The Geneva Reduction and Analysis Pipeline for High-contrast Imaging of planetary Companions

J. Hagelberg,[1,2]★ D. Ségransan,[2] S. Udry[2] and F. Wildi[2]

[1]*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*
[2]*Observatoire de Genève, Université de Genève, 51 Chemin des Maillettes, CH-1290 Versoix, Switzerland*

ABSTRACT

We present GRAPHIC, a new angular differential imaging reduction pipeline where all geometric image operations are based on Fourier transforms. To achieve this goal the entire pipeline is parallelized making it possible to reduce large amounts of observation data without the need to bin the data. The specific rotation and shift algorithms based on Fourier transforms are described and performance comparison with conventional interpolation algorithm is given. Tests using fake companions injected in real science frames demonstrate the significant gain obtained by using geometric operations based on Fourier transforms compared to conventional interpolation. This also translates in a better point spread function and speckle subtraction with respect to conventional reduction pipelines, achieving detection limits comparable to current best performing pipelines. Flux conservation of the companions is also demonstrated. This pipeline is currently able to reduce science data produced by Very Large Telescope (VLT)/NACO, Gemini/NICI, VLT/SPHERE, and Subaru/SCExAO.

Key words: methods: data analysis – techniques: high angular resolution – techniques: image processing – planets and satellites: detection – infrared: planetary systems.

1 INTRODUCTION

18 yr after the first discovery of an exoplanet around a sun-like star (Mayor & Queloz 1995) and the unambiguous detection of three brown dwarfs (Basri & Marcy 1995; Nakajima et al. 1995; Rebolo, Osorio & Martín 1995), thousands of planets and brown dwarfs have been discovered. These numbers are growing ever faster as the pace of new detections is increasing, thanks to newly built instruments purposely designed to search for substellar objects, but also to the optimization of data analysis techniques.

The vast majority of exoplanets are currently detected with the radial velocity (RV) or transit techniques. However, orbital periods longer than the time span of the observations will hardly be detected by these two techniques, inducing a sharp decrease in detectability beyond ≈5 au and leaving unprobed a large area in the mass-separation parameter space. Direct imaging on the opposite probes the outer orbital regions not accessible with the two previous techniques, but the high contrast at small separation which needs to be reached turns it into one of the most challenging exoplanet detection techniques. The main hurdle to detect companions by high-contrast imaging is to remove the stellar point spread function (PSF) without diminishing the signal from the faint companion. This can be achieved through instrumental improvements or by improving the observing and data reduction techniques, with efforts focusing on these two fronts concurrently.

Since the first planets around stars have been directly imaged (Marois et al. 2008; Lagrange et al. 2009), the rate of exoplanets detected by direct imaging is steadily increasing thanks to the progress made to overcome the many technical challenges and careful selection of the target samples. But the small number of detections in total contrasts with the many direct imaging surveys which generated only few or no detection at all (e.g. Masciadri et al. 2005; Biller et al. 2007; Lafrenière et al. 2007b; Chauvin et al. 2010; Heinze et al. 2010; Bowler et al. 2012; Crepp et al. 2012; Vigan et al. 2012; Janson et al. 2013; Nielsen et al. 2013; Wahhaj et al. 2013).

The technical challenge of subtracting the host star PSF is currently based on two complementary differential imaging methods, with the same core idea of generating a PSF as similar as possible to the one which should be subtracted, but without having any potential companion signal in it. The difficulty being that the speckle structure of the PSF evolves in time, with many speckles in the stellar halo having a similar shape and intensity to what would be expected from a companion. Two main differential techniques exist to detect thermal emission of a companion. The first method called simultaneous differential imaging (SDI) is based on simultaneous observations in multiple bands. One can then either take advantage of specific absorption bands of the companion so that it is visible in one band and not the other, thus making it possible to subtract

★ E-mail: janis.hagelberg@unige.ch

the PSF while keeping the companion signal, or use the fact that because speckles are chromatic, their pattern scales with the wavelength but the potential companion stays on the same spot (Racine et al. 1999; Lenzen et al. 2004). The other differential method known as angular differential imaging (ADI) is based on the rotation of the field (Schneider & Silverstone 2003; Liu 2004; Marois et al. 2006), and has proven to be currently the most efficient method for PSF subtraction when searching for companions. These two methods do not require the use of a coronograph even though their use can increase the detection limits in certain cases. Finally, the two methods can be combined by letting the field rotate while observing simultaneously in multiple bands. Nearly every survey developed its own reduction pipeline most often based on either the Locally Optimized Combination of Images (LOCI; Lafrenière et al. 2007a) or more recently Principal Component Analysis (PCA; Amara & Quanz 2012; Soummer, Pueyo & Larkin 2012).

Here we present the Geneva Reduction and Analysis Pipeline for High-contrast Imaging of planetary Companions (GRAPHIC) based on ADI for PSF subtraction, which makes intensive use of Fourier analysis. It was specifically developed for the Geneva high-contrast imaging search of companions revealed by RV trends in the High Accuracy Radial velocity Planet Searcher (HARPS) and CORALIE survey.

2 THE GENEVA HIGH-CONTRAST IMAGING SEARCH OF COMPANIONS REVEALED BY RADIAL VELOCITY TRENDS IN THE HARPS AND CORALIE SURVEYS

Our campaign aims at detecting with direct imaging companions revealed by the RV trend they are causing, based on data from our two CORALIE and HARPS RV planet-search surveys. The RV data span over more than a decade with a precision reaching below 1 m s^{-1} in the case of HARPS so that trends induced by substellar companions on wide orbits can readily be detected. The selected targets are observed using Very Large Telescope (VLT)/NAOS-CONICA (NACO) and the ADI technique with deep observations of up to 4 h on target in order to reach the faint companions which had time to cool down. Our targets are all bright which result in integration times below 1 s to reach saturation, and in order not to resort to frame binning we are using the cube mode offered by NACO where frames are stacked into a data cube. Each cube containing hundreds of frames is then saved into a single FITS file with the benefit of reducing readout overheads during observations.

2.1 Parallelization

The NACO observation sequences of up to 4 h used in our campaign lead to roughly 100 GB of data and 100 000 frames. A straightforward single core reduction would take an extremely long time and would run out of memory before finishing, due to the many complex operations involved in the data reduction mostly based on fast Fourier transforms (FFT). The most widely used and easiest solution to this large data handling issue would be to average bin the data.

By suitably binning the frames, one can decrease the total amount of data to a quantity which fits the hardware limitations. The drawback is that valuable information gets lost in the binning process. First of all, the characteristics of atmospheric turbulence are not constant in time as well as the quality of the adaptive optics (AO) turbulence correction. The Strehl ratio for binned frames is the mean ratio of the frames in the bin, so that if half of the frames have poor AO correction the final binned frame will also have below average

Strehl, even though the other half of the frames had good Strehl. Furthermore, binning frames before re-centring and correcting for the field rotation smears out the companion PSF which in turn decreases its signal in the final product. This is why we decided not to rely on binning.

The different algorithms of our pipeline fit very well to a *data parallelism* scheme, which focuses on distributing the data across different parallel computing nodes. A master node shares the data between the slave nodes, which only have a fraction of the data to process. Once the slaves have finished, the data are gathered by the master and reassembled. Two different types of parallelization are used, which differ on the way the data are shared between the nodes. If the operations are pixel based, the spatial parallelization scheme is used. In this scheme the data cube is cut in pieces along the time axis, which means that each node receives one specific region of all the frames (see Fig. 1a). The other scheme, temporal parallelization, is used when the whole frame is needed for a specific operation. This is for example the case for shifts and rotations. The data cube is then separated in frame packages, and each node receives a different package containing full frames (see Fig. 1b).

The pipeline is implemented in PYTHON using C and FORTRAN libraries for the calculation intensive parts. The parallelization is achieved using the *Open Message Passing Interface* (OPENMPI; Gabriel et al. 2004). Parallelization can be distributed transparently among many different nodes, independently of their architecture. The interface between the PYTHON code and OPENMPI is handled by the MPI4PY module (Dalcín et al. 2008). All the data reduction steps given in this paper are parallelized, and based on the specificity of the process either in spatial or time parallelization.

3 ADI DATA REDUCTION

The pipeline was initially developed to reduce non-coronographic saturated ADI observations in L' band (3.8 μm) from the Coude Near Infrared Camera (CONICA) 1024 × 1024 InSb Aladdin 3 detector which is part of the NACO instrument at the VLT (Lenzen et al. 2003). It was later extended to process coronographic data from NACO, Gemini/Near-Infrared Coronagraphic Imager (NICI; Chun et al. 2008), VLT/Spectro-Polarimetric High-Contrast Exoplanet Research (SPHERE; Beuzit et al. 2010), and Subaru/Subaru Coronagraphic Extreme Adaptive Optics (SCExAO; Jovanovic et al. 2015b) in any band.

The concept of this pipeline is to preserve the companion photometry by relying on techniques which have the least impact on the signal. Conservation of the noise structure is also a priority in order to carry out an efficient noise subtraction. This is obtained by applying the geometric transformations for centring and rotation in Fourier space. Another important aspect of the pipeline is the scalable parallelization which makes it possible to run it on anything from a cluster for high performance computing (HPC) down to any modern laptop.

The reduction procedure is partitioned in four main parts which will be described hereafter:

    (i) registration (Section 3.1);
    (ii) image pre-processing (Section 3.2);
    (iii) PSF subtraction (Section 3.3);
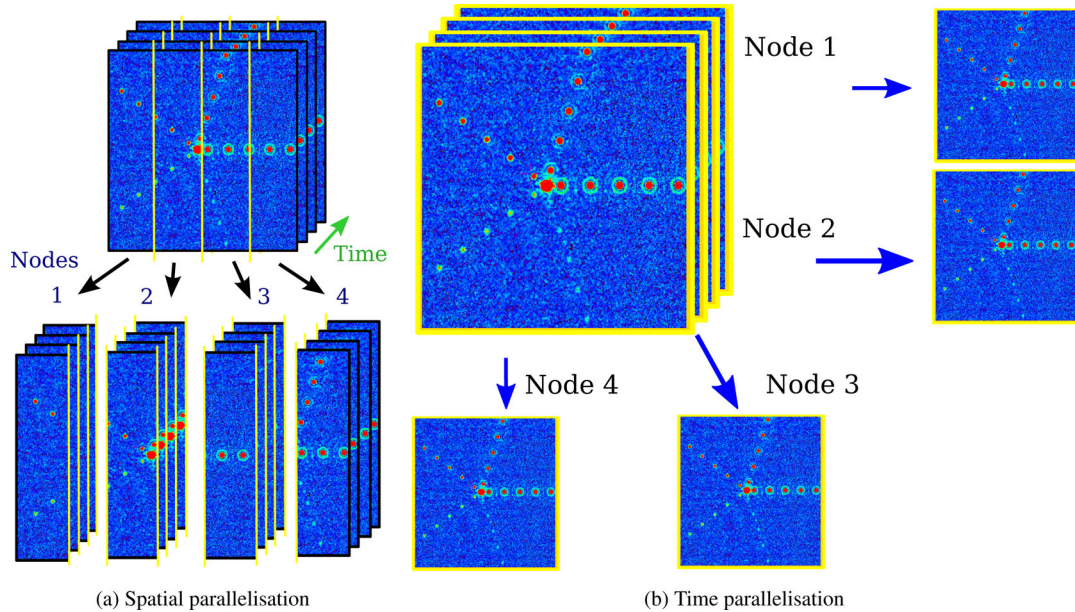   (iv) derotation (Section 3.4).

Figure 1. The two different data parallelism schemes used by the pipeline. In spatial parallelism (a), the observation frames are cut along the time axis. Each node is then processing a different region of the frame. In time parallelization (b), the frames are distributed among the nodes. In that case each node processes a different frame, or rather frame group.

3.1 Registration

The very first step of the reduction process is the registration. Every data frame is analysed and a table is generated containing for every frame the time of exposure, the parallactic angle, the star centre, and other PSF characteristics needed for the different reduction steps. Even though the data cubes are not modified at this step, it is a key element of the reduction as any error on the angle or star centre determination smears out the companion signal.

The registration process is also used to assess the quality of the AO correction for every single frame. This quality estimate is then later used to keep only the best frames by manually setting the constraints on the frame quality selection. This results in a kind of lucky AO.

3.1.1 Parallactic angle

The parallactic angle varies in time as the star moves in the sky, and its hour angle changes. Based on the 'Local Sidereal Time' (LST) given in seconds in the FITS header, and the right ascension α of the star (given in degrees), the hour angle h is given by

$$h = \frac{15\,\mathrm{LST}}{3600} - \alpha.$$

When observing with NACO in cube mode the frames are stored in data cubes of 30 to more than 500 frames, with only one single header. This implies that the hour angle for each single frame cannot directly be derived from the header, since solely one LST value is given. Thus the observing time for every single frame has to be interpolated in order to get the correct parallactic angles, which can become problematic when frame loss occurs.

3.1.2 Star point spread function registration

The determination of the star centre position is a complex task particularly in the case of saturated PSF and/or coronographic imaging

because the information at the core of the PSF is often lost. The determination of the centring method accuracy is also limited by the fact that the exact PSF centre is unknown. In order to make the PSF registration work on any combination of saturated/unsaturated, full/coronographic, and AO/non-AO modes the algorithm is split in a two stage process. First a basic centroid search is done, and once the centroid is found a two-dimensional function is fitted to the PSF, which takes into account possible coronographic or saturated cores by masking out pixels. This method also works if there is more than one star in the field of view, as long as no other star has exactly the same flux as the target.

The first step which is the centroiding algorithm searches for a patch of contiguous pixels above a given threshold value. If the patch size is within the range given by the user, the centre-of-mass of the patch is calculated. These values are then fed into the PSF fitting algorithm as initial values.

The pipeline uses the following form of the Moffat (1969) function to fit the PSF centred on x_0, y_0:

$$I(x, y, \alpha, \beta) = I_0 \frac{\beta - 1}{\pi \alpha^2} \left(1 + \frac{(x_0 - x)^2 + (y_0 - y)^2}{\alpha^2} \right)^{-\beta} + \mathrm{BG},$$

with α the seeing radius parameter, β the wing shape parameter, and BG the value of the background flux (Trujillo et al. 2001). The full width at half-maximum (FWHM) of the PSF is then given by the two fitting variables α and β:

$$\mathrm{FWHM}(\alpha, \beta) = 2\alpha \sqrt{2^{1/\beta} - 1}.$$

The Moffat fitting on the PSF is achieved using lmdif's modified Levenberg–Marquardt algorithm which is part of the FORTRAN MINPACK library (Moré, Garbow & Hillstrom 1980). The registration procedure achieves an accuracy better than 0.5 pixels on the star centre position, but its efficiency is highly dependent on the observational set-up.

With the growing tendency to highly saturate the PSF core or to use new generation coronographs, it becomes increasingly difficult to fit a Moffat function. A solution to this is the usage of satellite

speckles pioneered by Sivaramakrishnan & Oppenheimer (2006) using a reticulated wire grid, and recently brought to a new state of sophistication by Jovanovic et al. (2015a) using incoherent speckles generated with a pattern on the AO deformable mirror. Because of the chromaticity of speckles the satellites appear elongated on broad-band imaging giving them a shape which is difficult to fit with a conventional PSF model. On the other hand, the centroiding algorithm is a good candidate to implement a registration algorithm based on such satellite speckles.

3.1.3 Frame selection

In order to increase the signal-to-noise ratio of the companion, we keep only the frames with good AO correction. A first rough selection is based on the centroiding algorithm, by simply discarding frames where no centroid has been found.

The second step is based on the PSF geometry. When the detector integration time is more than a few seconds it can happen that bad tip-tilt correction create a sharp elongated PSF specially when in coronographic mode. Simple selection on maximum signal strength hardly detects this kind of frames, even though the PSF shape is asymmetric. Using the values of the PSF fitting has shown to be a robust way to verify the quality of the AO correction. Alternatively the satellite speckles position and intensity can be used when available.

Furthermore, by using individual frame PSF fitting instead of the widely used cross-correlation method, we can ascertain that only good and optimally centred frames are used. The independent centring also ensures that poor centring on one frame will not affect centring on the other frames.

Once the frame selection has been done an optional quick-look algorithm can be run which bins the recentred good frames to decrease the total number of frames to process in the following steps.

3.2 Image quality pre-processing

Image quality plays a key role when using Fourier transforms, because of the Gibbs (1898) phenomenon. This phenomenon causes a chequerboard pattern to appear with pixels alternately overshooting in the positive and negative values around an image discontinuity, such as bad pixels, saturated PSF or sharp image borders (Fig. 2). The reason for these oscillations is that a discontinuity would need infinite Fourier series to be correctly characterized, but since we are dealing with finite discrete Fourier transforms the discontinuity is not well approximated.
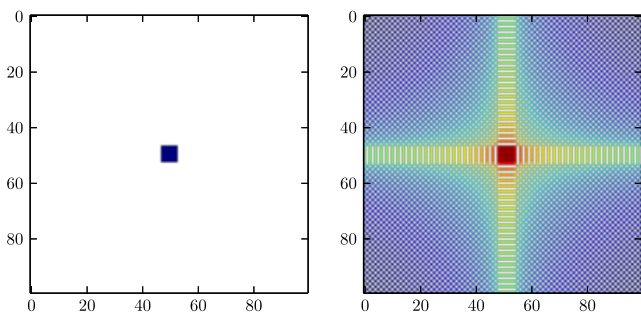

Figure 2. The single non-zero pixel surrounded by a background at zero (left-hand figure) is shifted in Fourier space by $\Delta x = 0.3$ pixels, and $\Delta y = 1.5$ pixels (right-hand figure), resulting in the typical pattern of the Gibbs phenomenon.

The frame preparation is thus a key step in order to use Fourier-based operations, where no deviant pixel should be left out. The two key steps for the image pre-processing are the sky generation and subtraction and the bad pixel correction.

3.2.1 Master sky

The sky background varies very rapidly when observing in the near-infrared especially in L' band. Observations are done in dithering mode in order to directly use the science frames to determine the sky background. To decrease the effect of the star PSF on the sky determination, a mask is applied on the region containing the star. A median of N time-contiguous cubes with N different masked dithering positions is then calculated. As dithering positions change every 1 or 2 s, and we are using four to five dithering positions, a sky frame is produced for every 5–10 s of observation. Using this method we manage to have a median sky as close as possible to the sky background of the science frames. Once the master skies have been generated, they are subtracted on each frame, using each time the master sky frame closest in time.

3.2.2 Bad pixel correction

First a bad pixel map is generated using either a master dark or a master sky. A simple σ clipping routine flags the deviant pixels, where bad pixels are defined as a pixels which values varies more than $C\sigma$ from the median value of the frame. By changing the value of the C coefficient the selection criteria can be adapted to ensure that all bad pixels are cleaned.

Since each bad pixel will be spread over many pixels as a consequence of the subpixel recentring and the derotation process, it is not possible to simply leave out a bad pixel. Not to mention the Gibbs phenomenon such pixels would induce. To clean the bad pixels, they are all first set to *NaN*. The bad pixel values are then replaced by the median of the neighbours, ignoring any *NaN* pixels. This ensures that bad pixel clumps are properly corrected.

3.3 Point spread function subtraction

The 'PSF subtraction' is the core of the reduction process. It is based on the well-established ADI algorithm which aims at subtracting the stellar PSF and speckles by using the field rotation in order to increase the sensitivity to surrounding point sources (Marois et al. 2006). We are generating a specific PSF for every single frame, in a similar way to LOCI (Lafrenière et al. 2007a) but without using any combination coefficients in an effort to preserve the flux.

For an observing sequence composed of C data cubes containing each N frames. The total number of frames is then $T = CN$. The frame $f_i = (c_j; n_k)$ will have a parallactic angle α_i given by

$$\alpha_i = \arctan\left(\frac{\cos\phi\sin h_i}{\cos\delta\sin\phi - \sin\delta\cos\phi\cos h_i}\right),$$

with ϕ the observatory latitude, δ the target declination, and h_i the hour angle at observing time t_i.

A PSF is generated for frame f_i using the frames f_x fulfilling the conditions on maximum time separation t_{max},

$$|t_k - t_i| < t_{max},$$

and minimum field rotation α_{min} such that

$$|\alpha_k - \alpha_i| > \alpha_{min} = 2\sin\left(\frac{n_{FWHM}\,FWHM}{2r_{min}}\right),$$

where $r_{\min}$ is the minimum radius in pixels to consider, FWHM is the PSF full width at half-maximum from the fitting (Section 3.1), and n_{FWHM} is the minimum number of PSF displacements to prevent companion self-subtraction.

3.3.1 Fourier shift

All the geometric operations on the image are based on one- and two-dimensional Fourier transforms. As a short reminder, we give the definition of Fourier transforms in two dimensions of a function $f(x, y)$:

$$\hat{f}(\nu_x, \nu_y) = \iint_{-\infty}^{\infty} f(x, y)\, e^{-i2\pi(\nu_x x + \nu_y y)}\, dx\, dy,$$

and its inverse Fourier transform:

$$f(x, y) = \hat{f}(\nu_x, \nu_y)^{\vee} = \iint_{-\infty}^{\infty} \hat{f}(\nu_x, \nu_y)\, e^{i2\pi(\nu_x x + \nu_y y)}\, d\nu_x\, d\nu_y.$$

The one-dimensional case is a trivial simplification of the two-dimensional case, and the notation for a one-dimensional Fourier transform along the x-axis will be noted by $\hat{f}(\nu_x, y)$, similarly the transform along the y-axis will be noted by $\hat{f}(x, \nu_y)$.

To perform a shift of the image we use the translation property of Fourier transforms. If $\hat{f}(\nu)$ is the Fourier transform of the one-dimensional function $f(x)$, then the Fourier transform of $f(x + a)$ is $\exp(-i2\pi\nu a)\hat{f}(\nu)$. Thus a spatial shift is equivalent to multiply the Fourier transform $\hat{f}(\nu)$ by a phasor $e^{-i2\pi\nu a}$. A shift along the x-axis in the two-dimensional case can thus be expressed as

$$f_x(x + a, y) = \int_{-\infty}^{\infty} e^{-i2\pi\nu_x a}\,\hat{f}(\nu_x, y)\, e^{i2\pi(\nu_x x)}\, d\nu_x,$$

and the more general case of a shift a in x and b in y is then obtained by a multiplication by a phasor $e^{-i2\pi(\nu_x a + \nu_y b)}$:

$$f(x + a, y + b)$$
$$= \iint_{-\infty}^{\infty} e^{-i2\pi(\nu_x a + \nu_y b)}\,\hat{f}(\nu_x, \nu_y)\, e^{i2\pi(\nu_x x + \nu_y y)}\, d\nu_x\, d\nu_y.$$

The frames we are recentring are defined on a finite area, whereas the shift property holds for infinite domains. We can nonetheless apply this operation to the frames provided we introduce a zero padding which also prevents apparition of Gibbs oscillations at the borders, but this implies that the operations have to be applied on at least double sized frames, which impacts significantly computational time.

3.4 Derotation

The final step of the ADI processing is to correct each frame for the field rotation and merge all the frames. In order to preserve the companion signal, and also to keep the noise structure unchanged we perform the rotation using Fourier transforms.

The rotation algorithm we are using is implemented by applying to Fourier transforms the property that a rotation matrix can be decomposed in three shear matrices (e.g. Unser, Thevenaz & Yaroslavsky 1995; Eddyy, Fitzgerald & Noll 1996; Welling, Eddy & Young 2006). This method is largely used in satellite imagery of the Earth and medical imaging. The galaxy image decomposition tool GALPHAT is an example of its previous use in astronomical imaging (Yoon, Weinberg & Katz 2011). We will only give a brief description of the method based on the detailed description of the algorithm given by Larkin, Oldfield & Klemm (1997).

Any rotation matrix $\mathbf{R}_\theta$ of a given angle θ can be expressed as the product of three shear matrices:

$$\underbrace{\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}}_{\mathbf{R}_\theta} = \underbrace{\begin{pmatrix} 1 & -\tan\frac{\theta}{2} \\ 0 & 1 \end{pmatrix}}_{\mathbf{S}_x} \underbrace{\begin{pmatrix} 1 & 0 \\ \sin\theta & 1 \end{pmatrix}}_{\mathbf{S}_y} \underbrace{\begin{pmatrix} 1 & -\tan\frac{\theta}{2} \\ 0 & 1 \end{pmatrix}}_{\mathbf{S}_x},$$

where $\mathbf{S}_x$ and $\mathbf{S}_y$ are shear matrices on the x- and y-axis, respectively.

To shear by a factor $a = \tan\frac{\theta}{2}$ in the x direction an image described by the function $f(x, y)$ we apply the transformation $s_x(x, y) = f(x + ay, y)$, which can be readily adapted to Fourier transforms using their shift property. The shear matrix $\mathbf{S}_x$ applied to the image $f(x, y)$ can then be expressed in terms of Fourier transforms as the function s_x:

$$s_x(x, y) = \int_{-\infty}^{\infty} \overbrace{e^{-i2\pi\nu_x ay}\, \underbrace{\hat{f}(\nu_x, y)}_{\mathrm{FT}}}^{\mathrm{IFT}}\, e^{i2\pi\nu_x x}\, d\nu_x,$$

the product $\mathbf{S}_y\mathbf{S}_x$ becomes by noting $b = -\sin\theta$,

$$s_{yx}(x, y) = \int_{-\infty}^{\infty} e^{-i2\pi\nu_y bx}\,\hat{s}_x(x, \nu_y)\, e^{i2\pi\nu_y y}\, d\nu_y,$$

and the rotation $\mathbf{S}_x\mathbf{S}_y\mathbf{S}_x$,

$$s_{xyx}(x, y) = \int_{-\infty}^{\infty} e^{-i2\pi\nu_x ay}\,\hat{s}_{yx}(\nu_x, y)\, e^{i2\pi\nu_x x}\, d\nu_x.$$

Similarly to the shift case, the frames need to be padded. This implies that in order to rotate one frame six FFT have to be applied on a double-sized frame, resulting in a significant increase of computation time compared to standard interpolation methods. This rotation technique can only be applied to such a large amount of frames thanks to parallelization.

4 PIPELINE PERFORMANCE

Contrast curves obtained for HD 142527 observations with NICI using GRAPHIC and PCA are displayed in Fig. 3. At small separation we reach a higher contrast than PCA, while the low-pass filtering
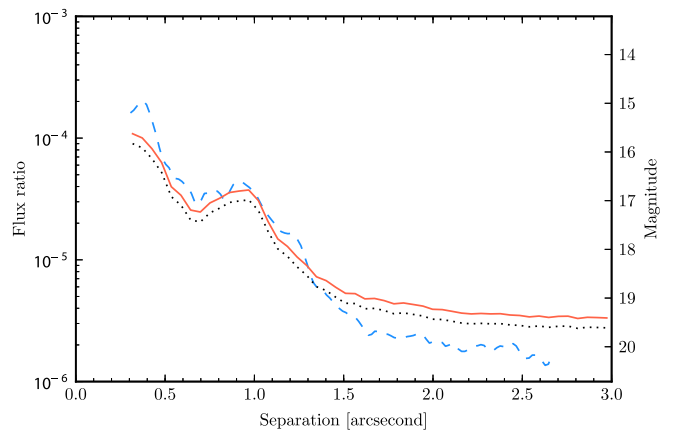

Figure 3. ADI detection limits for Gemini/NICI data in *CH4-K5%S* band, using a 0.22 arcsec semitransparent coronographic mask with $\approx 40°$ of field rotation. The three lines are detection limits at 5σ. The black dotted line is the detection limit achieved by GRAPHIC before correction for self-subtraction while the red solid line is the detection limit of GRAPHIC corrected in order to take into account the flux loss. The blue dashed line is the detection limit using a principal component analysis pipeline which was used in Casassus et al. (2013).

from interpolation results in better contrasts at higher separation where the noise is mainly Gaussian.

To test the pipeline performance we developed an algorithm to inject fake companions. These companions are generated by first including their signal into a plane wavefront which is then convoluted with a pupil based on the main optical characteristic of the used telescope. Using this technique we have a precise control on the companion flux, and furthermore the PSF scales precisely with wavelength. Poisson noise is finally added to these nearly perfect PSFs before adding them to the real science frame.

4.1 Performance of geometric transformations

To characterize the performance of the shift and rotation algorithms, we generated a test image with fake companions. This image is composed of an L' short-exposure image with a saturated PSF, which has been sky subtracted, cleaned from bad pixels, and median filtered. To this image we added fake companions with magnitudes differences to the star reaching stepwise from 1 to 8 and with separations from 0.5 to 6.5 arcsec with arcseconds steps, and Poisson noise was included in the process.

To test the performance of the shift algorithm, we shifted the original image in a Δx, Δy direction and then shifted it back in the opposite direction $(-\Delta x, -\Delta y)$. This double shifted image can then be compared with the original non-shifted image. By subtracting the original image from the double shifted, the effects induced by the different shifting algorithms become visible. The dotted black line in Fig. 4 shows the normalized root mean square of the original test frame, calculated in concentric annuli. The injected fake companions are causing the peaks at 0.5, 1.5, and 2.5 arcsec. The two additional lines show the root mean square of the difference between the original test frame and the double-shifted frames, where the interpolation and Fourier shift are represented by the blue dashed and red solid lines, respectively. These two lines would be flat with no root mean square if the shift algorithms were perfect. For the spline interpolation this is clearly not the case, with the curve reflecting two phenomenons in and out of the peaks which are both caused
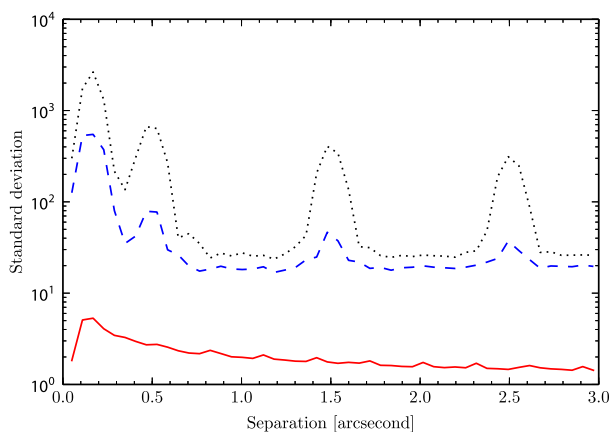
by the fact that interpolations in image space act as an uncontrolled low-pass filter. When the image is interpolated, the structure of the noise is modified so that the high-frequency noise is not removed by the subtraction, as it is missing from the double-shifted frame. This effect leaves an overall noise continuum. The effect on the peaks comes from the fact that the fake companions are flattened by the interpolation, so that part of the fake companion signal is not removed.

For the rotation algorithm we applied a similar test. We first rotated a test image by an angle $\alpha = 11.3$ followed by a rotation by an angle $-\alpha$. Ideally such a double rotation should return the original image. Changes in the image structure induced by the rotation can be found by subtracting the original image from the double-rotated one. The root mean square as a function of separation to the central PSF of the original image is plotted in Fig. 5 with a black dotted line, and the subtracted rotations using interpolation and Fourier shears are represented on the same figure by the blue dashed and red solid lines, respectively.

The images from the rotation test are represented in Fig. 6, where Fig. 6(a) is the original image, Figs 6(b) and (c) are the resulting image from the rotation followed by an inverse rotation using interpolation and the three-shear algorithm, respectively. The most striking difference between the two rotated images is the residual noise structure, the three-shear algorithm preserving the noise structure while the interpolation algorithm acts as an uncharacterized low-pass filter.

The different effects of the two rotation algorithms are even more evident when the original image is subtracted from the rotated images, as can be seen in Fig. 6(d) for the interpolation and in Fig. 6(e) for the three-shear algorithm. In the case of interpolation, the companions become dark blue points surrounded by signal which means that the companion PSF is spread, with signal being transferred from the centre of the PSF into the wings. The three-shear residuals show no structure, indicating that the star and companion signals are not altered by the rotation. Some numerical noise can be noted, but due to its very high frequency it can be filtered out without affecting the companion signal.
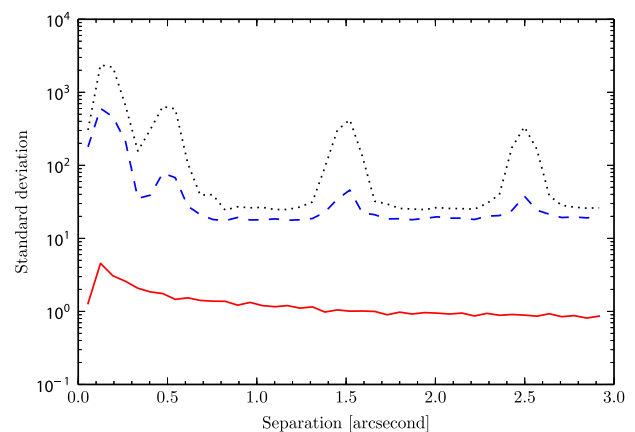

Figure 4. Root mean square representing the noise caused by the shift algorithms. As a reference the root mean square of the original image as a function of separation to the centre is given by the black dotted line. To test the algorithms we shift an image first by $\Delta x = 3.5$ pixels, $\Delta y = 2.7$ pixels, and then back. The difference between the original image and the one shifted back to the initial position is then plotted for the interpolation and Fourier shift, blue dashed and red solid lines, respectively.

Figure 5. Root mean square representing the noise caused by the rotation algorithms. As a reference the root mean square of the original image as a function of separation to the centre is given by the black dotted line. To test the algorithms we rotate the image first by an angle $\alpha = 11.3$ and then back. The difference between the original image and the one rotated back to the initial position is then plotted for the interpolation and Fourier rotation, blue dashed and red solid lines, respectively.
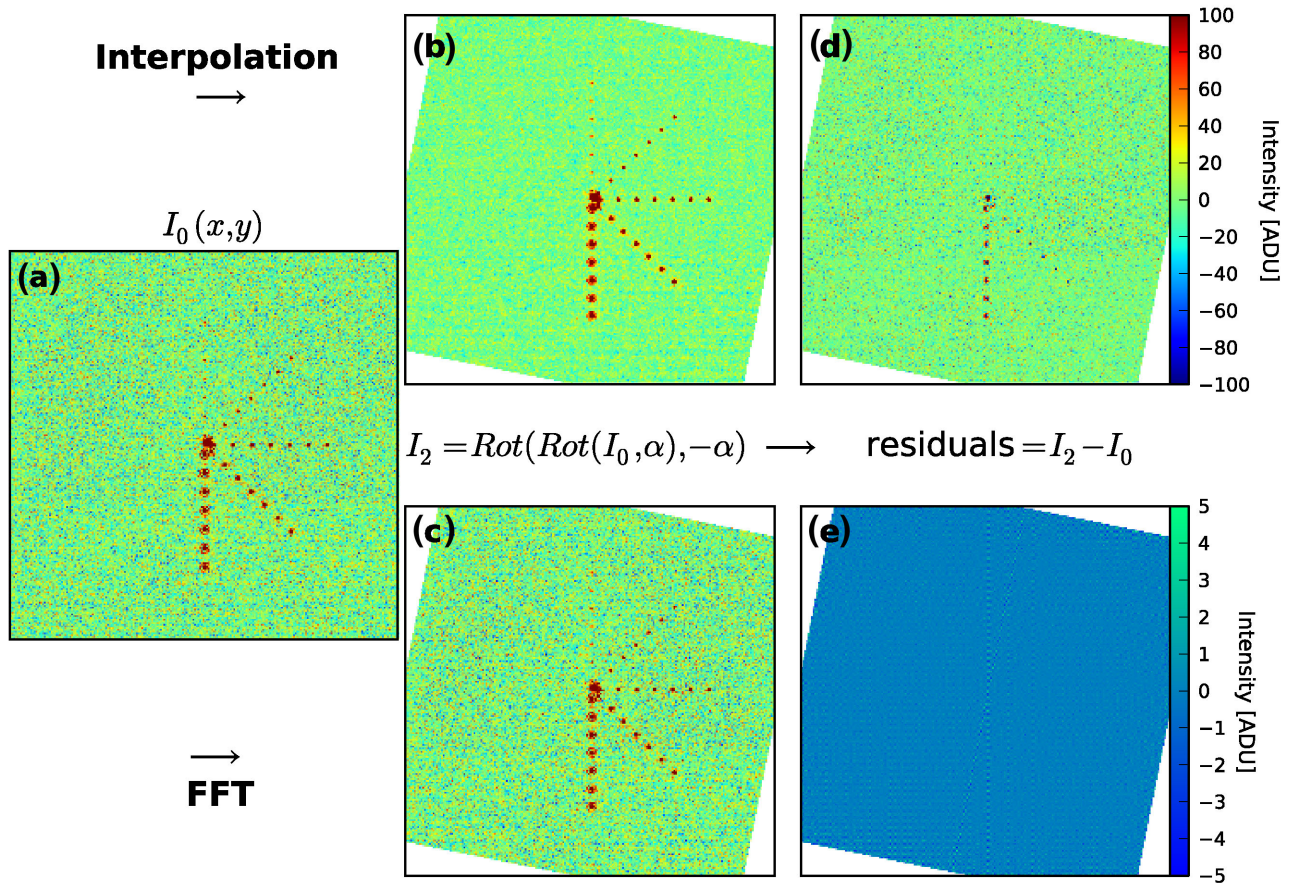
Interpolation
⟶

$I_0(x,y)$

(a)

(b)

(d)

$I_2 = Rot(Rot(I_0,\alpha),-\alpha) \longrightarrow$ residuals $= I_2 - I_0$

(c)

(e)

⟶
FFT

Figure 6. The original non-rotated test image is given in (a). The results of two consecutive 11°.3 and −11°.3 rotations of the original figure (a), using a third-order spline interpolation, and the three-shear algorithm are shown in (b) and (c), respectively. The residuals are computed by subtracting the original image from the interpolation rotated (d), and three-shear rotated (e) images. All panels have the same intensity scale, except for (e) where we reduced the cuts by a factor of 20 to reveal some of the induced noise.

Table 1. Magnitude difference of the injected companions and percentage of the recovered flux as a function of separation and bin size.

	1 arcsec	3 arcsec	5 arcsec
Injected Δmag	8	8–10	8–10
Recovered no binning	60 per cent	81–106 per cent	71–72 per cent
Recovered 10 frame bins	52 per cent	63–93 per cent	59–20 per cent
Recovered 50 frame bins	61 per cent	66–101 per cent	60–52 per cent

4.2 Photometric accuracy

Test data sets are created by injecting fake companions into all the raw frames with an angle following the field rotation. By injecting the companions into the raw frames we are able to take into account nearly all the steps of the reduction, namely image pre-processing, recentring, PSF subtracting, derotating, and final collapse. Centre and parallactic angle determination are the only two operations we cannot test with this method, as the companion injection already relies on these two parameters.

To illustrate the photometric accuracy we generated such a data set using a 2 h L' NACO observation sequence with a 120° field rotation, centred on the transit through meridian. The *no binning* line in Table 1 shows that in the worst case we have a flux loss of 40 per cent, which is similar to what is achieved with the PCA pipeline PYNPOINT but better by a factor of 2 compared to LOCI (Amara & Quanz 2012).

Our ability to inject fake companions in the raw frames is thus a key component to recover accurate photometry. By injecting the companions at the very beginning we can take into account all the effects that could decrease the companion flux during the various data reduction steps.

4.3 Effect of binning

The data set used to quantify the photometric accuracy was also used to analyse the effect of binning. To study this effect we binned the initial data set by combining each set of ten 0.2-s frames into a single median frame resulting in a smaller data set. We did the same by also combining 50 frames into a single one representing 10 s total integrations per frame.

The three data sets were then reduced using exactly the same parameters. The resulting detection limits along with retrieved photometry of the fake companions are plotted in Fig. 7. The detection limits increases beyond 3.5 arcsec because the result of dithering is that less frames are available at these separations which decreases the signal-to-noise ratio.

The percentage of recovered flux of the fake companions is given in Table 1. Flux recoveries above 100 per cent are caused by low signal-to-noise ratio, for companions at or below the detection limit.
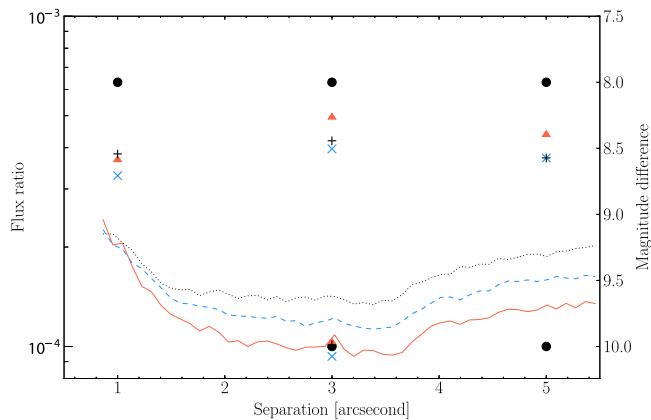
Figure 7. ADI detection limits for the same initial 2 h observation sequence (120° rotation). The injected fake companions are represented by the black circles. The non-binned detection limit and recovered companion flux are given by the red solid line and triangles, respectively. The 10 frame bins are represented by the blue dashed line and triangles, while the 50 bin results are given by the black dotted line and plus signs.

4.4 Performance versus observation duration

In order to test our observing strategy of long observations, we defined a specific test case. Taking one of our nearly 3 h observation data set, we added fake companions to the raw images. The observation was then reduced using three different data set subsamples. For the first one we used the whole data set trimmed in order to have as much observing time before and after transit at meridian. This results in a 2 h data set, with 1 h before and 1 h after meridian transit. We did the same for a 1 h subsample and a 30 min subsample from the same initial data set, each time centred on the meridian transit. The selected data set is of a star at a declination of −17°, which results in a smooth rotation rate. Targets that are crossing the meridian at zenith will have the entire field rotation happening in only a few minutes. For such specific targets the gain in sensitivity at small separation for longer observations will be very limited.

The resulting detection limits obtained by reducing the three subsamples with exactly the same reduction parameters are plotted in Fig. 8. The detection limits clearly show that the reduction was tuned for the inner region within 0.5 arcsec. At 0.3 arcsec separation,
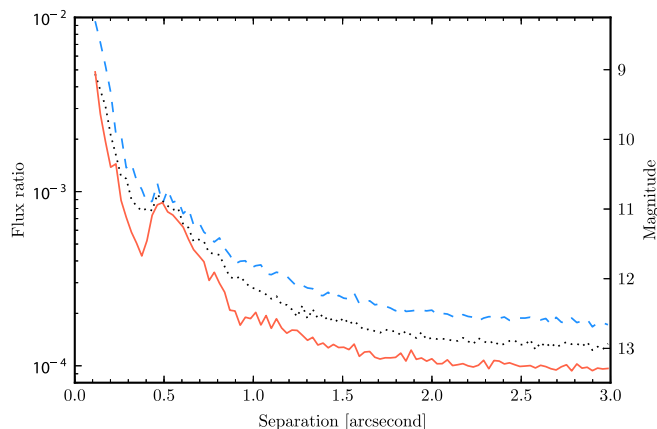

Figure 8. ADI detection limits for an observation in L' band of a 6.5-mag star at −17° declination. The data sets taken from this sequence are all centred on meridian transit using 30 min ($\approx$50° rotation), 1 h ($\approx$90°), and 2 h subsamples ($\approx$140°), plotted in blue dashed, black dotted, and solid red line, respectively.

the achieved magnitude differences are 7, 7.6, and 8 for 30 min, 1, and 2 h, respectively. At 1.5 arcsec separation these limits become 9, 9.3, and 9.7. With 2 h observations we thus gain 1 mag in sensitivity with respect to a short 30 min sequence, and half a magnitude with respect to a conventional 1 h observation. A 1 magn difference in L' band is what separates a $13M_J$ from a $20M_J$ companion, based on Allard ([2014](#)) at 1 Gyr.

5 CONCLUSIONS

We have presented a new pipeline for ADI which is the first, to our knowledge, to perform the critical steps of image recentring and field derotation in the Fourier domain and on individual frame. By doing so, we reduce most of the image smoothing introduced by interpolation and frame binning. The resulting image noise characteristic is also preserved. This theoretical approach is validated by several test cases, where we show that excellent dynamic range and photometry retrieval are obtained at separations shorter than 1.5 arcsec compared to a standard PCA analysis.

In addition, we validated the observing strategy that consists in observing the same target for at least two consecutive hours with a clear boost in performances of 1 mag at short separations in the case of a target with a 120° rotation in 2 h, representative of a typical ADI target. GRAPHIC is coded in PYTHON with a few C-modules and was designed to run on HPC clusters, with a minimum requirement of two cores and 4 GB RAM per core. Provided the right environment (48 cores, 192 GB RAM), GRAPHIC is able to process up to 100 000 2048 × 2048 frames with no binning thanks to massive parallelism in 9 h. This parallelization of the pipeline makes it also possible to implement further computationally demanding algorithms. Further development is planned in order to use graphics processing units (GPU) for a gain in processing time and wavelet filtering. GRAPHIC has also very recently been adapted to process VLT/SPHERE and Subaru/SCExAO data. GRAPHIC has also been applied on an extended source with results published in Casassus et al. ([2013](#)) under the development name PADIP.

REFERENCES

Allard F., 2014, in Booth M., Matthews B. C., Graham J. R., eds, Proc. IAU Symp. 299, Exploring the Formation and Evolution of Planetary Systems. Cambridge Univ. Press, Cambridge, p. 271
Amara A., Quanz S. P., 2012, MNRAS, 427, 948
Astropy Collaboration et al., 2013, A&A, 558, A33
Basri G., Marcy G. W., 1995, AJ, 109, 762
Beuzit J.-L. et al., 2010, in Coudé du Foresto V., Gelino D. M., Ribas I., eds, ASP Conf. Ser. Vol. 430, Pathways Towards Habitable Planets. Astron. Soc. Pac., San Francisco, p. 231
Biller B. A. et al., 2007, ApJS, 173, 143
Bowler B. P., Liu M. C., Shkolnik E. L., Dupuy T. J., Cieza L. A., Kraus A. L., Tamura M., 2012, ApJ, 753, 142
Casassus S. et al., 2013, Nature, 493, 191

Chauvin G. et al., 2010, A&A, 509, A52

Chun M. et al., 2008, Proc. SPIE, 7015, 70151V

Crepp J. R. et al., 2012, ApJ, 761, 39

Dalcín L., Paz R., Storti M., D'Elía J., 2008, J. Parallel Distributed Comput., 68, 655

Eddyy W. F., Fitzgerald M., Noll D. C., 1996, Magn. Resonance Medicine, 36, 923

Gabriel E. et al., 2004, in Kranzlmueller D., Kacsuk P., Dongarra J. J., eds, Lecture Notes in Computer Science, Vol. 3241, Recent Advances in Parallel Virtual Machine and Message Passing Interface. Springer-Verlag, Berlin, p. 97

Gibbs J. W., 1898, Nature, 59, 200

Heinze A. N., Hinz P. M., Kenworthy M., Meyer M., Sivanandam S., Miller D., 2010, ApJ, 714, 1570

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

Janson M. et al., 2013, ApJ, 773, 73

Jones E. et al., 2001, SciPy: open source scientific tools for Python. http://www.scipy.org/

Jovanovic N., Guyon O., Martinache F., Pathak P., Hagelberg J., Kudo T., 2015a, ApJ, 813, L24

Jovanovic N. et al., 2015b, PASP, 127, 890

Lafrenière D., Marois C., Doyon R., Nadeau D., Artigau É., 2007a, ApJ, 660, 770

Lafrenière D. et al., 2007b, ApJ, 670, 1367

Lagrange A.-M. et al., 2009, A&A, 493, L21

Larkin K. G., Oldfield M. A., Klemm H., 1997, Opt. Commun., 139, 99

Lenzen R. et al., 2003, Proc. SPIE, 4841, 944

Lenzen R., Close L., Brandner W., Biller B., Hartung M., 2004, Proc. SPIE, 5492, 970

Liu M. C., 2004, Science, 305, 1442

Marois C., Lafrenière D., Doyon R., Macintosh B., Nadeau D., 2006, ApJ, 641, 556

Marois C., Macintosh B., Barman T., Zuckerman B., Song I., Patience J., Lafrenière D., Doyon R., 2008, Science, 322, 1348

Masciadri E., Mundt R., Henning T., Alvarez C., Barrado y Navascués D., 2005, ApJ, 625, 1004

Mayor M., Queloz D., 1995, Nature, 378, 355

Moffat A. F. J., 1969, A&A, 3, 455

Moré J. J., Garbow B. S., Hillstrom K. E., 1980, User guide for MINPACK-1. Argonne National Laboratory Report ANL-80-74, Argonne, IL

Nakajima T., Oppenheimer B. R., Kulkarni S. R., Golimowski D. A., Matthews K., Durrance S. T., 1995, Nature, 378, 463

Nielsen E. L. et al., 2013, ApJ, 776, 4

Oliphant T. E., 2007, Comput. Sci. Eng., 9, 10

Pérez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21

Racine R., Walker G. A. H., Nadeau D., Doyon R., Marois C., 1999, PASP, 111, 587

Rebolo R., Osorio M. R. Z., Martín E. L., 1995, Nature, 377, 129

Schneider G., Silverstone M. D., 2003, Proc. SPIE, 4860, 1

Sivaramakrishnan A., Oppenheimer B. R., 2006, ApJ, 647, 620

Soummer R., Pueyo L., Larkin J., 2012, ApJ, 755, L28

Trujillo I., Aguerri J. A. L., Cepa J., Gutiérrez C. M., 2001, MNRAS, 328, 977

Unser M., Thevenaz P., Yaroslavsky L., 1995, IEEE Trans. Image Processing, 4, 1371

Vigan A. et al., 2012, A&A, 544, A9

Wahhaj Z. et al., 2013, ApJ, 773, 179

Welling J. S., Eddy W. F., Young T. K., 2006, Graphical Models, 68, 356

Yoon I., Weinberg M. D., Katz N., 2011, MNRAS, 414, 1625

This paper has been typeset from a TEX/LATEX file prepared by the author.