20] from the analysis of detection maps produced by several detection algorithms of the field. The analysis conducted by Mesa *et al.* 2021[20] of the spectral energy distribution of this point-like feature from SPHERE-IFS observations showed that this feature would be more likely part of the internal disk. The `robust REXPACO` reconstruction from the 2018's dataset clearly argues in favor of this hypothesis. Finally, the `robust REXPACO` reconstruction exhibits additional structures, at the best of our knowledge unrevealed by state-of-the-art algorithms of the field, in particular between the outer and internal disks[‡].

Figure 2 shows images of the theoretical accuracy on the reconstructed flux distribution for the two datasets of PDS 70. The distribution $\boldsymbol{\gamma}$ is obtained from Eqs. (22) and (24) considering the actual quality of the datasets while the distribution $\boldsymbol{\gamma}^{\text{best}}$ is obtained from Eqs. (23) and (24) by assuming that the observing conditions have been as good as that of the best temporal frame. This figure emphasizes that the overall quality of the 2018's dataset is significantly better than the quality of the 2019's dataset. Besides, this figure emphasizes that

[‡]A detailed astrophysical analysis based on the reconstruction of multiple observations of PDS 70 with state-of-the-art reconstruction algorithms will be carried out in a paper (Langlois *et al.*, in prep.) currently in preparation.

observing conditions were quite stable during acquisition of the 2018's dataset since the best expected accuracy $\boldsymbol{\gamma}^{\text{best}}$ is not too far from the actual one $\boldsymbol{\gamma}$. The 2019's dataset was acquired under highly variable observing conditions since the uncertainty $\boldsymbol{\gamma}^{\text{best}}$ is significantly lower than $\boldsymbol{\gamma}$ at each location of the field-of-view.

We now focus on the method proposed in Sect. 3 for the reconstruction of the circumstellar environnement from ASDI datasets. Figure 3 gives reconstructions obtained with the proposed method comparatively to the `PACO` and cADI algorithms. Similar conclusions than in the two previous paragraphs can be drawn: the proposed method is able to produce images of the flux distribution with a significant improvement compared to cADI and `PACO`. The gain is particularly promising for disks exhibiting a circular symmetry such as the spiral-like feature around SAO 206462. For this type of disks, the unmixing between the nuisance component and the disk is all the more difficult given that parts of the nuisance component are always superimposed to disk structures despite the temporal diversity brought by ADI. The additional spectral diversity combined with an accurate modeling of the nuisance component and of the instrumental effects are the keys to perform the reconstruction of the sought objects.

5. CONCLUSION

In this contribution, we have introduced `robust REXPACO`, an extension of the `REXPACO` algorithm, dedicated to the reconstruction of the intensity distribution in the circumstellar environment from ADI datasets. It encompasses a statistical modeling of the nuisance component with a scaled mixture of multi-variate Gaussian whose parameters are learnt in a data-driven fashion at the scale of small patches of a few tens of pixels. The statistics of the nuisance component and the modeling of the instrumental effects are embedded in a regularized reconstruction framework following an inverse problem approach. In a second part, we also discuss how this algorithm can be extended to process jointly the different spectral channels of ASDI datasets. We applied our methods on several ADI and ASDI observations from the IRDIS and IFS imagers of the VLT/SPHERE instrument and we showed that the proposed algorithms significantly reduce the typical artifacts produced by state-of-the-art algorithms. By also taking into account the instrumental point spread function, our algorithms yield deblurred estimates of the object of interest without the artifacts resulting from other methods. Besides, we have shown that `robust REXPACO` has an improved robustness than the state-of-the-art `REXPACO` algorithm. Finally, our experiments demonstrate that the fidelity of the (statistical) model with respect to the observations is the key to improve the overall quality of the reconstructions. In that context, we plan to work on refinements of the statistical model on which our algorithms are based, e.g. by capturing covariances at a larger spatial scale.

ACKNOWLEDGMENTS

This work has made use of the SPHERE Data Center, jointly operated by OSUG/IPAG (Grenoble, France), PYTHEAS/LAM/CESAM (Marseille, France), OCA/Lagrange (Nice, France), Observatoire de Paris/LESIA (Paris, France), and Observatoire de Lyon/CRAL (Lyon, France).

REFERENCES

[1] Flasseur, O., Denis, L., Thiébaut, É., Langlois, M., et al., "REXPACO: An algorithm for high contrast reconstruction of the circumstellar environment by angular differential imaging," *Astronomy & Astrophysics* **651**, A62 (2021).

[2] Traub, W. A. and Oppenheimer, B. R., "Direct imaging of exoplanets," *Exoplanets* , 111–156 (2010).

[3] Bowler, B. P., "Imaging extrasolar giant planets," *Publications of the Astronomical Society of the Pacific* **128**(968), 102001 (2016).

[4] Pueyo, L., "Direct imaging as a detection technique for exoplanets," *Handbook of Exoplanets* , 705–765 (2018).

[5] Beuzit, J.-L., Vigan, A., Mouillet, D., Dohlen, K., Gratton, R., Boccaletti, A., Sauvage, J.-F., Schmid, H. M., Langlois, M., Petit, C., et al., "SPHERE: the exoplanet imager for the Very Large Telescope," *Astronomy & Astrophysics* **631**, A155 (2019).

[6] Macintosh, B., Graham, J. R., Ingraham, P., Konopacky, Q., Marois, C., Perrin, M., Poyneer, L., Bauman, B., Barman, T., Burrows, A. S., et al., "First light of the Gemini Planet Imager," *Proceedings of the National Academy of Sciences* **111**(35), 12661–12666 (2014).

[7] Morzinski, K. M., Close, L. M., Males, J. R., Kopon, D., Hinz, P. M., Esposito, S., Riccardi, A., Puglisi, A., Pinna, E., Briguglio, R., et al., "MagAO: Status and on-sky performance of the Magellan adaptive optics system," in [*Adaptive Optics Systems IV*], **9148**, 914804, International Society for Optics and Photonics (2014).

[8] Jovanovic, N., Martinache, F., Guyon, O., Clergeon, C., Singh, G., Kudo, T., Garrel, V., Newman, K., Doughty, D., Lozi, J., et al., "The Subaru coronagraphic extreme adaptive optics system: enabling high-contrast imaging on solar-system scales," *Publications of the Astronomical Society of the Pacific* **127**(955), 890 (2015).

[9] Follette, K. B., Rameau, J., Dong, R., Pueyo, L., Close, L. M., Duchêne, G., Fung, J., Leonard, C., Macintosh, B., Males, J. R., et al., "Complex spiral structure in the HD 100546 transitional disk as revealed by gpi and MagAO," *The Astronomical Journal* **153**(6), 264 (2017).

[10] Langlois, M., Pohl, A., Lagrange, A.-M., Maire, A.-L., Mesa, D., Boccaletti, A., Gratton, R., Denneulin, L., Klahr, H., Vigan, A., et al., "First scattered light detection of a nearly edge-on transition disk around the T Tauri star RY lupi," *Astronomy & Astrophysics* **614**, A88 (2018).

[11] Mesa, D., Langlois, M., Garufi, A., Gratton, R., Desidera, S., D'Orazi, V., Flasseur, O., Barbieri, M., Benisty, M., Henning, T., et al., "Determining mass limits around HD 163296 through SPHERE direct imaging data," *Monthly Notices of the Royal Astronomical Society* **488**(1), 37–46 (2019).

[12] Garufi, A., Podio, L., Bacciotti, F., Antoniucci, S., Boccaletti, A., Codella, C., Dougados, C., Ménard, F., Mesa, D., Meyer, M., et al., "Sphere view of the jet and the envelope of RY Tauri," *Astronomy & Astrophysics* **628**, A68 (2019).

[13] Perrot, C., Thebault, P., Lagrange, A.-M., Boccaletti, A., Vigan, A., Desidera, S., Ausgereau, J.-C., Bonnefoy, M., Choquet, É., Kral, Q., et al., "First resolved observations of a highly asymmetric debris disc around HD 160305 with VLT/SPHERE," *Astronomy & Astrophysics* **626**, A95 (2019).

[14] Sissa, E., Gratton, R., Garufi, A., Rigliaco, E., Zurlo, A., Mesa, D., Langlois, M., De Boer, J., Desidera, S., Ginski, C., et al., "High-contrast study of the candidate planets and protoplanetary disk around HD 100546," *Astronomy & Astrophysics* **619**, A160 (2018).

[15] Cugno, G., Quanz, S., Hunziker, S., Stolker, T., Schmid, H., Avenhaus, H., Baudoz, P., Bohn, A., Bonnefoy, M., Buenzli, E., et al., "A search for accreting young companions embedded in circumstellar disks," *Astronomy and Astrophysics* **622**, A156 (2019).

[16] Rich, E. A., Wisniewski, J. P., Currie, T., Fukagawa, M., Grady, C. A., Sitko, M. L., Pikhartova, M., Hashimoto, J., Abe, L., Brandner, W., et al., "Multi-epoch direct imaging and time-variable scattered light morphology of the HD 163296 protoplanetary disk," *The Astrophysical Journal* **875**(1), 38 (2019).

[17] Keppler, M., Benisty, M., Müller, A., Henning, T., Van Boekel, R., Cantalloube, F., Ginski, C., Van Holstein, R., Maire, A.-L., Pohl, A., et al., "Discovery of a planetary-mass companion within the gap of the transition disk around PDS 70," *Astronomy & Astrophysics* **617**, A44 (2018).

[18] Müller, A., Keppler, M., Henning, T., Samland, M., Chauvin, G., Beust, H., Maire, A.-L., Molaverdikhani, K., van Boekel, R., Benisty, M., et al., "Orbital and atmospheric characterization of the planet within the gap of the PDS 70 transition disk," *Astronomy & Astrophysics* **617**, L2 (2018).

[19] Haffert, S., Bohn, A., de Boer, J., Snellen, I., Brinchmann, J., Girard, J., Keller, C., and Bacon, R., "Two accreting protoplanets around the young star PDS 70," *Nature Astronomy* **3**(8), 749–754 (2019).

[20] Mesa, D., Keppler, M., Cantalloube, F., Rodet, L., Charnay, B., Gratton, R., Langlois, M., Boccaletti, A., Bonnefoy, M., Vigan, A., et al., "VLT/SPHERE exploration of the young multiplanetary system PDS70," *Astronomy & Astrophysics* **632**, A25 (2019).

[21] Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., and Nadeau, D., "Angular differential imaging: A powerful high-contrast imaging technique," *The Astrophysical Journal* **641**(1), 556 (2006).

[22] Racine, R., Walker, G. A., Nadeau, D., Doyon, R., and Marois, C., "Speckle noise and the detection of faint companions," *Publications of the Astronomical Society of the Pacific* **111**(759), 587 (1999).

[23] Pairet, B., Cantalloube, F., and Jacques, L., "MAYONNAISE: a morphological components analysis pipeline for circumstellar discs and exoplanets imaging in the near-infrared," *Monthly Notices of the Royal Astronomical Society* **503**(3), 3724–3742 (2021).

[24] Flasseur, O., Denis, L., Thiébaut, É., and Langlois, M., "Robustness to bad frames in angular differential imaging: a local weighting approach," *Astronomy & Astrophysics* **634**, A2 (2020).

[25] Wainwright, M. J. and Simoncelli, E. P., "Scale mixtures of gaussians and the statistics of natural images," in [*Advances in Neural Information Processing Systems*], 855–861 (2000).

[26] Portilla, J., Strela, V., Wainwright, M. J., and Simoncelli, E. P., "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans Image Processing* **12**(11) (2003).

[27] Flasseur, O., Denis, L., Thiébaut, É., and Langlois, M., "Exoplanet detection in angular differential imaging by statistical learning of the nonstationary patch covariances - The PACO algorithm," *Astronomy & Astrophysics* **618**, A138 (2018).

[28] Flasseur, O., Denis, L., Thiébaut, É., and Langlois, M., "An unsupervised patch-based approach for exoplanet detection by direct imaging," in [*IEEE International Conference on Image Processing*], 2735–2739 (2018).

[29] Flasseur, O., Denis, L., Thiébaut, É. M., and Langlois, M., "Exoplanet detection in angular and spectral differential imaging: local learning of background correlations for improved detections," in [*SPIE Astronomical Telescopes + Instrumentation*], **10703**, 107032R, International Society for Optics and Photonics (2018).

[30] Ledoit, O. and Wolf, M., "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis* **88**(2), 365–411 (2004).

[31] Chen, Y., Wiesel, A., Eldar, Y. C., and Hero, A. O., "Shrinkage algorithms for MMSE covariance estimation," *IEEE Transactions on Signal Processing* **58**(10), 5016–5029 (2010).

[32] Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M., "Deterministic edge-preserving regularization in computed imaging," *IEEE Transactions on Image Processing* **6**(2), 298–311 (1997).

[33] Girard, D. A., "A fast 'monte-carlo cross-validation' procedure for large least squares problems with noisy data," *Numerische Mathematik* **56**, 1–23 (1989).

[34] Stein, C. M., "Estimation of the mean of a multivariate normal distribution," *The Annals of Statistics* , 1135–1151 (1981).

[35] Thiébaut, É., "Optimization issues in blind deconvolution algorithms," in [*SPIE Astronomical Telescopes + Instrumentation*], **4847**, 174–183, International Society for Optics and Photonics (2002).

[36] Kendall, M. G., Stuart, A., and Ord, J. K., [*The advanced theory of statistics*], vol. 1, JSTOR (1948).

[37] Pavlov, A., Möller-Nilsson, O., Feldt, M., Henning, T., Beuzit, J.-L., and Mouillet, D., "SPHERE data reduction and handling system: overview, project status, and development," in [*SPIE Astronomical Telescopes + Instrumentation*], **7019**, 701939, International Society for Optics and Photonics (2008).

[38] Delorme, P., Meunier, N., Albert, D., Lagadec, E., Le Coroller, H., Galicher, R., Mouillet, D., Boccaletti, A., Mesa, D., Meunier, J. C., Beuzit, J. L., Lagrange, A. M., Chauvin, G., Sapone, A., Langlois, M., Maire, A. L., Montargès, M., Gratton, R., Vigan, A., and Surace, C., "The SPHERE Data Center: a reference for high contrast imaging processing," in [*Annual meeting of the French Society of Astronomy and Astrophysics*], (Dec. 2017).