

Exoplanet Imaging Data Challenge, phase II: Comparison of algorithms in terms of characterization capabilities

Cantalloube F.^a, Christiaens V.^b, Cantero C.^c, Cioppa A.^d, Nasedkin E.^e, Absil O.^b,
Delorme P.^a, Wang J.-J.^{f,g}, Bonse J. M.^h, Daglayan H.ⁱ, Dahlgqvist C.-H.^b, Guyot N.^b,
Juillard S.^b, Mazoyer J.^j, Samland M.^e, Sabalbal M.^b, Ruffio J.-B.^k, and
Van Droogenbroeck M.^b

^aUniv. Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France

^bSpace Sciences, Technologies & Astrophysics Research (STAR) Institute, Université de Liège,
Allée du Six Août 19c, B-4000 Liège, Belgium

^cGeneva Observatory, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland

^dMontefiore Institute, Université de Liège, 4000 Liège, Belgium

^eMax-Planck-Institut für Astronomie, Königstuhl 17, Heidelberg 69117, Germany

^fCenter for Interdisciplinary Exploration and Research in Astrophysics (CIERA),
Northwestern University, 1800 Sherman Ave, Evanston, IL, 60201, USA

^gDepartment of Physics and Astronomy, Northwestern University, 2145 Sheridan Rd,
Evanston, IL 60208, USA

^hInstitute for Particle Physics and Astrophysics, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093,
Zurich, Switzerland

ⁱICTEAM Institute, UCLouvain, Louvain-la-Neuve, Belgium

^jLESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de
Paris, 5 place Jules Janssen, F-92195 Meudon, France

^kDepartment of Astronomy Astrophysics, University of California, San Diego, La Jolla, CA
92093, USA

ABSTRACT

In this communication, we report on the results of the second phase of the *Exoplanet Imaging Data Challenge* started in 2019. This second phase focuses on the characterization of point sources (exoplanet signals) within multispectral high-contrast images from ground-based telescopes. We collected eight data sets from two high-contrast integral field spectrographs (namely Gemini-S/GPI and VLT/SPHERE-IFS) that we calibrated homogeneously and in which we injected a handle of synthetic planetary signals to be characterized by the data challenge participants (ground truth). The tasks of the participants consist of (1) extracting the precise astrometry of each injected planetary signals, and (2) extracting the precise spectro-photometry of each injected planetary signal. Additionally, the participants may also provide with the 1-sigma uncertainties on their estimation for further analyses. When available, the participants can also provide the posterior distribution used to estimate the position/spectrum and uncertainties. The data are permanently available on a Zenodo repository and the participant can submit their results through the EvalAI platform. The EvalAI submission platform opened on April 2022 and closed on the 31st of May 2024. In total, we received 4 valid submissions for the astrometry estimation and 4 valid submissions for the spectrophotometry (each submission, corresponding to one pipeline, has been submitted by a unique participant). In this communication, we present an analysis and interpretation of the results comparison.

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

Further author information: (Send correspondence to F.C.)
F.C.: E-mail: faustine.cantalloube@univ-grenoble-alpes.fr

1. INTRODUCTION

The *Exoplanet Imaging Data Challenge* is an initiative led by several members of the exoplanet high-contrast imaging community. The goal of this initiative is to benchmark and compare the various existing post-processing techniques that are dedicated to exoplanet direct detection via high-contrast imaging. Indeed, the last ten years have seen the implementation of a number of advanced image processing techniques whose goals is to further carve out the starlight residuals in the images so as to reach the highest contrast possible, especially at close angular separation from the star (where most planets are expected), towards detecting faint planetary signals.

Current high-contrast images delivered by state-of-the-art instruments are affected by bright starlight residuals hindering the presence of exoplanets. These residuals come either from adaptive optics residuals¹ (usually short-lived so showing as smooth structures), low order residuals such as jitter and low wind effect² (usually varying fast and affecting the inner region), non-common path aberrations³ (fixed aberrations to calibrate, internal turbulence showing smooth structures and quasi-static aberrations), and any other error related to the detection procedure (e. g. bad pixels, miscalibration). The ultimate limit is given by the photon noise of the target star residuals. Usually, before applying dedicated post-processing technique to gain one or two orders of magnitude in contrast, we apply some *pre-processing* (also called cosmetic, or calibration) including the dark subtraction, flat fielding, distortion correction, wavelength calibration, centering* etc. The task of post-processing techniques is to deal with the effect of all the residual optical aberrations, which distribute starlight across the field of view on different spatial and temporal scales,⁴ in order to identify and characterize circumstellar signals.

The *Exoplanet Imaging Data Challenge* (EIDC) is organized in several phases, and is evolving to include more features and adapt to the latest findings. The first phase⁵ (2020), was focused on the detection of planetary signals in the field of view. This second phase is focused on the characterization of planetary signals. Upcoming phases will include extended features, the use of high resolving power spectroscopy, the exploitation of large data base and/or multi-epoch, telemetry and/or environmental data, and simulated images for the Extremely Large Telescope suites of instrument. The EIDC working group plans to publish its work and various results regularly, in particular every two years in the form of an *SPIE Astronomical Telescope + Instrumentation proceeding*.

Data from the various phases of the *Exoplanet Imaging Data Challenge* are permanently available on a Zenodo repository, so that anyone can use the ensemble of curated data to test and compare the capabilities of different algorithms. We are also providing analysis tools, including the metrics and comparison set of the corresponding phase. The *Exoplanet Imaging Data Challenge* intends to make the process of testing new algorithms more straightforward and robust, via the use of standard metrics and benchmark data set offered to the community. People are more than welcome to use the various resources from the *Exoplanet Imaging Data Challenge* platform for test and scientific publication (with due acknowledgement).

The data offered for this phase of the data challenge as well as the metrics used for the leaderboard are previously described⁶ in a 2022 SPIE proceeding. More information about this community-led initiative can be found on the dedicated website[†] (context, data, submission procedure, results and extended bibliography).

2. EXOPLANET CHARACTERIZATION WITH HIGH-CONTRAST IMAGING

Direct imaging of exoplanets, among other techniques, offers a unique window on planetary systems: (1) first, it provides a full picture of the system to study its architecture (interaction with circumstellar disks, multiplicity etc.); (2) second, the brightness estimation of the substellar companions provide with estimates of planet properties, such as its mass (from the mass luminosity relationships, linked to evolutionary track models), and if multi-wavelength images are additionally offered, spectral forward models and spectra retrievals unveil information on the atmospheric structure and composition; (3) precise astrometric follow-up may provide orbital constraints that can help us understanding the formation and evolution processes at stake such as migration, scattering and resonances, as well as dynamical mass estimate. The target dataset detection limit is computed and used to set statistical constraints on the occurrence rate of exoplanets of a given mass, at a given separation, around this type of target star. With current instruments, direct imaging of exoplanets is sensitive to the population of young giant gaseous planets at relatively large orbit that appear to be quite rare.^{7,8}

*See e.g. the review talk summarizing the importance of these steps <https://www.youtube.com/watch?v=0v4q6VN4b9M>.

†<https://exoplanet-imaging-challenge.github.io/>

In terms of image processing, the characterization of a planetary signal detection consists in extracting two observables: (I) the *astrometry*, which is the estimated position of the planetary companion(s) in the image, and (II) the *spectro-photometry*, which is the estimated luminosity in terms of contrast of the planetary companion(s) in the images. From these two estimates, observers can analyze and interpret the physical nature of the planet and, on larger scales, infer the formation and evolution processes of planetary systems. Obtaining accurate and precise observable is therefore of acute importance in the science of exoplanets.

2.1 Reminder on ADI post-processing techniques

As a reminder the most common method used for processing high-contrast images is based on differential imaging (DI), in particular making use of the angular diversity⁹ offered by pupil-tracking observations from telescopes having an alt-az mount. Indeed, in such a configuration, the orientation of the telescope pupil (where most optical aberrations arise) is kept along a given direction in the image cube, whereas the orientation of the field of view in the focal plane rotates with the parallactic angles as the Earth rotates during the observation sequence. In the resulting temporal cubes of images, the starlight residuals (nuisance term, also called stellar glare) is kept more or less stable in the field of view, while the off-axis astrophysical objects rotate around the image center (where the target star is located) with the parallactic angle during the observation.

Differential Imaging can be summarized in four steps: (1) estimating the starlight distribution in the field of view, also called *reference PSF* or *model PSF*; (2) Subtracting it to the science image(s), also called *PSF subtraction*; (3) repeat the subtraction procedure for each frame of the image cube then rotate the frames in order to align the circumstellar signals and combine them to obtain a *residual map*; (4) build a *detection map* (e.g. an SNR map) giving the probability of detecting a signal in the image field of view[‡]. The essential differences between current DI-based techniques are focused on the step (1), critical to obtaining an effective stellar glare subtraction. For instance, classical ADI⁹ consists in using the temporal median of the cube as the reference PSF. Another widely used method is LOCI,¹⁰ which consists in finding the linear combination of images that minimizes the variance of the residuals between the reference PSF and the science image. At last, another mainstream method is to make a principal component analysis (PCA) of the data cube to build the reference PSF as the sum of the first few principal components.

2.2 Characterization methods with ADI post-processing techniques

As of today, there are mainly two ways of characterizing exoplanets detected within high-contrast images processed with ADI-based techniques: (i) the injection of Negative Fake Companion^{11,12} (NEGFC) in the data set, and (ii) the Forward Modeling (FM) approaches.^{13,14}

The NEGFC technique consists in injecting a synthetic planetary signal (i.e. the non-coronagraphic / non-saturated PSF of the instrument, usually acquired during calibration before or after the coronagraphic / saturated image cube) with negative intensity in the raw image cube, at the rough position of the detection and with a rough flux estimate of the detection. The image cube with the negative injection is then processed with exactly the same user-parameters as for the planetary detection (e.g. with PCA, using exactly the same number of principal components to build the PSF model to be subtracted to the images). Using various methods (such as Nelder-Mead simplex algorithm, MCMC sampler or nested-sampling), this procedure is repeated with various positions and fluxes (around the first guess) for the negative injection, until the residuals are minimized at the location of the planet, providing with the astrometry (position of the negative injection that minimizes the residuals) and the photometry (intensity of the negative injection that minimizes the residuals).

The FM technique usually models the planetary signature (again as the non-coronagraphic / non-saturated PSF of the instrument) and tracks it in the data within an inverse problem framework. At the exception of the PACO algorithm,¹⁵ which does not perform any subtraction, the FM technique is done after step (2). It means that under a certain assumption of the residual noise distribution in the residual frames (2) - and in the case of PACO, the raw frames at (1) - the model of the planetary signature is fitted to the actual data. By construction, as the flux of the model planetary signature is the optimized parameter, the SNR map, used as a detection map, is a by-product (defined as $\hat{\phi}/\sigma(\hat{\phi})$, with $\hat{\phi}$ the estimated flux and $\sigma(\hat{\phi})$ the uncertainties of the estimated flux).

[‡]For a review of post-processing technique and concepts of ADI, please visit our website <https://exoplanet-imaging-challenge.github.io/biblio/>.

In the previous EIDC publication,⁶ we compared these two basic approaches on the so-called *training dataset* (dubbed *sphere0*): one using a PCA-SADI procedure, as implemented in the VIP package,¹⁶ followed by the NEGFC approach, and the other using ANDROMEDA¹⁷ in its classical flavor (using a L2-norm minimization because the noise distribution is assumed to be white and Gaussian). In the following, we consider the PCA-SADI+NEGFC procedure as the baseline to analyze and compare the results, which includes the uncertainties (under the assumption of a Gaussian distribution of the residuals). For this baseline, the uncertainties are estimated by injecting hundreds of synthetic planetary signals in the raw images at the same angular separation as the considered one, estimate their parameters, and then fitting a Gaussian to the deviation between the estimated parameters and the injected ones (ground-truth).

We received a total of 4 valid submissions on the *EvalAI* platform, from four different participants. None of the participants included neither the uncertainties nor the posterior in their results. The four submissions were obtained using: **RSM**, the Regime Switching Model,¹⁸ implemented with a forward modeling approach, following the description in Dahlqvist et al.¹⁹ (2021); **pyKLIP-FM**, the Karhunen-Loève Image Processing,²⁰ as implemented in the pyKLIP²¹ pipeline, using Forward Modeling for the characterization described in Pueyo et al.²² (2016); **AMAT**, the Alternating Minimization Algorithm with Trajectory,²³ which uses a L1-norm low rank approximation to enhance the PCA-based PSF subtraction, associated to a forward model approach similar to ANDROMEDA but under the assumption of a Laplacian residual noise distribution (so using a L1-norm minimization to estimate the flux of the planet signature) that corresponds better to the actual distribution of the residuals after the subtraction; **ANDROMEDA**, in its classical¹⁷ ADI flavour, that is to say using a Gaussian residual assumption for the flux estimation and running every wavelength cube separately to combine them at the end (note that we only received the astrometry estimations from the latter but no spectrophotometry).

3. RESULTS: ASTROMETRY

As a reminder, for the astrometry estimation of the Exoplanet Imaging Data Challenge second phase, we asked the participants to provide with (1) the estimated position of each injected planetary signals in the provided data set, (2) optionally the corresponding 1σ uncertainties of the estimations, and (3) optionally the corresponding posterior distribution used to estimate the position and its uncertainties. As for the metric used for the ranking of the submitted results, the EIDC phase 2 working group decided to compute the distance (in the sense of the L2-norm, euclidian distance) between the estimated value and the ground truth value. All the details about the provided data sets, submission process and chosen metrics for comparison and ranking can be found in our previous SPIE publication Cantalloube et al.,²⁴ (2022).

3.1 Astrometry: baseline results as a function of the dataset

The provided data sets come from two different instruments (namely GPI²⁵ and SPHERE-IFS²⁶). For each instrument, we provided 4 data sets under very different observing conditions (see previous publication and website for more details about it). As a first qualitative interpretation, we compare the astrometry estimated by the baseline algorithm for each data set. The astrometry is inferred with PCA-ADI in individual channels (in the top 5 channels in terms of SNR) at the first guess location found with PCA-SADI. The results on the GPI images are shown in Fig. 1 and on the SPHERE-IFS images in Fig. 2 (actual position of the injections in the images, residuals map, SNR map and astrometry error to the ground-truth). Quantitative results, including the comparison metric used, are presented in Tab. 1 and Tab. 2 for the GPI and SPHERE-IFS data respectively.

For both GPI and SPHERE-IFS, the observing conditions have an important impact on the quality and accuracy of the estimated astrometry. As expected, the baseline SNR of the injection also has an impact on both the estimation and their uncertainties (being larger with lower SNR). The detections close to the coronagraphic inner working angle (IWA) usually introduce a small bias in the estimation, except in the case of excellent observing conditions. When the low wind effect (LWE) is strong, as in the *sphere2* data set, we observe a small bias in the estimated astrometry for the injection located on the diffraction patterns due to the LWE; with the exception of the one injection 'c' located in the middle of a diffraction pattern. Overall, the PCA-ASDI+NEGFC technique provides excellent results: the ground-truth falls within the 1σ uncertainties for about 50% of the injections, and always within the 3σ uncertainties except for 2 injections (*sphere2* data set affected by LWE). The astrometry estimated for the 21 injected planetary signals includes the ground truth within the 5σ uncertainties. In addition, the estimated astrometry is always within one resolution element from the ground-truth.

Table 1. Results of the baseline method (PCA-ASDI+NEGFC) on the four GPI data sets, for each injection. The color of the data set ID (first column) indicates the observing conditions as good (green), medium (orange) and bad (red). See the website for more details on the data. The *Injection* column corresponds to the injections as labeled in the figures all along this paper. For each injected planetary signal, SNR_{bsl} , δ_{bsl} , and θ_{bsl} , are respectively the signal-to-noise ratio, the angular separation, and the position angle estimated with the baseline method. The *Location* column indicates qualitatively where the injection is placed in the field of view. D_{astro}^{GT} is the metric chosen for the astrometry (L2-norm distance between the estimation and the ground-truth). D_{photo}^{GT} is the metric chosen for the spectrophotometry (normalized L1-norm distance between the estimation and the ground-truth.)

| ID | Injection | SNR_{bsl} | δ_{bsl} [mas] | θ_{bsl} [deg] | Location | D_{astro}^{GT} | D_{photo}^{GT} |
|------|-----------|--------------------|----------------------|----------------------|---------------------------|------------------|------------------|
| gpi1 | b | 8.3 | 265.58 ± 0.11 | 94.28 ± 0.033 | Coronagraph IWA | 0.004 | 0.559 |
| | c | 5.5 | 486.75 ± 3.80 | -154.17 ± 0.17 | sitting in bright speckle | 0.354 | 0.423 |
| | d | 32.3 | 978.50 ± 0.52 | 132.22 ± 0.06 | Satellite spot | 0.053 | 0.023 |
| gpi2 | b | 17.0 | 436.66 ± 1.11 | 26.20 ± 0.06 | AO dark-hole | 0.099 | 0.067 |
| | c | 5.2 | 166.69 ± 4.18 | -179.07 ± 1.65 | Coronagraph IWA | 0.609 | 0.209 |
| gpi3 | b | 10.0 | 284.32 ± 0.60 | -63.24 ± 0.21 | AO dark-hole, close | 0.043 | 0.073 |
| | c | 4.9 | 144.89 ± 1.53 | 160.24 ± 0.92 | Coronagraph IWA | 0.138 | 0.18 |
| | d | 47.7 | 793.40 ± 0.23 | -168.57 ± 0.02 | AO dark-hole, far | 0.004 | 0.013 |
| gpi4 | b | 16.1 | 645.06 ± 0.23 | 12.95 ± 0.02 | AO dark-hole | 0.024 | 0.212 |
| | c | 4.2 | 151.22 ± 1.40 | -153.84 ± 0.25 | Coronagraphic IWA | 0.134 | 3.003 |

Table 2. Results of the baseline method (PCA-ASDI+NEGFC) on the four SPHERE-IFS data sets, for each injection. The *sphere3* data set is taken under good observing conditions but the target star is faint. See description of Tab. 1.

| ID | Injection | SNR_{bsl} | δ_{bsl} [mas] | θ_{bsl} [deg] | Location | D_{astro}^{GT} | D_{photo}^{GT} |
|---------|-----------|--------------------|----------------------|----------------------|-------------------------|------------------|------------------|
| sphere1 | b | 4.8 | 428.47 ± 2.22 | 56.63 ± 0.09 | AO dark-hole, far | 0.0351 | 1.618 |
| | c | 7.7 | 144.34 ± 0.51 | -18.59 ± 0.34 | Coronagraph IWA | 0.0504 | 0.028 |
| | d | 11.7 | 185.91 ± 0.94 | -149.43 ± 0.02 | AO dark-hole, close | 0.0503 | 0.023 |
| sphere2 | b | 4.6 | 137.17 ± 0.97 | 46.37 ± 0.01 | Low Wind Effet, close | 0.4398 | 1.01 |
| | c | 11.0 | 218.54 ± 0.80 | -41.84 ± 0.42 | Low Wind Effet, middle | 0.0487 | 0.542 |
| | d | 6.8 | 291.85 ± 0.79 | -116.63 ± 0.09 | Low Wind Effet, further | 0.4801 | 0.411 |
| sphere3 | b | 3.2 | 161.59 ± 2.08 | 130.86 ± 1.47 | Coronagraph IWA | 0.0741 | 0.812 |
| | c | 7.0 | 456.35 ± 1.01 | -79.78 ± 0.18 | AO ring of fire | 0.1083 | 4.282 |
| sphere4 | b | 6.5 | 235.62 ± 0.81 | -21.08 ± 0.34 | AO dark-hole close | 0.1413 | 0.226 |
| | c | 2.0 | 430.10 ± 0.32 | 119.28 ± 0.08 | AO dark-hole far | 0.0130 | > 1000 |
| | d | 5.9 | 162.39 ± 2.03 | -129.55 ± 0.10 | Coronagraph IWA | 0.2661 | 8.549 |

3.2 Astrometry: results as a function of the submitted algorithms

The results for the baseline method and the four submitted methods are presented in Tab. 3 (GPI datasets) and Tab. 4 (SPHERE-IFS datasets). For each dataset and injection, we show the chosen metric⁶ as the euclidean distance between the estimation and the ground-truth. The last column averages the results over all injections to show an overall quantity for each method. A graphical view of the results is presented in Fig. 5 and Fig. 6 for the GPI and SPHERE-IFS data respectively.

The presented numbers and appendix figures lead to the following conclusions:

- Overall, higher SNR lead to a better estimation of the position, as expected;
- Apart of ANDROMEDA, the estimations are rarely off by more than one resolution element;
- Observing conditions do not seem to be the major factor in the accuracy of the position estimation. Instead the localisation seem to be the leading factor. Indeed, the proximity to any bright feature due to a diffraction effect (such as the AO-correction ring, satellite spots, apodized-Lyot coronagraph inner-working angle, low wind effect or so) varying with the wavelength might affect the overall estimated position of the injection by distorting its flux distribution depending on the spectral channel.

Table 3. Astrometry results on the four GPI data: euclidean distance between the estimated position and the ground-truth (both in Cartesian coordinates) for each injection within each data set. For the baseline method and the four submitted methods, the metric is shown for (i) each injection, (ii) averaged for each data set (bold values), and (iii) averaged on the four GPI data sets (green bold values).

| Method | gpi1 | | | gpi2 | | gpi3 | | | gpi4 | | Average |
|----------|--------------|-------|-------|--------------|-------|--------------|-------|-------|--------------|-------|--------------|
| | b | c | d | b | c | b | c | d | b | c | |
| Baseline | 0.004 | 0.354 | 0.053 | 0.099 | 0.609 | 0.043 | 0.138 | 0.004 | 0.024 | 0.134 | 0.146 |
| | 0.137 | | | 0.354 | | 0.062 | | | 0.079 | | |
| RSM | 0.453 | 0.779 | 0.053 | 0.127 | 0.534 | 0.049 | 0.727 | 0.119 | 0.126 | 0.324 | 0.329 |
| | 0.428 | | | 0.33 | | 0.298 | | | 0.225 | | |
| pyKLIP | 0.088 | 0.1 | 0.035 | 0.018 | 0.302 | 0.137 | 0.325 | 0.169 | 0.176 | 1.05 | 0.24 |
| | 0.074 | | | 0.16 | | 0.21 | | | 0.613 | | |
| AMAT | 0.16 | 0.268 | 0.079 | 0.102 | 0.69 | 0.275 | 0.175 | 0.03 | 0.075 | 0.131 | 0.198 |
| | 0.169 | | | 0.396 | | 0.16 | | | 0.103 | | |
| ANDRO | 3.928 | 2.563 | 4.387 | 5.28 | 4.393 | 8.827 | 6.569 | 4.617 | 1.389 | 5.819 | 4.777 |
| | 3.626 | | | 4.837 | | 6.671 | | | 3.604 | | |

Table 4. Astrometry results on the four SPHERE-IFS data: euclidean distance between the estimated position and the ground-truth (both in Cartesian coordinates) for each injection within each data set. For the baseline method and the four submitted methods, the metric is shown for (i) each injection, (ii) averaged for each data set (bold values), and (iii) averaged on the four SPHERE-IFS data sets (green bold values).

| Method | sphere1 | | | sphere2 | | | sphere3 | | sphere4 | | | Average |
|----------|--------------|-------|-------|--------------|-------|-------|--------------|-------|--------------|-------|-------|--------------|
| | b | c | d | b | c | d | b | c | b | c | d | |
| Baseline | 0.035 | 0.05 | 0.05 | 0.44 | 0.049 | 0.48 | 0.074 | 0.108 | 0.141 | 0.013 | 0.266 | 0.155 |
| | 0.045 | | | 0.323 | | | 0.091 | | 0.14 | | | |
| RSM | 0.343 | 0.191 | 0.128 | 0.828 | 0.297 | 0.297 | 0.254 | 1.835 | 0.185 | 0.036 | 0.638 | 0.457 |
| | 0.221 | | | 0.474 | | | 1.045 | | 0.286 | | | |
| pyKLIP | 0.184 | 0.071 | 0.293 | 0.38 | 0.23 | 0.554 | 0.231 | 0.054 | 0.092 | 0.068 | 0.052 | 0.201 |
| | 0.183 | | | 0.388 | | | 0.142 | | 0.071 | | | |
| AMAT | 0.223 | 0.58 | 0.037 | 0.506 | 0.226 | 0.325 | 0.362 | 0.01 | 0.085 | 0.109 | 0.238 | 0.246 |
| | 0.28 | | | 0.353 | | | 0.186 | | 0.144 | | | |
| ANDRO | 2.292 | 3.236 | 2.746 | 1.102 | 0.871 | 3.394 | 5.103 | 2.07 | 2.702 | 5.389 | 6.166 | 3.188 |
| | 2.758 | | | 1.789 | | | 3.587 | | 4.752 | | | |

-On the contrary, even at low SNR, but at a favorable location (in the AO-corrected area, at large separation from the star), the estimations are very accurate (e.g. *sphere4*, injection 'c').

-We note here the specific case of strong low wind effect in the *sphere2* data set. The inner ('b') and outer ('d') injections, both sitting on an intense spider diffraction pattern, are not as well estimated as the other injection ('c') although located right on a spider diffraction pattern. The better estimation for the latter may also be accounted for a higher SNR.

In terms of methods, most trends are similar from one injection to another and from one data set to another (at a few exceptions). Overall, PCA-SADI+NEGFC, AMAT, RSM and pyKLIP-FM all succeed in retrieving the injection position with a very good accuracy, usually within one resolution element. Only the classical version of ANDROMEDA, fitting a 2-D Gaussian at the location of a detection in the SNR map, shows inaccurate astrometry estimation often exceeding a few pixels.

4. RESULTS: SPECTROPHOTOMETRY

As a reminder, for the spectro-photometry estimation of the Exoplanet Imaging Data Challenge second phase, we asked the participants to provide with (1) the estimated contrast of each injected signal at each wavelength, (2) optionally the corresponding 1σ uncertainties of the estimations, and (3) optionally the corresponding posterior distribution used to estimate the contrast and its uncertainties. As for the metric used for the ranking of the submitted results, the EIDC phase 2 working group decided to compute the normalized absolute distance (in the sense of the L1-norm) between the estimated value and the ground truth value for each spectral channel separately. Normalizing the absolute distance by the ground-truth contrast ensures that low-contrast signals do not penalize smaller true contrasts in the averaged score. More details about the data, submission process and metrics for comparison and ranking can be found in our previous SPIE publication Cantalloube et al., 2022.²⁴

4.1 Spectro-photometry: baseline results as a function of the dataset

As a first qualitative interpretation, we compare the spectro-photometry estimated by the baseline PCA-SADI+NEGFC algorithm in each data set. The results on the GPI images are shown in Fig. 3 and those on the SPHERE-IFS images are shown in Fig. 4. On the figure we show on the left the actual position of the injections overlaid in the first coronagraphic image of the cube and on the right the extracted spectra for each injection (2 to 3 injection per dataset) including the 3σ errorbars compared to the ground truth (dark solid line), as well as the residuals between the extracted spectra and the corresponding ground-truth (shaded areas represent the corresponding 3σ uncertainties). Quantitative results, including the comparison metric used, are presented in Tab. 1 and Tab. 2 for the GPI and SPHERE-IFS data respectively.

Overall, there are a few singular outliers at some random wavelength of the estimated spectrum (probably due to some convergence problems in the NEGFC method). Seemingly, there is no systematic wavelength affected by this problem. Regardless of these outliers, the spectro-photometric estimations always include the ground-truth within 5σ . Uncertainties are matching the spectrum depth: lower contrast yield larger uncertainties, as expected (clearly visible in e.g. companion 'c' of data *gpi1*, *gpi3* and *sphere1*).

Globally, the further the injected companion, the better its spectro-photometric estimation. Higher SNR yield better estimations. Equivalently, injections at lower contrast are better estimated (see e.g. *sphere1* 'b' and *gpi1* 'c', both having a very high contrast, under exquisite observing conditions and located right in the AO-corrected zone). In general, the structure of the spectrum is correctly retrieved.

There are two notable exceptions to these conclusions, due to poor observation conditions. Low-wind effect (LWE) is present in the *sphere2* data set, in which we placed three injections on top of the characteristics spider diffraction peaks. As a result, the three corresponding spectra are underestimated (lower than the ground-truth): the presence of low-wind effect biases the spectrum estimation. Wind-driven halo (WDH) is present in the *gpi2* and the *sphere4* data set and we placed injections *gpi2* 'b' and *sphere4* 'd' within the characteristic butterfly shape. The estimated spectrum of *gpi2* 'b' (located beyond 600mas), does not seem to be much affected, but *sphere4* 'd' estimated spectrum is overestimated despite a relatively good SNR: the presence of wind-driven halo seems to affect close-in companions, but this needs to be verified in a more systematic fashion to be confirmed.

4.2 Spectro-photometry: results as a function of the submitted algorithms

The results for the baseline method and the three submitted methods are presented in Tab. 5 (GPI datasets) and Tab. 6 (SPHERE-IFS datasets). For each dataset and injection, we show the chosen metric⁶ as the normalized L1-norm distance between the estimation to the ground-truth. The last column averages the results for each injection to present an overall quantity for each algorithm. A graphical view of the results is presented in Fig. 7 and Fig. 8 for the GPI and SPHERE-IFS data respectively.

On average over all the injections, at all wavelength and for all the data set, it appears that PCA-NEGFC and pyKLIP-FM are better at retrieving the spectro-photometry, followed closely by AMAT, then RSM. These trends remain the same with various observing conditions. For the GPI data set, all methods give metrics below one but RSM and PCA-NEGFC for the closest companion of the *gpi4* data set (under bad observing conditions) showing a SNR below 5. For the SPHERE-IFS data set, the very high values obtained for *sphere4* 'c' are due to a very low SNR of the injection (less than 2 for the baseline).

However, when we look at the detailed extractions shown on Fig. 7 and Fig. 8 (for GPI and SPHERE-IFS respectively), we may notice a few additional effects hidden in the averaged metric. Some algorithms are good for faint planets (e.g. pyKLIP and AMAT), but suffer from significant biases when it comes to relatively bright planets, and reversely the baseline can yield significant outliers for faint planets but is very reliable for bright planets. The adopted procedure for the **PCA-NEGFC** (baseline), which selects the number of principal components that optimize the SNR may be prone to overestimating the flux in some channels, by including a residual speckle signal. **pyKLIP-FM** seems to overestimate the flux at lower contrast values (bright planets): the oversubtraction may not be well estimated for channels where the companion is bright. In general the structure of the extracted spectra appear much smoother compared to e. g. the baseline. **AMAT** is relatively good in general, but slightly underestimates the flux, in particular under good observation conditions and/or for bright planets. **RSM** is still good on average but suffers from many outliers and the spectrum structure is jagged. Note that the results concerning pyKLIP-FM and VIP/PCA-NEGFC are similar to the conclusions presented in Nasedkin et al.²⁷ (2023) for pyKLIP-FM vs Pynpoint/PCA-NEGFC.

In the specific case of low-wind effect (*sphere2* data set), all algorithm are underestimating the spectra mainly at larger wavelengths. Only RSM (green lines) and AMAT (blue lines) are estimating correctly at short wavelengths for only one or two in three injections.

Table 5. Spectro-photometry results on the four GPI data set: absolute distance between the estimated photometry at each spectral channel and the corresponding ground-truth for each injection within each data set, normalized by the ground-truth. For the baseline method and the four submitted methods, the metric is shown for (i) each injection, (ii) averaged over each injection (bold values), and (iii) averaged on the four GPI data sets (green bold values).

| Method | gpi1 | | | gpi2 | | gpi3 | | | gpi4 | | Median |
|----------|--------------|-------|-------|--------------|-------|--------------|-------|-------|--------------|-------|--------------|
| | b | c | d | b | c | b | c | d | b | c | |
| Baseline | 0.559 | 0.423 | 0.023 | 0.067 | 0.209 | 0.073 | 0.18 | 0.013 | 0.212 | 3.003 | 0.237 |
| | 0.335 | | | 0.138 | | 0.089 | | | 1.607 | | |
| RSM | 0.822 | 0.307 | 0.077 | 0.116 | 0.474 | 0.091 | 0.213 | 0.068 | 0.094 | 3.456 | 0.349 |
| | 0.402 | | | 0.295 | | 0.124 | | | 1.775 | | |
| pyKLIP | 0.251 | 0.033 | 0.226 | 0.449 | 0.204 | 0.34 | 0.183 | 0.308 | 0.25 | 0.204 | 0.252 |
| | 0.17 | | | 0.327 | | 0.277 | | | 0.227 | | |
| AMAT | 0.663 | 0.11 | 0.213 | 0.165 | 0.46 | 0.085 | 0.276 | 0.026 | 0.122 | 0.105 | 0.221 |
| | 0.329 | | | 0.312 | | 0.129 | | | 0.114 | | |

Table 6. Spectro-photometry results on the four SPHERE-IFS data: absolute distance between the estimated photometry at each spectral channel and the corresponding ground-truth for each injection within each data set, normalized by the ground-truth. For the baseline method and the four submitted methods, the metric is shown for (i) each injection, (ii) averaged over each injection (bold values), and (iii) averaged on the four GPI data sets (green bold values).

| Method | sphere1 | | | sphere2 | | | sphere3 | | sphere4 | | | Median |
|----------|--------------|-------|-------|--------------|-------|-------|--------------|-------|------------------|-----------|-------|--------------|
| | b | c | d | b | c | d | b | c | b | c | d | |
| Baseline | 1.618 | 0.028 | 0.023 | 1.01 | 0.542 | 0.411 | 0.812 | 4.282 | 0.226 | 29981.619 | 8.549 | 1.601 |
| | 0.556 | | | 0.654 | | | 2.547 | | > 1000 | | | |
| RSM | 1.779 | 0.079 | 0.053 | 6.323 | 0.104 | 0.3 | 0.361 | 14.49 | 0.272 | 1065.896 | 0.408 | 4.834 |
| | 0.637 | | | 2.242 | | | 7.426 | | 355.525 | | | |
| pyKLIP | 0.199 | 0.239 | 0.099 | 0.572 | 0.599 | 0.524 | 0.356 | 6.08 | 0.044 | 622.754 | 0.114 | 1.891 |
| | 0.179 | | | 0.565 | | | 3.218 | | 207.637 | | | |
| AMAT | 1.476 | 0.029 | 0.118 | 0.175 | 0.242 | 0.294 | 0.273 | 8.568 | 0.139 | 707.266 | 0.123 | 2.481 |
| | 0.541 | | | 0.237 | | | 4.42 | | 235.843 | | | |

5. CONCLUSION & LEGACY

The *Exoplanet Imaging Data Challenge* is a community-led project that aims at offering tools for a better understanding of the current capacities of the various high-contrast imaging post-processing techniques, and to foster developments of new methods. To that end, we offer several resources and an extensive bibliography gathered on our website. As part of the data challenge, we provide with a curated set of data that are accessible on a *Zenodo* repository, as well as a set of assorted metrics to gauge the results of a given algorithm that are available on a *Github* repository. In addition, to ease the comparison to other methods, we wrote a *jupyter notebook* that processes any given input and provides the comparison graphics and ranking with respect to the results already published in the 1st phase of the data challenge (Cantalloube et al.,⁵ 2020). This tool can be found at: <https://github.com/exoplanet-imaging-challenge/phase2/tree/main/eidc2>.

Since its creation in 2019, several publications have used the data set and metric set to support the results of new post-processing methods. After several years of running this data challenge, we have gathered some *lessons learned* to improve the engagement of participants. This lack of engagement turns out to be the major breakpoint preventing us from collecting a sufficient number of inputs in order to compile a consistent comparative conclusion to communicate the community. One important piece of advice is to organize a special hands-on session to help with the basics of the data format. Another idea we tried is to involve students during lab sessions. We also reached out to networks that run data challenges in other scientific fields to get feedback and ensure a smooth management of the EIDC.

To enhance this legacy prospect, in the future we would like to gather more data from the state-of-the-art high-contrast imaging instruments, such as Subaru/CHARIS, Magellan-Clay/MagAO-X, VLT/NEAR, LBT/SHARK-NIR and LBT/SHARK-VIS, under various observing conditions. These data would come with a set of tools to inject planetary signals and circumstellar disks, as well as metrics for comparison. As a community project based on volunteer work, we are doing our best to facilitate the use of the EIDC resources. For any suggestion or specific request, please contact us at exoimg.datachallenge@gmail.com.

APPENDIX A. GALLERY: BASELINE RESULTS

In this appendix, we first show the results of the astrometric estimations using the baseline method (PCA-SADI+NEGFC) on the four GPI datasets (Fig. 1) and on the four SPHERE-IFS datasets (Fig. 2). We then show the results of the spectro-photometric estimations using the baseline method (PCA-SADI+NEGFC) on the four GPI datasets (Fig. 3) and on the four SPHERE-IFS datasets (Fig. 4).

APPENDIX B. GALLERY: RESULTS ASTROMETRY

In this appendix, we show the results of the astrometric estimations using the baseline method (red), as well as the four received results using pyKLIP-FM (gold), RSM (green), AMAT (blue), and ANDROMEDA (purple). For each injection, the ground-truth is represented in the middle (black diamond), where the horizontal and vertical black dashed lines intersect. The two centered grey shaded areas represent the size of a half resolution element ($0.5\lambda/D$) at the shortest and largest wavelengths λ of the spectro-imager, whose effective diameter is D . For the baseline method (in red), the size of the symbol corresponds to the 1-sigma uncertainties (averaged in x- and y- directions) and the shaded area to the 3-sigma uncertainties. To ease the reading, we added on top of each plots the SNR of the injected planetary signal obtained with the baseline as well as the separation of the considered injection. The four GPI datasets are shown in Fig. 5 and the four SPHERE-IFS datasets in Fig. 6.

APPENDIX C. GALLERY: RESULTS SPECTRO-PHOTOMETRY

In this appendix, we show the results of the spectro-photometric estimations using the baseline method (red), as well as for the three submitted results with pyKLIP-FM (gold), RSM (green), and AMAT (blue). The top panel show the retrieved spectra by each method, with the ground-truth shown as a black solid line. The bottom panel shows the residuals (difference) between the ground-truth and the estimated spectra. The spectra corresponding to the four GPI datasets are shown in Fig. 7 and those of the four SPHERE-IFS datasets are shown in Fig. 8.

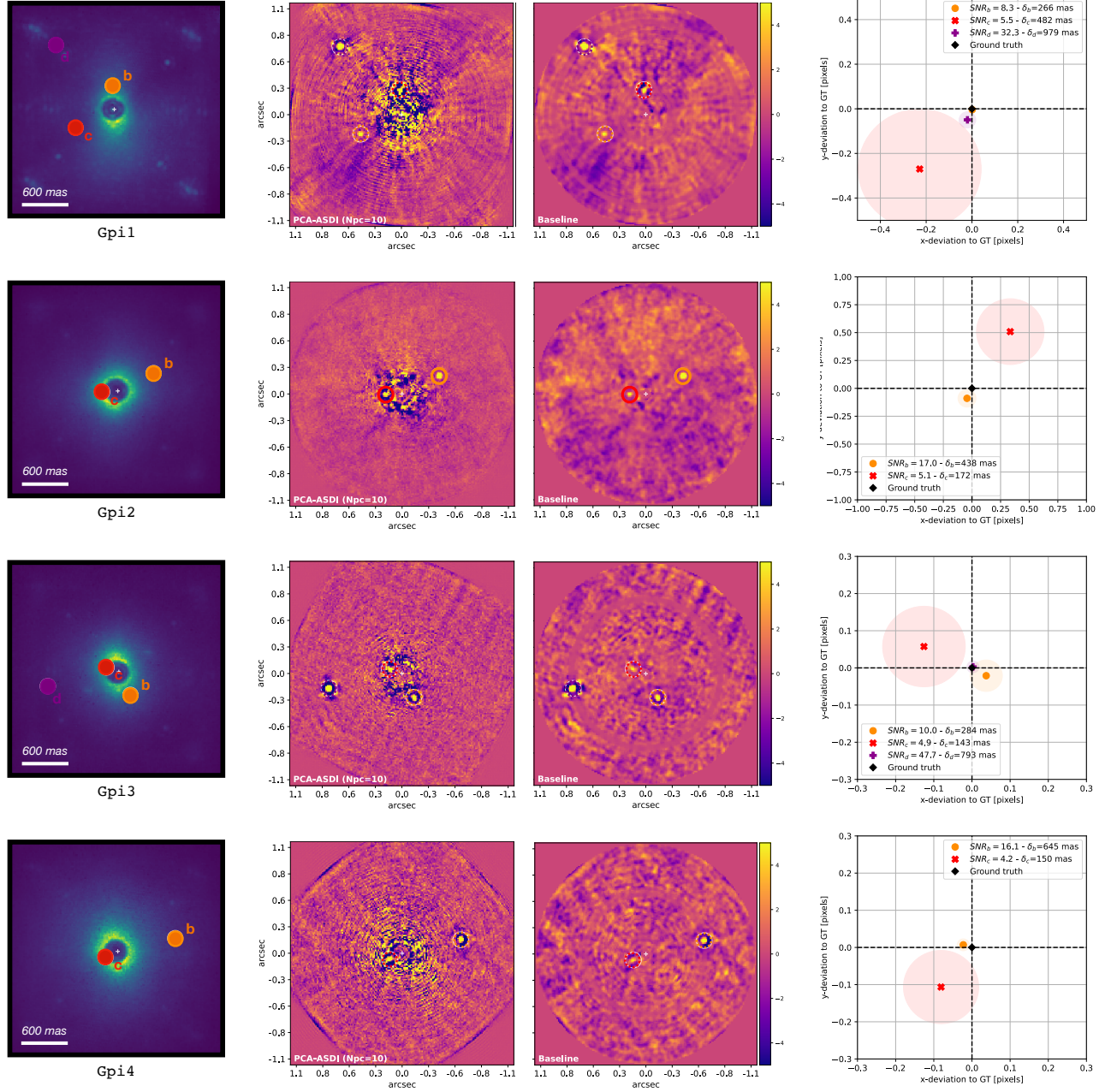


Figure 1. Astrometry: GPI data baseline results for each injection. On each image, the location of the injected synthetic planetary signals are shown with the orange, red and purple circles. From left to right: Temporal median image at the shortest wavelength (left); residual map after applying a PCA-SADI post-processing, as implemented in the VIP package and using the first 10 principal components to build the reference image for subtraction (middle-left); corresponding SNR map, as implemented in the VIP package (middle-right); and astrometry estimations for each injections, using the baseline NEGFC technique (right). The latter highlights the relative astrometry errors (in pixels) for each injection (color coded) compared to the ground-truth shown as a black diamond in the center of the image. The shaded area correspond to the 1-sigma uncertainty on the astrometry estimation. From top to bottom: data set *gpi1*, *gpi2*, *gpi3*, and *gpi4*.

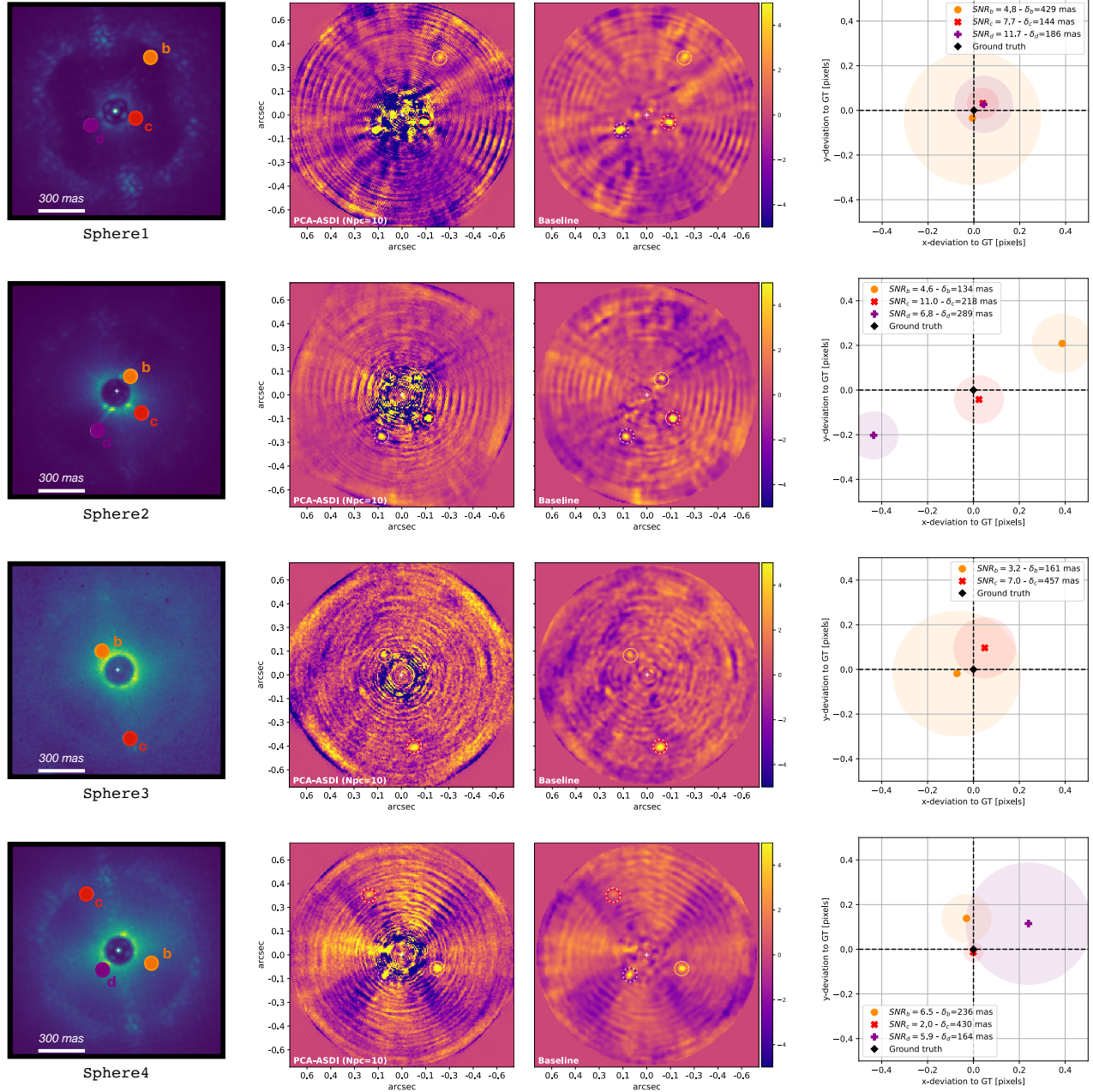


Figure 2. Astrometry: SPHERE-IFS data baseline results for each injection. On each image, the location of the injected synthetic planetary signals are shown with the orange, red and purple circles. From left to right: Temporal median image at the shortest wavelength (left); residual map after applying a PCA-SADI post-processing, as implemented in the VIP package and using the first 10 principal components to build the reference image for subtraction (middle-left); corresponding SNR map, as implemented in the VIP package (middle-right); and astrometry estimations for each injections, using the baseline NEGFC technique (right). The latter highlights the relative astrometry errors (in pixels) for each injection (color coded) compared to the ground-truth shown as a black diamond in the center of the image. The shaded area correspond to the 1-sigma uncertainty on the astrometry estimation. From top to bottom: data set *sphere1*, *sphere2*, *sphere3*, and *sphere4*.

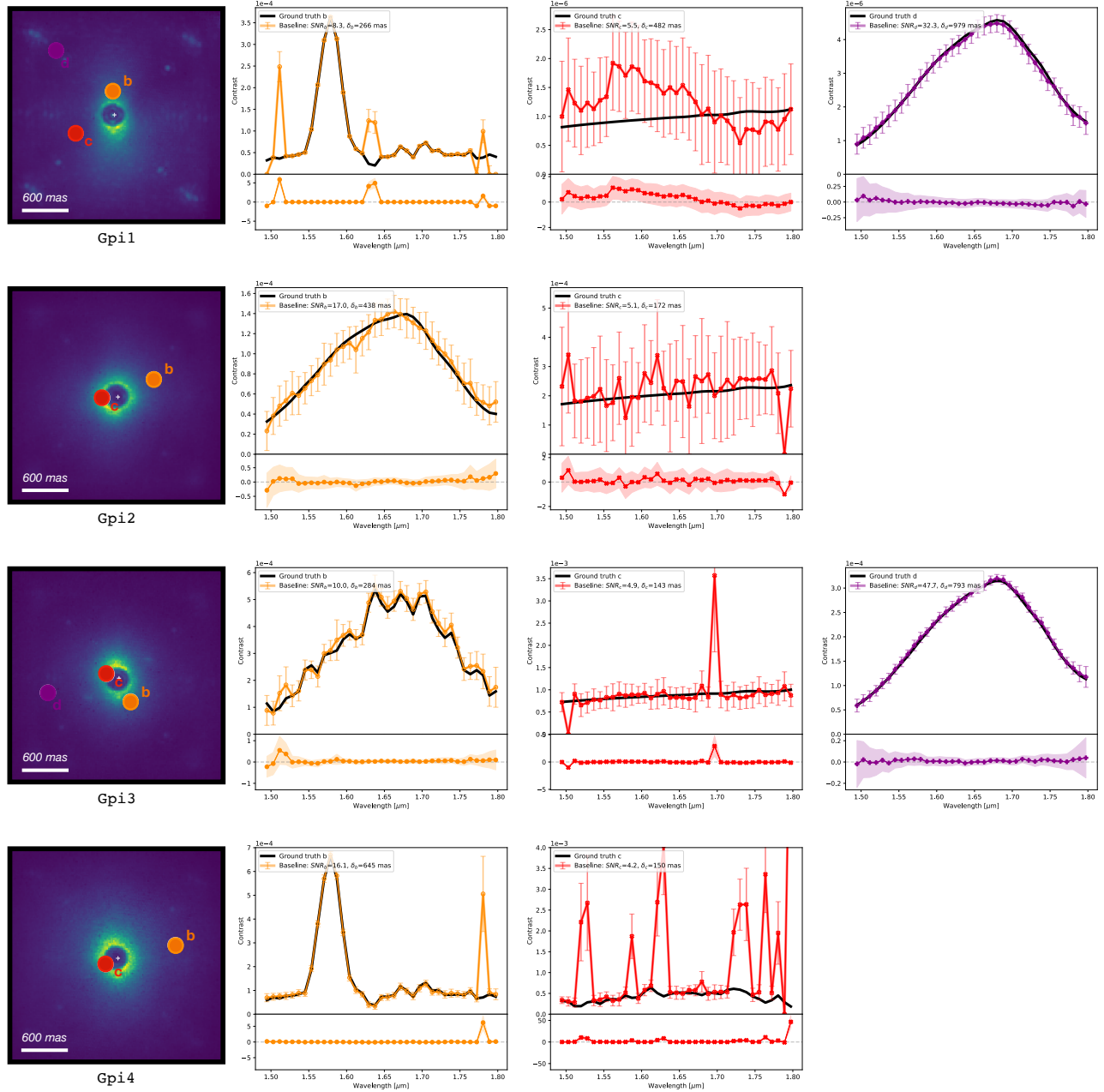


Figure 3. Spectro-photometry: baseline results for each injection in the GPI data sets. From left to right: Temporal median image at the shortest wavelength with the position of the injected planetary signals (colored circles); Corresponding estimated spectrum with 3σ uncertainties (colored lines), compared to the injected spectrum (dark line). The bottom panel shows the residuals between the estimated spectra and the ground-truth (the shaded area corresponds to the 3σ uncertainties on the estimations). From top to bottom: data set *gpi1*, *gpi2*, *gpi3*, and *gpi4*.

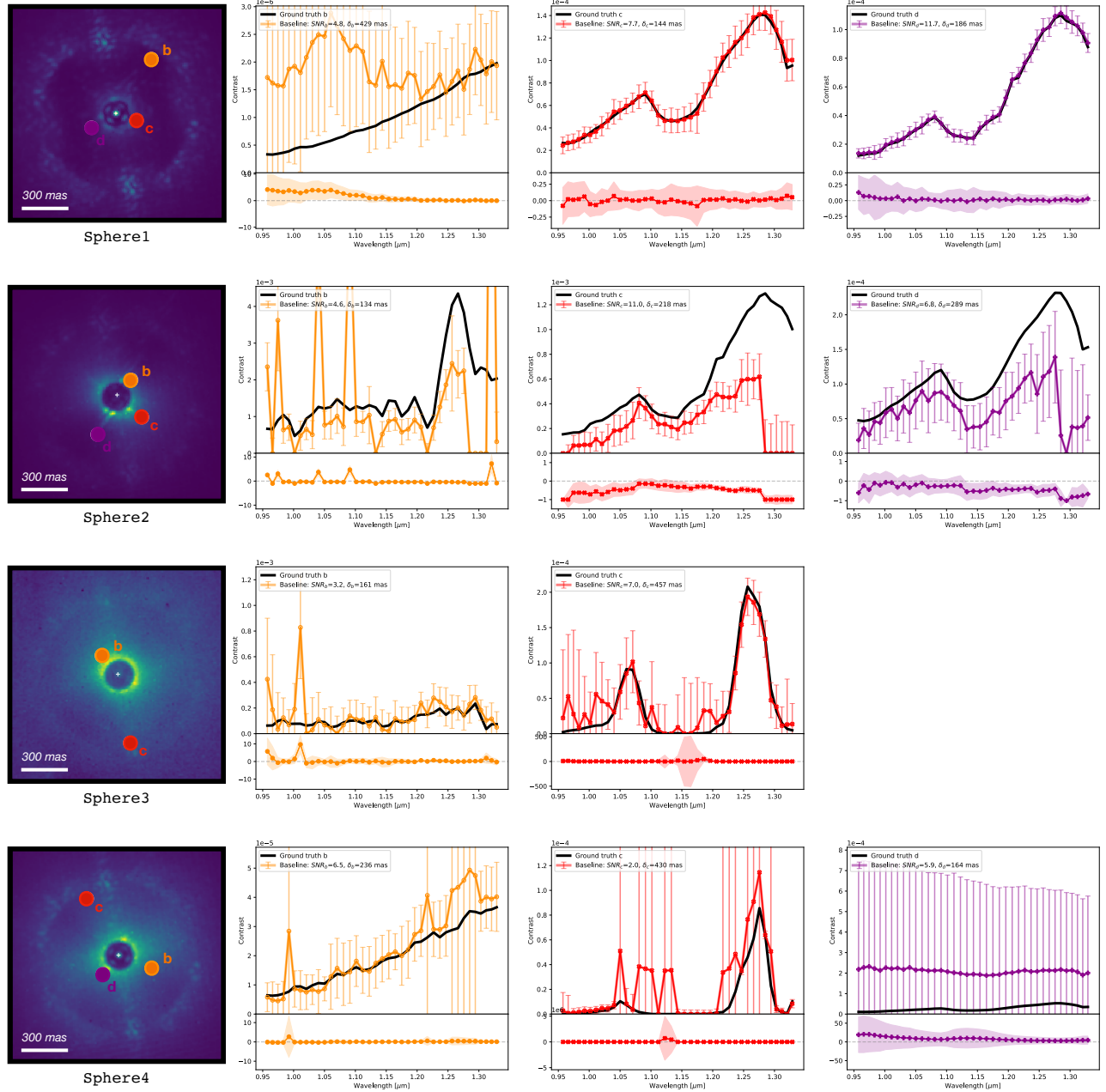


Figure 4. Spectro-photometry: baseline results for each injection in the SPHERE-IFS data sets. From left to right: Temporal median image at the shortest wavelength with the position of the injected planetary signals (colored circles); Corresponding estimated spectrum with 3σ uncertainties (colored lines), compared to the injected spectrum (dark line). The bottom panel shows the residuals between the estimated spectra and the ground-truth (the shaded area corresponds to the 3σ uncertainties on the estimations). From top to bottom: data set *sphere1*, *sphere2*, *sphere3*, and *sphere4*.

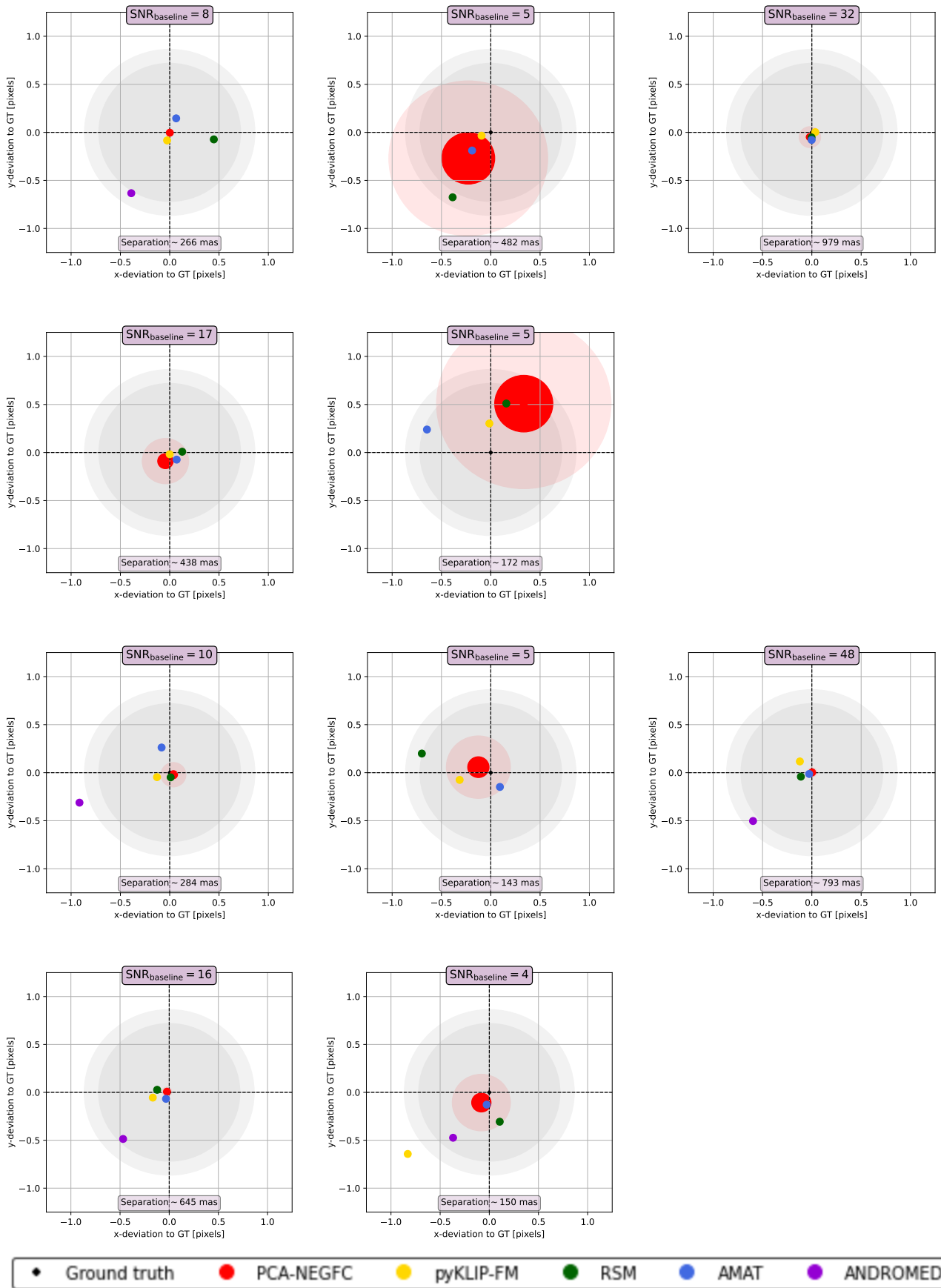


Figure 5. GPI relative astrometry errors for each injection (in pixels). The ground-truth is shown as a black diamond in the center of the grid. From top to bottom: data set *gpi1*, *gpi2*, *gpi3*, and *gpi4*. The platescale of GPI is 14.16 mas/px.

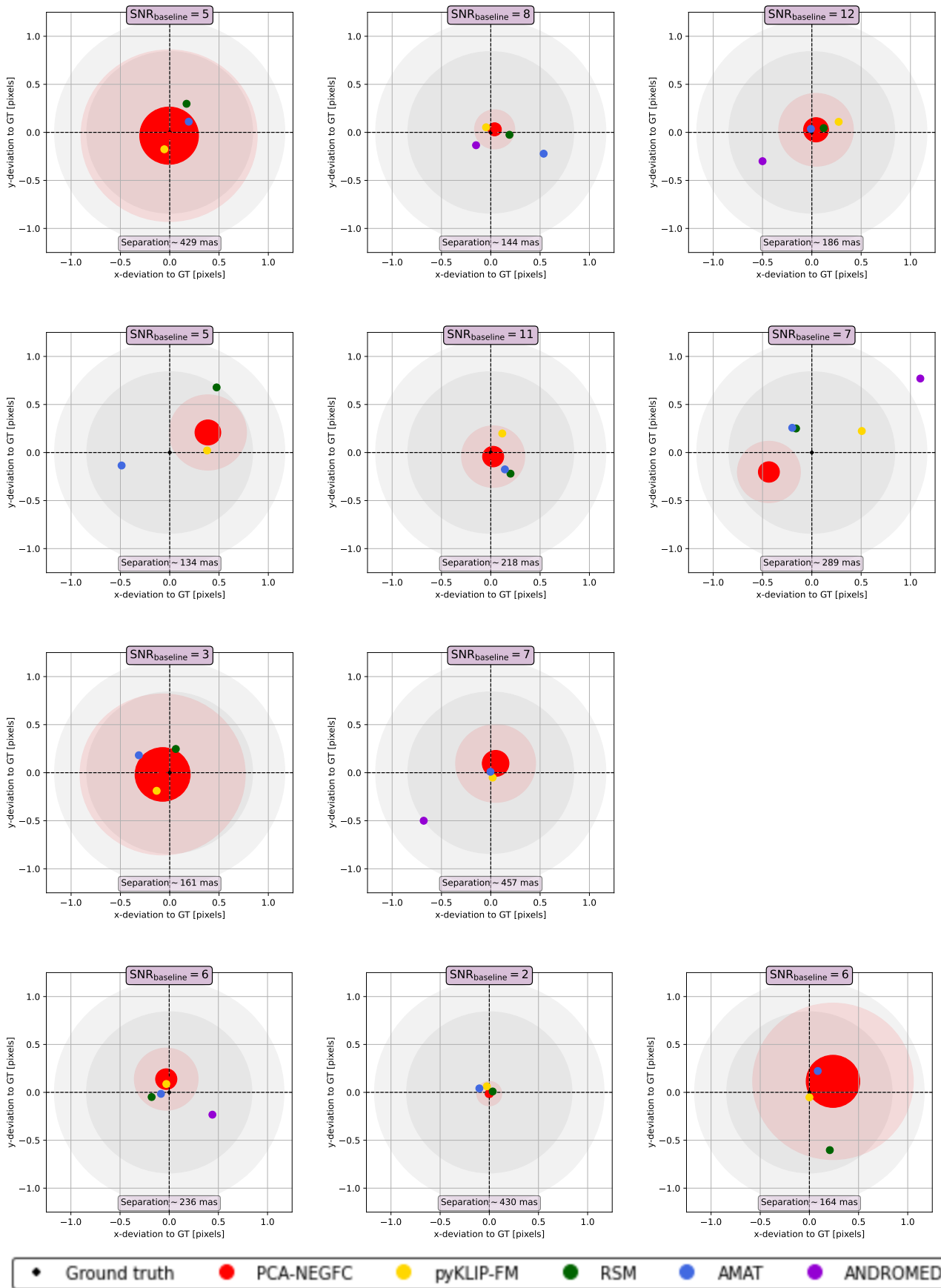


Figure 6. SPHERE-IFS relative astrometry errors for each injection (in pixels). The ground-truth is shown at the center. From top to bottom: data set *sphere1*, *sphere2*, *sphere3*, and *sphere4*. The platescale of SPHERE-IFS is 7.46 mas/px.

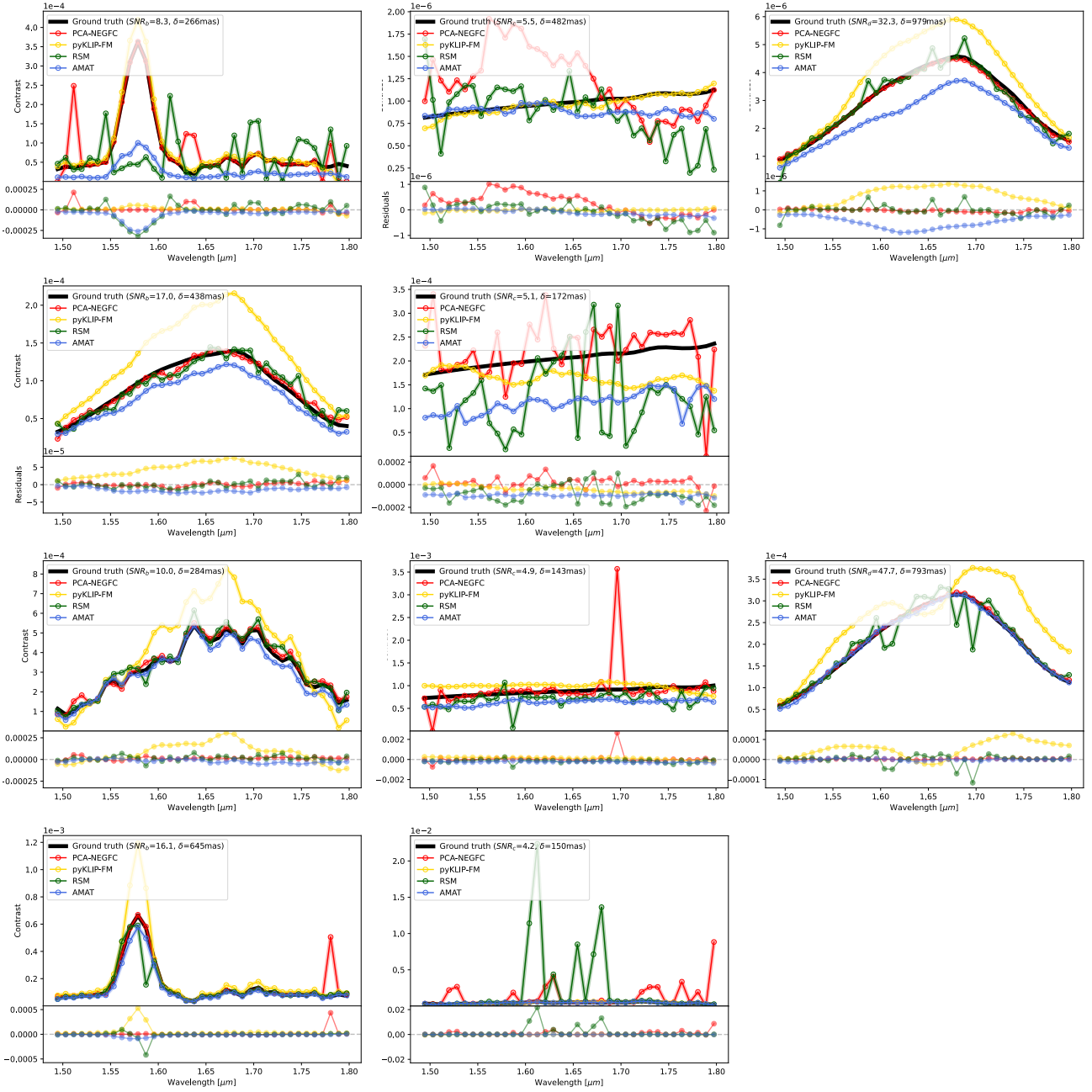


Figure 7. GPI estimated spectra for each injected planetary signal (in contrast to the star). The ground-truth is shown as a black solid line. The top panel shows the estimated spectra by each method, while the bottom panel highlights the difference between the ground-truth spectrum and the estimated spectra. For reference, the spectra extracted with the baseline method are shown with red solid lines (PCA-SADI with NEGFC). The spectra extracted with the three submissions are shown in gold (pyKLIP-FM), green (RSM), and blue (AMAT). From top to bottom: data set *gpi1*, *gpi2*, *gpi3*, and *gpi4*. The wavelength range of GPI is $[1.495 - 1.797]\mu\text{m}$ (around the H-band), shared in 37 spectral channels.

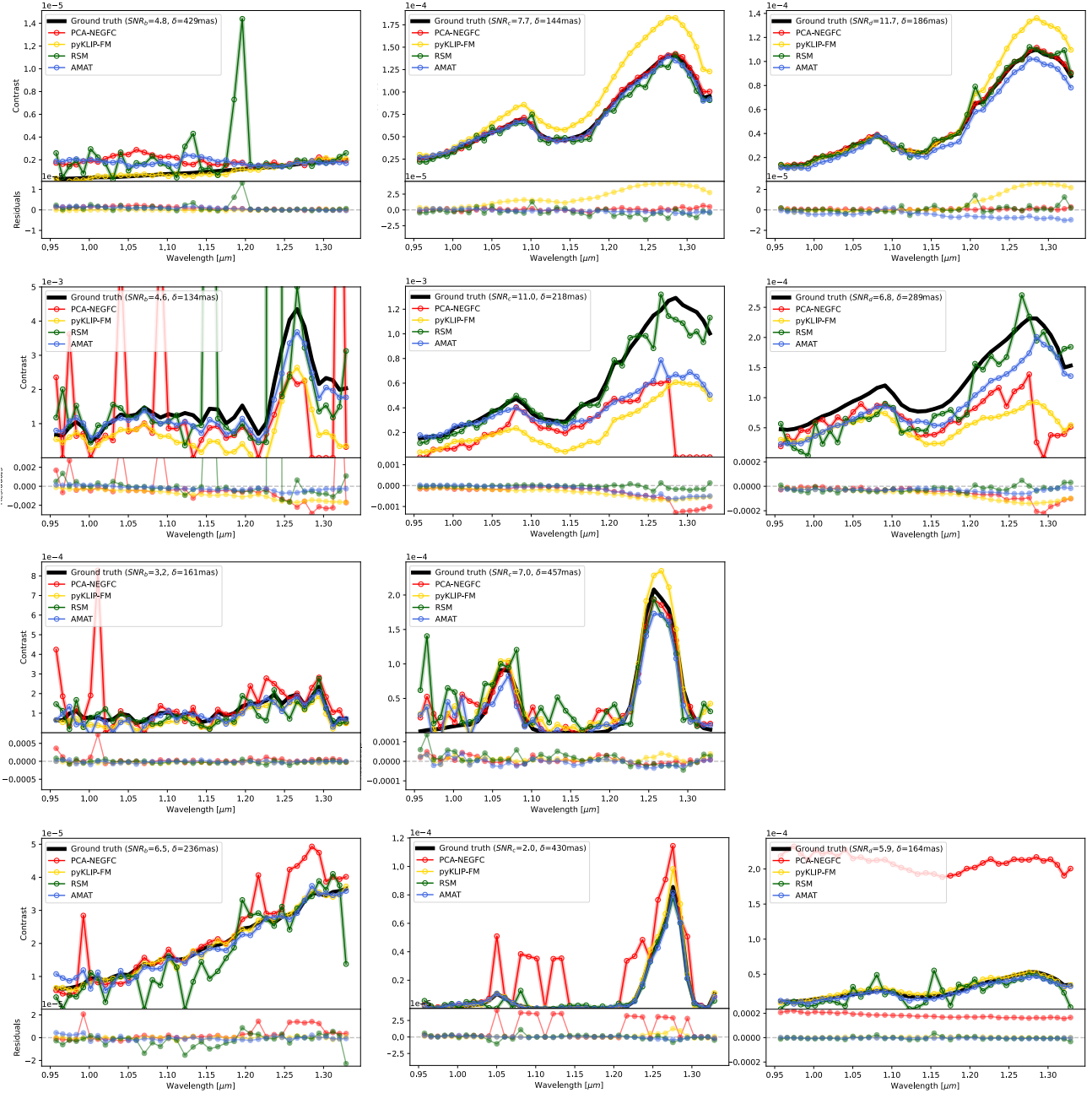


Figure 8. SPHERE-IFS estimated spectra for each injected planetary signal (in contrast to the star). The ground-truth is shown as a black solid line. The top panel shows the estimated spectra by each method, while the bottom panel highlights the difference between the ground-truth spectrum and the estimated spectra. For reference, the spectra extracted with the baseline method are shown with red solid lines (PCA-SADI with NEGFC). The spectra extracted with the three submissions are shown in gold (pyKLIP-FM), green (RSM), and blue (AMAT). From top to bottom: data set *sphere1*, *sphere2*, *sphere3*, and *sphere4*. The wavelength range of SPHERE-IFS is $[0.957 - 1.329]\mu\text{m}$ (YJ-bands), shared in 39 spectral channels.

ACKNOWLEDGMENTS

This work has been partially supported by the LabEx FOCUS ANR-11-LABX-0013. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 819155), and from the Wallonia-Brussels Federation (grant for Concerted Research Actions). VC and OA thank the Belgian F.R.S.-FNRS, and the Belgian Federal Science Policy Office (BELSPO) for the provision of financial support in the framework of the PRODEX Programme of the European Space Agency (ESA) under contract number 4000142531. This work has made use of the High Contrast 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 (Investissements d’avenir – ANR10 LABX56).

REFERENCES

- [1] Cantalloube, F., Farley, O. J., Milli, J., Bharmal, N., Brandner, W., Correia, C., Dohlen, K., Henning, T., Osborn, J., Por, E., et al., “Wind-driven halo in high-contrast images-i. analysis of the focal-plane images of sphere,” *Astronomy & Astrophysics* **638**, A98 (2020).
- [2] Sauvage, J.-F., Fusco, T., Lamb, M., Girard, J., Brinkmann, M., Guesalaga, A., Wizinowich, P., O’Neal, J., N’Diaye, M., Vigan, A., et al., “Tackling down the low wind effect on sphere instrument,” in [*Adaptive Optics Systems V*], **9909**, 408–416, SPIE (2016).
- [3] Vigan, A., Dohlen, K., N’diaye, M., Cantalloube, F., Girard, J., Milli, J., Sauvage, J.-F., Wahhaj, Z., Zins, G., Beuzit, J.-L., et al., “Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a zernike phase-mask sensor-iv. temporal stability of non-common path aberrations in vlt/sphere,” *Astronomy & Astrophysics* **660**, A140 (2022).
- [4] Cantalloube, F., Dohlen, K., Milli, J., Brandner, W., and Vigan, A., “Peering through SPHERE Images: A Glance at Contrast Limitations,” *The Messenger* **176**, 25–31 (June 2019).
- [5] 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).
- [6] Cantalloube, F., Christiaens, V., Cantero, C., Nasedkin, E., Cioppa, A., Absil, O., Bonse, J., Delorme, P., Gomez-Gonzalez, C., Juillard, S., et al., “Exoplanet imaging data challenge, phase ii: characterization of exoplanet signals in high-contrast images,” in [*Adaptive Optics Systems VIII*], **12185**, 8–24, SPIE (2022).
- [7] Nielsen, E. L., De Rosa, R. J., Macintosh, B., Wang, J. J., Ruffio, J.-B., Chiang, E., Marley, M. S., Saumon, D., Savransky, D., Ammons, S. M., et al., “The gemini planet imager exoplanet survey: giant planet and brown dwarf demographics from 10 to 100 au,” *The Astronomical Journal* **158**(1), 13 (2019).
- [8] Vigan, A., Fontanive, C., Meyer, M., Biller, B., Bonavita, M., Feldt, M., Desidera, S., Marleau, G.-D., Emsenhuber, A., Galicher, R., et al., “The sphere infrared survey for exoplanets (shine)-iii. the demographics of young giant exoplanets below 300 au with sphere,” *Astronomy & Astrophysics* **651**, A72 (2021).
- [9] 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).
- [10] Lafreniere, D., Marois, C., Doyon, R., Nadeau, D., and Artigau, É., “A new algorithm for point-spread function subtraction in high-contrast imaging: a demonstration with angular differential imaging,” *The Astrophysical Journal* **660**(1), 770 (2007).
- [11] 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).
- [12] 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).
- [13] 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).

- [14] Pueyo, L., “Detection and characterization of exoplanets using projections on karhunen–loève eigenimages: Forward modeling,” *The Astrophysical Journal* **824**(2), 117 (2016).
- [15] Flasseur, O., Denis, L., Thiébaud, 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).
- [16] Christiaens et al., V., “Special: A python package for the spectral characterization of directly imaged low-mass companions,” *subm. to JOSS* (2022).
- [17] 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).
- [18] Dahlgvist, C.-H., Cantalloube, F., and Absil, O., “Regime-switching model detection map for direct exoplanet detection in adi sequences,” *Astronomy & Astrophysics* **633**, A95 (2020).
- [19] Dahlgvist, 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).
- [20] Soummer, R., Pueyo, L., and Larkin, J., “Detection and Characterization of Exoplanets and Disks Using Projections on Karhunen-Loève Eigenimages,” **755**, L28 (Aug. 2012).
- [21] Wang, J. J., Ruffio, J.-B., De Rosa, R. J., Aguilar, J., Wolff, S. G., and Pueyo, L., “pyKLIP: PSF Subtraction for Exoplanets and Disks.” *Astrophysics Source Code Library*, record ascl:1506.001 (June 2015).
- [22] Pueyo, L., “Detection and characterization of exoplanets using projections on karhunen–loève eigenimages: Forward modeling,” *The Astrophysical Journal* **824**(2), 117 (2016).
- [23] Daglayan Sevim, H., Vary, S., and Absil, P.-A., “An alternating minimization algorithm with trajectory for direct exoplanet detection,” in [*ESANN 2023-European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*], (2023).
- [24] Cantalloube, F., Christiaens, V., Cantero, C., Nasedkin, E., Cioppa, A., Absil, O., Bonse, J., Delorme, P., Gomez-Gonzalez, C., Juillard, S., et al., “Exoplanet imaging data challenge, phase ii: characterization of exoplanet signals in high-contrast images,” in [*Adaptive Optics Systems VIII*], **12185**, 8–24, SPIE (2022).
- [25] 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).
- [26] 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).
- [27] Nasedkin, E., Mollière, P., Wang, J., Cantalloube, F., Kreidberg, L., Pueyo, L., Stolker, T., and Vigan, A., “Impacts of high-contrast image processing on atmospheric retrievals,” *Astronomy & Astrophysics* **678**, A41 (2023).