Exoplanet Imaging Data Challenge, phase II: Characterization of exoplanet signals in high-contrast images

Cantalloube F.[a], Christiaens V.[b], Cantero C.[b,c], Nasedkin E.[d], Cioppa A.[c], Absil O.[b], Bonse J. M.[e], Delorme P.[f], Gomez-Gonzalez C.[g], Juillard S.[b], Mazoyer J.[h], Samland M.[d], Ruffio J.-B.[i], and Van Droogenbroeck M.[c]

[a]Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[b]Space Sciences, Technologies & Astrophysics Research (STAR) Institute, Université de Liège, Allée du Six Août 19c, B-4000 Liège, Belgium
[c]Montefiore Institute, Université de Liège, 4000 Liège, Belgium
[d]Max-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg 69117, Germany
[e]Institute for Particle Physics and Astrophysics, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093, Zurich, Switzerland
[f]Univ. Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France
[g]Barcelona Supercomputing Center, Spain
[h]LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, F-92195 Meudon, France
[i]Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

ABSTRACT

Today, there exists a wide variety of algorithms dedicated to high-contrast imaging, especially for the detection and characterisation of exoplanet signals. These algorithms are tailored to address the very high contrast between the exoplanet signal(s), which can be more than two orders of magnitude fainter than the bright starlight residuals in coronagraphic images. The starlight residuals are inhomogeneously distributed and follow various timescales that depend on the observing conditions and on the target star brightness. Disentangling the exoplanet signals within the starlight residuals is therefore challenging, and new post-processing algorithms are striving to achieve more accurate astrophysical results. The Exoplanet Imaging Data Challenge is a community-wide effort to develop, compare and evaluate algorithms using a set of benchmark high-contrast imaging datasets. After a first phase ran in 2020 and focused on the detection capabilities of existing algorithms, the focus of this ongoing second phase is to compare the characterisation capabilities of state-of-the-art techniques. The characterisation of planetary companions is two-fold: the astrometry (estimated position with respect to the host star) and spectrophotometry (estimated contrast with respect to the host star, as a function of wavelength). The goal of this second phase is to offer a platform for the community to benchmark techniques in a fair, homogeneous and robust way, and to foster collaborations.

Keywords: Exoplanet imaging; Exoplanet characterisation; High-contrast imaging; Adaptive Optics; Coronagraphy; Post-processing techniques; Data challenge

1. INTRODUCTION

High-contrast imaging (HCI) is a very demanding observation technique that requires dedicated state-of-the-art instrumentation in the visible to infrared on the largest ground-based telescopes. Even if other direct exoplanet detection/characterization techniques come into being, high-contrast imaging remains a technique of choice to obtain images of extrasolar systems, for which many archival data are publicly available to the community. When a point source is detected within high-contrast images, it primarily needs to be characterized, that is to

say to estimate its relative projected *position* and *brightness* with respect to the target star (for one or more wavelengths). This characterisation step is essential to confirm that the detected signal is gravitationally bound to the star and is of a (sub)stellar nature, making it a firm companion detection. The characterization step is also fundamental to deduce the orbital and atmospheric properties of the companion, which allows to trace its formation history. In particular, the extracted spectrum, allows one to estimate the chemical composition, cloud properties, formation and migration histories, and in the future may allow the detection of biosignatures in the spectrum of rocky exoplanets.[1,2]

Post-processing techniques are an essential part of HCI whose goal is to gain several additional orders of magnitude in contrast in the images delivered by high-contrast instruments. These images are heavily contaminated by bright starlight residuals that follow various timescales and have an inhomogeneous spatial distribution. Indeed, typical observing sequences last about 1 hour during which observing conditions are likely to change and internal aberrations to evolve. In that context, extracting accurate information about the position and flux of an exoplanet signal that is more than one order of magnitude fainter than the starlight residuals is arduous.

Therefore, exoplanet characterization in HCI is crucial both on the technical and astrophysical point of view. In this context, the goal of the second phase of the *Exoplanet Imaging Data Challenge* (EIDC) is to focus on the characterization capabilities of most post-processing techniques dedicated to HCI that exist to date. Within this phase of the EIDC, we provide the community with four types of resources that are described in this paper: (1) an ensemble of HCI datasets with injected planetary signals, (2) a platform for participants to submit their results, (3) an assortment of metrics to evaluate and compare the submitted results, and (4) various software tools to help the HCI community to make the most out of this challenge. The main objective of this second phase of the EIDC is to get insights about the characterizations accuracy of each algorithm depending on the provided datasets. This way the EIDC will (i) provide guidance on the most appropriate techniques to be used by observers, depending on the scientific context, (ii) identify potential systematic errors in these techniques, (iii) clear new exploration pathways for improving our current tools and (iv) offer a set of data and evaluation metrics to assess the performance of the next generation of post-processing algorithms that can be used for publications.

The second phase of the EIDC has been launched on the 25th of April 2022, and the closing date has been postponed to 31st of December 2022. The main information about the *Exoplanet Imaging Data Challenge* initiative can be found on the dedicated website*, which contains two tabs dedicated to the first and second phases (top-right drop-down menu). The ensemble of data sets for this characterization phase are hosted on a Zenodo† open-access repository. The participants are invited to submit their results via the EvalAI‡ competition platform, which computes a basic evaluation metrics that we previously defined to publicly display the obtained score on a leaderboard.

In the following, we first describe the data sets provided to the participants, containing 21 injected planetary signals, that are available on the Zenodo repository (Sect. 2). We then describe the evaluation procedure used for the leaderboard ranking on the evalAI platform for this characterization phase (Sect. 3). We then give two examples of retrieval performed by two traditional post-processing algorithms, ANDROMEDA and PCA-NEGFC (Sect. 4). At last, we propose a path-forward for the next phases of the EIDC and for leveraging the limitations noted during the first and second phases. The co-authors of the present paper are all part of the working group preparing this second phase of the EIDC and/or kindly provided with the pre-reduced HCI datasets used in this second phase (contact: exoimg.datachallenge@gmail.com).

2. DATA SETS

For this characterization challenge (EIDC, phase 2), we gathered 8 data sets from two of the latest generation of high-contrast low-resolution integral field spectrographs: VLT/SPHERE-IFS[3,4] and Gemini-S/GPI.[5] The pre-processing applied and the different files provided to the participants are described in Sect. 2.1 and Sect. 2.2, respectively. In each data set we injected 2 to 3 synthetic planetary signals with a realistic spectrum, making a total of 21 planetary signals to retrieve. The injection procedure, along with its limitations and its application to

---

a SPHERE test data set, is described in Sect. 2.3. We describe in Sect. 2.4 the various resources publicly available on a dedicated Github repository[§], including a training dataset and Python utilities. At last we illustrate an example injection and retrieval on the training dataset.

2.1 Pre-reduced data sets

The most common observing mode for high-contrast imaging of exoplanets and circumstellar disk consists in carrying out the observations in *pupil-tracking* mode by carefully tuning the derotator of the alt-az mount telescope. This way, the pupil plane is kept in the same orientation, which means that the manifestation of optical aberrations in the image plane remains more or less at the same position. This makes it possible to model and subtract the stellar residuals. Meanwhile, the circumstellar objects rotate in the image field around the target star (centered behind the coronagraphic mask - if any) at a deterministic rate, following the parallactic angles. This so-called *angular diversity* is exploited by the *Angular Differential Imaging* technique,[6] which consists of three steps: (1) estimating a model of the starlight residuals without circumstellar signals, (2) subtracting this model from each of the frames of the temporal image cube, and (3) combining the subtracted images into a single image therefore containing the subtraction residuals and the circumstellar objects. From this residual map, most post-processing algorithms build a detection map (or probability map, or signal-to-noise ratio map) along with a detection criterion (e.g. a threshold).[7] Because the so-called *ADI sequences* provided by the high-contrast instruments are very common and represents hundreds of available datasets, the second phase of the EIDC is based on such images.

For this second phase we provide eight multispectral temporal image cubes taken with the integral field units (IFU) GPI, installed at the Gemini-South telescope, and SPHERE-IFS, installed at the Very Large Telescope. Both instruments benefit from an exquisite extreme adaptive optics correction, bringing the Strehl ratio up to 90% in the H-band[8–10] and an apodized Lyot coronagraph[11–13] that makes it possible to reach a raw contrast of 10^{-4} at a separation of 500 mas to the star. Each data set is taken under different, yet representative, observing conditions, ranging from very good observing conditions of a bright target star to very bad observing conditions of a faint target star. Table 1 gathers the main parameters of the observation sequence and observing conditions of each of the 8 data sets. Additionally, we note that the effective telescope diameter (downstream of the coronagraph) is of 7.57 m for GPI and 7.87 m for SPHERE-IFS and that the plate-scale (the pixel scale of the detector used) is of 14.16 mas/px for GPI and 7.46 mas/px for SPHERE-IFS images. All the images are of odd-number dimension and the center is located on the central pixel. The centering of the host star is performed as part of the pre-processing.

Table 1. Description of the eight data sets provided for the EIDC characterization phase: N_{img} is the size of the coronagraphic images; N_λ is the number of channels along the spectral axis; Δ_λ is the spectral coverage range; N_t is the number of frames along the temporal axis; Δ_{field} is the total field rotation of circumstellar objects. The target star is labeled as 'bright' for magnitude in the J-band below 7 and as 'faint' above. The last column qualitatively indicates the observing conditions: 'LWE' stands for low-wind effect and 'WDH' stands for wind-driven halo.

Instrument	ID	N_{img} [px] × [px]	N_λ	Δ_λ [μm]	N_t	Δ_{field} [°]	Target	Obs. Conditions
GPI	gpi1	169 × 169	37	1.495 − 1.797	37	14.4	bright	very good
	gpi2	169 × 169	37	1.495 − 1.797	36	156.6	bright	medium (WDH)
	gpi3	169 × 169	37	1.495 − 1.797	38	13.3	bright	medium
	gpi4	169 × 169	37	1.495 − 1.797	37	14.2	faint	poor (LWE)
SPHERE-IFS	sphere1	189 × 189	39	0.957 − 1.329	144	26.8	bright	very good
	sphere2	189 × 189	39	0.957 − 1.329	80	14.4	bright	poor (LWE)
	sphere3	189 × 189	39	0.957 − 1.329	64	25.8	faint	very good
	sphere4	189 × 189	39	0.957 − 1.329	90	144.8	bright	poor (WDH)

The pre-processing of the raw images delivered by the instruments (i.e. dark subtraction, flat fielding, bad pixel correction, re-centering, cropping, flux normalization and frame selection) were performed by the GPI and

---

SPHERE consortia via the custom-tools they developed: the GPI data cruncher[14, 15] and the SPHERE data center,[16, 17] respectively. The pre-processed data are kindly provided by the two consortia. Before injecting synthetic planetary signals, we homogenized the data (centering, cropping and setting of a parallactic angles direction convention).

2.2 Files provided to the participants

The 8 final complete data sets, containing the injected planetary signals, are provided to the participants on the Zenodo repository in *.fits* file format.[18] Table 2 gathers the names of the available files and their content. The header of each file contains general information about the instrumental set-up and observing conditions.

Table 2. Description of the five files included for each data set. The 'ID' in the filename corresponds to the ones provided in Tab. 1. N_{inj} is the number of injected planetary signals in the considered coronagraphic image cube.

Name	Content	Dimension
$image\_cube\_ID.fits$	Multispectral temporal coronagraphic image cube	4
$parallactic\_angles\_ID.fits$	Vectors of parallactic angles and airmass variation	2
$wavelength\_vect\_ID.fits$	Vector of the central wavelength of each channel of the IFU	1
$psf\_cube\_ID.fits$	Multispectral non-coronagraphic image cube of the target	3
$first\_guess\_astrometry\_ID.fits$	First guess position of the injected signals (within $1\lambda/D$)	$N_{inj} \times 2$

2.3 Injection procedure

The goal of this second phase is not to detect the companions but rather to recover an accurate position and spectrum of the injected companions. In each multispectral image cube, two to three synthetic planetary signals are injected at various locations corresponding to specific starlight residual noise distribution (e.g. at short or large separation, close to low wind effect or wind-driven halo features, etc.). In addition, the signal-to-noise ratio is chosen to assess how it impacts the estimations. At last, because the final goal of planetary spectrum retrieval is to estimate the atmospheric physical parameters of the planet, we injected planets following an emission spectra modeled with a radiative transfer code. In the following we describe the injection procedure that we used in practice. This injection procedure is based on functions included in the Vortex Image Processing (VIP¶) and the SPEctral Characterization of directly ImAged Low-mass companions (SPECIAL‖) open-source packages.[19–22]

2.3.1 Position of the injected signals

The position of the synthetic planetary signals is of subpixel precision. Fourier-transform based methods are used for shifts and rotations in order to avoid interpolation errors.[23] The reference position injected (and to be retrieved) is the signal location in the derotated image. In order to avoid any real astrophysical signal to interfere, we injected the synthetic planetary signals using the opposite parallactic angle variation: it preserves the temporal correlation of the starlight residuals while temporally smearing out any potential true astrophysical signal present in the image cube.

2.3.2 Spectrum models

As illustrated in Fig. 1, an error in the characterisation stemming from the post-processing has a non-negligible consequence on the physical interpretation of a given planetary emission spectrum. To assess and interpret the effect of the accuracy of the algorithms spectro-photometric estimations on the physical parameters, we injected spectra simulated with the radiative transfer tool petitRADTRANS[24, 25] that is available in open source**. We chose realistic physical parameters (L- or T-dwarf typical spectra, at the transition, cloudy etc.) and sometimes added contribution from circumplanetary material emission by adding a black-body contribution using SPECIAL, and sometimes a bit of noise. The latter also blends the advantages that some algorithms could have because they
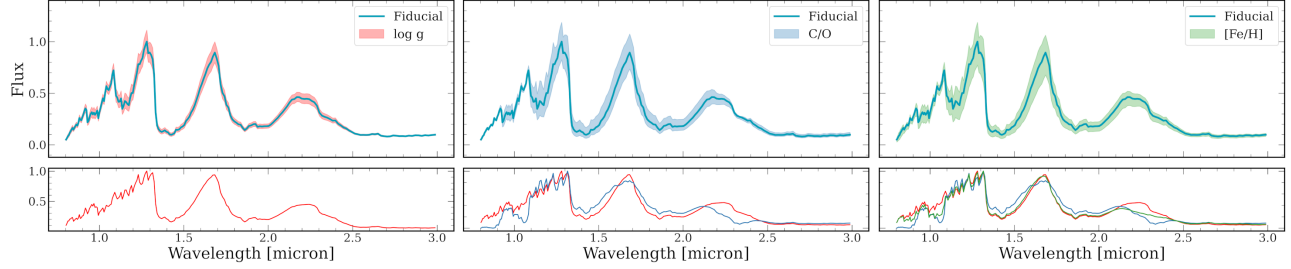
Figure 1. Typical T-dwarf spectrum obtained with the radiative transfer model PetitRadTrans accross the R- to L-band (blue line) for which key physical parameters are varied. The bottom panels focus on the absolute value of the difference between the highest and lowest part of the parametrized spectra. Left: the surface gravity, log(g), is varied around $3.5^{+0.1}_{-0.1}$ cm/s^2. Middle: the carbon-oxygen ratio, C/O, is varied around $0.5^{+0.1}_{-0.1}$. Right: the metallicity, [Fe/H], is varied around $0.5^{+0.3}_{-0.2}$ dex.

are taking into account priors on the template spectra for characterization. We remind here that the injected spectrum is unknown for the participants since it has to be estimated.

In order to assess the flux of the host star we performed a 2D-gaussian fit to the non-coronagraphic image cube (file *psf_cube_ID.fits*). We also took into account the airmass variation during the observing sequence such that the injected planet flux in frame t is scaled by $\epsilon_t = \exp(-(X_t - X_{\mathrm{med}}))$, where X_t is the airmass value when frame t was acquired, and X_{med} is the median airmass over the observation. Because the measured spectrum of the star is also affected by both instrumental and atmospheric transmissions, we additionally take these into account for the injection of the synthetic planetary signals. To do so, we use the *NextGen* SED models[26] for the different observed target stars of the challenge (see Tab. 3 for the assumed stellar parameters), convolved and resampled following the IFUs spectral resolving power and wavelength range, respectively, as implemented in SPECIAL.

Table 3. Stellar spectra used for each of the 8 data sets. The spectral type, the effective temperature (T_{eff}) and surface gravity (log(g)) are retrieved from the *NextGen* SED models of the 8 observed target stars.

ID	Spectral type	T_{eff} (K)	log(g)
gpi1	A3V	8.750	4.5
gpi2	A0V	9.600	4.5
gpi3	F2V	7.050	4.5
gpi4	G5V	5.660	4.5
sphere1	B9V	11.400	4.5
sphere2	K7V	4.000	4.5
sphere3	M0V	3.750	4.5
sphere4	A3V	8.750	4.5

The planetary signals are injected sequentially into the image cube. To choose the signal-to-noise of the injected planetary signals, we run an annular PCA[27,28] (as implemented in VIP) and extract the average contrast over all wavelengths that we define for the injection. Our injection procedure then takes as input: the multispectral coronagraphic image cube in which the injection is made, the normalized non-coronagraphic PSF used as planetary signal model, the parallactic angle and airmass variation through time, the subpixel position for the injection, the planetary spectrum and mean contrast of the injection, as well as the fitted stellar spectrum, the model SED stellar spectrum and finally the instrument spectral resolving power. The planet spectrum is first convolved using the spectral FWHM (inferred from the mean spectral resolution of the instrument), and then

---

¶https://github.com/vortex-exoplanet/VIP

‖https://github.com/VChristiaens/special

**https://gitlab.com/mauricemolli/petitRADTRANS

resampled to match the IFU wavelength sampling. This is also done with the stellar model SED, before taking the ratio with the measured stellar flux in order to estimate the atmospheric and instrumental transmission to be applied to the planet spectrum. The injected signal P_{inj} for planet i at a given wavelength λ, position (x_i, y_i) and timestamp t then writes:

$$P_{inj}(x_i, y_i, \lambda, t) = PSF(x_i, y_i, \lambda) \times S_i(\lambda) \times \overline{C_i} \times T_{atm+instr}(\lambda) \times \epsilon_t, \tag{1}$$

were not known to me (but not necessary for me to inject the spectra at given mean contrast levels) - just in case someone asks you where $PSF(x_i, y_i, \lambda)$ is the non-coronagraphic image of the target star at the wavelength λ placed at the position (x_i, y_i), $S_i(\lambda)$ is the chosen spectrum of the planet i (simulated by PetitRADTRANS and SPECIAL), $\overline{C_i}$ is the mean contrast of the planet i over all the wavelengths probed by the IFU, $T_{atm+instr}$ the transmission of the atmosphere and instrument (i.e. the measured stellar spectrum via Gaussian fit of the non-coronagraphic spectral images, divided by the convolved+resampled *NextGen* model SED), and ϵ_t the airmass factor for frame t.

2.3.3 Limitation to the injection procedure

For the injection procedure, we did not include the following:

- No diffraction effect from the coronagraph focal plane mask at close separation;

- No smearing effect at large separation (due to long integration time and field of view rotation);

- No temporal binning is included;

- No additional type of noise is considered on the injection;

- In particular, we did not include any other source of flux variability with time (neither intrinsic nor instrumental) apart from the airmass variation;

- No jitter (tip-tilt) off-centering the star behind the coronagraph focal plane mask during the exposure: the center of the image is assumed fixed over the whole observation sequence.

Note that according to Biller et al., 2021[29] the second last point is a correct assumption considering our temporal sampling. For the last point however, it artificially boosts the signal-to-noise ratio compared to reality where it is the major limitations when exploiting angular diversity.

2.4 Other available resources

The 8 data sets provided with blind injections are what we call the challenge set, for which the ground-truth remains private and from which the final score of a submission will be computed and displayed in the EvalAI platform leaderboard. Besides this challenge set, we additionally provide one training set consisting of an empty image data set (without injections) from the SPHERE-IFS instrument, along with a procedure to inject synthetic planetary signals. This training set is for the users to estimate hyper-parameters of algorithms when necessary and perform various tests before their final submission on the EvalAI platform. The ID of this empty data set is *sphere0*, a bright target star observed under average conditions, with a total field rotation of 26.8° (144 image frames in Y-J band).

On the dedicated Github repository, one can find a toolkit to help participants with processing and submission. Within this toolkit, a folder called *tutorial* contains the tutorial to perform the injections into the training data set (non-blind) following the steps described in Sect. 2.3. Another folder called *data* contains: (i) the stellar spectra used for the planet injections in each of the 9 data set provided, and (ii) examples of two planet spectra used for the example injection in the non-blind data set (*sphere0*). This tutorial contains a first part to visualise the input data and a last part to process the data and obtain a quick-look using full-frame PCA-ASDI[30,31] (using, in this case, both temporal and spectral dimensions and 10 principal components to build the model to be subtracted), using its implementation in the VIP package.

2.5 Illustration with the training dataset (sphere0)

For the rest of this paper, we consider the example of two injected planetary signals in the *training data set*, dubbed 'sphere0'. For this SPHERE-IFS data set, the observing conditions were average (mean seeing 0.99 ± 0.22 arcsec and mean coherence time 5.8 ± 1.3 ms) and the target star rather bright (magnitude $H_{mag} = 6.22$ and $V_{mag} \simeq 9.4$), providing with a Strehl ratio of $68 \pm 10\%$ in H-band. The median image of the cube at the shortest wavelength (0.957 μm) is shown in Fig. 2 (top-left panel), along with the location of the two injected planets.

The target was observed for about 1 hour, with an exposure time of 64 s, giving 65 images with a field rotation $\Delta_{field} = 22.2°$. The airmass variation during the observing time is shown in Fig. 2 (bottom-left panel). The model stellar SED used for calculation of the atmospheric and instrumental transmission (see Eq. 1) is a K7V spectral type, with an effective temperature $T_{\text{eff}} = 4000$ K and a surface gravity $log(g) = 4.5$ cm/m^2, whose shape is shown in Fig. 2 (bottom-right panel). The PSF flux extracted from Gaussian fits to the non-coronagraphic multispectral images is shown in Fig. 2 (bottom-central panel).

The planet 'b' signal is injected close to the star at a separation of 155.3 mas (position angle of 167.2°, measured counter-clockwise from the positive y axis in the image), while planet 'c' is injected outside of the AO correction zone, at a separation of 615.9 mas (PA of 122.3°). The model planet spectra considered for the two injections are shown in Fig. 2 (top-central panel). These correspond to a 1800 K black body (blue line) and the best-fit model to planet HR 8799 e (red line),[32] for 'b' and 'c' respectively. These spectra are also shown in Fig. 2 after convolution and resampling considering the SPHERE-IFS spectral resolving power and wavelength sampling (top-right panel). The actual spectra injected in the data cube further consider atmospheric+instrumental transmission, and the effect of airmass (see Eq. 1). The mean contrast over all the spectral channels chosen for the injections is set to 1.8×10^{-4} and 2.0×10^{-4} for planets b and c, respectively. After a simple full frame ASDI PCA processing, using 10 principal components to build the model PSF that is subtracted, it corresponds to Signal-to-Noise Ratio values of 5σ and 25σ in the post-processed image, respectively. In Sect. 4.3, we compare the results obtained on this data set with two algorithms.

3. EVALUATION AND RANKING PROCEDURE

With these 8 data sets (challenge set), there are two tasks to be completed by each participant. The first task is to accurately estimate the astrometry (distance to the star in pixels, using Cartesian coordinates) of all the injected exoplanet signals. The second task is to accurately estimate their spectra (contrast to the star, estimated at each wavelength). Due to the various output that every algorithm may provide, we have decided to focus on distance-based metrics for the leaderboard of the challenge platform EvalAI. This simple metric is detailed in this section.

3.1 First task: Astrometry

For this first task, the EvalAI leaderboard displays the distance (in the sense of the L2-norm, euclidian distance) between the estimated value and the ground truth value. The final result is the mean of this distance for all the 21 injections:

$$
\begin{aligned}
D_{astrometry} &= \frac{1}{N_{inj}} \cdot \sum_i |\hat{p}_i - p_i^{GT}|_2 \\
&= \frac{1}{N_{inj}} \cdot \sum_i \sqrt{(\hat{x}_i - x_i^{GT})^2 + (\hat{y}_i - y_i^{GT})^2},
\end{aligned}
\tag{2}
$$

with $\hat{p}_i$ the estimated position for the injection i, p_i^{GT} its corresponding ground truth, and N_{inj} the total number of injections. In this challenge the position of the planetary signals are to be provided as the number of pixels, in Cartesian coordinates, with respect to the star at the center of the image frame. The leaderboard displays this value, the closer to 0 the better.
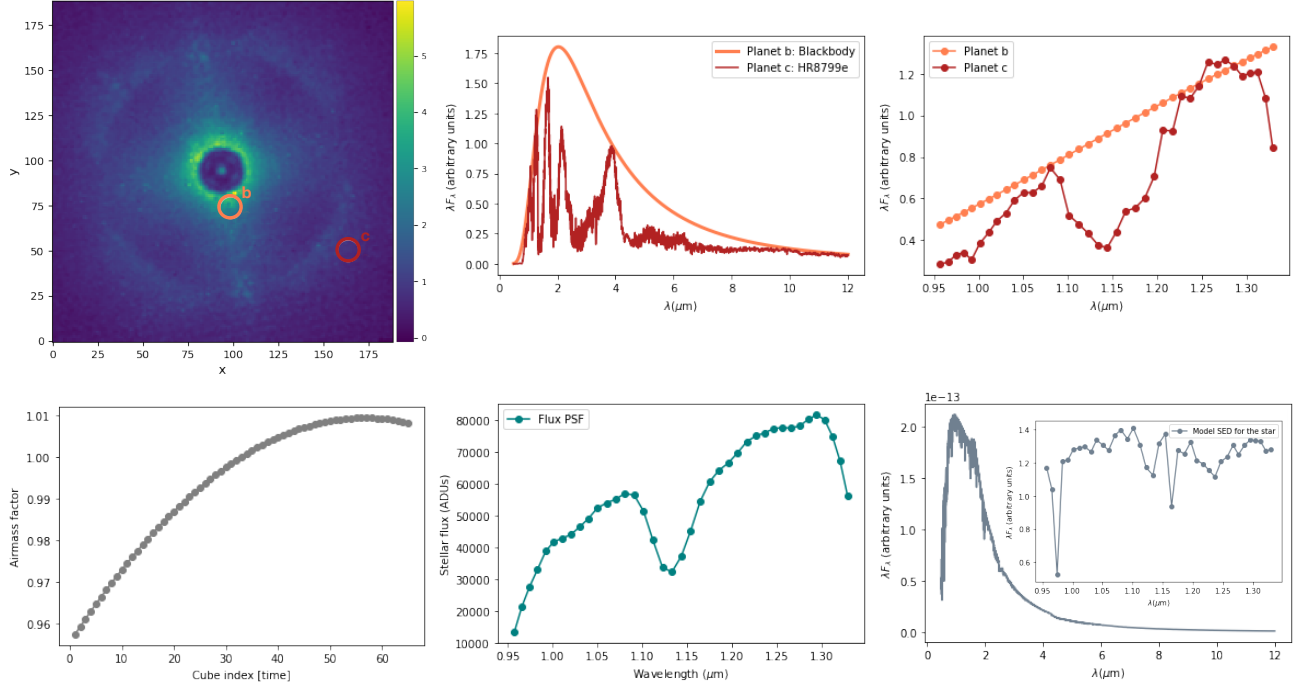
Figure 2. Example injection procedure for the training data set 'sphere0'. Top panel: Median of the image cube at the shortest wavelength with the position of the two injections 'b' and 'c' indicated by colored circles (left) and their injected spectra of the two planetary signals (middle), sampled at the SPHERE-IFS spectral resolving power (right). Bottom panel: Airmass factor used for the whole observing sequence (left); extracted spectrum of the non-coronagraphic PSF (middle); stellar SED model used for the target star (right), sampled at the SPHERE-IFS spectral resolving power.

3.2 Second task: Spectro-photometry

For this second task, the EvalAI leaderboard displays the distance (in the sense of the L1-norm) between the estimated value and the ground truth value for each spectral channel separately. We then average the results over all the spectral channels. The final result is the mean of this distance for all the 21 injections:

$$D_{photometry} = \frac{1}{N_{inj}} \cdot \frac{1}{N_\lambda} \cdot \sum_i \sum_\lambda \frac{|\hat{c}_\lambda - c_\lambda^{GT}|_1}{c_\lambda^{GT}}, \tag{3}$$

with $\hat{c}_\lambda$ the estimated contrast for the injection i at wavelength λ, c_λ^{GT} its corresponding ground truth, and N_{inj} the total number of injections and N_λ the total number of spectral channels of the IFU. The absolute value of the difference is normalized by the ground-truth contrast in order to consider the same weights for same relative differences in inferred contrasts (be it in shallow or high contrast parts of the spectra) in the final metric. The leaderboard displays this value, the closer to 0 the better.

3.3 Limitation to the evaluation procedure

This distance-based metric does not take into account two important aspects: (i) the uncertainties on the estimations and (ii) the difficulty of each retrieval per injection. We made this decision because some algorithms (e.g. algorithms based on supervised machine learning from the artificial intelligence field) do not provide any uncertainties on the estimations. In addition the uncertainties depend on an underlying probability density function that could either be estimated and provided by the participants or assumed Gaussian (whereas we know this assumption is often not respected[33]). Taking into account this uncertainty is therefore not trivial.

For the second point, we decided to ignore it in a first step as most algorithms should naturally adjust to it. For the next communication giving the complete results of this second phase during the year 2023, we plan to

use more evolved assessment methods taking into account those two aspects by e.g. weighting the grade with the signal-to-noise ratio and uncertainties (if any).

3.4 Submission guidelines

To participate, users must submit on the EvalAI platform ZIP files containing their predictions. To participate in both tasks, two ZIP files must be uploaded under the following name: *username_astrometry.zip* and *username_photometry.zip*. Each of these ZIP files will contain as many Multiple Extension FITS (MEF) files as data cubes to evaluate, i.e eight files per task. The name of each MEF .fits file must follow the format of *astrometry_ID.fits* and *photometry_ID.fits* (with 'ID' provided in Table 1). Each MEF fits file should contain three extensions, one for the estimated parameters, one for the uncertainties on these estimated parameters [optional] and one for posterior distributions over the estimated parameters [optional]. The second and third dimensions (uncertainties and posterior respectively) are optional input that will be used for further analyses in order to better understand the performance of a given algorithm.

In the Github repository dedicated to the second phase of the exoplanet imaging data challenge, a folder called *eidc2* contains the code that the EvalAI platform runs to compute the leaderboard ranking, as well as the code computing the metrics chosen for the leaderboard ranking. In the *tutorial* folder of the EIDC phase 2 Github repository, there is a tutorial to create a multi-extension *.fits* file.

4. RESULTS USING TWO METHODS: PCA-NEGFC AND ANDROMEDA

In this section, we show the results obtained with two techniques widely used in the community to infer the position and contrast of the two planets injected in the training data set (Sect. 2.5). We compare the estimated astrometry and contrasts to the ground truth, and apply the metrics defined in Secs. 3.1 and 3.1 to evaluate the score achieved by each method. In our previous paper,[7] we reviewed the main detection algorithms that currently exist and sorted them in four main families. Similarly, available techniques in the community for characterization of planetary signals in high-contrast images can be sorted into two broad families: (a) the fake-companion injection methods, and (b) inverse-problem methods.

The fake companion injection methods are based on the injection of synthetic planetary signals using the non-coronagraphic PSF of the star so as to make a forward model. Different implementations have been proposed. The imprint of the companion on the calculated principal components can be considered to build a forward model which captures the effects of over- and self-subtraction and is then directly fitted to the stellar PSF subtracted image (as e.g. in KLIP-FM[34]). Alternatively negative synthetic planetary signals can be injected in the image cube before applying the differential imaging technique, with the flux and position of the negative injection varied until the residuals in the post-processed images are minimized at the location of the detection (i.e. NEGFC[35,36]). In either cases, the minimization can be performed via statistical sampling methods, e.g. by running a Markov Chain Monte Carlo (MCMC) fitting algorithm[37] in order to estimate the parameters and their uncertainties.[38]

The inverse problem approach consists in building a model of the planetary signal (after differential imaging or without differential imaging) by using the normalized PSF (non-coronagraphic or unsaturated image of the star) and estimating the contrast of any planetary signal present in the field of view via a maximum likelihood estimation.[39] Algorithms such as ANDROMEDA,[40] FMMF,[41] PACO,[42] and RSM-FM[43] are part of this family.

4.1 Results with PCA-NEGFC

We used the simplex algorithm of the NEGFC implementation in VIP in order to find the optimal position and contrast relative to the star of the companions injected in the training data set.[21,44] More specifically we proceeded as follows. We obtained a first guess estimate of the injected position of the companions in a PCA-ASDI processed image (10 principal components used). We then inferred the optimal number of principal components which maximized the SNR at that position in each individual spectral channel, using the `pca_annulus` function of VIP – this function along with these optimal number of principal components are then used for all subsequent steps. Next, we considered the top 5 channels in terms of SNR for the extraction of the astrometry in a combined NEGFC involving 7 free parameters: 2 for the position and 5 for the individual fluxes in each of these spectral
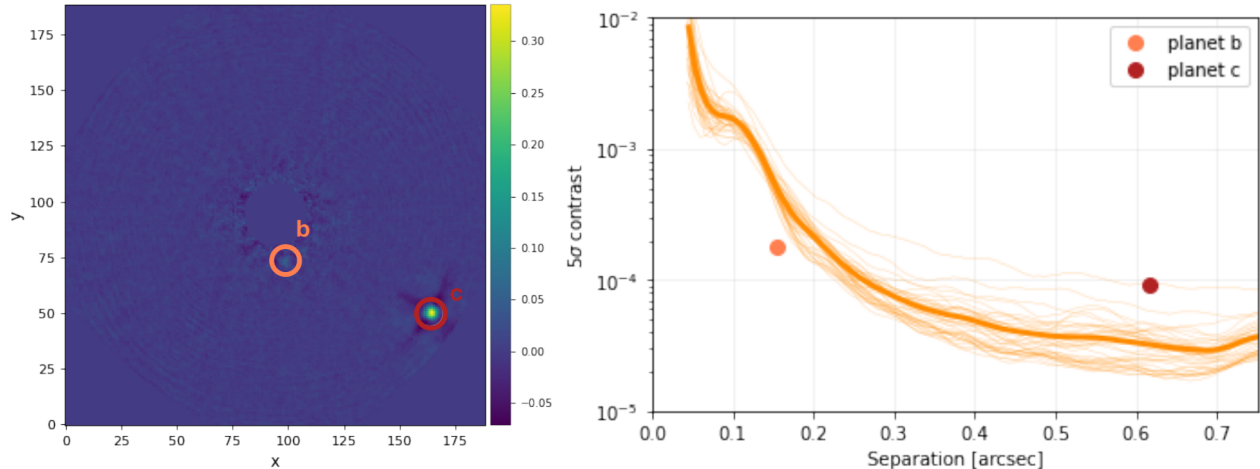
Figure 3. Results of the 'sphere0' data set processed with principal component analysis. Left: Final residual map obtained with PCA-ASDI using 10 principal components. Right: Detection limit as 5-sigma contrast curves obtained with PCA-ADI. The 39 thin lines are the contrast curve for each channel of the spectro-imager. The thick line is the mean of all the thin lines. The two dots show the radial position and mean contrast of the two injected planetary signals.

channels. Once the astrometry was inferred, we fixed the position of the companion and ran the NEGFC simplex algorithm to derive the flux of the companion individually in each of the 39 channels.

The astrometric uncertainties on 'b' and 'c' were estimated with the injection of 360 fake companions at the same radius as 'b' and 'c', and with the same estimated flux as 'b' and 'c', respectively. The injections were performed individually, 1 deg apart from each other, and the same parameters were used for the retrieval of their position and contrast using the NEGFC simplex algorithm as for 'b' and 'c'. We fitted a Gaussian to the distribution of deviations between the retrieved positions and the known injected positions, to infer the 1 sigma uncertainties on the radial separation and azimuth.[44] A similar method was adopted to infer the 1-sigma uncertainty on the contrast of 'b' and 'c' in each channel.

Figure 3 shows the final PCA-ASDI image (left) and the 5-sigma contrast curves obtained with PCA-ADI (10 PCs) in each spectral channel (right). The characterization results obtained PCA-NEGFC are shown in Fig 5 and Fig 6, where they are compared to the ground truth astrometry and contrast values, respectively (orange).

4.2 Results with ANDROMEDA

ANDROMEDA is based on an inverse problem approach:[39] it includes a model of the expected planetary signal (as the normalized non-coronagraphic image of the star) and fits its position and flux. A first step is to perform a pairwise subtraction of a couple of images for which the field has rotated enough to not subtract the planetary signal but that are temporally close in time so that the starlight residuals are still correlated enough. This first step allows on the one hand to whiten the residual starlight and on the other hand to create a specific signature if a planetary companion is present. The key idea of ANDROMEDA is to estimate the flux and position of any planetary signal in the field of view by using a maximum likelihood estimation, under the hypothesis that the residuals after the pairwise subtraction is white and Gaussian.[40]

In order to retrieve the astrometry of the two injected companions, we ran ANDROMEDA with the default parameters in SADI mode. This mode first performs a spectral differential imaging after rescaling a reference image at a different wavelength (set to the first spectral channel in our case) before running the classical ADI-based ANDROMEDA. ANDROMEDA provides one signal-to-noise ratio (SNR) map per spectral channel. The final detection map is obtained by summing the squared SNR maps and then taking its square root before normalizing it to obtain an equivalent SNR map. The astrometry is estimated directly on this latter detection map by performing a 2D-Gaussian fit on the signals that are above a threshold set to 5σ. The associated uncertainties are driven by the errors of the 2D-Gaussian fit performed on the planet detection signal. The spectro-photometry is estimated at the position of the planetary signals extracted on the final detection map,
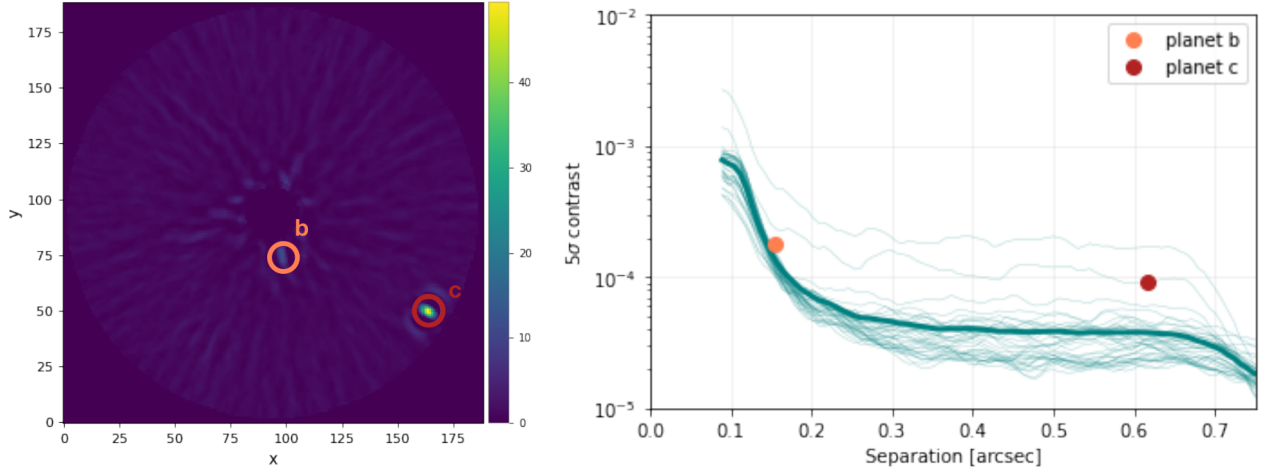
Figure 4. Results of the 'sphere0' data set processed using ANDROMEDA performing first a spectral image difference (after wavelength rescaling) and then an angular pairwise difference (SADI mode). Left: Final signal-to-noise ratio map on which the detection and astrometry estimation is performed. Right: Corresponding detection limit as 5-sigma contrast curves. The 39 thin lines are the contrast curve for each channel of the spectro-imager. The thick line is the mean of all the thin lines. The two dots show the radial position and mean contrast of the two injected planetary signals.

showing the highest SNR. The flux is read for each spectral channel directly on each map of the maximum likelihood estimation of the flux (classical ADI-based ANDROMEDA). The associated uncertainties are therefore driven by the signal-to-noise ratio of the planetary signal.

Figure 4 shows the final residual map (left) and the associated 5-sigma contrast curves (right). The SNR for the planet 'b' is of 10, and for the planet 'c' is 48. The characterization results obtained with ANDROMEDA are shown in Fig 5 and Fig 6 with respect to the ground truth (blue).

4.3 Comparison

We applied both PCA-NEGFC and ANDROMEDA on the training data of the data challenge (sphere0) with the injection described in Sect. 2.5. The results of the metrics are presented in Tab. 4. Note that when the planetary signal has not been detected, the first guess provided in the data can be used to extract the flux.

Table 4. Results of the astrometry and spectrophotometry metrics described in Sect. 3 obtained with PCA-NEGFC and ANDROMEDA for the two injections.

Method	PCA-NEGFC			ANDROMEDA		
	planet b	planet c	all	planet b	planet c	all
Astrometry	0.37	0.03	0.20	1.95	0.06	1.01
Spectro-photometry	16.98	1.03	9.00	35.67	2.29	18.99

The first injected planet 'b' is very close to the star. Current algorithms, based on ADI and/or SDI are limited in this region due to the poor amount of diversity and the larger starlight residuals. This explains why, as a general rule planet b has a higher metrics in both astrometry and spectrophotometry compared to planet c that is very bright and far from the star, in a region with diversity and less affected by starlight residuals.

Astrometry estimation

The results of the astrometry estimations analysis are shown in Fig. 5 for both planets (left and right). Note that the grid is 10 times smaller for the right panel (planet c) than the left panel (planet b).

In addition to being located at short separation, planet b lies on a bright speckle, which significantly affects the estimation by ANDROMEDA. Indeed the SNR spot at the location of planet b is quite elongated radially
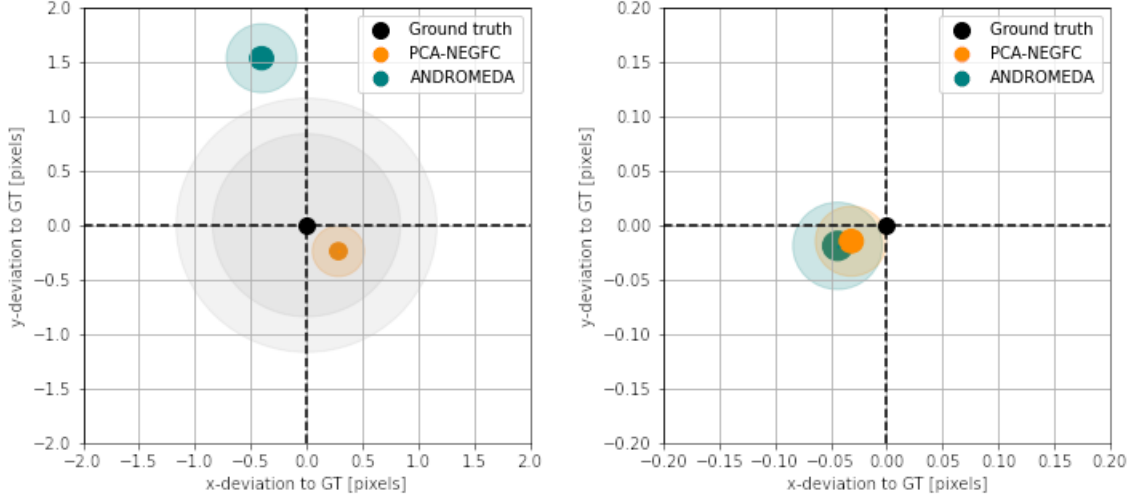
Figure 5. Astrometry estimations of the injected planets using PCA-NEGFC (orange) and ANDROMEDA (blue), relative to the ground truth (black) in Cartesian coordinates. The size of the symbols corresponds to the 1-sigma uncertainty (averaged in x- and y- direction) and the shaded area to the 3-sigma uncertainty. The two centered grey shaded areas represent the size of a half resolution element $(0.5\lambda/D)$ for the shortest and largest wavelengths of the spectro-imager. Left: Planet 'b' (155 mas). Right: Planet 'c' (616 mas).

(see Fig. 4, left), which biases the estimation towards a closer estimated separation (by about $7 - 8$ mas, which is one pixel). The NEGFC technique appears to be less affected by the bright speckle, possibly owing to the astrometry estimate being made only in the 5 channels with maximum SNR. However, for neither method does the ground truth fall within the 3-sigma uncertainties provided by the NEGFC method.

As for planet c, both methods provide estimations within $\sim$0.05 pixel from the ground truth, although it is worth mentioning that the associated uncertainties appear slightly underestimated in both cases.

Spectrophotometry estimation

The results of the photometry estimations analysis are shown in Fig. 6 for both planets (top and bottom). Note that the cuts are not the same for the two panels.

For planet b, we note that ANDROMEDA provides the correct trend but seems to be always slightly higher than the ground truth, in particular in the redder spectral channels (i.e. the residuals are flat with wavelength). This can again be explained by the presence of the bright speckle that lies below the injection and that crosses the injection location at larger wavelength. On the contrary, PCA-NEGFC slightly overestimates the bluer spectral channels, while extracting very accurate values at longer wavelengths (i.e. the residuals are decreasing with wavelength). PCA-NEGFC provide a rather flat spectrum. This can be understood by the residual minimization procedure performed by NEGFC that is more aggressive towards residual speckles. However, we can see that even though both methods give results within 5-sigma from the ground truth, this is still suboptimal as many points are not within 3-sigma (for both methods), demonstrating that there is still room for improvement.

For planet c, PCA-NEGFC provides an almost perfect estimate of the spectrum. ANDROMEDA provides slightly worst results at bluer wavelengths. This is probably because the starlight residuals distribution is more intense and shaped in this part of the spectrum and ANDROMEDA heavily relies on the assumption that the subtraction residuals are Gaussian and white, which we now know is not a perfect assumption.[33] Both methods have larger uncertainties at short wavelength due to the inherent shape of the spectrum.

In conclusion, we can already see that this comparison, based on the same dataset and the same injected signals, provides a lot of information about the fundamental limitations of the two methods used here, which are based on very different concepts. Having more data, for various instruments, under various observing conditions and for injections of varying complexity (in terms of contrast and localization) will hopefully provide us with
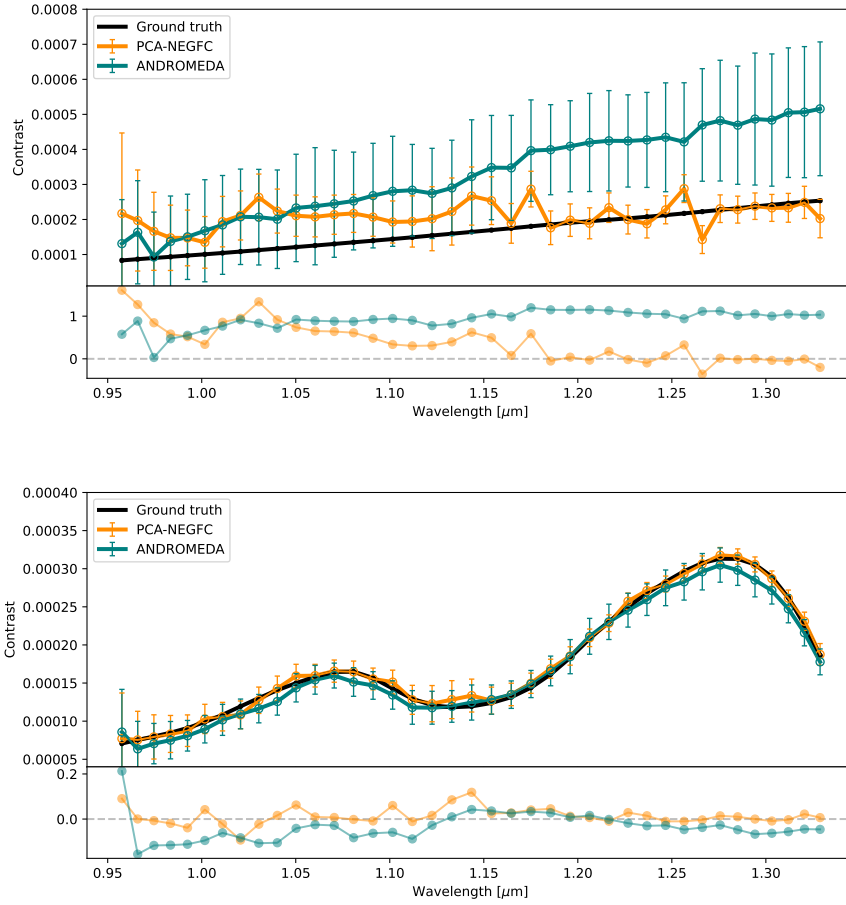
Figure 6. Spectro-photometry estimations of the injected planets using PCA-NEGFC (orange line) and ANDROMEDA (blue line), compared to the ground truth (black line), as well as the residuals (bottom panel). The error bars displayed correspond to the 3-sigma uncertainties on the photometry estimation. Top: Planet 'b' (Blackbody spectrum). Bottom: Planet 'c' (HR8799e-like spectrum).

valuable information on the functioning of the algorithms and their limitations. This work can therefore offer ways to improve the various methods that exist today and open avenues for the development of new methods.

5. CONCLUSION & PERSPECTIVES

Thanks to this second phase of the challenge, we hope to understand better the differences we often observe in the estimation of flux and position of the planet candidates, as a function of the chosen post-processing technique along with its user-parameters. This full exercise will provide us with an insight on the global error made due to the intrinsic limitations of all the algorithms and subsequently on how it affects the retrieval of physical parameters. We therefore highly encourage the high-contrast community to participate in the challenge.

For the next phase of the *Exoplanet Imaging Data Challenge*, we intend to focus on circumstellar disk imaging. Indeed, new dedicated methods were recently developed (e.g. MAYONNAISE,[45] NMF-DI,[46] REXPACO[47]) and classical methods adapted (e.g. iterative PCA,[48] NMF[49]) to this purpose. The goal of this third phase is to detect the circumstellar signal but also to restore accurately its total intensity distribution for structures such as rings, spirals and shadows. Within this phase, we intend to make multi-reference differential imaging (mRDI[50, 51]) technique possibles by using SPHERE and GPI data, for which a large collection of archival data is available (making sure the data set of the challenge is not included in the library).

As a fourth phase, we envision to focus on High Resolution Spectroscopy (HRS), with instruments such as VLT/SINFONI,[52] VLT/MUSE,[53] VLT/CRIRES+[54] and Keck/NIRCSPEC,[55] as techniques based on molecular mapping are currently under development.[56, 57] Another focus would be to simulate ELT high contrast images to see which post-processing technique is the most adapted and get ready with the image processing tools before first light.[58] As a conclusion, the *Exoplanet Imaging Data Challenge* community initiative has various tasks to carry out in order to support the community of exoplanetary system observers and image processing developers to make the most out of the data delivered by the latest generation of instruments.

ACKNOWLEDGMENTS

REFERENCES

[1] Madhusudhan, N., "Exoplanetary atmospheres: Key insights, challenges and prospects," *arXiv preprint arXiv:1904.03190* (2019).

[2] Mollière, P., Molyarova, T., Bitsch, B., Henning, T., Schneider, A., Kreidberg, L., Eistrup, C., Burn, R., Nasedkin, E., Semenov, D., et al., "Interpreting the atmospheric composition of exoplanets: sensitivity to planet formation assumptions," *arXiv preprint arXiv:2204.13714* (2022).

[3] Beuzit, J. L., Vigan, A., Mouillet, D., Dohlen, K., Gratton, R., Boccaletti, A., Sauvage, J. F., Schmid, H. M., Langlois, M., Petit, C., Baruffolo, A., Feldt, M., Milli, J., Wahhaj, Z., Abe, L., Anselmi, U., Antichi, J., Barette, R., Baudrand, J., Baudoz, P., Bazzon, A., Bernardi, P., Blanchard, P., Brast, R., Bruno, P., Buey, T., Carbillet, M., Carle, M., Cascone, E., Chapron, F., Charton, J., Chauvin, G., Claudi, R., Costille, A., De Caprio, V., de Boer, J., Delboulbé, A., Desidera, S., Dominik, C., Downing, M., Dupuis, O., Fabron, C., Fantinel, D., Farisato, G., Feautrier, P., Fedrigo, E., Fusco, T., Gigan, P., Ginski, C., Girard, J., Giro, E., Gisler, D., Gluck, L., Gry, C., Henning, T., Hubin, N., Hugot, E., Incorvaia, S., Jaquet, M., Kasper, M., Lagadec, E., Lagrange, A. M., Le Coroller, H., Le Mignant, D., Le Ruyet, B., Lessio, G., Lizon, J. L., Llored, M., Lundin, L., Madec, F., Magnard, Y., Marteaud, M., Martinez, P., Maurel, D., Ménard, F., Mesa, D., Möller-Nilsson, O., Moulin, T., Moutou, C., Origné, A., Parisot, J., Pavlov, A., Perret, D., Pragt, J., Puget,

P., Rabou, P., Ramos, J., Reess, J. M., Rigal, F., Rochat, S., Roelfsema, R., Rousset, G., Roux, A., Saisse, M., Salasnich, B., Santambrogio, E., Scuderi, S., Segransan, D., Sevin, A., Siebenmorgen, R., Soenke, C., Stadler, E., Suarez, M., Tiphène, D., Turatto, M., Udry, S., Vakili, F., Waters, L. B. F. M., Weber, L., Wildi, F., Zins, G., and Zurlo, A., "SPHERE: the exoplanet imager for the Very Large Telescope," **631**, A155 (Nov 2019).

[4] Claudi, R. U., Turatto, M., Gratton, R. G., Antichi, J., Bonavita, M., Bruno, P., Cascone, E., De Caprio, V., Desidera, S., Giro, E., Mesa, D., Scuderi, S., Dohlen, K., Beuzit, J. L., and Puget, P., "SPHERE IFS: the spectro differential imager of the VLT for exoplanets search," in [*Ground-based and Airborne Instrumentation for Astronomy II*], McLean, I. S. and Casali, M. M., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7014**, 70143E (July 2008).

[5] Macintosh, B. A., Graham, J. R., Palmer, D. W., Doyon, R., Dunn, J., Gavel, D. T., Larkin, J., Oppenheimer, B., Saddlemyer, L., Sivaramakrishnan, A., Wallace, J. K., Bauman, B., Erickson, D. A., Marois, C., Poyneer, L. A., and Soummer, R., "The Gemini Planet Imager: from science to design to construction," in [*Adaptive Optics Systems*], **7015**, 701518 (July 2008).

[6] Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., and Nadeau, D., "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique," **641**, 556–564 (Apr. 2006).

[7] Cantalloube, F., Gomez-Gonzalez, C., Absil, O., Cantero, C., Bacher, R., Bonse, M., Bottom, M., Dahlqvist, C.-H., Desgrange, C., Flasseur, O., et al., "Exoplanet imaging data challenge: benchmarking the various image processing methods for exoplanet detection," in [*Adaptive Optics Systems VII*], **11448**, 1027–1062, SPIE (2020).

[8] Fusco, T., Sauvage, J. F., Petit, C., Costille, A., Dohlen, K., Mouillet, D., Beuzit, J. L., Kasper, M., Suarez, M., Soenke, C., Fedrigo, E., Downing, M., Baudoz, P., Sevin, A., Perret, D., Barufolo, A., Salasnich, B., Puget, P., Feautrier, F., Rochat, S., Moulin, T., Deboulbé, A., Hugot, E., Vigan, A., Mawet, D., Girard, J., and Hubin, N., "Final performance and lesson-learned of SAXO, the VLT-SPHERE extreme AO: from early design to on-sky results," in [*Adaptive Optics Systems IV*], Marchetti, E., Close, L. M., and Vran, J.-P., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **9148**, 91481U (Aug. 2014).

[9] Bailey, V. P., Poyneer, L. A., Macintosh, B. A., Savransky, D., Wang, J. J., De Rosa, R. J., Follette, K. B., Ammons, S. M., Hayward, T., Ingraham, P., et al., "Status and performance of the gemini planet imager adaptive optics system," in [*Adaptive Optics Systems V*], **9909**, 291–305, SPIE (2016).

[10] Poyneer, L. A., Palmer, D. W., Macintosh, B., Savransky, D., Sadakuni, N., Thomas, S., Véran, J.-P., Follette, K. B., Greenbaum, A. Z., Ammons, S. M., et al., "Performance of the gemini planet imager's adaptive optics system," *Applied Optics* **55**(2), 323–340 (2016).

[11] Soummer, R., "Apodized Pupil Lyot Coronagraphs for Arbitrary Telescope Apertures," **618**, L161–L164 (Jan. 2005).

[12] Martinez, P., Dorrer, C., Carpentier, E. A., Kasper, M., Boccaletti, A., Dohlen, K., and Yaitskova, N., "Design, analysis, and testing of a microdot apodizer for the apodized pupil lyot coronagraph," *Astronomy & Astrophysics* **495**(1), 363–370 (2009).

[13] Savransky, D., Thomas, S. J., Poyneer, L. A., Dunn, J., Macintosh, B. A., Sadakuni, N., Dillon, D., Goodsell, S. J., Hartung, M., Hibon, P., et al., "Automated alignment and on-sky performance of the gemini planet imager coronagraph," in [*Ground-based and Airborne Instrumentation for Astronomy V*], **9147**, 1331–1341, SPIE (2014).

[14] Perrin, M. D., Ingraham, P., Follette, K. B., Maire, J., Wang, J. J., Savransky, D., Arriaga, P., Bailey, V. P., Bruzzone, S., Chilcote, J. K., De Rosa, R. J., Draper, Z. H., Fitzgerald, M. P., Greenbaum, A. r. Z., Hung, L.-W., Konopacky, Q., Macintosh, B., Marchis, F., Marois, C., Millar-Blanchaer, M. A., Nielsen, E., Rajan, A., Rameau, J., Rantakyro, F. T., Ruffio, J.-B., Ward-Duong, K., Wolff, S. G., and Zalesky, J., "Gemini Planet Imager observational calibrations XI: pipeline improvements and enhanced calibrations after two years on sky," in [*Ground-based and Airborne Instrumentation for Astronomy VI*], Evans, C. J., Simard, L., and Takami, H., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **9908**, 990837 (Aug. 2016).

[15] Maire, J., Perrin, M. D., Doyon, R., Artigau, E., Dunn, J., Gavel, D. T., Graham, J. R., Lafrenière, D., Larkin, J. E., Lavigne, J.-F., Macintosh, B. A., Marois, C., Oppenheimer, B., Palmer, D. W., Poyneer, L. A., Thibault, S., and Véran, J.-P., "Data reduction pipeline for the Gemini Planet Imager," in [*Ground-based and Airborne Instrumentation for Astronomy III*], McLean, I. S., Ramsay, S. K., and Takami, H., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7735**, 773531 (July 2010).

[16] Delorme, P., Meunier, N., Albert, D., Lagadec, E., Coroller, H. L., Galicher, R., Mouillet, D., Boccaletti, A., Mesa, D., Meunier, J.-C., et al., "The sphere data center: a reference for high contrast imaging processing," *arXiv preprint arXiv:1712.06948* (2017).

[17] Pavlov, A., Möller-Nilsson, O., Feldt, M., Henning, T., Beuzit, J.-L., and Mouillet, D., "SPHERE data reduction and handling system: overview, project status, and development," in [*Advanced Software and Control for Astronomy II*], **7019**, 701939 (July 2008).

[18] Wells, D. C., Greisen, E. W., and Harten, R. H., "FITS - a Flexible Image Transport System," **44**, 363 (June 1981).

[19] Gonzalez, C. A. G., Wertz, O., Absil, O., Christiaens, V., Defrère, D., Mawet, D., Milli, J., Absil, P.-A., Van Droogenbroeck, M., Cantalloube, F., et al., "Vip: Vortex image processing package for high-contrast direct imaging," *The Astronomical Journal* **154**(1), 7 (2017).

[20] Christiaens et al., V., Gomez Gonzalez, C. A., Farkas, R., Dahlqvist, C. H., Nasedkin, E., Milli, J., Absil, O., Ngo, H., Cantero, C., Rainot, A., Hammond, I., Cantalloube, F., and Vigan, A., "Vip: A python package for high-contrast imaging," *subm. to JOSS* (2022).

[21] Christiaens, V., Ubeira-Gabellini, M. G., Cánovas, H., Delorme, P., Pairet, B., Absil, O., Casassus, S., Girard, J. H., Zurlo, A., Aoyama, Y., Marleau, G. D., Spina, L., van der Marel, N., Cieza, L., Lodato, G., Pérez, S., Pinte, C., Price, D. J., and Reggiani, M., "A faint companion around CrA-9: protoplanet or obscured binary?," **502**, 6117–6139 (Apr. 2021).

[22] Christiaens et al., V., "Special: A python package for the spectral characterization of directly imaged low-mass companions," *subm. to JOSS* (2022).

[23] Larkin, K. G., Oldfield, M. A., and Klemm, H., "Fast Fourier method for the accurate rotation of sampled images," *Optics Communications* **139**, 99–106 (Feb. 1997).

[24] Mollière, P., Wardenier, J. P., van Boekel, R., Henning, T., Molaverdikhani, K., and Snellen, I. A. G., "petitRADTRANS. A Python radiative transfer package for exoplanet characterization and retrieval," **627**, A67 (July 2019).

[25] Mollière, P., Stolker, T., Lacour, S., Otten, G., Shangguan, J., Charnay, B., Molyarova, T., Nowak, M., Henning, T., Marleau, G.-D., et al., "Retrieving scattering clouds and disequilibrium chemistry in the atmosphere of hr 8799e," *Astronomy & Astrophysics* **640**, A131 (2020).

[26] Hauschildt, P. H., Allard, F., and Baron, E., "The NextGen Model Atmosphere Grid for 3000¡=T_{eff}¡=10,000 K," **512**, 377–385 (Feb. 1999).

[27] Soummer, R., Pueyo, L., and Larkin, J., "Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages," **755**, L28 (Aug. 2012).

[28] Amara, A. and Quanz, S. P., "PYNPOINT: an image processing package for finding exoplanets," **427**, 948–955 (Dec. 2012).

[29] Biller, B. A., Apai, D., Bonnefoy, M., Desidera, S., Gratton, R., Kasper, M., Kenworthy, M., Lagrange, A.-M., Lazzoni, C., Mesa, D., Vigan, A., Wagner, K., Vos, J. M., and Zurlo, A., "A high-contrast search for variability in HR 8799bc with VLT-SPHERE," **503**, 743–767 (May 2021).

[30] Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., and Marois, C., "Speckle Noise and the Detection of Faint Companions," **111**, 587–594 (May 1999).

[31] Sparks, W. B. and Ford, H. C., "Imaging Spectroscopy for Extrasolar Planet Detection," **578**, 543–564 (Oct. 2002).

[32] Nasedkin, E., "Hr8799," *Astronomy & Astrophysics* (in prep.).

[33] Pairet, B., "Gravity as a purely quantum effect," *arXiv e-prints* (Feb. 2019).

[34] Pueyo, L., "Detection and characterization of exoplanets using projections on karhunen–loeve eigenimages: Forward modeling," *The Astrophysical Journal* **824**(2), 117 (2016).

[35] Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., and Barman, T., "Images of a fourth planet orbiting hr 8799," *Nature* **468**(7327), 1080–1083 (2010).

[36] Lagrange, A.-M., Bonnefoy, M., Chauvin, G., Apai, D., Ehrenreich, D., Boccaletti, A., Gratadour, D., Rouan, D., Mouillet, D., Lacour, S., et al., "A giant planet imaged in the disk of the young star β pictoris," *Science* **329**(5987), 57–59 (2010).

[37] Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J., "emcee: the mcmc hammer," *Publications of the Astronomical Society of the Pacific* **125**(925), 306 (2013).

[38] Bottom, M., Kuhn, J., Mennesson, B., Mawet, D., Shelton, J. C., Wallace, J. K., and Serabyn, E., "Resolving the delta andromedae spectroscopic binary with direct imaging," *The Astrophysical Journal* **809**(1), 11 (2015).

[39] Mugnier, L. M., Cornia, A., Sauvage, J.-F., Rousset, G., Fusco, T., and Védrenne, N., "Optimal method for exoplanet detection by angular differential imaging," *Journal of the Optical Society of America A* **26**, 1326 (May 2009).

[40] Cantalloube, F., Mouillet, D., Mugnier, L., Milli, J., Absil, O., Gonzalez, C. G., Chauvin, G., Beuzit, J.-L., and Cornia, A., "Direct exoplanet detection and characterization using the andromeda method: Performance on vlt/naco data," *Astronomy & Astrophysics* **582**, A89 (2015).

[41] Ruffio, J.-B., Macintosh, B., Wang, J. J., Pueyo, L., Nielsen, E. L., De Rosa, R. J., Czekala, I., Marley, M. S., Arriaga, P., Bailey, V. P., et al., "Improving and assessing planet sensitivity of the gpi exoplanet survey with a forward model matched filter," *The Astrophysical Journal* **842**(1), 14 (2017).

[42] Flasseur, O., Denis, L., Thiébaut, E., and Langlois, M., "An unsupervised patch-based approach for exoplanet detection by direct imaging," in [*2018 25th IEEE international conference on image processing (ICIP)*], 2735–2739, IEEE (2018).

[43] Dahlqvist, C.-H., Louppe, G., and Absil, O., "Improving the rsm map exoplanet detection algorithm-psf forward modelling and optimal selection of psf subtraction techniques," *Astronomy & Astrophysics* **646**, A49 (2021).

[44] Wertz, O., Absil, O., Gómez González, C. A., Milli, J., Girard, J. H., Mawet, D., and Pueyo, L., "VLT/SPHERE robust astrometry of the HR8799 planets at milliarcsecond-level accuracy. Orbital architecture analysis with PyAstrOFit," **598**, A83 (Feb. 2017).

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

[46] Ren, B., Pueyo, L., Chen, C., Choquet, É., Debes, J. H., Duchêne, G., Ménard, F., and Perrin, M. D., "Using data imputation for signal separation in high-contrast imaging," *The Astrophysical Journal* **892**(2), 74 (2020).

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

[48] Pairet, B., Cantalloube, F., and Jacques, L., "Reference-less algorithm for circumstellar disks imaging," *arXiv preprint arXiv:1812.01333* (2018).

[49] Ren, B., Pueyo, L., Zhu, G. B., Debes, J., and Duchêne, G., "Non-negative matrix factorization: robust extraction of extended structures," *The Astrophysical Journal* **852**(2), 104 (2018).

[50] Lafrenière, D., Marois, C., Doyon, R., and Barman, T., "Hst/nicmos detection of hr 8799 b in 1998," *The Astrophysical Journal* **694**(2), L148 (2009).

[51] Ruane, G., Mawet, D., Kastner, J., Meshkat, T., Bottom, M., Castellá, B. F., Absil, O., Gonzalez, C. G., Huby, E., Zhu, Z., et al., "Deep imaging search for planets forming in the tw hya protoplanetary disk with the keck/nirc2 vortex coronagraph," *The Astronomical Journal* **154**(2), 73 (2017).

[52] Eisenhauer, F., Abuter, R., Bickert, K., Biancat-Marchet, F., Bonnet, H., Brynnel, J., Conzelmann, R. D., Delabre, B., Donaldson, R., Farinato, J., et al., "Sinfoni: integral field spectroscopy at 50-milli-arcsecond resolution with the eso vlt," in [*Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*], **4841**, 1548–1561, SPIE (2003).

[53] Bacon, R., Accardo, M., Adjali, L., Anwand, H., Bauer, S., Biswas, I., Blaizot, J., Boudon, D., Brau-Nogue, S., Brinchmann, J., et al., "The muse second-generation vlt instrument," in [*Ground-based and Airborne Instrumentation for Astronomy III*], **7735**, 131–139, SPIE (2010).

[54] Follert, R., Dorn, R., Oliva, E., Lizon, J., Hatzes, A., Piskunov, N., Reiners, A., Seemann, U., Stempels, E., Heiter, U., et al., "Crires+: a cross-dispersed high-resolution infrared spectrograph for the eso vlt," in [*Ground-based and Airborne Instrumentation for Astronomy V*], **9147**, 476–485, SPIE (2014).

[55] McLean, I. S., Becklin, E. E., Bendiksen, O., Brims, G., Canfield, J., Figer, D. F., Graham, J. R., Hare, J., Lacayanga, F., Larkin, J. E., et al., "Design and development of nirspec: a near-infrared echelle spectrograph for the keck ii telescope," in [*Infrared astronomical instrumentation*], **3354**, 566–578, SPIE (1998).

[56] Snellen, I. A., Brandl, B. R., De Kok, R. J., Brogi, M., Birkby, J., and Schwarz, H., "Fast spin of the young extrasolar planet β pictoris b," *Nature* **509**(7498), 63–65 (2014).

[57] Rameau, J., Chanussot, J., Carlotti, A., Bonnefoy, M., and Delorme, P., "Spectral unmixing for exoplanet direct detection in hyperspectral data," *Astronomy & Astrophysics* **649**, A143 (2021).

[58] Houllé, M., Vigan, A., Carlotti, A., Choquet, É., Cantalloube, F., Phillips, M. W., Sauvage, J.-F., Schwartz, N., Otten, G. P., Baraffe, I., et al., "Direct imaging and spectroscopy of exoplanets with the elt/harmoni high-contrast module," *arXiv preprint arXiv:2104.11251* (2021).