

PIC: a data reduction algorithm for integral field spectrographs

Application to the SPHERE instrument

Anthony Berdeu¹, Ferréol Soulez², Loïc Denis¹, Maud Langlois², and Éric Thiébaud²

¹ Université de Lyon, UJM-Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, 42023 Saint-Etienne, France
e-mail: a.berdeu@gmail.com

² Université de Lyon, Université Lyon1, ENS de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon, UMR 5574, 69230 Saint-Genis-Laval, France
e-mail: ferreol.soulez@univ-lyon1.fr

Received 10 October 2019 / Accepted 8 January 2020

ABSTRACT

Context. The improvement of large size detectors permitted the development of integral field spectrographs (IFSs) in astronomy. Spectral information for each spatial element of a two-dimensional field of view is obtained thanks to integral field units that spread the spectra on the 2D grid of the sensor.

Aims. Here we aim to solve the inherent issues raised by standard data-reduction algorithms based on direct mapping of the 2D+ λ data cube: the spectral cross-talk due to the overlap of neighbouring spectra, the spatial correlations of the noise due to the re-interpolation of the cube on a Cartesian grid, and the artefacts due to the influence of defective pixels.

Methods. The proposed method, Projection, Interpolation, and Convolution (PIC), is based on an “inverse-problems” approach. By accounting for the overlap of neighbouring spectra as well as the spatial extension in a spectrum of a given wavelength, the model inversion reduces the spectral cross-talk while deconvolving the spectral dispersion. Considered as missing data, defective pixels undetected during the calibration are discarded on-the-fly via a robust penalisation of the data fidelity term.

Results. The calibration of the proposed model is presented for the Spectro-Polarimetric High-contrast Exoplanet REsearch instrument (SPHERE). This calibration was applied to extended objects as well as coronagraphic acquisitions dedicated to exoplanet detection or disc imaging. Artefacts due to badly corrected defective pixels or artificial background structures observed in the cube reduced by the SPHERE data reduction pipeline are suppressed while the reconstructed spectra are sharper. This reduces the false detections by the standard exoplanet detection algorithms.

Conclusions. These results show the pertinence of the inverse-problems approach to reduce the raw data produced by IFSs and to compensate for some of their imperfections. Our modelling forms an initial building block necessary to develop methods that can reconstruct and/or detect sources directly from the raw data.

Key words. techniques: image processing – techniques: imaging spectroscopy – methods: numerical

1. General context

The pertinence of single-field spectroscopy has been demonstrated in numerous fields in astrophysics, such as stellar classification, chemical analysis, velocity measurement, magnetic field probing, and so on. In recent decades, a requirement has emerged to complement this method with increased sampling on extended objects. While the use of long-slit spectrographs is possible by scanning the field of view in one direction, this technique is time consuming and requires precise control of the position in the field being sampled.

An integral field spectrograph (IFS) combines this spatial sampling of an object with its spectral information resolved in different points of the field of view in a single exposure to obtain 2D + λ data. This new technique became available with the rise of large detectors. Indeed, if the 2D + λ information is three-dimensional, the detectors remain bi-dimensional. The technique consists in using an integral field unit (IFU) coupled with a dispersing element to “unfold” the 3D spatio-spectral information onto the 2D surface of a sensor of sufficient size. Different unfolding techniques with dedicated optical components have been deployed in different kinds of IFSs: fibre bundles, lenslet arrays and image slices.

Firstly, Vanderriest (1980) proposed the use of an IFU composed of fibre bundles, eventually coupled with a lenslet array to maximise the flux entering in the cores of the fibres, to conjugate positions in the field of view with a long slit that disperses their spectrum. Mapping Nearby Galaxies at Apache Point Observatory (MaNGA) implements this design (Drory et al. 2015).

Secondly, to improve the spatial resolution and limit the light loss, Courtes (1982) proposed the concept of the Traitement Intégral des Galaxies par l'Étude de leurs Raies (TIGER, Bacon et al. 1995), which uses a grism behind an IFU composed of a lenslet array. Many IFSs are based on this principle, such as for example the Optically Adaptive System for Imaging Spectroscopy (OASIS), the Spectrographic Areal Unit for Research on Optical Nebulae (SAURON, Bacon et al. 2001), the Supernovae Integral Field Spectrograph (SNIFS, Lantz et al. 2004), and the OH-Suppressing Infrared Integral Field Spectrograph (OSIRIS, Larkin et al. 2006).

Thirdly, to go beyond the limited spectral resolution of the lenslet arrays, recent instruments based on image slicing IFU, as proposed by Bowen (1938), were developed such as the Spectrograph for INtegral Field Observations in the Near Infrared (SINFONI, Eisenhauer et al. 2003), the Multi Unit Spectroscopic

Explorer (MUSE, [Bacon et al. 2006](#)), and the Near-Infrared Spectrograph (NIRSpec, [Bagnasco et al. 2007](#)).

These instruments provide a new window onto many astrophysical phenomena, such as planets and comets of the solar system, stellar jets, stellar population dynamics, protostars, galactic X-ray sources, the nuclei of galaxies, active nuclei, interacting galaxies (active galactic nuclei (AGNs) and starbursts), radiogalaxies, quasar environments, gravitational lenses, and so on ([Vanderriest et al. 1994](#)).

With the development of extreme adaptive optics ([Jovanovic et al. 2015](#)), in the last few years we have witnessed the results of a new generation of IFSs, such as those incorporated into the Project 1640 IFS ([Hinkley et al. 2011](#), Hale Telescope), the Gemini Planet Imager (GPI, [Macintosh et al. 2008](#), Gemini South Telescope), the Spectro-Polarimetric High-contrast Exoplanet REsearch instrument (SPHERE, [Beuzit et al. 2019](#), Very Large Telescope; VLT), and the Coronagraphic High Angular Resolution Imaging Spectrograph (CHARIS, [Groff et al. 2016](#), Subaru Telescope). These are based on lenslet arrays that limit the non-common path aberrations compared to slicer concepts. Coupled with a coronagraph, they are dedicated to high-contrast and high-resolution imaging of the surroundings of stars. These IFSs acquire a large number of monochromatic images simultaneously, which are then used to retrieve the spectral features of substellar companions, but also to calibrate and then to reduce the impact of the speckle noise that is induced by the stellar leakages through the coronagraph. Such instruments consequently represent a powerful technique to detect and characterise exoplanets ([Vérinaud et al. 2006](#)) by achieving higher contrasts ([Sparks & Ford 2002](#); [Marois et al. 2014](#)).

A common point of all these IFSs is their global optical scheme: the target object is sampled at specific points in the field of view, each point producing a spectrum that is dispersed onto the sensor. The raw data consequently consist of hundreds to tens of thousands of spectra rearranged compactly on one or several 2D detectors. Dedicated codes are therefore needed to reduce this large amount of 2D data into a spatio-spectral $2D + \lambda$ datacube exploitable for astrophysical analyses.

In this paper we describe PIC (Projection, Interpolation, and Convolution), a reduction technique of the $2D + \lambda$ object of interest based on a forward model of the instrument. The reduction performed by PIC amounts to solving the inverse problem of retrieving the $2D + \lambda$ object given the instrument model and the data. Such an “inverse-problems” approach is fairly general and can readily be adapted to any kind of IFS. In addition, PIC automatically detects damaged pixels and copes with them in a consistent way. Compared to existing methods, PIC produces $2D + \lambda$ reduced data of better quality with almost no artifacts. As shown by our work, this improved quality is also beneficial to post-processing methods.

Section 2 details the motivation behind the proposed method and details the mathematical formalism of the forward model of an IFS instrument as well as the detector model. The inversion uses adequately chosen regularisations and constraints as well as robust penalisation to identify on-the-fly defective pixels missed during the calibration or cosmic rays hitting the detector during the science acquisitions.

In Sect. 3, PIC is applied to SPHERE-IFS data for two typical kinds of target: extended objects (Ganymede and Io, Sect. 3.1) and coronagraphic data (dedicated to the exoplanets and disc detection and characterisation around HR 8799 and RY Lup; Sect. 3.2). Comparisons with the reductions performed with the SPHERE pipeline as well as angular differential imaging (ADI) post-processing on the reduced datacubes (Sect. 3.5) highlight the benefits of the proposed approach.

Appendices A–D focus on the tuning of the method with the available SPHERE calibration files associated to the reduced science data. Special care is taken on the detector characterisation (Appendix A) and on the chromatic model of the spectral dispersion (Appendix B).

2. Proposed method

2.1. Motivation of the inverse-problems approach

Beyond calibration considerations, the main feature of data-reduction algorithms dedicated to IFSs is the mapping of the raw data measured on the 2D sensor to the targeted $2D + \lambda$ hyperspectral cube. Different approaches are implemented depending on the instrument.

A straightforward extraction consists in a direct read of the raw data. Indeed, whatever the optical design of the IFS (fibre bundle, slicer, or lenslet array), the spectra are generally dispersed along an axis of the 2D sensor. After a wavelength calibration step, it is possible to link, via the so-called “pixel association map” (SPHERE, [Pavlov et al. 2008](#); [Mesa et al. 2015](#); GPI, [Perrin et al. 2016](#)), or “pixel table” (SINFONI, [Modigliani et al. 2007](#); MUSE, [Weilbacher et al. 2012](#)), each spectrum on the 2D sensor to a given position in the $2D + \lambda$ hyperspectral cube. Each spectrum is then independently extracted by integrating pixels in small segments perpendicular to the dispersion direction.

While this direct mapping is conceptually easy to grasp and computationally fast, it suffers from several limitations. In particular, pixels flagged as defective in the calibration step are replaced by interpolating the values of their neighbours on the sensor, which introduces spurious correlations in the reconstructed hyperspectral cube. Furthermore, many pixels not detected as defective during calibration are corrupted and lead to aberrant values in the reconstructed cube. In addition, the method does not account for the spectral cross-talk-induced overlapping of neighbouring spectra. And finally, in order to produce a hyperspectral cube, it is necessary to re-interpolate spatial locations on a regular rectangular grid and according to a common spectral sampling. The interpolation operations introduce spatial and spectral correlations due to the linear combination of the spatially and spectrally closest measurements. These spatial and spectral correlations are generally ignored by the downstream processing methods.

For spectrographs with high spectral resolution such as Elodie ([Baranne et al. 1996](#)) and SpeX ([Cushing et al. 2004](#)), the general approach for the extraction of the spectra is the optimal method proposed by [Horne \(1986\)](#). Most of the IFS reduction pipelines (TIGER and OASIS, [Bacon et al. 1995](#), and their successors, SAURON and SNIFS, [Bacon et al. 2001](#); MaNGA, [Law et al. 2016](#); and OSIRIS, [Lockhart et al. 2019](#)), also implement this technique. It is based on a chromatic model of the transverse profile of each spectrum. Combining this line-spread function with the measurement errors to weight the pixels, it solves the first previously mentioned problem by avoiding the interpolation of the defective pixels. It is also possible to use additional weighting coefficients to partially correct the effect of the cross-talk among neighbouring spectra (SPHERE, [Pavlov et al. 2008](#)). Nonetheless, this approach assumes that the pixel variance is known and that the defective pixels are correctly identified. Furthermore, this model of adjacent 1D line-spread functions does not account for the spectral response of the instrument, which is along the direction of the spectral dispersion.

For the GPI instrument, [Draper et al. \(2014\)](#) proposed an improvement of this method, which was further refined by

Brandt et al. (2017) for the CHARIS instrument. The general method comprises a χ^2 -based extraction of the spectra, previously proposed by Zimmerman et al. (2011) for the calibration of the Project 1640 IFS. Each spectrum is described by a super-resolved model combining wavelength-dependent point spread functions (PSFs) found by weighted least-squares fitting with weights given by the inverse variance of each pixel. By setting their weights to zero, this approach can correctly deal with defective pixels. As the individual spectrum model can be computed with a wavelength sampling common to all wavelengths, this method also solves the spectral interpolation issue and performs a deconvolution by the instrument line spread function. Finally, the spectral cross-talk is reduced by running the extraction twice after data cleaning by the model after the first estimation. Nonetheless, by extracting the spectra separately, this method does not address the previously raised problem of spatial interpolations that are still needed to obtain a datacube on a regular grid.

In this work, we implement a method that goes a step further by jointly reconstructing the whole $2D + \lambda$ datacube using a general framework based on an inverse-problems approach to reduce IFS 2D raw data. In this approach, the reconstructed datacube \mathbf{x} is the minimum of a global cost function:

$$\tilde{\mathbf{x}} = \underset{\mathbf{x} \in \mathcal{D}}{\operatorname{argmin}} \mathcal{D}(\mathbf{y}, \mathbf{M}\mathbf{x}, \mathbf{W}) + \mathcal{R}(\mathbf{x}), \quad (1)$$

where $\mathcal{D}(\mathbf{y}, \mathbf{M}\mathbf{x}, \mathbf{W})$ is the data-fidelity term that penalises how “far” the modelled intensity $\mathbf{M}\mathbf{x}$, which is predicted by the linear model of the instrument \mathbf{M} for a given object \mathbf{x} , is from the data \mathbf{y} according to the confidence on the data, which itself is characterised by the precision matrix \mathbf{W} . Further, $\mathcal{R}(\mathbf{x})$ denotes the regularisation term implemented to enforce spatial and spectral priors on the hyperspectral cube \mathbf{x} , and \mathcal{D} is the admissible domain for the spatio-spectral cube \mathbf{x} to reconstruct (e.g. $\mathbf{x} \geq 0$). In terms of notation, the object $\mathbf{x} \in \mathbb{R}^{n_1 \times n_2 \times n_\lambda}$ is a datacube in arbitrary units defined by the value $x_{\theta, \ell}$ at each discrete element of this volume; each element is referred to as a voxel in the following. Each voxel has a 2D angular position of $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \llbracket 1, n_1 \rrbracket \times \llbracket 1, n_2 \rrbracket$ and a wavelength index of $\ell \in \llbracket 1, n_\lambda \rrbracket$. The raw data, denoted by $\mathbf{y} \in \mathbb{R}^{m_1 \times m_2}$, are the intensities in analog-digital units (adu), the same units as given by the detector, sampled at the 2D positions of the sensor pixels y_k indexed by $\mathbf{k} = (k_1, k_2) \in \llbracket 1, m_1 \rrbracket \times \llbracket 1, m_2 \rrbracket$.

2.2. Forward model

The critical point of any inverse-problems approach is the forward model \mathbf{M} that predicts the data \mathbf{y} for a given object \mathbf{x} , or in the present case, the pixelated intensities measured by the sensor and the $2D + \lambda$ hyperspectral cube of the object. Figure 1 is the general optical scheme of an IFS, exemplified using the SPHERE instrument (SPHERE-IFS, Claudi et al. 2008), presented in Fig. 1b. It emphasises the different steps to convert the coloured image of the object presented in Figs. 1a,d to the spectral projection on the 2D sensor of each of its spatial elements, presented in Figs. 1c,g:

- \mathbf{C} : $\mathbb{R}^{n_1 \times n_2 \times n_\lambda} \rightarrow \mathbb{R}^{n_1 \times n_2 \times n_\lambda}$, a convolution operator: the entrance of the instrument is composed of an IFU array that performs a convolution \mathbf{C} of the input signal by the pupil entrance of its sub-elements, as illustrated in Fig. 1e;
- \mathbf{I} : $\mathbb{R}^{n_1 \times n_2 \times n_\lambda} \rightarrow \mathbb{R}^{n_h \times n_\lambda}$, a sampling operator: the integration of the signal by the IFU induces a loss of its spatial information, performing its interpolation \mathbf{I} on the positions of the n_h hexagonal lenslets, as illustrated in Fig. 1f;
- \mathbf{P} : $\mathbb{R}^{n_h \times n_\lambda} \rightarrow \mathbb{R}^{m_1 \times m_2}$, a sparse matrix projection operator: the sampled coloured signal is dispersed over the sensor

onto neighbouring spectra following the IFU array geometry modelled by a projection matrix \mathbf{P} . As the elementary monochromatic PSF of a single lenslet impacts only a portion of the detector, which is as small as possible to reduce inter-spectra cross-talk, this projection matrix is sparse.

In the present work, \mathbf{C} is a convolution by the hexagonal pupil shape of the lenslets and is performed by a simple component-wise multiplication in the Fourier space. Here, \mathbf{I} is a linear interpolation of the Cartesian grid of the hyperspectral cube on the lenslet positions, assumed to be identical to the spectra positions on the sensor given by the dispersion model. Finally, for a given interpolated spectrum sampled at the wavelengths of the reconstructed hyperspectral cube, the projection matrix \mathbf{P} consists in a linear combination of axisymmetric Gaussian patterns whose position and size depend on the considered lenslet and wavelength. Considering a super-resolved model such as the one proposed by Brandt et al. (2017) would be straightforward in our framework but would require further calibration steps that are not available in the European Southern Observatory (ESO) calibration framework of SPHERE, as discussed in Appendix B. Thus, the global optical forward model of an IFS is linear and is synthesised by the following equation, using a compact matrix notation:

$$\mathbf{M}\mathbf{x} = \mathbf{P}\mathbf{I}\mathbf{C}\mathbf{x}. \quad (2)$$

This gives the name of the proposed method, mentioned as PIC in the following. These operators are tuned from calibration exposures in different calibration steps described in Appendix B for the SPHERE instrument.

We would like to mention here that one could use this model to inject artificial planet signals directly in the raw data rather than in the reduced hyperspectral cubes in order to provide an end-to-end characterisation of the IFS performance that includes the reduction and post-processing algorithms.

2.3. Detector model

Beyond the effect due to the optical elements of the IFS, the recorded raw signal \mathbf{y}^{raw} presents additional contributions. Indeed, in practice, each pixel of the detector records only an affine transform of the modelled signal $[\mathbf{M}\mathbf{x}]_k$ corrupted by additive noise:

$$y_k^{\text{raw}} = g_k [\mathbf{M}\mathbf{x}]_k + b_k + \epsilon_k, \quad (3)$$

where $g_k \simeq 1$ is a correction factor accounting for the variability of the quantum efficiency and of the electronic gain, and accounting for the optical throughput such as vignetting or masking by dust particles; b_k is a bias term accounting for a constant offset set by the detector and all the background signals that are proportional to the integration time: the dark current of the sensor, the thermal background of the instrument, and the contribution of the sky; and ϵ_k is a centred additive noise of variance v_k which accounts for the detector read-out noise of variance v_k^0 at null flux, and the photon noise of the incident flux. Assuming that the forward model $\mathbf{M}\mathbf{x}$ correctly predicts the intensity on the sensor in adu, the variance of the photon noise is proportional to the incident flux and the total pixel variance is

$$v_k = \eta_k \left([\mathbf{M}\mathbf{x}]_k + g_k^{-1} (b_k - y_k^b) \right) + v_k^0, \quad (4)$$

where η_k , in adu, accounts for the unit conversion from the number of photons to adu and y_k^b is the bias of the detector (see Eq. (A.1)) at pixel k .

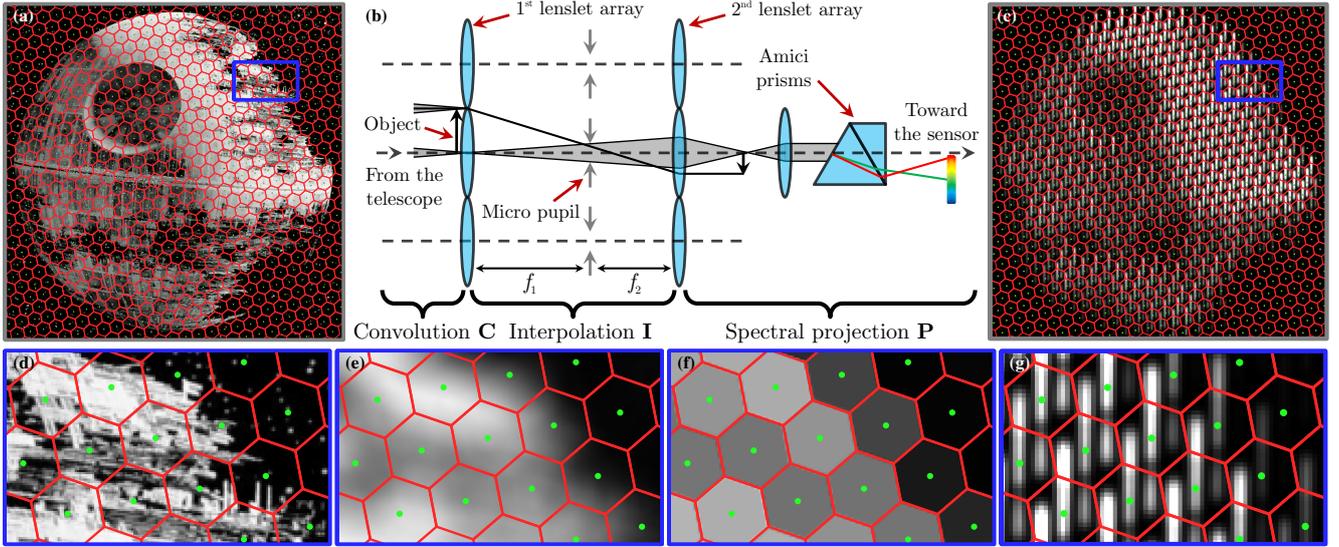


Fig. 1. Forward model $\mathbf{M}\mathbf{x} = \mathbf{P}\mathbf{I}\mathbf{C}\mathbf{x}$ of an integral field spectrograph with the example of the SPHERE-IFS instrument (b). The telescope forms an image of the targeted object \mathbf{x} (a,d) on the instrument that projects a spectrum of each of its spatial elements onto the sensor (c,g). The low-crosstalk BIGRE design of the IFU (b) proposed by Antichi et al. (2009) is based on two hexagonal lenslet arrays whose effect is first modelled as a convolution \mathbf{C} by their pupil shape, averaging the light received by a single lenslet (d→e). An interpolation \mathbf{I} then accounts for the sampling performed by the first array of lenslets of the IFU (red) at the lenslet positions (green dots) (e→f). Finally, their spectrum is dispersed by prisms on the sensor, corresponding to a sparse matrix \mathbf{P} linking the different wavelengths seen by a given lenslet to the location of their projection on the detector (f→g). Photo Credit: Lucasfilm Ltd.

Like any other inverse approach, the proposed method relies on precise modelling of the recorded raw signal \mathbf{y}^{raw} to trustfully retrieve all the information contained in the data. It is therefore mandatory to have good knowledge of the gain \mathbf{g} , the bias \mathbf{b} , and the noise parameters η_k and v_k^0 for each pixel k . These parameters are estimated from calibration exposures in different calibration steps described in Appendix A for the SPHERE instrument.

After the calibration step, to perform the data reduction, the raw data are preprocessed, leading to the calibrated data \mathbf{y} considered in Eq. (1):

$$\mathbf{y}_k = \mathbf{g}_k^{-1} (\mathbf{y}_k^{\text{raw}} - \mathbf{b}_k). \quad (5)$$

2.4. Robust penalisation for the data fidelity

The role of the data fidelity term $\mathcal{D}(\mathbf{y}, \mathbf{M}\mathbf{x}, \mathbf{W})$ in Eq. (1) is to ensure that the retrieved object \mathbf{x} produces via the forward model \mathbf{M} an intensity that is as “close” as possible to the calibrated data \mathbf{y} . In this section, we describe what we mean by “close”.

The precision matrix \mathbf{W} is the inverse of the covariance of the errors due to noise and approximations of the model (see e.g., Mugnier et al. 2004). Assuming uncorrelated noise, this matrix is diagonal and is comprised of coefficients (or weights) \mathbf{w} that are the inverse of the pixel variance: $w_k = v_k^{-1}$. These weights represent the confidence on each measured pixel k . Nonetheless, during the instrument calibration, some defective pixels, such as those that are hot or dead, can be identified and discarded. Their precision is set to $w_k = 0$, that is, their variance is considered infinite. We cope with these discarded pixels by introducing a mask δ defined by:

$$\delta_k \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } y_k \text{ is discarded} \\ 1 & \text{otherwise} \end{cases}, \quad (6)$$

and rewrite the weighting coefficients \mathbf{w} as:

$$w_k = \delta_k v_k^{-1}, \quad (7)$$

where the variance v_k of the k th pixel is given by Eq. (4).

Owing to the high number of photons per pixel in a frame, independent but non-homogeneous Gaussian noise can be assumed. The data fidelity is then defined as the sum of the weighted square residuals between the data and the predicted intensity:

$$\mathcal{D}(\mathbf{y}, \mathbf{M}\mathbf{x}, \mathbf{w}) = \sum_{k=(1,1)}^{m_1, m_2} w_k (y_k - [\mathbf{M}\mathbf{x}]_k)^2. \quad (8)$$

Nonetheless, this method is not robust to outliers. Indeed, some defective pixels that have an erratic behaviour are missed during the sensor calibration. In addition, energetic particles may hit the detector during an exposure and impact several pixels.

To reduce the incidence of these erroneous measurements, robust penalisation approaches have been developed (Hogg 1979; Huber 1996). The general idea is to replace the quadratic penalisation of the residuals of Eq. (8) by a function ρ that is approximately quadratic around zero and that grows sub-quadratically for large values. Adapting Eq. (8) to account for a robust penalisation leads to:

$$\mathcal{D}(\mathbf{y}, \mathbf{M}\mathbf{x}, \mathbf{w}) \stackrel{\text{def}}{=} \sum_{k=(1,1)}^{m_1, m_2} \rho(\sqrt{w_k} s_k^{-1} (y_k - [\mathbf{M}\mathbf{x}]_k)), \quad (9)$$

where s_k is a scaling factor on the residuals that characterises how “far” from a normal Gaussian distribution the whitened residual statistics can be.

With $s_k = 1$ and $\rho(r) = r^2$, Eq. (9) is equivalent to the standard least-squares cost of Eq. (8).

Several robust penalisation functions ρ exist in the literature for which the equivalent weight w_ρ of a least-squares minimisation can be computed (Holland & Welsch 1977). In the following, the Cauchy function is used:

$$\rho(r) \stackrel{\text{def}}{=} \frac{\gamma}{2} \log(1 + r^2/\gamma^2) \Rightarrow w_\rho(r) = (1 + r^2/\gamma^2)^{-1}, \quad (10)$$

with $\gamma = 2.385$.

Such a robust penalisation is pertinent in the case of IFS imaging since according to our forward model, a given voxel $x_{\theta,\ell}$ has an impact on several regions of the detector. It is unlikely that these regions are all polluted by many defective pixels. This redundancy coupled with the robust penalisation avoids most artefacts due to outliers.

2.5. Regularisations and constraints

There is no guarantee that the model \mathbf{M} defined in Sect. 2.2 is a well-conditioned and full-rank matrix. This implies that minimising Eq. (9) may be an ill-posed and ill-conditioned problem leading to noise amplification. Introducing the regularisation function $\mathcal{R}(\mathbf{x})$ in Eq. (1) is a means of coping with these issues by accounting for some prior knowledge about the observed object \mathbf{x} (Titterton 1985). Two kinds of regularisations are implemented in the proposed inversion method.

A first regularisation favours sharp-edged objects to control the spatial continuity and smoothness of the reconstruction. Enforcing such an edge-preserving smoothness is done by encouraging the sparsity of spatial gradients (Rudin et al. 1992; Charbonnier et al. 1997). In addition, it can be expected that the edges of the reconstructed objects are co-localized according to the wavelength. Similarly, we are more likely to have a full spectrum in a given position than fragmented spectra spread on neighbouring positions. This prior about the spectral behaviour is implemented by means of a structured norm (Blomgren & Chan 1998):

$$\mathcal{R}_{2D}(\mathbf{x}) = \sum_{\theta=(1,1)}^{n_1, n_2} \left[\sqrt{\sum_{\ell=1}^{n_\lambda} ([\nabla_{\theta_1} \mathbf{x}]_{\theta,\ell}^2 + [\nabla_{\theta_2} \mathbf{x}]_{\theta,\ell}^2) + \epsilon^2} - \epsilon \right], \quad (11)$$

where ∇_{θ_1} and ∇_{θ_2} correspond to finite difference operators along the first and second spatial dimensions, respectively, and $\epsilon > 0$ is a threshold corresponding to the smallest gradient at a sharp edge and ensures that this hyperbolic approximation of the Euclidean norm is differentiable at zero. Because of the summation over ℓ under the square root, hyperspectral images with edges co-localised at all wavelengths are favoured.

A second regularisation ensures that the reconstructed spectra are smooth. Indeed, with the limited spectral resolution of the instruments, no extremely thin emission or absorption lines are expected to be reconstructed. This also prevents the apparition of the high-frequency noise mentioned by Draper et al. (2014) and Brandt et al. (2017) in the super-resolved reconstructed spectra. We chose to implement this spectral regularisation by the following quadratic penalty:

$$\mathcal{R}_\lambda(\mathbf{x}) = \sum_{(\theta,\ell)=(1,1,1)}^{n_1, n_2, n_\lambda} [\nabla_\lambda \mathbf{x}]_{\theta,\ell}^2, \quad (12)$$

where ∇_λ is the finite difference operator along the spectral dimension.

The two chosen regularisations are combined in the term $\mathcal{R}(\mathbf{x})$ in Eq. (1) with respective hyperparameters μ_{2D} and μ_λ to weight the priors relative to the data fidelity term:

$$\mathcal{R}(\mathbf{x}) = \mu_{2D} \mathcal{R}_{2D}(\mathbf{x}) + \mu_\lambda \mathcal{R}_\lambda(\mathbf{x}). \quad (13)$$

Combined with the robust penalisation, the chosen regularisation helps to identify corrupted pixels that were not yet detected as such. Indeed, to obtain a good fit with an outlier, the estimated hyperspectral cube would necessarily involve a value at a given spatial location and wavelength that significantly differs from its spatial and spectral neighbours. Such a situation

is penalised by the regularisation in order to reject outliers and obtain a more meaningful estimation.

In addition to the ‘‘loose’’ constraints imposed by the regularisation, two ‘‘tight’’ constraints are applied to \mathbf{x} . Firstly, the spectra can only have positive intensities. Secondly, to ensure that the whole spectral extension is accounted for by the forward model, the reconstructed wavelengths go slightly beyond the instrument spectral sensitivity range. Thus, the extreme wavelengths are forced to be null. Combined with the previously defined regularisations, this forces the spectra to smoothly drop to zero at their edges.

In summary, the feasible domain \mathbb{D} is:

$$\mathbb{D} = \{\mathbf{x} \in \mathbb{R}^{n_1 \times n_2 \times n_\lambda} \mid \forall (\theta, \ell), x_{\theta,\ell} \in \mathbb{R}^+, x_{\theta,1} = x_{\theta,n_\lambda} = 0\}. \quad (14)$$

2.6. Reconstruction algorithm

After the calibration of the detector (to get a preliminary list of defective pixels and determine η_k , g_k , and b_k irrespective of k) and of the instrument (to code $\mathbf{M} = \mathbf{PIC}$), the reduction algorithm consists in recovering \mathbf{x} by minimising the objective function defined in Eq. (1) and updating the factors δ_k . This is implemented by the iterative Algorithm 1.

Algorithm 1 Data-reduction algorithm

1:	$\mathbf{x} \leftarrow \mathbf{0}$	▷ Initialisation of the object
2:	$\forall k, y_k \leftarrow_{\text{Eq. (5)}} g_k^{-1} (y_k^{\text{raw}} - b_k)$	▷ Preprocessing
3:	$\mathbf{y}^m \leftarrow \text{medfilter}(\mathbf{y})$	▷ Initialisation of the model
4:	for i from 1 to n_{it} do	▷ n_{it} optimisations
5:	$\forall k, \delta_k \leftarrow \delta_k^{\text{cal}}$	▷ Initialisation of δ in \mathbf{w}
6:	if $i \geq 3$ and $w_\rho(r_k)_{\text{Eq. (10)}} < w_{th}$ then	▷ Equivalent weight
7:	$\delta_k \leftarrow 0$	▷ δ update in \mathbf{w}
8:	$\forall k, w_k \leftarrow_{\text{Eq. (7)}} \delta_k (\eta_k (y_k^m + g_k^{-1} b_k) + v_k^0)^{-1}$	▷ \mathbf{w} update
9:	$\mathbf{x} \leftarrow_{\text{Eq. (1)}} \underset{\mathbf{x} \in \mathbb{D}}{\text{argmin}} \mathcal{D}(\mathbf{y}, \mathbf{M} \mathbf{x}, \mathbf{w}) + \mathcal{R}(\mathbf{x})$	▷ \mathbf{x} update, \mathbf{w} fixed
10:	$\mathbf{y}^m \leftarrow_{\text{Eq. (2)}} \mathbf{M} \mathbf{x}$	▷ Update of the model
11:	$\forall k, r_k \leftarrow_{\text{Eq. (9)}} \sqrt{w_k} s_k^{-1} (y_k - y_k^m)$	▷ Weighted residuals
12:	return \mathbf{x}	

Strictly, looking at Eq. (7), the weights \mathbf{w} in Eq. (9) depend on the current value of \mathbf{x} during the optimisation of the cost function Eq. (1), increasing its complexity and raising potential convergence issues. To overcome this limitation, the reconstruction of \mathbf{x} is subdivided into a loop of n_{it} optimisations. At a given iteration of the loop, the weighting coefficients are assumed to be constant, computed via Eq. (7) based on the forward model given by the solution of the previous iteration $\mathbf{y}^m = \mathbf{M} \mathbf{x}$.

At the initialisation, for the first estimation of the weighting coefficients \mathbf{w} , the predicted model is approximated by a median filter of the preprocessed data \mathbf{y} to diminish the influence of the defective pixels. If some of them have been detected during the calibration, the map δ is initialised with the flagged pixels δ^{cal} . Otherwise, it is set to $\delta_k = 1$, for all k .

At each loop after the second loop $i \geq 3$, the map of discarded pixels δ^{cal} is updated using a threshold w_{th} on the equivalent weight w_ρ given by the robust penalisation on the residuals of the previous loop.

3. Results

SPHERE (Beuzit et al. 2019) is a second-generation VLT instrument optimised for high-contrast imaging. Its IFS is based on a new optical concept proposed by Antichi et al. (2009) to

overcome the limitations affecting a 3D spectrograph working in diffraction-limited regime (Claudi et al. 2011). BIGRE is an evolution of the classical TIGER design, optimised for seeing-limited conditions in the SAURON instrument (Bacon et al. 1995, 2001). The micro-lenses composing the lenslet array of the IFU are smaller than the full width at half maximum (FWHM) of the projected PSF of the telescope, and therefore provide resolved images of the PSF and the speckle patterns produced by the stellar leakage in the coronagraphic mode of the instrument. To limit the cross-talk between the lenslets, the BIGRE design improves the TIGER concept by adding a second lenslet array as schemed in Fig. 1b.

The BIGRE SPHERE-IFS provides low-resolution spectroscopy ($R \sim 50$) in the near-infrared (0.95–1.7 microns), an ideal spectral range for the detection of exoplanets and the characterisation of their features. Its IFU is composed of 145×145 lenslets and covers a field of view of about 1.8×1.8 square arcseconds (Claudi et al. 2008; Dohlen et al. 2016). The spatial dimension of the raw data taken on the Hawaii-2RG sensor is equal to $m_1 \times m_2 = 2048 \times 2048$ pixels.

This section presents the results of the proposed PIC method on representative datasets of the SPHERE-IFS instrument compared to the reduction obtained with the SPHERE reduction pipeline (Delorme et al. 2017) used to date. Two situations are considered: the direct imaging of extended objects, and the coronagraphic mode of SPHERE dedicated to exoplanet detection and characterisation. The output cubes of the SPHERE reduction pipeline are composed of $n_1 \times n_2 \times n_\lambda = 291 \times 291 \times 39$ voxels. The proposed reduction pipeline provides cubes of $n_1 \times n_2 \times n_\lambda = 300 \times 300 \times n_\lambda$ voxels with $n_\lambda = 44$ for the YH-mode and $n_\lambda = 47$ for the YJ-mode, which represents a ratio of approximately eight between the pitch of the angular position θ of the lenslets in the reduced cube and the pitch of their position k on the sensor in terms of pixel. An important component of our method is the calibration of the parameters of the forward model. We describe the methodology to estimate those parameters in Appendices A and B.

The code is implemented in MatlabTM and uses the open-source GlobalBioIm framework (Soubies et al. 2019). The optimisation problem (1) is solved by the VMLM-B algorithm (Thiébaud 2002), a limited-memory quasi-Newton method with BFGS updates (Nocedal 1980) that handles bound constraints. The version of the code used in this paper can be found at <https://doi.org/10.5281/zenodo.3585632>. All the data and reductions presented in this papers can be downloaded at <https://doi.org/10.5281/zenodo.3585656>.

3.1. Extended-object imaging

Even if SPHERE-IFS is mainly used in its coronagraphic mode and the exoplanet search, it is also suitable for the observation of extended objects (see e.g., Hanus et al. 2017; King et al. 2019). To highlight the pertinence of the proposed method for the imaging of such targets, we reduced two Jovian moons with the SPHERE reduction pipeline (Delorme et al. 2017) and the proposed PIC method: Ganymede (60.A-9372 – 1 Feb. 2015, 298th out of 300 acquisitions) and Io (60.A-9357 – 6 Dec. 2014, 2nd out of 8 acquisitions), both acquired in YH mode. The results are given in Figs. 2 and 3.

For each reduced cube, spectra are extracted along three positions θ of interest. For comparison, the transmission t_{atm} of the Earth’s atmosphere, computed via the software “HITRAN on the Web” proposed by Rothman et al. (2013), is given in green along with the extracted spectra.

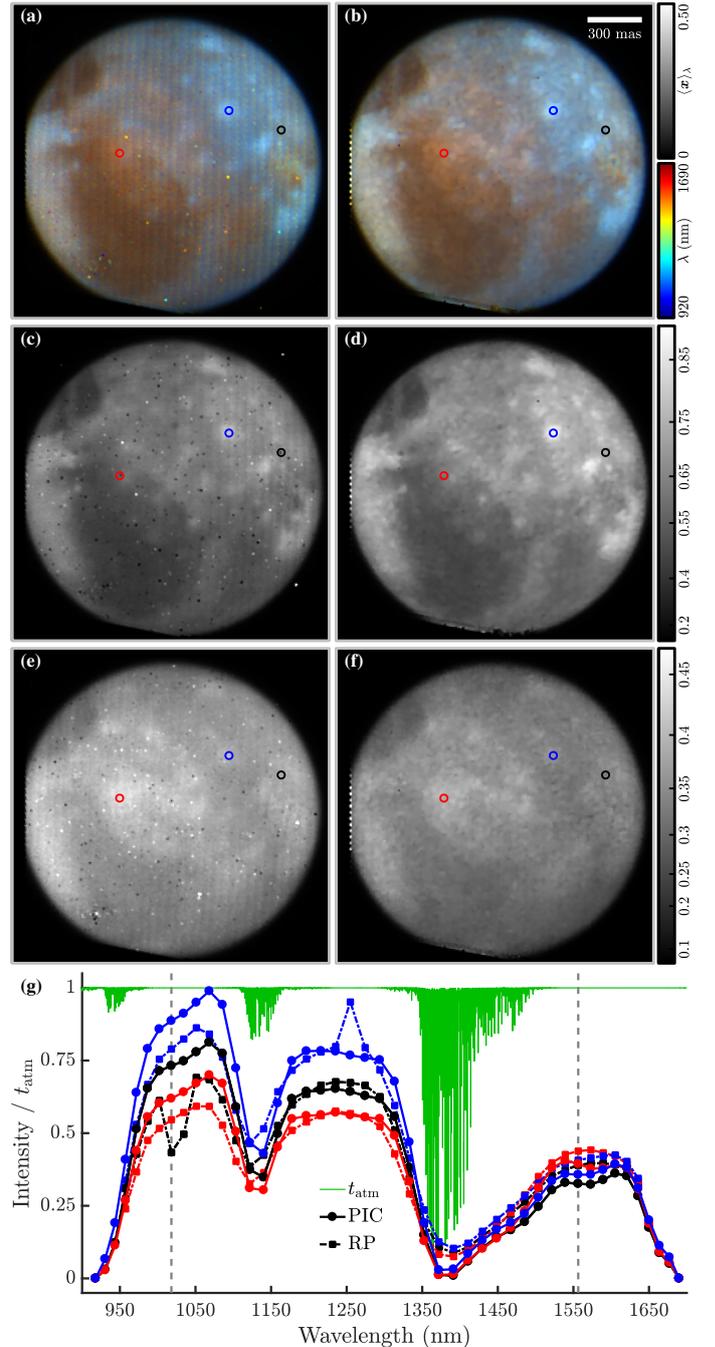


Fig. 2. Comparison between the SPHERE reduction pipeline (RP, Delorme et al. 2017) (a,c,e) and the proposed method (PIC) (b,d,f) on the Jovian moon Ganymede (square stretch of the intensity, colour bar: arbitrary unit). a,b: coloured projection $\langle x \rangle_\lambda$ of the reconstructed hyperspectral cube (square stretch of the intensity). c–f: slices of the reconstructed hyperspectral cube (at $\lambda = 1018$ nm and $\lambda = 1556$ nm). g: spectra extracted from the reduced hyperspectral cubes at the circled positions. The two vertical dashed lines indicate the wavelengths of the monochromatic images shown in (c–f).

Several remarks can be made about the colour coding in Figs. 2a,b and 3a,b. The colour intensity reflects the spectral average $\langle x \rangle_\lambda$, the colour hue indicates the wavelengths that deviate from this projection, and the colour saturation represents the magnitude of this deviation.

The colour variations at the Ganymede’s surface shown by Figs. 2a,b suggest that the reconstructed cubes contain informa-

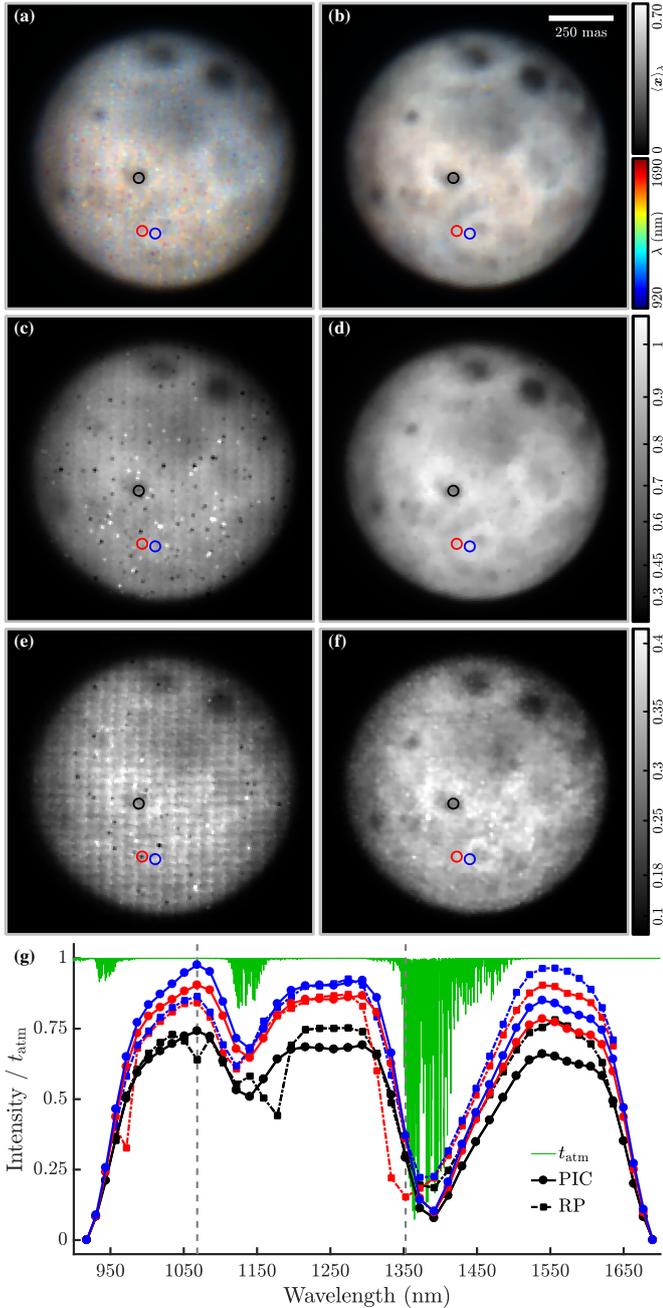


Fig. 3. Comparison between the SPHERE reduction pipeline (RP, Delorme et al. 2017) (a,c,e) and the proposed method (PIC) (b,d,f) on the Jovian moon Io (square stretch of the intensity, colour bar: arbitrary unit). a,b: coloured projection $\langle x \rangle_\lambda$ of the reconstructed hyperspectral cube (square stretch of the intensity). c–f: slices of the reconstructed hyperspectral cube (at $\lambda = 1068$ nm and $\lambda = 1353$ nm). g: spectra extracted from the reduced hyperspectral cubes at the circled positions. The two vertical dashed lines indicate the wavelengths of the monochromatic images shown in (c–f).

tion of spatial variations of the surface spectrum. While smooth regions are favoured, the proposed regularisation does not prevent spatial structures with small contrasts and spectral variations from being restored.

This spectral diversity can be seen in the extracted spectra in Fig. 2g. The black spectrum is close to the average spectrum of the reduced cube. The red (resp. blue) spectrum is extracted in an area where the spectrum is attenuated (resp. enhanced) in the shorter wavelengths and enhanced (resp. attenuated) at longer

wavelengths. This diversity is also visible on the two extracted slices at $\lambda = 1018$ nm in Figs. 2c,d and $\lambda = 1556$ nm in Figs. 2e,f and which show very different spatial structures. These spectral contrasts are less pronounced on Io’s surface.

In terms of reconstruction quality, the cubes reduced by the SPHERE reduction pipeline exhibit vertical artefacts that are visible in the spectral projections in Figs. 2a and 3a and on the slices in Figs. 2c and 3c. At low fluxes, horizontal artefacts are also visible in the Io cube in Fig. 3e.

In addition to these large-scale structures, the reduced cubes produced by the SPHERE pipeline are corrupted by many artefacts due to defective pixels, most of them leading to dark voxels, as visible in the slices in Figs. 2c,e and 3c,e. These artefacts leave a strong distortion in the extracted spectra, as visible in the black spectra in Figs. 2g and 3g. A defective pixel on the sensor plane impacts several wavelengths in the cube reduction. This can also lead to misinterpretations, as in the case of the red spectrum in Fig. 3g where the absorption bands of the Earth’s atmosphere are shifted from their actual positions, for instance, at ~ 1100 nm for the blue spectrum and at ~ 1350 nm for the red spectrum. With the proposed calibration and regularised reconstruction, these artefacts are corrected in the reduced cubes as visible in Figs. 2a,c,e and 3a,c,e. The striped pattern has disappeared and no strong artefacts are visible. Moreover, qualitatively, the spatial regularisation reduces the noise in the reconstruction and improves the contrast of the geological formations, even at very low flux such as in Figs. 3e,f.

3.2. Coronagraphic imaging

SPHERE-IFS is mainly used for high-contrast observation with a coronagraph. To directly image an extrasolar planet, the light from its host star must be reduced by several orders of magnitude. Associated with dedicated detection algorithms, SPHERE as well as other instruments of its generation such as GPI and CHARIS, allow the spectral characterisation of the imaged exoplanets at low resolution ($R \sim 30\text{--}50$) and achieve typically planet-to-star contrast ratios down to 10^{-6} at $0.2''$ angular separation (see e.g., Currie et al. 2015; Macintosh et al. 2015; Ruffio et al. 2018; Mesa et al. 2019).

To illustrate the present method when SPHERE-IFS operates in its coronagraphic mode, we considered two datasets: one on the HR 8799 star that hosts a cortege of exoplanets (Marois et al. 2008) and that was observed in YH mode (095.C-0298 – 4 Jul. 2015, 38th out of 56 acquisitions), and one on the RY lup star that hosts a bright transition disc (Langlois et al. 2018) and was observed in YJ mode (097.C-0865 – 15 Apr. 2016, 50th out of 80 acquisitions). A comparison between the reduced cubes with the SPHERE reduction pipeline is given in Figs. 4 and 5.

The reduced HR 8799 cube is dominated by the speckles produced by the stellar leakage. The coloured spectral projection $\langle x \rangle_\lambda$ emphasises that this is a chromatic effect with the speckle structure scaling with the wavelength from the shorter wavelengths represented in blue to the longer wavelengths represented in red. Without further processing, no exoplanet is visible. On the contrary, the signal of the reduced RY lup cube is dominated by the signal of the accretion disc that is visible at all wavelengths as a well-contrasted structure. Contrary to the speckles, the location of an object imaged by the IFS does not scale with the wavelength and thus appears in white.

Concerning the reduction quality, similar conclusions for the reduction of extended objects can be drawn. The HR 8799 cube presents the same striped artefacts that are corrected by the proposed method. The important noise of the RY lup cube due to

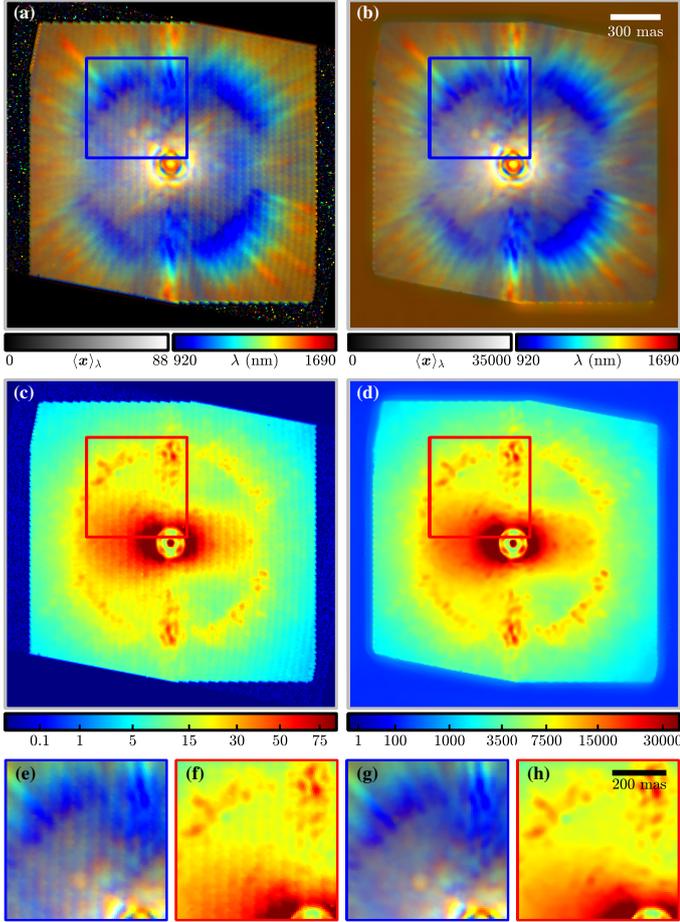


Fig. 4. Comparison between the SPHERE reduction pipeline (Delorme et al. 2017) (a,c,e,g) and the proposed PIC method (b,d,f,h) on the HR 8799 dataset from a single IFS exposure. The proposed method removes artificial background structures. a,b: coloured projection of the reconstructed hyperspectral cube (cubic root stretch of the intensity). c,d: slice of the reconstructed hyperspectral cube (at $\lambda = 1068$ nm, cubic root stretch of the intensity, colour bar: arbitrary unit). e–h: zooms on (a–d).

the small exposure time is smoothed by the regularisation of the proposed method.

The zooms of Figs. 5a,c given in Figs. 5e,f highlight the necessity to correct the “remnant signal”, as presented in Fig. A.2. The subtraction of this unwanted signal, a ghost of a previous high dynamic acquisition on the sensor, in the sky background correction step produces the artificial dark spot visible in the cube reduced by the standard method. With the proposed method, this region is correctly reconstructed.

These same zooms show the effectiveness of the robust penalisation in discarding defective pixels or pixels hit by energetic particles based on the spatial and spectral regularisations of the reconstructed cube. Indeed, as visible in Figs. 6a,e, such a particle leaves a long trace on the sensor. Looking at the equivalent weights w_p in Figs. 6c,g, these aberrant measures are clearly identified as detector pixels with very low statistical weights below the threshold w_{th} . These can be discarded by the reduction procedure. They are identified in blue in Figs. 6d,h. In addition, defective pixels missed during the calibration are also flagged in red among the already discarded pixels in green.

The comparison between the data in Figs. 6a,e and the intensity predicted by the model in Figs. 6b,f show that the coupling of the redundancy of the hyperspectral information on an IFS

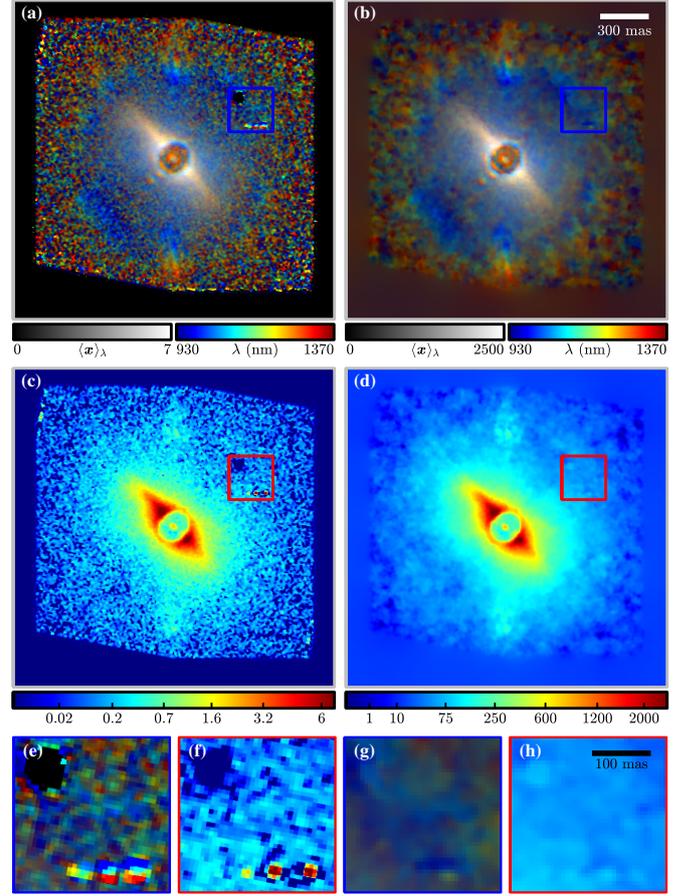


Fig. 5. Comparison between the SPHERE reduction pipeline (Delorme et al. 2017) (a,c,e,g) and the proposed PIC method (b,d,f,h) on the RY lup dataset from a single IFS exposure. The proposed sky background estimation removes the host star remnant and the proposed reconstruction algorithm is robust to aberrant measurements (see Fig. 6). a,b: coloured projection of the reconstructed hyperspectral cube (cubic root stretch of the intensity). c,d: slice of the reconstructed hyperspectral cube (at $\lambda = 1165$ nm, cubic root stretch of the intensity, colour bar: arbitrary unit). e–h: zooms on (a–d).

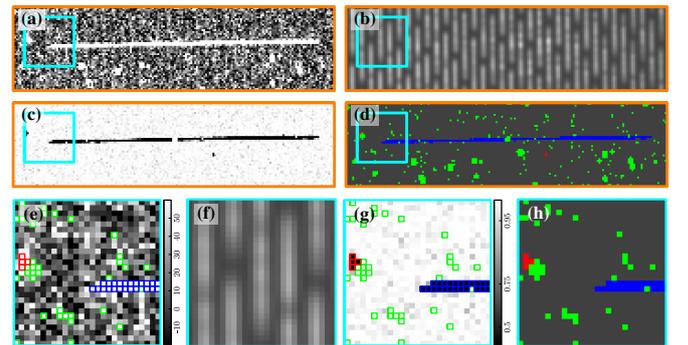


Fig. 6. Robust defective-pixel identification on the sensor region corresponding to the region of interest of Figs. 5 and 7b. In addition to the known defective pixels (green), the robust penalisation can flag additional unusable pixels in a given acquisition that are blinking (red) or hit by an energetic particle (blue). a,b: raw data (a) and intensity (b) predicted by the model (colour bar: adu). c: equivalent weights obtained by the robust penalisation (cubic stretch). d: new map of the unconsidered pixels. e–h: zooms on (a–d).

field with adequate regularisation allows the estimation of the signal even under high levels of noise.

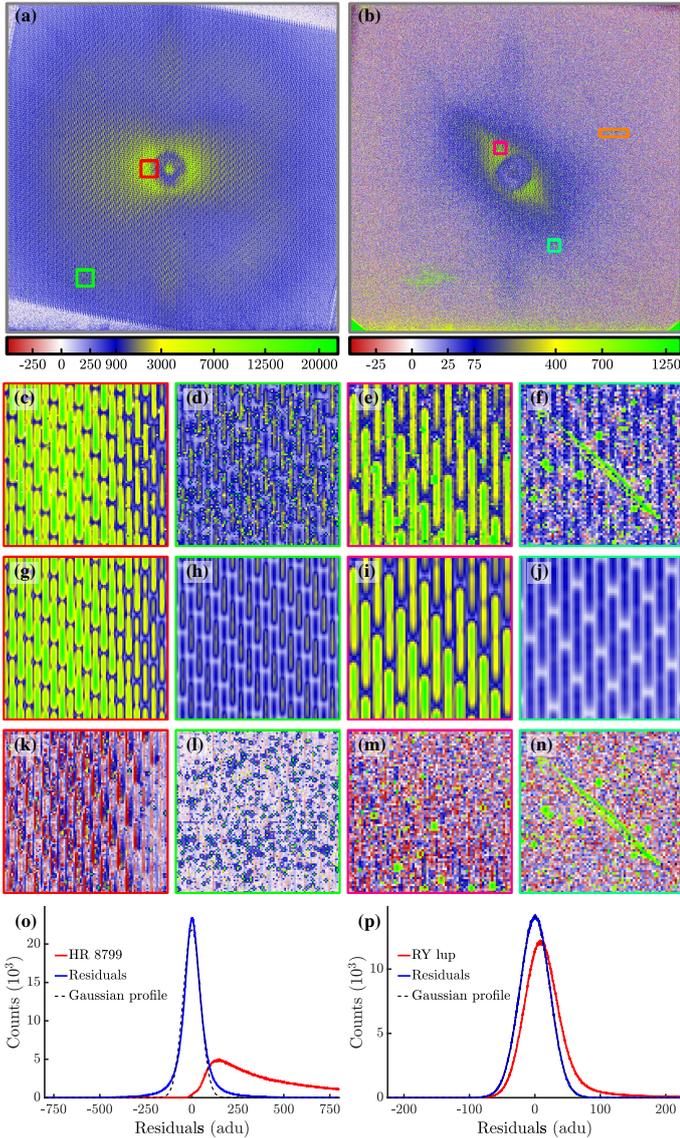


Fig. 7. Residuals of the model on the HR 8799 (on the *left column*) and the RY Lup (on the *right column*) datasets (square root stretch, colour bar: adu). *a,b*: raw data. *c-f*: zooms on the regions framed in (*a,b*). *g-j*: model prediction on the same regions of interest. *k-n*: residuals on the same regions of interest. *o,p*: statistics of the residuals. For comparison, Gaussian profiles of standard deviation $\sigma = 47$ adu (HR 8799) and $\sigma = 23$ adu (RY Lup) are plotted. The region of interest framed in orange in (*b*) is the region detailed in Fig 6.

3.3. Statistics of the residuals

The reconstructed model compared to the raw data is further detailed in Fig. 7 for the two datasets. The zooms in Figs. 7d,h,l,f,j,n correspond to regions of interest previously identified in the calibration because of their high density of defective pixels (Fig. A.4d) or because of the presence of a dust particle (Fig. A.4c).

Despite the numerous discarded pixels in the raw data flagged in Figs. 7d,f, the algorithm predicts a model, as shown in Figs. 7h,j, where no structure is visible in the residuals displayed in Figs. 7l,n, except for the defective pixels and the pixels masked by a dust particle.

The zooms in Figs. 7c,g,k,e,i,m focus on regions closer to the coronagraph mask where the flux is the highest. No structure is visible in the residuals of Fig. 7m. On the contrary, on the limit of

the coronagraph mask in Fig. 7k, vertical blue and red structures can still be discerned. This is because of a slight lateral shift of the spectra in the raw data compared to the positions predicted by the calibrated forward model. This phenomenon is also visible on the reduction of the star PSF acquisitions (not presented here). In these two cases, the object inducing these shifts is a star, that is, an unresolved and coherent object inducing stronger coherent cross-talk effects that are not totally suppressed by the BIGRE design, as described by Antichi et al. (2009). These effects cannot be estimated by the calibration acquisitions or accounted for by the forward model which considers only incoherent cross-talk.

Figures 7o,p compare the histogram of the raw data (red), the histogram of the residuals between the data and the model prediction (blue), and a Gaussian profile (dashed black). The statistics of the raw data are correctly matched by the statistics of the model prediction. The residuals have approximately Gaussian distributions.

3.4. Image registration and pixel scale calibration

In coronagraphic mode, the centre of the field of view is given by the position of the masked star. In order to determine the star position behind the coronagraph's mask, specific acquisitions are obtained by applying a periodic phase pattern on the deformable mirror of the instrument. This induces four first-order diffraction patterns of the masked star as presented in Fig. 8b. The star is at the centre of the square formed by these four satellite images, as presented by the black dots and dashed lines on the figure.

With the SPHERE reduction pipeline, the centre of the field of view is determined after the reduction of all the acquisitions. Once the centre is known, all the reduced cubes are re-interpolated to centre the star on a given pixel. With the proposed method, this shift between the centre of the field and the centre of the IFS can be taken into account by the forward model. This is done by adapting the sampling positions of the interpolation operator \mathbf{I} in the cube \mathbf{x} . Doing so, the reduced cube is by construction centred on the host star. This avoids performing any further interpolation of the reduced data that would imply a slight loss in spatial resolution and stronger noise correlations due to the re-interpolation operation.

In the proposed method, the centre of the field of view is determined by the centres of the square diffraction patterns fit at each wavelength of the cube. To be robust to outliers, we first take the centroid of all the square centres and then iteratively take the centroid of the subset of square centres that are at a distance from the previous centroid of less than 0.4 voxels (≈ 2.6 mas). The results are shown in Fig. 8d for the two datasets and for two different acquisitions in the same night. Most of the centres are within 0.25 voxels (≈ 1.7 mas) of the estimated centre. No noticeable difference can be seen in the estimated parameters according to the dataset or the acquisition date.

Being a diffraction-driven phenomenon, these square patterns should scale with the wavelength. This can be seen by the four radial rainbow-coloured features on the coloured projection $\langle \mathbf{x} \rangle_\lambda$ in Fig. 8a. To study this scaling law, the square best-fitting the four spots in the reconstructed cube is estimated at each wavelength. The slope ξ of the linear law that best fits the diagonal s according to the wavelength $s \approx \xi \lambda$ is given in Fig. 8c for the two datasets (blue and red) for each acquisition date in the observation night (squares and circles). Table 1 summarises the fitted parameters for the two datasets and their two acquisition dates. The standard deviation of the chromatic slope is $0.02 \text{ voxel } \mu\text{m}^{-1}$ for all datasets. This linear law is used

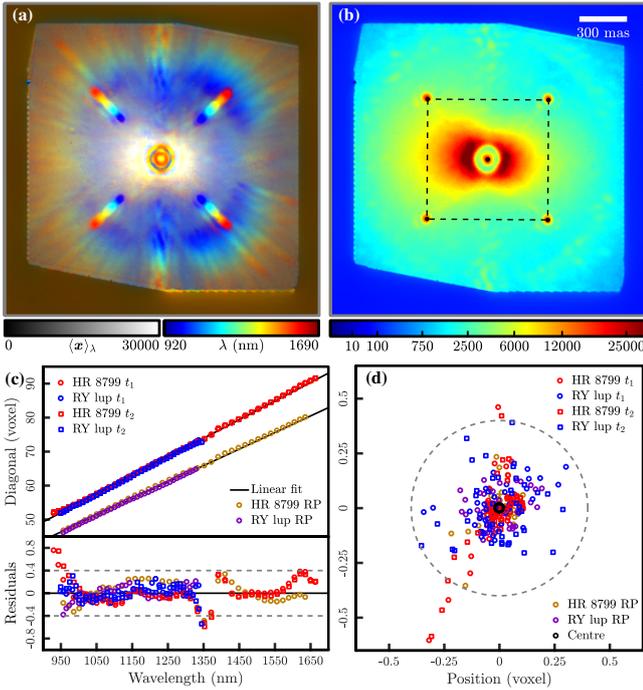


Fig. 8. Study of the chromatic scaling of the diffraction images of the host star for the two datasets. *a*: coloured projection of the HR 8799 reconstructed hyperspectral cube (cubic root stretch of the intensity). *b*: slice of the HR 8799 reconstructed hyperspectral cube ($\lambda = 1486$ nm, cubic root stretch of the intensity, colour bar: arbitrary unit). The five black dots represent the fitted positions of the four images of the star and of the centre of the fitted square (dashed lines). *c*: linear fits of the square diagonals according to the wavelength for each dataset at two different times in the night with the proposed PIC method and, for the first time in the night, with the SPHERE reduction pipeline (RP, Delorme et al. 2017). *d*: fitted positions of the squares at each wavelength compared to the estimation of the star position. Only the wavelengths whose position falls inside the grey circle of 0.4 voxels radius are used to fit the linear laws.

Table 1. Estimation of the chromatic slope ξ of the linear law linking the diagonal s of the diffracted square pattern with the wavelength λ and the corresponding relative error on the spectral calibration σ_λ for the different datasets and their acquisition dates with the proposed method and, for the first acquisition date, with the SPHERE reduction pipeline (RP, Delorme et al. 2017).

Dataset	HR 8799			RY lup		
	t_1	t_2	RP	t_1	t_2	RP
voxel μm^{-1}	54.83	54.95	49.00	54.70	54.56	48.57
σ_λ (%)	0.25	0.26	0.33	0.16	0.22	0.25

to estimate a relative error on the spectral calibration given by the sample standard deviation σ_λ of $\left(\frac{s_\ell}{\xi \lambda_\ell} - 1\right)$ computed on all spectral channels ℓ .

As for the centre determination, no noticeable difference can be seen in the estimated parameters according to the dataset or the acquisition dates.

The differences in the chromatic slope are due to the choice of smaller pixels in the reconstructions in the proposed method compared to the SPHERE reduction pipeline. We note here that this is not a measure of the spatial resolution of the reduced datacubes that would be different with the SPHERE reduction pipeline or

PIC. Indeed, the SPHERE-IFS instrument resolution is fixed and is given by the pitch of the lenslets in the array. With the chosen reduction parameters, there are roughly two voxels per lenslet.

The small standard deviations on the estimated chromatic slope (≈ 0.02 voxel μm^{-1} for all datasets) reflect the good spectral calibration of both PIC and the SPHERE pipeline. The residuals in Fig. 8c were used to estimate the relative spectral calibration errors given in Table 1 that are well under 1%. This supports the quality of the wavelength calibration procedure.

Nevertheless, the residuals in Fig. 8c worsen around $\lambda = 1.4 \mu\text{m}$. This corresponds to the main atmosphere absorption band in Figs. 2g and 3g. As a consequence, the signal-to-noise ratio to estimate the positions of the four diffraction patterns is worse and degrades the square estimation.

On the spectral range edges, in addition to the limited reconstructed flux, the wavelengths are extrapolated beyond the calibration laser wavelengths, as mentioned in the description of the model calibration in Appendix B. This may explain why the residuals are also worse at short and large wavelengths.

3.5. Post-processing examples

As seen above, residual starlight is present in the form of speckles which limit the detection of faint point-source objects. Such residuals can be suppressed by adopting different strategies based for example on the polarisation of the residual speckle pattern (differential polarimetric imaging; DPI) or assuming that it scales with the wavelength (spectral differential imaging; SDI) or that it is stable in time (angular differential imaging; ADI). These techniques are used in dedicated post-processing algorithms such as cADI (Marois et al. 2006), LOCI (Lafreniere et al. 2007), KLIP (Soummer et al. 2012), ANDROMEDA (Cantalloube et al. 2015), or PACO (Flasseur et al. 2018), which are used to detect the faint signal of the exoplanets hidden in this speckle noise.

We chose to run a Template Locally Optimized Combination of Images (TLOCI) detection algorithm (Marois et al. 2013, 2014) on the hyperspectral cubes of the 46 best dates (among the 56 acquisitions) reduced with the SPHERE reduction pipeline and the proposed method.

TLOCI is an evolved version of the LOCI algorithm that is based on ADI where the intrinsic field-of-view rotation during the acquisitions is used to separate the residual light from the star masked by the coronagraph and the light coming from off-axis sources, that is, the objects of interest, for each wavelength of the reduced hyperspectral cube. Further, TLOCI uses a template for the spectrum of the sought-after point sources in the hyperspectral cube to combine images along the spectral dimension.

Figure 9 gives the comparison between the results obtained with the SPHERE reduction pipeline and PIC. The global detection maps presented in Figs. 9e,f combine the S/N maps obtained at all wavelengths as well as the S/N derived for the detection of each planet as given in Table 2. If σ^+ is the list of the n_λ positive S/N of a planet or a position in the detection map, the global S/N is computed as follows:

$$\tilde{\sigma} = \sqrt{\langle \sigma^{+2} \rangle_\lambda}, \text{ with } \begin{cases} \sigma_\lambda^+ = \sigma_\lambda & \text{if } \sigma_\lambda \geq 0 \\ \sigma_\lambda^+ = 0 & \text{if } \sigma_\lambda < 0 \end{cases}, \quad (15)$$

where $\langle \cdot \rangle_\lambda$ denotes spectral averaging.

At short wavelengths, as in Figs. 9a,b, the planet fluxes are too low to be seen. The suppression of the vertically striped artefacts with the proposed method leads to a better speckle removal by the TLOCI algorithm. This is especially striking close to the

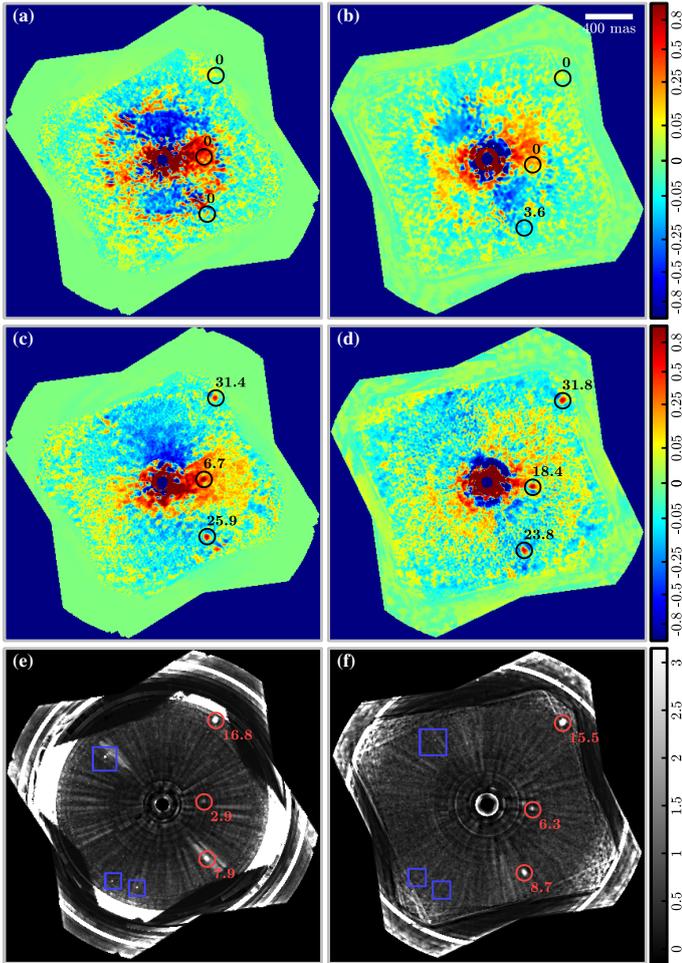


Fig. 9. Comparison between a TLOCI post-processing of the SPHERE reduction pipeline (Delorme et al. 2017) (left column) and the proposed PIC method (right column) on the HR 8799 dataset. The contrast of the planet closest to the star is increased by the PIC method and several false positives are avoided. *a–d*: monochromatic images of the post-processed hyperspectral cube and planet detection S/N (square root stretch of the intensity, colour bar: arbitrary unit) at $\lambda = 972$ nm (*a,b*) and $\lambda = 1636$ nm (*c,d*). *e,f*: maps of planet detection S/N combining all wavelengths (red circles: true positives, blue boxes: false positives with the SPHERE pipeline, colour bar: S/N).

Table 2. Comparison between the S/N of the innermost planet at large wavelengths for the SPHERE reduction pipeline (RP, Delorme et al. 2017) and the proposed method (PIC).

λ (nm)	RP	PIC	λ (nm)	RP	PIC
1556	4.4	<3	1605	5.8	12.1
1573	4.9	10.5	1621	6.4	15.0
1589	5.5	11.1	1636	6.7	18.4

host star at longer wavelengths, as in Figs. 9c,d where the S/N of the innermost planet is strongly increased. This S/N gain can be observed at all wavelengths (except the shortest); see Table 2. In addition, the better reduction of the datacubes reduces the artefacts of post-processing algorithms such as TLOCI, removing potential false detections (in the blue squares) and enlarging the usable area of the field of view.

The reductions on the RY lup dataset are compared by averaging the 80 temporal acquisitions after rotating them to

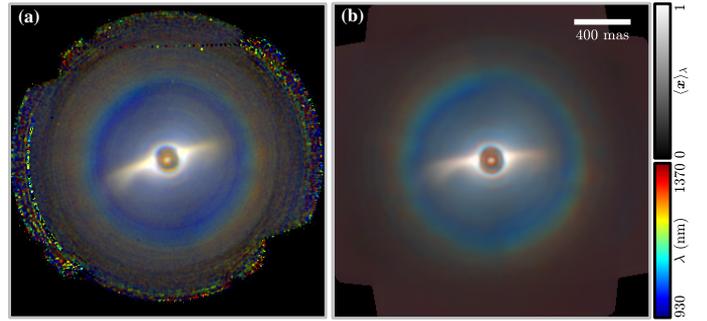


Fig. 10. Coloured projection of the post-processing of the SPHERE reduction pipeline (Delorme et al. 2017) (*a*) and of the proposed PIC method (*b*) on the RY lup dataset (cubic root stretch of the intensity).

compensate for the time-varying field rotation. These averaged reduced data are presented in Fig. 10. As a consequence of the imposed angular correction, the speckle pattern rotates and is averaged out while the accretion disc remains fixed and is enhanced. The coloured projections show that the disc is globally white without notable spectral variations. As previously noted, the edge of the field of view is better reconstructed with the proposed method (absence of outliers).

4. Discussion and conclusions

Here, we propose a new algorithm named PIC designed to reduce the raw data of an IFS in order to produce a 2D + λ cube. Based on an inverse problem approach implementing a robust penalisation, the proposed method directly reconstructs the full hyperspectral cube.

Our method amounts to solving an inverse problem which provides a number of advantages compared to previous methods. The direct object extraction avoids a posteriori interpolation of the extracted spectra on a Cartesian spatial grid or a common wavelength sampling. Modelled by the forward model, the spectral cross-talk between neighbouring spectra is reduced and the spectral response of the instrument is deconvolved in the reduction. This framework also provides a natural and consistent way to deal with corrupted measurements (e.g. defective pixels or cosmic particles) so as to avoid artefacts in the result.

Contrary to other methods, the proposed algorithm explicitly introduces a spatio-spectral regularisation. Our results demonstrate that this regularisation improves the quality of the reduced data by limiting the incidence of the noise and of outlier pixels which are automatically detected during the reduction.

In addition to this new reduction algorithm, we also describe an improved calibration procedure for the SPHERE-IFS instrument. A model of the statistics of the noise in the raw data provided by the sensor is obtained from the calibrations files. Defective pixels are also flagged to be discarded in the reduction procedure. Polynomial laws for the positions and the dispersion are robustly fitted for each lenslet, providing a model that is as close as possible to the instrument physics. An autocalibration step is added to correct for possible shifts between the calibration and the science acquisitions.

The reconstructed spectra, in particular from the calibration observations presented in Appendix B, show a slight improvement of the spectral resolution due to the spectral deconvolution performed by the proposed data-reduction method. The results on extended objects show that such a regularised method and careful calibration enhances the spatial contrast while reducing

the impact of defective pixels along the spectral dimension. Some artefacts visible in some SPHERE reductions in the form of large-scale structures (vertical bands) are also suppressed.

Applied to coronagraphic acquisitions, the method provides reconstructed hyperspectral cubes of improved quality and with fewer artefacts, which in turn lead to detection maps with better separation between sources (like exoplanets) and background speckle. In the quest to detect very faint exoplanets by direct imaging, the proposed reconstruction may contribute to improving the achievable contrast.

The forward model could be further refined while keeping the same general structure in order to account for non-Gaussian microlense PSFs. Another refinement would be to account for the small spatial shifts of the spectra in the area that surrounds the coronagraph mask. However, this would, require a more complex calibration procedure, with a step where a refinement is computed directly on the science data. Finally, the method should be tested and certified routinely on a large amount of data.

This work can be considered as a preliminary building block from which more accurate and more complete reconstruction algorithms can be built. Currently, only the IFS instrument is accounted for in the forward model. The next step is to include the telescope response via its chromatic PSF. This would allow an end-to-end deconvolution of the reduced cube directly from the raw data. Our model could also be included in detection or imaging algorithms to separate the stellar leakages from the off-axis objects in high-contrast imaging.

Acknowledgements. This work was performed within the framework of the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program “Investissements d’Avenir” (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). It has been supported by the (project DIAGHOLO). This work has made use of the SPHERE Data Centre, jointly operated by OSUG/IPAG (Grenoble), PYTHEAS/LAM/CeSAM (Marseille), OCA/Lagrange (Nice), Observatoire de Paris/LESIA (Paris), and Observatoire de Lyon/CRAL, and supported by a grant from Labex OSUG@2020, within the program “Investissements d’Avenir” (ANR-10-LABX-56). All the reductions are based on public data provided by the ESO Science Archive Facility and acquired for different programs: 60.A-9372 (Fletcher L., 1 Feb. 2015), 60.A-9357 (Carter J., 6 Dec. 2014), 095.C-0298 (Beuzit J.-L., et al., 4 Jul. 2015) and 097.C-0865 (Beuzit J.-L., et al., 15 Apr. 2016). The authors would like to thank Gratton R., Dohlen K., and Delorme P. for their fruitful discussions on the IFS principle and on the SPHERE reduction pipeline. We are grateful to Carter J. and Guiot P. for their end user’s view on the reduction of Io and Ganymede observations as well as to Vigan A. for having shared the data of Fig. C.2e. We are also thankful to the referee and its pertinent feedback.

References

- Antichi, J., Dohlen, K., Gratton, R. G., et al. 2009, *ApJ*, **695**, 1042
- Bacon, R., Adam, G., Baranne, A., et al. 1995, *A&AS*, **113**, 347
- Bacon, R., Copin, Y., Monnet, G., et al. 2001, *MNRAS*, **326**, 23
- Bacon, R., Bauer, S., Boehm, P., et al. 2006, *Ground-based and Airborne Instrumentation for Astronomy*, 6269, 62690J
- Bagnasco, G., Kolm, M., Ferruit, P., et al. 2007, *Cryogenic Optical Systems and Instruments XII*, 6692, 66920M
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, *A&AS*, **119**, 373
- Beuzit, J. L., Vigan, A., Mouillet, D., et al. 2019, *A&A*, **631**, A155
- Blomgren, P., & Chan, T. 1998, *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, **7**, 304
- Bowen, I. S. 1938, *ApJ*, **88**, 113
- Brandt, T. D., Rizzo, M. J., Groff, T. D., et al. 2017, *J. Astron. Telesc. Instrum. Syst.*, **3**, 048002
- Cantaloube, F., Mouillet, D., Mugnier, L. M., et al. 2015, *A&A*, **582**, A89
- Charbonnier, P., Blanc-Feraud, L., Aubert, G., & Barlaud, M. 1997, *Trans. Image Process.*, **6**, 298
- Claudi, R. U., Turatto, M., Gratton, R. G., et al. 2008, in *Ground-based and Airborne Instrumentation for Astronomy II*, Proc. SPIE, 7014, 70143E
- Claudi, R. U., Giro, E., Anselmi, U., et al. 2011, in *Optical Design and Engineering IV*, Proc. SPIE, 8167, 81671S
- Courtes, G. 1982, *Int. Astron. Union Colloq.*, **67**, 123
- Currie, T., Cloutier, R., Brittain, S., et al. 2015, *ApJ*, **814**, L27
- Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, *PASP*, **116**, 362
- Delorme, P., Meunier, N., Albert, D., et al. 2017, in *SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, eds. C. Reylé, P. Di Matteo, F. Herpin, et al., 347
- Dohlen, K., Vigan, A., Mouillet, D., et al. 2016, *Ground-based and Airborne Instrumentation for Astronomy VI*, 9908, 99083D
- Draper, Z. H., Marois, C., Wolff, S., et al. 2014, *Ground-based and Airborne Instrumentation for Astronomy V*, 9147, 91474Z
- Drory, N., MacDonald, N., Bershad, M. A., et al. 2015, *AJ*, **149**, 77
- Eisenhauer, F., Abuter, R., Bickert, K., et al. 2003, *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*, 4841
- Flasseur, O., Denis, L., Thiébaud, É., & Langlois, M. 2018, *A&A*, **618**, A138
- Groff, T. D., Chilcote, J., Kasdin, N. J., et al. 2016, *Ground-based and Airborne Instrumentation for Astronomy VI*, 9908, 99080O
- Hanus, J., Marchis, F., Viikinkoski, M., Yang, B., & Kaasalainen, M. 2017, *A&A*, **599**, A36
- Hinkley, S., Oppenheimer, B. R., Zimmerman, N., et al. 2011, *PASP*, **123**, 74
- Hogg, R. V. 1979, *Am. Stat.*, **33**, 108
- Holland, P. W., & Welsch, R. E. 1977, *Commun. Stat. Theory Methods*, **6**, 813
- Horne, K. 1986, *PASP*, **98**, 609
- Huber, P. 1996, *Robust Statistical Procedures: Second Edition*, CBMS-NSF Regional Conference Series in Applied Mathematics (Society for Industrial and Applied Mathematics)
- Huber, P. J. 2011, *Robust Statistics* (Springer)
- Jovanovic, N., Martinache, F., Guyon, O., et al. 2015, *PASP*, **127**, 890
- King, O. R., Fletcher, L. N., & Ligier, N. 2019, *Compositional Mapping of Europa with VLT/SPHERE*, EPSC-DPS2019-596-1, 13
- Lafreniere, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, E. 2007, *ApJ*, **660**, 770
- Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. 1998, *J. Optim.*, **9**, 112
- Langlois, M., Pohl, A., Lagrange, A.-M., et al. 2018, *A&A*, **614**, A88
- Lantz, B., Aldering, G., Antilogus, P., et al. 2004, *Optical Design and Engineering*, 5249
- Larkin, J., Barczys, M., Krabbe, A., et al. 2006, *Ground-based and Airborne Instrumentation for Astronomy*, 6269, 62691A
- Law, D. R., Cherinka, B., Yan, R., et al. 2016, *AJ*, **152**, 83
- Lockhart, K. E., Do, T., Larkin, J. E., et al. 2019, *AJ*, **157**, 75
- Macintosh, B. A., Graham, J. R., Palmer, D. W., et al. 2008, *Adaptive Optics Systems*, 7015, 701518
- Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Science*, **350**, 64
- Marois, C., Lafreniere, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, *ApJ*, **641**, 556
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Science*, **322**, 1348
- Marois, C., Correia, C., V´eran, J.-P., & Currie, T. 2013, *Proc. Int. Astron. Union*, **8**, 48
- Marois, C., Correia, C., Galicher, R., et al. 2014, in *Adaptive Optics Systems IV*, Int. Soc. Opt. Photon., 9148, 91480U
- Mesa, D., Gratton, R., Zurlo, A., et al. 2015, *A&A*, **576**, A121
- Mesa, D., Langlois, M., Garufi, A., et al. 2019, *MNRAS*, **488**, 37
- Modigliani, A., Hummel, W., Abuter, R., et al. 2007, ArXiv e-prints [arXiv:astro-ph/0701297]
- Mugnier, L. M., Fusco, T., & Conan, J.-M. 2004, *J. Opt. Soc. Am. A*, **21**, 1841
- Nocedal, J. 1980, *Math. Comput.*, **35**, 773
- Pavlov, A., Moller-Nilsson, O., Feldt, M., et al. 2008, in *Advanced Software and Control for Astronomy II*, Proc. SPIE, 7019, 701939
- Perrin, M. D., Ingraham, P., Follette, K. B., et al. 2016, in *Ground-based and Airborne Instrumentation for Astronomy VI*, Proc. SPIE, 9908, 990837
- Rothman, L., Gordon, I., Babikov, Y., et al. 2013, *J. Quant. Spectr. Rad. Transf.*, **130**, 4 hTRAN2012 special issue
- Rudin, L. I., Osher, S., & Fatemi, E. 1992, *J. Phys. D*, **60**, 259
- Ruffio, J.-B., Mawet, D., Czekala, I., et al. 2018, *AJ*, **156**, 196
- Soubies, E., Soulez, F., McCann, M. T., et al. 2019, *Inverse Probl.*, **35**, 104006
- Soummer, R., Pueyo, L., & Larkin, J. 2012, *ApJ*, **755**, L28
- Sparks, W. B., & Ford, H. C. 2002, *ApJ*, **578**, 543
- Thi´ebaut, E. 2002, *Proc. SPIE*, 4847, 174
- Titterton, D. M. 1985, *A&A*, **144**, 381
- Vanderriest, C. 1980, *PASP*, **92**, 858
- Vanderriest, C., Bacon, R., Georgelin, Y., LeCoarer, E., & Monnet, G. J. 1994, *Instrumentation in Astronomy VIII*, 2198
- V´erinaud, C., Hubin, N., Kasper, M., et al. 2006, *Proc. SPIE*, 6272, 62720M
- Vigan, A., Gry, C., Salter, G., et al. 2015, *MNRAS*, **454**, 129
- Weilbacher, P. M., Streicher, O., Urrutia, T., et al. 2012, *Software and Cyberinfrastructure for Astronomy II*, 8451, 84510B
- Zimmerman, N., Brenner, D., Oppenheimer, B. R., et al. 2011, *PASP*, **123**, 746

Appendix A: Detector calibration

Each SPHERE-IFS scientific acquisition comes with a set of calibration files described in Pavlov et al. (2008). These files are split into two subsets: one dedicated to the calibration of the sensor and the other dedicated to the calibration of the lenslet array, that is, the spectral dispersion. In each configuration, a set of n_a acquisitions are performed for various exposure times $\tau \in \mathbb{R}^n$. The calibration files that we use in our calibration of the sensor are the following:

- detector flat-field acquisitions \mathbf{y}^f performed with an integrating sphere producing a uniform light in the optical system after the lenslet array for different light sources (lasers or broadband sources);
- background acquisitions \mathbf{y}^{bg} performed without any flux of interest reaching the sensor. The recorded signal is then mainly due to the thermal background. These measurements are performed in three different configurations: (a) $\mathbf{y}^{\text{bg}/c} \in \mathbb{R}^{m_1 \times m_2 \times n_i \times n_a}$ with part of the instrument optical path closed (i.e. same optical path as for the flat-field acquisitions, without the lenslet array); (b) $\mathbf{y}^{\text{bg}/o} \in \mathbb{R}^{m_1 \times m_2 \times n_i \times n_a}$ with opened optical path (i.e. complete optical path of the instrument from the entrance focus, including the lenslet array); (c) $\mathbf{y}^{\text{bg}/s} \in \mathbb{R}^{m_1 \times m_2 \times n_i \times n_a}$ with the shutter of the instrument opened, the telescope pointing at a dark region of the sky to record the thermal emission of the atmosphere.

In the following, we describe how these different images are combined in order to estimate the parameters that characterise the detector. We first introduce notations for generic operations on images:

- $\langle \mathbf{a} \rangle_i$ is the empirical mean of \mathbf{a} along the dimension indexed by the letter i ,
- $\mathcal{V}_i(\mathbf{a})$ is the empirical variance of \mathbf{a} along the dimension indexed by the letter i ,
- $\mathcal{M}_i(\mathbf{a})$ is the median of \mathbf{a} along the dimension indexed by the letter i ,
- $\mathcal{A}_i(\mathbf{a}) = \mathcal{M}_i(|\mathbf{a} - \mathcal{M}_i(\mathbf{a})|)$ is the median absolute deviation of \mathbf{a} along the dimension indexed by the letter i . It is often used as a robust estimator of the standard deviation, see Huber (2011).

The calibration procedures are illustrated using the two datasets HR 8799 and RY lup introduced in the main text.

A.1. Bias calibration

The bias term b_k in Eq. (3) is the sum of different contributions:

$$\mathbf{b} = \mathbf{y}^b + \mathbf{y}^{\text{dc}} + \mathbf{y}^{\text{th}} + \mathbf{y}^{\text{sky}}, \quad (\text{A.1})$$

with \mathbf{y}^b the bias of the detector, \mathbf{y}^{dc} the dark current, \mathbf{y}^{th} the thermal emission of the instrument and \mathbf{y}^{sky} the thermal emission of the sky. All these contributions but the detector bias \mathbf{y}^b are proportional to the exposure time. Hence the bias in the raw data shall follow an affine law of the exposure time. The coefficients of this law can be estimated from the background calibration data given by:

$$y_{k,a,t}^{\text{bg}} = y_k^b + f_k^{\text{bg}} \tau_t + \epsilon_{k,a,t}, \quad (\text{A.2})$$

where f_k^{bg} is the background flux in adu s^{-1} accounting for the dark current and the thermal emission, τ_t is the exposure time and $\epsilon_{k,a,t}$ is the noise supposed to be independent.

Due to the high level of instrumental thermal emission, the term $f_k^{\text{bg}} \tau_t$ is predominant even at short exposure times and the constant bias is negligible. In the following, we therefore assume

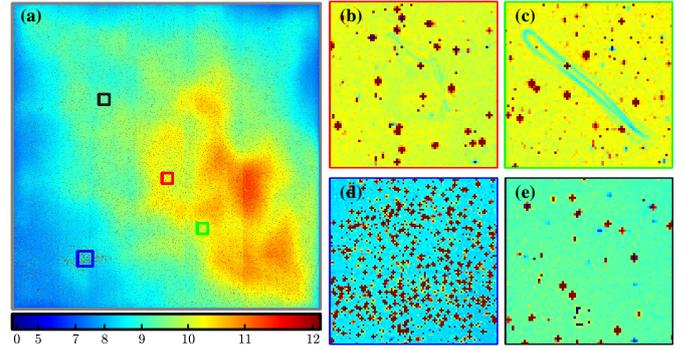


Fig. A.1. Estimation of the temporal background flux $\tilde{f}_k^{\text{bg}/o}$ on the RY lup dataset, with the shutter opened (cubic stretch, colour bar: adu s^{-1}). *b–e*: regions of interest of (a).

that: $\mathbf{y}^b = \mathbf{0}$. The estimated flux \tilde{f}_k^{bg} at a given pixel k is then obtained by a simple linear fit of the averaged background calibration data $\mathbf{q}^{\text{bg}} = \langle \mathbf{y}^{\text{bg}} \rangle_a \in \mathbb{R}^{m_1 \times m_2 \times n_i}$ weighted by their empirical precision $w_{k,t}^{\text{bg}} = [\mathcal{V}_a(\mathbf{y}^{\text{bg}})]_{k,t}^{-1} (\forall k, t)$:

$$\tilde{f}_k^{\text{bg}} = \underset{f}{\text{argmin}} \sum_{t=1}^{n_t} w_{k,t}^{\text{bg}} (q_{k,t}^{\text{bg}} - f \tau_t)^2 = \frac{\sum_{t=1}^{n_t} q_{k,t}^{\text{bg}} w_{k,t}^{\text{bg}} \tau_t}{\sum_{t=1}^{n_t} \tau_t^2 w_{k,t}^{\text{bg}}}. \quad (\text{A.3})$$

Finally, in Eq. (5), the term b_k is estimated by $\tilde{b}_k = \tau^{\text{exp}} \tilde{f}_k^{\text{bg}/s}$ or by $\tilde{b}_k = \tau^{\text{exp}} \tilde{f}_k^{\text{bg}/o}$ with τ^{exp} the exposure time and depending on whether the sky contribution must be removed or not.

Figure A.1 shows the estimated flux $\tilde{f}_k^{\text{bg}/o}$ for the RY lup dataset with the shutter opened. It is very similar to the estimated flux with the shutter closed (not presented here), meaning that the thermal background is mainly due to the internal components of the instrument. The figure also highlights some regions of the sensor where some dust masks the thermal emission, such as Figs. A.1b,c, and other regions with a high density of anomalous pixels, such as Fig. A.1d.

A.2. Remnant signal

During a night-time calibration procedure, it happens that the sky background $\mathbf{y}^{\text{bg}/s}$ is polluted by a remnant signal from a previous acquisitions \mathbf{f}^{pa} , as seen in Figs. A.2a,b. This is typically the case in high contrast imaging mode when the target star is decented in order to measure the off-axis PSFs and the spectral energy distribution of the star, just before a sky background is acquired, as shown in Fig. A.2c on RY lup.

As several sky background acquisitions are performed, the remnant signal flux \tilde{f}^{rs} decreases while the background flux \tilde{f}^{bg} remains constant. These two contributions are iteratively split under the following assumptions:

- $f^{\text{pa}} \approx \tilde{f}^{\text{bg}} + \tilde{f}^{\text{rs}}$,
- $\forall a \in \llbracket 1, n_a \rrbracket, f_a^{\text{bg}/s} \approx \tilde{f}^{\text{bg}} + \beta_a \tilde{f}^{\text{rs}}$,
- $\forall a \in \llbracket 1, n_a - 1 \rrbracket, \beta_a \geq \beta_{a+1} \geq 0$,
- $\forall k, \tilde{f}_k^{\text{bg}} \geq 0$,

where the \approx sign indicates that some discrepancy is expected because of noise.

Algorithm 2 implements a simple iterative method to estimate the background \tilde{f}^{bg} with a reduced contribution of the remnant signal \tilde{f}^{rs} . Figures A.2d,e show that this algorithm is effective and that the two contributions can be correctly identified. The remnant artefact is strongly attenuated in the estimated

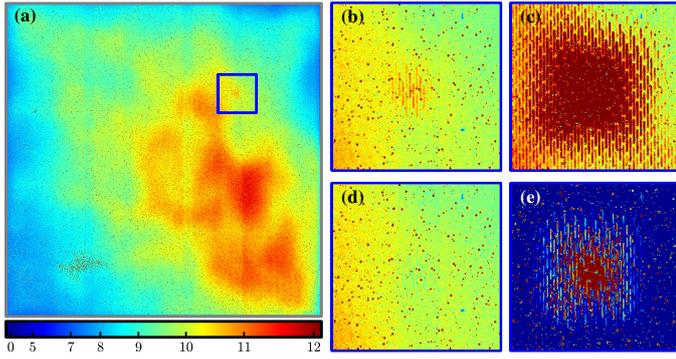


Fig. A.2. Illustration of the remnant signal suppression from the sky background acquisitions (RY lup dataset, colour bar: adu s^{-1}). *a, b*: averaged sky background fluxes $\langle f^{\text{bg}/s} \rangle_a$. *c*: previous acquisition flux f^{pa} . *d*: corrected sky background flux \tilde{f}^{bg} . *e*: estimated remnant signal \tilde{f}^{rs} in (c).

background flux \tilde{f}^{bg} while the remnant signal contribution \tilde{f}^{rs} is not polluted by an offset due to the background.

Algorithm 2 Remnant-free background flux \tilde{f}^{bg} estimate

- 1: $\tilde{f}^{\text{bg}} \leftarrow \langle f^{\text{bg}/s} \rangle_a$ ▷ Initial background estimate
 - 2: **for** i from 1 to n_{rs} **do**
 - 3: $\tilde{f}^{\text{rs}} \leftarrow f^{\text{pa}} - \tilde{f}^{\text{bg}}$ ▷ Update foreground estimate
 - 4: **for** a from 1 to n_a **do** ▷ Loop over calibration backgrounds
 - 5: $\beta_a \leftarrow \frac{(f_a^{\text{bg}/s} - \tilde{f}^{\text{bg}})^{\text{T}} \cdot \tilde{f}^{\text{rs}}}{\|\tilde{f}^{\text{rs}}\|_2^2}$ ▷ Least-squares estimate of β_a
 - 6: $\beta_a \leftarrow \begin{cases} \max(\min(\beta_a, \beta_{a-1}), 0) & \text{if } a > 1 \\ \max(\beta_a, 0) & \text{else} \end{cases}$ ▷ Apply constraints
 - 7: $\forall (k, a), r_{k,a} \leftarrow f_{k,a}^{\text{bg}/s} - \beta_a \tilde{f}_k^{\text{rs}}$ ▷ Subtract remnant signal
 - 8: $\tilde{f}^{\text{bg}} \leftarrow \mathcal{M}_a(r)$ ▷ Update background estimate
 - 9: $\forall k, \tilde{f}_k^{\text{bg}} \leftarrow \max(\tilde{f}_k^{\text{bg}}, 0)$ ▷ Apply constraints
 - 10: **return** \tilde{f}^{bg}
-

A.3. Gain calibration

The factor g_k in Eq. (3) is a gain converting the signal of interest $[\mathbf{M}\mathbf{x}]_k$ at pixel k into digital values. This gain accounts for the quantum efficiency, for the optical throughput and for the electronic gain. The pixelwise gain is estimated from flat-field acquisitions \mathbf{y}^f obtained under a flux that is made as uniform as possible over the whole sensor. These acquisitions must first be pre-corrected from the background corresponding to the situation where the shutter is closed:

$$\tilde{y}_{k,a,t}^f = y_{k,a,t}^f - \tau_t \tilde{f}_k^{\text{bg}/c}. \quad (\text{A.4})$$

To account for the fact that the flat-field calibrations were acquired under different illuminations, the estimated gain \tilde{g} is then given by fitting the averaged (background compensated) flat-field calibrations $\mathbf{q}^f = \langle \tilde{\mathbf{y}}^f \rangle_a \in \mathbb{R}^{m_1 \times m_2 \times n_t}$ as a linear function of their median intensity $\mathbf{p}^f = \mathcal{M}_k(\mathbf{q}^f) \in \mathbb{R}^{n_t}$ weighted by their empirical precisions $w_{k,t}^f = [\mathcal{V}_a(\tilde{\mathbf{y}}^f)]_{k,t}^{-1} (\forall k, t)$:

$$\tilde{g}_k = \underset{g}{\operatorname{argmin}} \sum_{t=1}^{n_t} w_{k,t}^f (q_{k,t}^f - g p_t^f)^2 = \frac{\sum_{t=1}^{n_t} q_{k,t}^f w_{k,t}^f p_t^f}{\sum_{t=1}^{n_t} (p_t^f)^2 w_{k,t}^f}. \quad (\text{A.5})$$

Figure A.3 shows the gain \tilde{g} calibrated on the RY lup dataset. Contributions from the optics and from the electronics can be

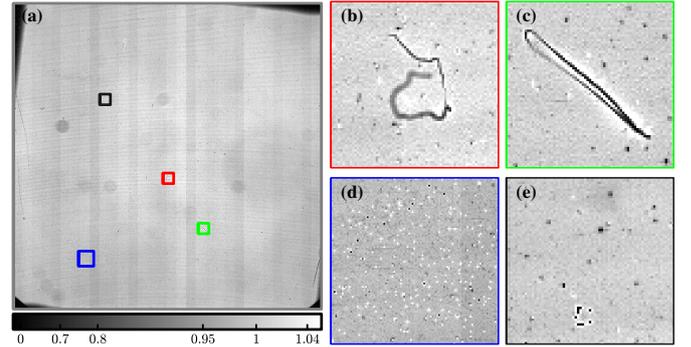


Fig. A.3. Estimation of the flat-field correction factor \tilde{g} on the RY lup dataset. *b–e*: regions of interest of (a).

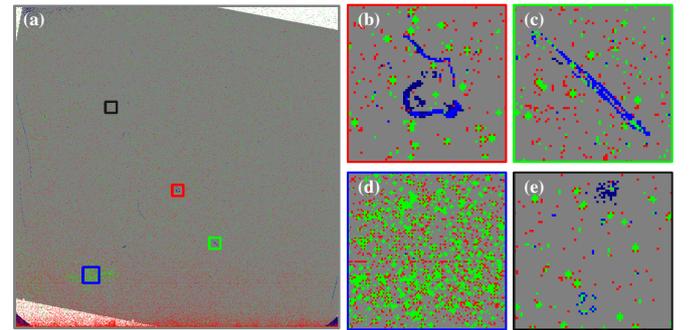


Fig. A.4. Identified defective pixels in the RY lup dataset. White: Pixels out of the lenslet array field. Green: pixels flagged by the median filter on the dark acquisitions. Red: pixels flagged by their anomalous variance in the dark acquisitions. Blue: pixels flagged by the median filter on the flat-field acquisitions. Dark blue: pixels flagged by their anomalous variance in the flat-field acquisitions. Grey: unflagged pixels. *b–e*: regions of interest of (a).

identified. The vertical stripes are due to the gain differences between the 32 reading channels of the sensor. The curved horizontal lines are probably artefacts due to parallel plates on the optical path. Large darker areas are attributable to dust particles on the optics. The dust particles directly laying on the sensor and already visible in the calibrated background flux, such as in Figs. A.1b,c, are now clearly visible, as shown in the zooms of Figs. A.3b,c.

A.4. Defective pixels identification

The different calibration files can be used to form a first estimate of the map of defective pixels, noted δ^{cal} in Algorithm 1. Figure A.4 represents the discarded pixels for the RY lup dataset as well as the reason of their flagging.

The spectral flat-field, which is discussed in more detail in Appendix B and in Fig. C.1a, is first used to identify the active area of the sensor (the grey area of Fig. A.4). Outside of this area the pixels are discarded. In addition we discard also the first four and the last four columns and rows of the sensor. This active area corresponds to the grey square region in the figure.

The defective pixels are then identified by analysing the statistics of the different calibration datasets $\mathbf{y}^c \in \mathbb{R}^{m_1 \times m_2 \times n_a \times n_t}$ and their model $y_{k,a,t}^c \approx p_k^c u_t^c (\forall a)$ with a pixel-wise map $\mathbf{p}^c \in \mathbb{R}^{m_1 \times m_2}$ and exposure-wise factors $\mathbf{u}^c \in \mathbb{R}^{n_t}$ respectively given by:

$$(\mathbf{y}^c, \mathbf{p}^c, \mathbf{u}^c) \in \left\{ (\mathbf{y}^{\text{bg}/o}, \tilde{f}^{\text{bg}/o}, \tau^{\text{bg}/o}), (\mathbf{y}^{\text{bg}/c}, \tilde{f}^{\text{bg}/c}, \tau^{\text{bg}/c}), (\mathbf{y}^f, \tilde{g}, \mathbf{p}^f) \right\},$$

where we recall that $\mathbf{p}^f = \mathcal{M}_k(\langle \tilde{\mathbf{y}}^f \rangle_a)$. For each of these three sets, a map of defective pixels is built as presented in Algorithm 3.

Algorithm 3 Defective pixel identification on a set $(\mathbf{y}^c, \mathbf{p}^c, \mathbf{u}^c)$

- 1: $\tilde{\mathbf{p}}^c \leftarrow \mathbf{p}^c - \text{medfilter}(\mathbf{p}^c)$ ▷ Difference with a 5×5 median filter
 - 2: $\forall \mathbf{k}, \delta_k^a \leftarrow \begin{cases} 0 & \text{if } |\tilde{p}_k^c| > \alpha^a \mathcal{A}_k(\tilde{\mathbf{p}}^c) \\ 1 & \text{else} \end{cases}$ ▷ Threshold on the tolerance
 - 3: $n_w \leftarrow \sum_{\mathbf{k}=(1,1)}^{m_1, m_2} (1 - \delta_k^a)$ ▷ Count of the discarded pixels
 - 4: $\forall (\mathbf{k}, t), r_{k,t} \leftarrow [\langle \mathbf{y}^c \rangle_a]_{k,t} - \mathbf{p}_k^c \mathbf{u}_t^c$ ▷ Residuals of the linear fit
 - 5: $\forall \mathbf{k}, c_k \leftarrow \sum_{t=1}^{n_t} r_{k,t}^2 / [\mathcal{V}_a(\mathbf{y}^c)]_{k,t}$ ▷ Weighted cost of the residuals
 - 6: $\forall \mathbf{k}, c_k \leftarrow c_k \delta_k^a$ ▷ Discard already flagged pixels
 - 7: $\forall \mathbf{k}, \delta_k^v \leftarrow 1$ ▷ Init. of the map on the statistics of \mathbf{y}^c
 - 8: **for** n_w worst c_k **do** ▷ Discarding n_w worst statistics
 - 9: $\delta_k^v \leftarrow 0$ ▷ δ^v update
 - 10: $\forall \mathbf{k}, \delta_k^{\text{cal}} \leftarrow \delta_k^a \delta_k^v$ ▷ Global map on v
 - 11: **return** δ^{cal}
-

A first set δ^a of n_w defective pixels is built by forming a residual map $\tilde{\mathbf{p}}^c$ by subtracting to \mathbf{p}^c its local median in a 5×5 pixels neighbourhood and identifying as defective the pixels whose absolute value in $\tilde{\mathbf{p}}^c$ is greater than a tolerance α^a multiplied by the median absolute deviation $\mathcal{A}_k(\tilde{\mathbf{p}}^c)$. In Fig. A.4, such defective pixels are marked in green if they are found via the background calibration files (shutter opened or closed) and in blue if they are found via the detector flat-field acquisitions.

In a second step, the statistics of the residuals of the linear fit \mathbf{p}^c in the averaged acquisitions $\langle \mathbf{y}^c \rangle_a$, according to \mathbf{u}^c weighted by their empirical variance $\mathcal{V}_a(\mathbf{y}^c)$, are analysed. The n_w pixels with the worst costs (n_w corresponds to the number of defective pixels previously detected), that is to say, an anomalous variance, are discarded among the remaining pixels not yet flagged as defective. This gives a second map δ^v of defective pixels. In Fig. A.4, such defective pixels are marked in red if they are found via the background calibration files (shutter opened or closed) and in dark blue, if they are found via the detector flat-field acquisitions.

The maps δ^a and δ^v built for the three sets are then combined to obtain the final map of discarded pixels δ^{cal} . This method clearly identifies the pixels impacted by the presence of dust particles laying on the sensor, as shown in blue the zooms of Figs. A.4b,c.

As one would expect, looking at the red-flagged pixels, it appears that some of the pixels with an anomalous variance cluster around defective pixels. This is because, at high fluxes, defective pixels overflow and corrupt surrounding pixels. This justifies the choice of discarding the n_w worst variances. Some areas, such as the one of Fig. A.4d, have a high density of discarded pixels.

The zooms of Figs. A.4b–e must be compared with the ones of Figs. A.1b–e and A.3b–e: some of the discarded pixels would have been hardly identifiable in the estimated background fluxes $\tilde{\mathbf{f}}^{\text{bg}}$ or in the gains $\tilde{\mathbf{g}}$.

A.5. Noise variance estimation

In this section, we describe how to estimate the variance v_k of the additive noise used in Eq. (4) on each pixel \mathbf{k} , namely the variance at null flux v_k^0 , the so-called readout noise, and η_k , which contains the conversion from the number of photons to adu in the Poisson noise term of this equation.

As the different calibration images are not sufficient to estimate these parameters independently for each pixel \mathbf{k} , they are

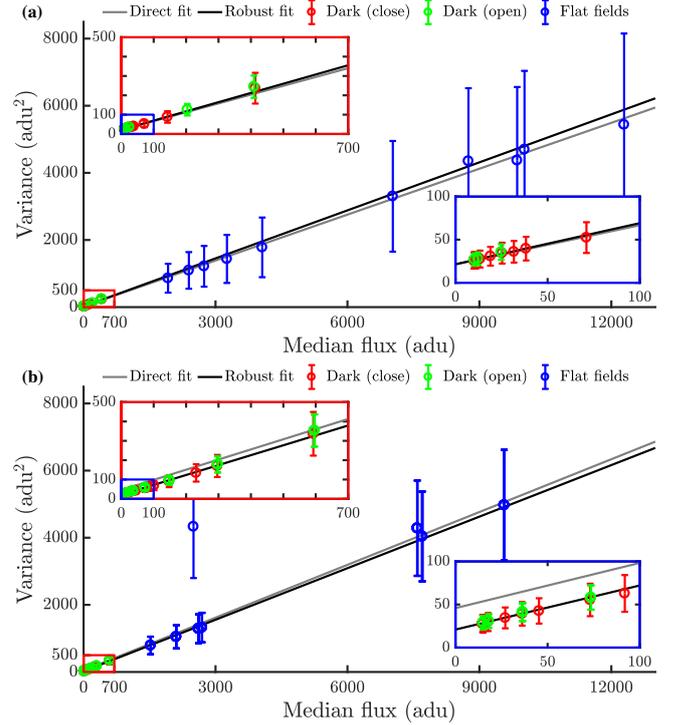


Fig. A.5. Estimation of the noise model $v = \tilde{\eta} \cdot I_{\text{adu}} + \tilde{v}^0$ for the HR 8799 (a) and RY lup (b) datasets. The model is fitted on the median fluxes and the median variances of the different calibration files: the dark currents with the shutter close (red) and open (green) and the flat-fields (blue). The error bars are obtained by the median absolute deviation on all the sensor pixels for each acquisition set. The grey law is fitted without taking into account the errors. The black law is fitted by accounting for the errors. The red and blue insets are zooms on regions of low incident fluxes.

assumed to be constant on the whole sensor:

$$\forall \mathbf{k}, \eta_k = \tilde{\eta} \text{ and } v_k^0 = \tilde{v}^0. \quad (\text{A.6})$$

This approximation holds if all the pixels behave similarly and if $g_k \approx 1$ whatever \mathbf{k} .

The coefficients $(\tilde{\eta}, \tilde{v}^0)$ are estimated using all the calibration files $\mathbf{y}^c \in \{\mathbf{y}^{\text{bg/o}}, \mathbf{y}^{\text{bg/c}}, \mathbf{y}^{\text{f}}\}$. An affine law is fitted to their median empirical variances $\mathbf{q}^c = \mathcal{M}_k(\mathcal{V}_a(\mathbf{y}^c)) \in \mathbb{R}^{n_t}$: $\mathbf{q}^c \approx \eta \mathbf{p}^c + v^0 \mathbf{1}$, where $\mathbf{p}^c = \mathcal{M}_k(\langle \mathbf{y}^c \rangle_a) \in \mathbb{R}^{n_t}$ is the median flux in the calibration images \mathbf{y}^c . Here, $\eta \mathbf{p}^c$ is consequently the variance associated with the Poisson noise induced by the median flux. Weights $\mathbf{w}^c = [\mathcal{A}_k(\mathcal{V}_a(\mathbf{y}^c))]^{-2} \in \mathbb{R}^{n_t}$, corresponding to the inverse of the square of the median absolute deviation are used to weight each sample of the fit:

$$(\tilde{\eta}, \tilde{v}^0) = \underset{\eta, v^0}{\text{argmin}} \sum_{c \in \{\text{bg/o, bg/c, f}\}} \|\mathbf{q}^c - (\eta \mathbf{p}^c + v^0 \mathbf{1})\|_{\mathbf{w}^c}^2, \quad (\text{A.7})$$

$$\tilde{\eta} = \frac{(\sum_{c,t} w_t^c) (\sum_{c,t} w_t^c p_t^c q_t^c) - (\sum_{c,t} w_t^c p_t^c) (\sum_{c,t} w_t^c q_t^c)}{(\sum_{c,t} w_t^c) (\sum_{c,t} w_t^c (p_t^c)^2) - (\sum_{c,t} w_t^c p_t^c)^2}, \quad (\text{A.8})$$

$$\tilde{v}^0 = \frac{(\sum_{c,t} w_t^c (p_t^c)^2) (\sum_{c,t} w_t^c q_t^c) - (\sum_{c,t} w_t^c p_t^c) (\sum_{c,t} w_t^c p_t^c q_t^c)}{(\sum_{c,t} w_t^c) (\sum_{c,t} w_t^c (p_t^c)^2) - (\sum_{c,t} w_t^c p_t^c)^2}. \quad (\text{A.9})$$

Results obtained on HR 8799 and RY lup calibration files are given in Fig. A.5 and in Table A.1.

Table A.1. Estimation of the coefficients $(\tilde{\eta}, \tilde{\nu}^0)$ for the two datasets, when accounting for the median absolute deviation of the empirical variances (weighted least squares) or not (standard least squares).

Dataset	HR 8799		RY lup	
	Weighted	Standard	Weighted	Standard
Fit				
$\tilde{\eta}$ (adu)	47.69	45.58	51.19	52.44
$\tilde{\nu}^0$ (adu ²)	21.48	21.24	20.91	45.80

Accounting for the median absolute deviation $\tilde{\sigma}$ of the empirical variance improves the estimation: the estimated line (in black) is less affected by values with large uncertainty than an unweighted least square fit (grey line).

It appears that the smaller errors on the low values help to counterbalance the *lever arm* effect on the large fluxes that are noisier. In the case of the RY lup dataset, the weighted fit reduces the influence of an outlier point around a flux of 2750 adu whose variance exceeds 4000 adu². This can be seen in the estimation of $\tilde{\nu}^0$ presented in Table A.1: as expected the weighted least squares gives more similar coefficients for the two datasets than the standard least square estimation.

Appendix B: Spectral dispersion calibration

This appendix describes the calibration of the lenslet array, that is, how to determine the operator $\mathbf{P}: \mathbb{R}^{n_h \times n_\lambda} \rightarrow \mathbb{R}^{m_1 \times m_2}$ that represents the spectral dispersion produced by the individual lenslets projected on the sensor.

To build this operator, several calibration files are used:

- Spectral flat-field acquisitions \mathbf{y}^s are obtained when an integrating sphere produces a uniform light upstream of the lenslet array with a broadband lamp. These images help to calibrate the 2D location of each spectrum and the lenslet transmissions as well as to identify defective lenslets.
- Laser acquisitions \mathbf{y}^w are obtained with $n_\lambda^{\text{cal}} = 4$ calibrated lasers placed upstream of the lenslet array. Their wavelengths are $\lambda^{\text{cal}} \in \{987.72, 1123.71, 1309.37, 1545.10\}$ nm. Only the first three lasers are used for the YJ band calibrations ($n_\lambda^{\text{cal}} = 3$).

For each of these files, a set of several acquisitions $a \in \llbracket 1, n_a \rrbracket$ are performed for a given exposure time τ . In the following, only their average, corrected for the background and the gain, is considered:

$$\forall l \in \{s, w\}, y_k^l \leftarrow g_k^{-1} \left([\langle \mathbf{y}^l \rangle_a]_k - \tau^l f_k^{\text{bg}/o} \right). \quad (\text{B.1})$$

B.1. Dispersion model

As previously described, after dispersion, the spectrum produced by a given lenslet is modelled as a weighted sum of elementary normalised Gaussian patterns, one for each reconstructed wavelength. For a pattern of size $\sigma \in \mathbb{R}^+$, centred on $\boldsymbol{\kappa} \in \mathbb{R}^2$ in pixel coordinates, the value at a given pixel \mathbf{k} is:

$$\xi_{\mathbf{k}}(\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(k_1 - \kappa_{1,1})^2 + (k_2 - \kappa_{2,1})^2}{2\sigma^2}\right). \quad (\text{B.2})$$

In practice, to save memory and computational time, the Gaussian patterns are truncated outside a region of a few pixels. As a consequence, \mathbf{P} is a sparse matrix.

The reduced hyperspectral cube produced by the standard reduction pipeline of SPHERE (Pavlov et al. 2008) possesses $n_\lambda = 39$ wavelengths. In our calibration method, the lists of reconstructed wavelengths λ are extended beyond the edges of the spectrum:

- for the YH-mode, $n_\lambda = 44$ wavelengths are reconstructed (and the $n_\lambda^{\text{cal}} = 4$ calibration lasers are used);
- for the YJ-mode, $n_\lambda = 47$ wavelengths are reconstructed (and the first $n_\lambda^{\text{cal}} = 3$ calibration lasers are used).

Thus, if $\boldsymbol{\kappa} \in \mathbb{R}^{2 \times n_\lambda}$ is the list of the positions of the elementary Gaussian patterns in a spectrum and $\boldsymbol{\sigma} \in \mathbb{R}^{n_\lambda}$ the list of their sizes, the light incoming from a given lenslet h is dispersed to form a spectrum \mathbf{u} on the sensor, whose value at pixel \mathbf{k} is:

$$u_{\mathbf{k}} = \sum_{\ell=1}^{n_\lambda} \Lambda_\ell \gamma_h \xi_{\mathbf{k}}(\boldsymbol{\kappa}_{1,\ell}, \boldsymbol{\kappa}_{2,\ell}, \sigma_\ell), \quad (\text{B.3})$$

where Λ is the spectrum of the lamp and γ_h is the transmission of lenslet h .

For each lenslet, the position $\boldsymbol{\kappa}$ and the width $\boldsymbol{\sigma}$ of each Gaussian pattern are assumed to follow a polynomial law of degree d as a function of the wavelength. At wavelength λ_ℓ , the value of the polynomial p_ℓ defined by the coefficients $\mathbf{c} \in \mathbb{R}^{d+1}$ is:

$$p_\ell(\mathbf{c}) = \sum_{j=1}^{d+1} c_j \lambda_\ell^{j-1}. \quad (\text{B.4})$$

The spectral projection operator \mathbf{P} is defined based on the list $\boldsymbol{\gamma} \in \mathbb{R}^{n_h}$ of the lenslet transmissions, the list $\mathbf{c}^\kappa \in \mathbb{R}^{n_h \times 2 \times (d_\kappa+1)}$ of the coefficients of the polynomial law of degree d_κ describing the positions of the elementary patterns for each lenslet, and the list $\mathbf{c}^\lambda \in \mathbb{R}^{n_h \times (d_\lambda+1)}$ describing the coefficients of the polynomial law of degree d_λ of the size of the elementary patterns for each lenslet. With these notations, the model of the spectrum at a given pixel \mathbf{k} produced by the lenslet h is:

$$u_{\mathbf{k}}(\mathbf{c}_h^\kappa, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda) = \sum_{\ell=1}^{n_\lambda} \Lambda_\ell \gamma_h \xi_{\mathbf{k}}(p_\ell(\mathbf{c}_{h,1}^\kappa), p_\ell(\mathbf{c}_{h,2}^\kappa), p_\ell(\mathbf{c}_h^\lambda)). \quad (\text{B.5})$$

In practice, the degree of the polynomial model is dictated by the number of calibration spots in each spectrum: $d_\kappa = n_\lambda^{\text{cal}} - 1$ i.e. $d_\kappa = 3$ in the YH-mode and $d_\kappa = 2$ in the YJ-mode. For both cases, we use $d_\lambda = 2$.

B.2. Dispersion calibration

With the notations defined for the robust penalisation, the estimation of these dispersion coefficients is obtained by minimising the following data fidelity term for a given lenslet h :

$$\mathcal{S}(\mathbf{y}, \mathbf{c}_h^\kappa, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda) \stackrel{\text{def}}{=} \sum_{\mathbf{k}=(1,1)}^{m_1, m_2} \rho\left(\delta_{\mathbf{k}}^{\text{cal}} s^{\text{cal}-1} (y_{\mathbf{k}} - u_{\mathbf{k}}(\mathbf{c}_h^\kappa, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda))\right), \quad (\text{B.6})$$

applied on the spectral flat-field $\{\mathbf{y}^s, \Lambda, \lambda\}$ or on the wavelength calibration $\{\mathbf{y}^w, \Lambda_h^{\text{cal}}, \lambda^{\text{cal}}\}$, where $\Lambda_h^{\text{cal}} \in \mathbb{R}^{n_h \times n_\lambda^{\text{cal}}}$ is the list of the amplitudes of the spots for each lenslet.

Estimating the dispersion parameters from the laser calibration images \mathbf{y}^w is not suitable because some laser spots may fall into regions with many defective pixels, or be totally hidden by a dust particle. Using the spectral flat-field images \mathbf{y}^s is more adapted since the signal covers many more pixels, but is very

limited in terms of wavelength characterisation. It is best to combine the two calibration files via a global data fidelity term:

$$\mathcal{F}(\mathbf{y}^s, \mathbf{y}^w, \mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \Lambda_h^{\text{cal}}) \stackrel{\text{def}}{=} \mathcal{S}(\mathbf{y}^s, \mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda) + \mu_w \mathcal{S}(\mathbf{y}^w, \mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda_h^{\text{cal}}, \lambda^{\text{cal}}), \quad (\text{B.7})$$

where the hyperparameter μ_w balances the influence of the wavelength calibration compared to the spectral flat-field.

Minimising Eq. (B.7) to determine the dispersion parameters of a given lenslet h implies knowledge of the spectrum Λ of the lamp. This spectrum can be robustly estimated from all the lenslets by minimising the following data fidelity term:

$$\mathcal{L}(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda) \stackrel{\text{def}}{=} \mu_\Lambda \mathcal{R}_\lambda(\Lambda) + \sum_{k=(1,1)}^{m_1, m_2} \rho \left(\delta_k^{\text{cal}} s^{\text{cal}-1} \left(\mathbf{y}_k^s - \sum_{h=1}^{n_h} u_k(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda) \right) \right), \quad (\text{B.8})$$

regularised by the smoothing term $\mu_\Lambda \mathcal{R}_\lambda(\Lambda)$ previously defined in Eq. (12) and constrained to the domain:

$$\mathbb{D}_\Lambda = \{\Lambda \in \mathbb{R}^{n_\lambda} / \forall \ell, \Lambda_\ell \in \mathbb{R}^+, \Lambda_1 = \Lambda_{n_\lambda} = 0\}. \quad (\text{B.9})$$

The spectrum of the lamp is initialised on a few neighbouring spectra at the centre of the sensor field. Each lenslet spectrum and its dispersion coefficients $(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda_h^{\text{cal}})$ are then identified neighbour after neighbour by minimising Eq. (B.7).

This first estimation can be refined to account for the overlap of neighbouring spectra. If the impact on the position coefficients \mathbf{c}^k is minor, it influences the size coefficients \mathbf{c}^λ in particular on the spectra edges. That is why the first set of parameters is refined n_c times by Algorithm 4 by iteratively working on the residuals.

Algorithm 4 Spectral dispersion calibration algorithm

- 1: **for** i from 1 to n_c **do** $\triangleright n_c$ refinement loops
 - 2: $\Lambda \stackrel{\text{Eq. (B.8)}}{\leftarrow} \underset{\Lambda \in \mathbb{D}_\Lambda}{\text{argmin}} \mathcal{L}(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda)$ $\triangleright \Lambda$ update, $(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h)$ fixed
 - 3: $\forall k, r_k \stackrel{\text{Eq. (B.5)}}{\leftarrow} \gamma_k^s - \sum_{h=1}^{n_h} u_k(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda)$ \triangleright Global residuals
 - 4: **for** h from 1 to n_h **do** \triangleright Loop on the lenslets
 - 5: $\forall k, r_k^h \stackrel{\text{Eq. (B.5)}}{\leftarrow} r_k + u_k(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda)$ $\triangleright h^{\text{th}}$ lenslet residuals
 - 6: $(\mathbf{c}_h^k, \Lambda_h^{\text{cal}}) \stackrel{\text{Eq. (B.7)}}{\leftarrow} \underset{\mathbf{c}_h^k, \Lambda_h^{\text{cal}}}{\text{argmin}} \mathcal{F}(\mathbf{r}^h, \mathbf{y}^w, \mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \Lambda_h^{\text{cal}})$
 $\triangleright (\mathbf{c}_h^k, \Lambda_h^{\text{cal}})$ update on $(\mathbf{r}^h, \mathbf{y}^w)$, $(\mathbf{c}_h^\lambda, \gamma_h, \Lambda)$ fixed
 - 7: $(\mathbf{c}_h^\lambda, \gamma_h) \stackrel{\text{Eq. (B.6)}}{\leftarrow} \underset{\mathbf{c}_h^\lambda, \gamma_h}{\text{argmin}} \mathcal{S}(\mathbf{r}^h, \mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda)$
 $\triangleright (\mathbf{c}_h^\lambda, \gamma_h)$ update on \mathbf{r}^h , $(\mathbf{c}_h^k, \Lambda)$ fixed
 - 8: **return** $(\mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \Lambda_h^{\text{cal}})$
-

At each iteration of the refinement loop, the parameters of the model are alternatively estimated. First, the lamp spectrum is robustly fitted on all the lenslets by minimising Eq. (B.8) with the VMLM-B algorithm.

The positions \mathbf{c}_h^k and amplitudes Λ_h^{cal} of each spot are then estimated, one lenslet after another, after the contributions of all other lenslets have been removed from the data. These estimations are also performed with VMLM-B. In a third step, the size \mathbf{c}_h^λ of the patterns and the lenslet transmissions γ_h are estimated using a simplex search method (Lagarias et al. 1998).

Doing so, the estimation of the position parameters is split from the pattern size law fit. Thus, \mathbf{c}_h^k is not influenced by the monochromatic spots size in \mathbf{y}^w whereas they are supposed to fit continuous spectra in \mathbf{y}^s .

B.3. A super-resolved model for the lenslet PSF?

Considering a super-resolved model, as proposed by Brandt et al. (2017), is straightforward in the proposed framework. It consists in replacing the Gaussian patterns ξ_k with a more accurate model of the lenslet PSFs. The real difficulty comes from the calibration of this model. In particular, the calibration files available for the SPHERE instrument are not sufficient to derive an accurate super-resolved PSF model for several reasons.

Firstly, the calibration lasers simultaneously illuminate the IFU. Thus, the foot of the PSF produced by one laser overlaps with the PSFs of the other wavelengths. A sequential acquisition (one wavelength at a time) would be necessary to directly recover the whole support of the PSFs without requiring to unmix the PSFs. Secondly, even with a single monochromatic illumination, all lenslets are illuminated at once, which leads to two issues. The spots produced by the different lenslets lie on a regular grid and slightly overlap, which creates a degeneracy of the super-resolved model. Some interference also occurs between patterns which are not representative of the PSFs obtained with incoherent light. The selective illumination of one lenslet at a time would solve these issues but would be time consuming and would raise technical instrumental issues. Finally, the number of distinct wavelengths used in the calibration procedure is too low to allow a fine parameterisation of a PSF model, and in particular to address the variability from one lenslet to another. A larger set of wavelengths, covering the whole spectrum (to prevent extrapolations on the edges of the spectral domain) would be necessary. For these different reasons, such a super-resolved model was not implemented in the present paper.

Appendix C: Results of the dispersion calibration

C.1. Dispersion

Results of Algorithm 4 are illustrated in Fig. C.1 for the calibration in the YH-mode of the HR 8799 dataset.

The $n_\lambda^{\text{cal}} = 4$ laser spots are visible in the wavelength calibration acquisition \mathbf{y}^w for each lenslet in Figs. C.1a,c. The coloured dots mark the positions of the spots fitted by the Algorithm 4. The radii of the dark circles indicate the size of the Gaussian patterns. The chromatic evolution of the fitted size is visible with a minimal spot extension around the third calibration wavelength.

The residuals in Fig. C.1g of the first iteration of the wavelength calibration present a strong diversity due to inaccurate values of the size of the Gaussian patterns and incorrect amplitudes for the spots Λ^{cal} . Some patterns are not well-aligned with the laser spots, as suggested by nonsymmetric residuals on some spots (which indicates a bad wavelength calibration). The final residuals in Fig. C.1h, after $n_c = 10$, are more homogeneous, with centred residuals on each spot. Nonetheless, all the spots present red ring-shaped residuals and seem to worsen. This suggests that the pattern size given by \mathbf{c}^λ is overestimated when applied to the laser spots. This is expected as these spots are monochromatic whereas the dispersion law is fitted on the continuous spectra of the spectral flat-field file \mathbf{y}^s . This also implies that a wavelength calibration only performed on the wavelength calibration file \mathbf{y}^w underestimates the size parameters.

The spectra of the $n_h \approx 19\,000$ lenslets of the spectral flat-field acquisition \mathbf{y}^s are visible in Figs. C.1b,d. The coloured dots represent the positions of the $n_\lambda = 44$ wavelengths of the forward model. It can be seen that the spectral range goes beyond the calibration laser spots, especially in the longer wavelengths

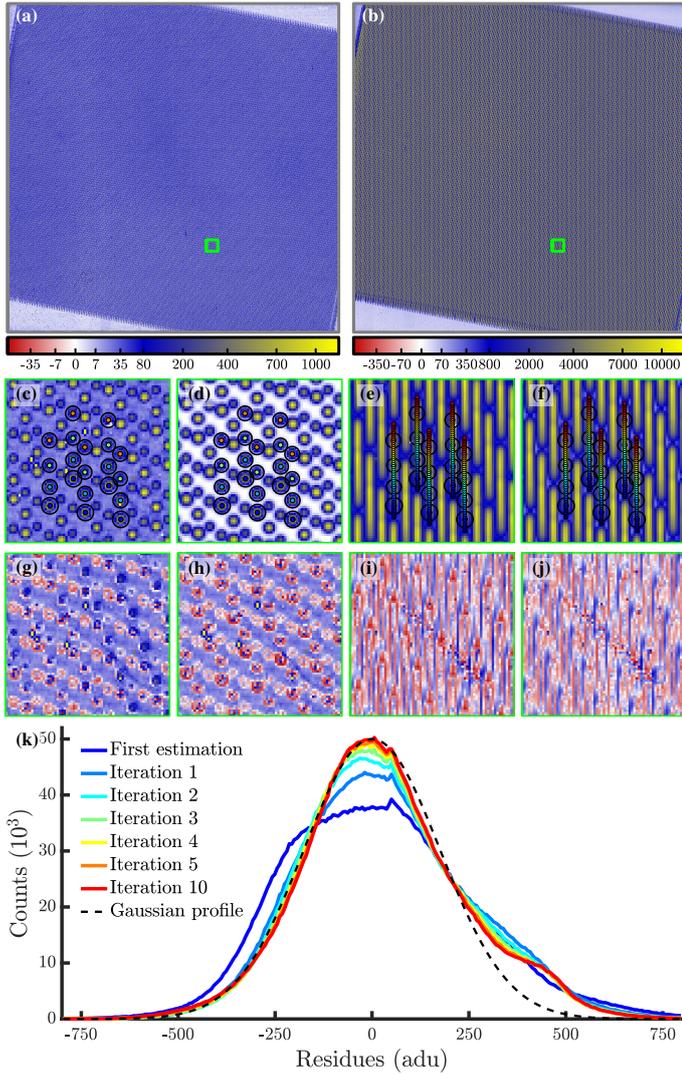


Fig. C.1. Calibration of the spectral projection for the HR 8799 dataset (square root stretch, colour bar: adu). *a*: averaged wavelength calibration acquisition \mathbf{y}^w . *b*: averaged spectral flat-field calibration acquisition \mathbf{y}^s . *c*: zoom of (*a*). The found positions for the calibration wavelengths are highlighted by coloured dots. The black circles are proportional to the size of the Gaussian patterns. *d*: simulation of the wavelength calibration by the projection model. *e*: zoom of (*b*). The positions of the wavelengths simulated in the forward model are highlighted by coloured dots. *f*: simulation of the spectral flat by the projection model. *g*: residuals of the first estimation of the wavelength calibration. *h*: final residuals of the wavelength calibration. *i*: residuals of the first estimation of the spectral flat calibration. *j*: final residuals of the spectral flat calibration. *k*: histogram of the residuals of the spectral flat calibration along the iterations. For comparison purposes, a Gaussian profile of standard deviation $\sigma = 175$ adu is plotted.

(red dots). The extrapolation of the positions beyond these calibrated spots leads to wider spatial steps in the short wavelengths (blue dots) while the positions are more packed in the longer wavelengths (red dots). This may lead to a degraded wavelength calibration in the outer regions of the spectral range (see Fig. 8 and associated discussion).

In the residuals of the first calibration of the spectral flat calibration in Fig. C.1i, it appears that the overall signal is overestimated. This may be due to the fact that the overlap of the spectra is not taken into account at this first iteration. The dust

presents in Fig. A.3c is clearly identifiable. After the $n_c = 10$ refinements, the residuals in Fig. C.1j improve, especially on the spectra edges, to fall below $\sim 2.5\%$ of the lamp spectrum intensity. This is confirmed by the curves of Fig. A.3k that show the evolution of the histogram of the residuals along with the refinement iterations. Starting from a spread distribution, the residuals squeeze towards a Gaussian curve with smaller dispersion.

Blue structures in Fig. C.1j indicate that some regions are still slightly underestimated by the model, leading to positive residuals. This corresponds to the small bump around 400 adu in the residue distribution. This is due to the fact that the Gaussian pattern model is an approximation of the spots diffracted by the hexagonal apertures of the lenslets. They have “feet” in privileged directions that are not fitted by an axisymmetric pattern, to which optical aberrations may be added. This also explains the blue positive background around the laser spots in Figs. C.1a,c.

In terms of wavelength centring of the forward model, as mentioned by Vigan et al. (2015), the SPHERE reduction pipeline (Pavlov et al. 2008) needs an extra-patch to correct its wavelength calibration. To check our calibration, the wavelength calibration file \mathbf{y}^w is reconstructed using Algorithm 1. The results are given in Fig. C.2 for the two datasets.

The coloured projections of Figs. C.2a,b and the monochromatic slices given in Figs. C.2c,d show that the reconstructed laser illuminations are not uniform but are characteristic of speckle noise.

The median intensity $\mathcal{M}_\theta(\mathbf{x})$, computed on each slice of the reconstructed hyperspectral cube, is shown in Fig. C.2e for each dataset. Maxima match well the wavelength of the calibration lasers, represented by vertical dashed lines. The effective resolution is estimated from the full width at half maximum of each laser. The average resolution of the YJ-mode and YH-mode is $R = 57$ and $R = 31$ respectively. These values are compared to the standard reduction method based on a direct mapping (Vigan et al. 2015) in the YH-mode from which the estimated average spectral resolution is $R = 20$. The better spectral resolution using PIC is close to the SPHERE-IFS theoretical resolution $R = 30$ of Beuzit et al. (2019) and shows the impact of the spectral deconvolution performed by the inverse problems approach (this deconvolution is possible because the forward model accounts for the spectral mixing).

C.2. Transmission

The calibration procedure also gives the achromatic transmission map γ of the IFU, including the lenslet array, the prism, and the other associated lenses; it is shown in Fig. C.3a, with a median transmission of $\mathcal{M}_h(\gamma) = 1$.

It can be seen that the transmission map γ further corrects the effect of the dust particles on the optical path that cannot be suppressed only by the sensor flat-field correction factor \tilde{g} . Indeed, comparing Figs. A.3a and C.3a, it appears that these dust particles have a counterpart in the transmission map γ . Even if the effect is minor, Fig. C.3b shows that the lamp spectrum Λ is refined with the iterations of the calibration loop. The corrections are mainly due to the refinement of the dispersion law c^λ . Indeed, the dilution of the intensity in a Gaussian pattern according to its size must be counterbalanced by its amplitude. Some features are visible in the lamp spectra that can participate to the robust wavelength calibration.

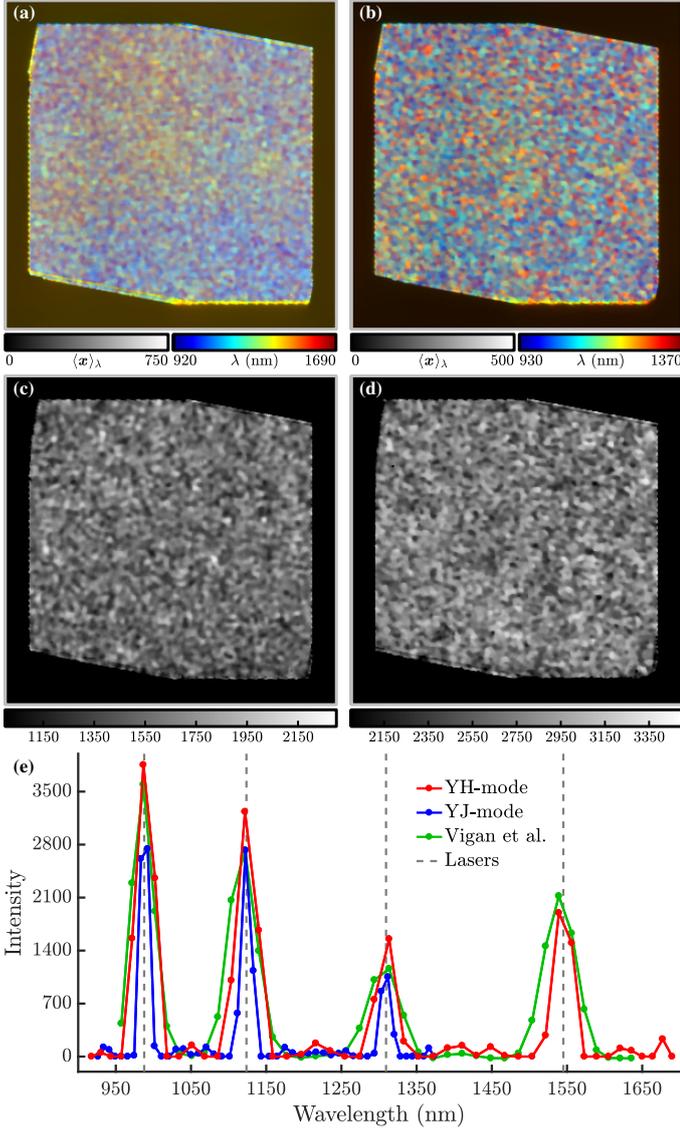


Fig. C.2. Wavelength calibration of the model on the HR 8799 (YH-mode, on the *left panel*) and the RY lup (YJ-mode, on the *right panel*) datasets (colour bar: arbitrary unit). *a,b*: coloured projection $\langle x \rangle_\lambda$ of the reconstructed hyperspectral cube. *c,d*: slice of the reconstructed hyperspectral cube ($\lambda = 1314$ nm for the HR 8799 dataset, $\lambda = 1122$ nm for the RY lup dataset). *e*: median intensity $M_\theta(x)$ computed on each slice of the reconstructed hyperspectral cube for each dataset compared with the reduction proposed by Vigan et al. (2015). The wavelengths of the calibration lasers are highlighted by dashed grey lines.

C.3. Autocalibration

A last correction must be added in the calibration of the forward model. Indeed, it appears that the data on the sky are not perfectly aligned with the calibration files. This is emphasised in Figs. C.4a–d that presents a region of interest of an acquisition on the sky (a), the fit of the model obtained before the autocalibration (b), and the corresponding residuals (d). The difference between the model before and after the alignment is shown in Fig. C.4c.

To perform this alignment, n_r regions of interest, equally spread on the sensor and of $m_r \times m_r$ pixels, are extracted from the raw science data to perform an autocalibration. The regions are small and the spectra $\Lambda^a \in \mathbb{R}^{n_r \times n_\lambda}$ of the lenslets are considered to be equal on these reduced fields.

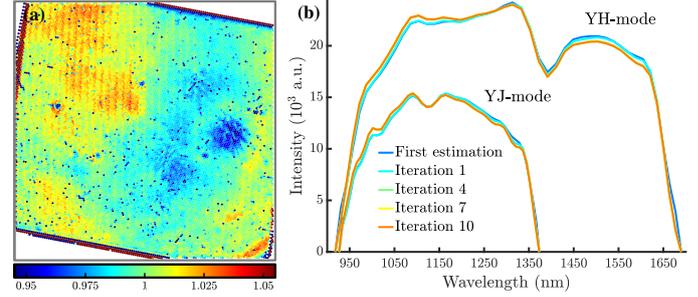


Fig. C.3. *a*: transmission map γ of the IFU estimated for the HR 8799 dataset. *b*: fitted spectrum of the lamp for the two datasets HR 8799 (YH-mode) and RY lup (YJ-mode) along the iterations of Algorithm 4.

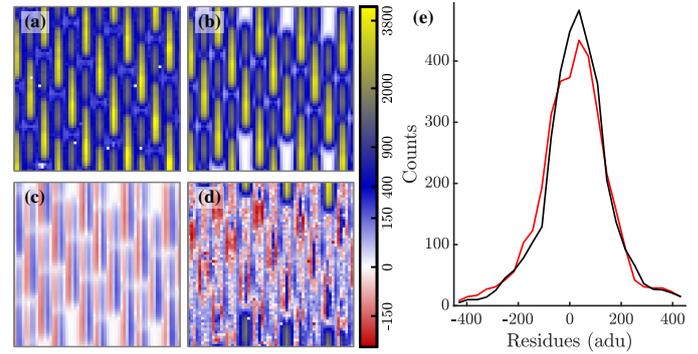


Fig. C.4. Autocalibration for the HR 8799 dataset (square root stretch, colour bar: adu). *a*: zoom on a region of interest. *b*: model before the alignment. *c*: difference between the model prediction before and after the alignment. *d*: residuals after the alignment. *e*: statistics before (red) and after (black) the autocalibration step.

Table C.1. Estimation of the shift δ^k in pixel to apply on the coefficients c^k for the two datasets.

Dataset	HR 8799	RY lup
δ_1^k (pix)	-2.55×10^{-2}	$+1.90 \times 10^{-2}$
δ_2^k (pix)	-5.90×10^{-2}	$+2.61 \times 10^{-1}$

Then an algorithm similar to Algorithm 4 is applied. A refinement loop is performed to compute the global shift $\delta^k \in \mathbb{R}^2$ that is assumed to be constant for all the coefficients c^k . Each iteration alternates two steps:

- the spectrum Λ_r^a of each of the n_r regions is fitted via Eq. (B.8), applied on the region of interest, assuming that the shift δ^k is known;
- the shift δ^k is fitted by combining all the regions of interest by adapting Eq. (B.6), for fixed spectra Λ_r^a .

The histograms of the residuals before and after the alignment are compared in Fig. C.4e. After the alignment they are less dispersed and more symmetrical. The estimated shifts for the two datasets are given in Table C.1. The shifts are small, of the order of a quarter of a pixel for the strongest correction in the case of the RY lup dataset. For the HR 8799 dataset, the correction is just a few percent of a pixel. Nonetheless, this shift corresponds to a measurable signal as seen in Fig. C.4c that compares the intensity predicted by the model before and after the autocalibration step.

Appendix D: Numerical value of the different parameters

Table D.1 summarises the different parameters and hyperparameters of the different algorithms. The numerical values were the ones used to obtain the results presented in this paper.

The scaling factors s_k applied on the pixels k of the data are chosen to be constant all over the field: $\forall k, s_k = s_0$. As the role of s_0 is to discriminate the data that deviate compared to the expected values, it depends on the overall data dynamics and must be adapted to each dataset. This can be done by looking at the distribution of the residuals such as the ones given in Figs. 7o,p: s_0 is proportional to their standard deviation. The RY lup dataset having a low dynamics, this explains why its s_0 is smaller.

The hyperparameters μ_λ , μ_{2D} and μ_Λ introduced in Eqs. (13) and (B.8) are normalised as follows:

$$\mu_\lambda = \frac{1}{\varrho} \frac{\mu_\lambda^0}{s_0^2}, \mu_{2D} = \frac{1}{\varrho} \frac{\mu_{2D}^0}{s_0^2} \text{ and } \mu_\Lambda = \frac{1}{\varrho_\Lambda} \frac{\mu_\Lambda^0}{s_0^2}, \quad (\text{D.1})$$

where ϱ is a normalisation factor to scale the number of unknowns in the regularisation terms with the measurements in the data fidelity term. In other words, it accounts for the number of reconstructed spectra and wavelengths compared to the number of pixels in the forward model. The normalisation by s_0^2 accounts for the quadratic development of the Cauchy function close to zero. It ensures that the regularisation scales coherently with the scaling coefficient s_0 . μ_λ^0 and μ_{2D}^0 are manually set.

Here, μ_λ^0 is chosen to be as small as possible while rejecting the high-frequency modes that appear in the reduced spectra due to the oversampling of the forward model as mentioned by Draper et al. (2014). Also, μ_{2D}^0 must be high enough to participate in the outlier pixel detection while avoiding adding too much smoothing in the reduced cubes. To do so, it is placed at the limit at which the acquisition noise starts to pollute the reduced cubes, and consequently depends on the data dynamic range.

For the refinement of the spectral dispersion model, to refine the parameters of each lenslet, the local spectra are extracted and analysed on their region of interest. These regions are defined as the window that encompasses the position of the different Gaussian patterns, extended on each side by 2 pixels, leading to a window of roughly 5×45 pixels for each spectrum. In addition, at each loop, of the calibration, the parameter μ_w in Eq. (B.7) is updated so that the cost on the spectral flat-field acquisition, defined as

$$\mathcal{J}_h^s = \mathcal{S}(\mathbf{y}^s, \mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda, \lambda), \quad (\text{D.2})$$

is balanced with the cost on the wavelength calibration file, defined as

$$\mathcal{J}_h^w = \mathcal{S}(\mathbf{y}^w, \mathbf{c}_h^k, \mathbf{c}_h^\lambda, \gamma_h, \Lambda_n^{\text{cal}}, \lambda^{\text{cal}}), \quad (\text{D.3})$$

to favor the wavelength calibration file as follows:

$$\mu_w = 50 \mathcal{M}_h \left(\frac{\mathcal{J}_h^s}{\mathcal{J}_h^w} \right). \quad (\text{D.4})$$

Table D.1. Numerical values of the parameters and hyperparameters used in the different algorithms.

Parameter	Value	Description
Algorithm 1: Data reduction		
medfilter	3×3	Window size of the median filter
n_{it}	6	Number of optimisation loops
–	20	Number of iterations of the VMLM-B algorithm to solve Eq. (1)
w_{th}	0.5	Threshold on the equivalent weights defined in Eq. (10) to identify outlier pixels
$\forall k, s_k = s_0$	5 for RY lup/25 for others	Scaling factor of the robust penalisation in Eq. (9) (\propto the standard deviation of the residuals as seen on Figs. 7o,p for the HR 8799 and RY lup datasets)
$\forall k, \eta_k, v_k^0$	Eq. (A.7), $\tilde{\eta}, \tilde{v}^0$	Noise model used in Eq. (4) (Readout noise and Poisson noise)
ϱ	$\frac{n_1 n_2 n_\lambda}{m_1 m_2}$	Normalisation factor of the hyperparameters to balance the unknowns and the measurements
μ_λ^0	5×10^{-6}	Hyperparameter of the spectral regularisation defined in Eq. (13)
μ_{2D}^0	2.5×10^{-2}	Hyperparameter of the spatial regularisation defined in Eq. (13)
ϵ	10^{-16}	Spatial regularisation threshold defined in Eq. (11)
Algorithm 2: Remnant signal estimation		
n_{rs}	100	Number of iterations of the algorithm for the HR 8799 and RY lup datasets (the algorithm is not run for Ganymede and Io, the background obtained with the shutter opened is used instead).
Algorithm 3: Defective pixel identification		
medfilter	5×5	Window size of the median filter
α^a	15	Tolerance factor
Algorithm 4: Spectral dispersion calibration		
n_c	10	Number of optimisation loops
–	30	Number of iterations of the VMLM-B algorithm to minimize Eq. (B.8)
–	100	Number of iterations of the VMLM-B algorithm to minimize Eqs. (B.6) and (B.7)
s^{cal}	350	Scaling factor of the robust penalisation in Eqs. (B.6)–(B.8)
$\forall k, \eta_k v_k^0$	0 1	No noise model used in Eqs. (B.6)–(B.8)
ϱ_Λ	$\frac{n_\lambda}{m_1 m_2}$	Normalisation factor of the hyperparameters to balance the unknowns and the measurements
μ_Λ^0	1.5×10^{-3}	Hyperparameter of the spectral regularisation defined in Eq. (B.8)
μ_w	$50 M_h \left(\frac{\mathcal{S}_h^s}{\mathcal{S}_h^w} \right)$	Hyperparameter to balance the spectral flat-field acquisition with the wavelength calibration file in Eq. (B.7)
Autocalibration algorithm		
n_r	3×3	Number of regions of interest
m_r	61	Size of the regions of interest
–	3	Number of optimisation loops
–	100	Number of iterations of the VMLM-B algorithm to estimate Λ_r^a via the adapted Eq. (B.8)
–	100	Number of iterations of a simplex search method to estimate δ^k via the adapted Eq. (B.6)
s^{cal}	350	Scaling factor of the robust penalisation in Eqs. (B.6) and (B.8)
$\forall k, \eta_k v_k^0$	0 1	No noise model used in Eqs. (B.6) and (B.8)
ϱ_Λ	$\frac{n_\lambda}{m_r m_r}$	Normalisation factor of the hyperparameters to balance the unknowns and the measurements
μ_Λ^0	8×10^{-4}	Hyperparameter of the spectral regularisation defined in Eq. (B.8)

Notes. The parameters that are unnamed in the text are noted “–”.