SPHERE non-common path aberrations measurement and pre-compensation with optimized phase diversity processes

J.-F. Sauvage[1a], T. Fusco[1,2], L. Mugnier[1], B. Paul[1], C. Petit[1], and K. Dohlen[2]

[1] ONERA-DOTA, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex
[2] Laboratoire d'Astrophysique de Marseille (UMR7326 - CNRS-INSU, Université d'Aix-Marseille)

Abstract. The SPHERE instrument is the 2nd generation instrument dedicated to exoplanet direct imaging and characterization. The extremely high imaging performance required by these observation mode calls for a high performance AO system. Particularly, this system has to provide a wavefront corrected from turbulent and internal defects. We present here the experimental results for the complete focal-plane calibration procedure of the SPHERE instrument internal defects. An optimized phase diversity method is applied allowing to deal with model uncertainties in the image formation (noise, residual background, amplitude fluctuation, sampling factor, defocalisation distance, object size and SH-model for reference slope modifications) The full procedure includes both Non-Common Path Aberrations (NCPA) compensation at the level of the coronagraphic mask using the eXtreme AO system itself (by the mean of modification of the filtered SH WFS reference slopes during close loop operations), but also additional measurements of IRDIS differential optical path aberrations for post processing of dual-band images. We validate the algorithm and the pre-compensation procedure using data obtained during the first eXtreme AO bench (SAXO) integration and tests. We also applied the Diversity tool on stand-alone IRDIS data obtained at LAM during its local integration. In both cases, we demonstrate the robustness and the ultimate performance (nanometric precision of the residual quasi-static pattern) of the phase diversity approach which will allow us to obtain nanometric accuracy on the final SPHERE system.

1 Introduction

The direct imaging of exoplanets is a challenging goal of today's astronomy. Photons emitted by exoplanets carry information about the chemical composition of the planet's atmosphere, and may testify the evidence of life possibility on it. Nevertheless, the fact that planets are often very faint with respect to their parent star and very close makes their observation very difficult. Exoplanet direct detection and characterization is the goal of systems currently being integrated, SPHERE[1] and GPI[2]. These systems combine extreme adaptive optics (AO) systems, coronagraphic devices, and sophisticated dedicated calibration and post-processing methods. The AO system allows a highly efficient correction of the wave-front induced by atmospheric turbulence. This correction performs a concentration of both star and planet photons and helps disentangling one from the other. The coronagraphic stage allows the rejection of most of the star signal, and its associated photon noise. The ultimate performance of such systems is limited by the level of the quasi-static speckles[3], which are due to residual quasi-static aberrations, and thus by the achievable precision in the system calibration. While a system calibration performed during daytime will eliminate most of the star image residuals which could be confused with planets, this calibration will nevertheless be imperfect, and the aberrations of the system will evolve between the calibration and the scientific acquisition. For SPHERE, the baseline calibration of the quasi-static aberrations will be done off-line (off-line meaning with an internal point-source in between scientific acquisitions), by means of phase diversity[4] in pseudo-closed-loop mode[5]. The detail of this calibration, as well as experimental results obtained during the AIT of the Sphere eXtreme AO system will be detailed in the first section. Although phase diversity has recently been extended to perform on-line calibration for AO-corrected telescopes[7], it has not yet been extended to coronagraphic imaging as it assumes a shift-invariant imaging model. The second section presents an alternative approach, where the classical phase diversity (classical meaning without coronagraph) accounts for a coronagraphic PSF.

[a] jean-francois.sauvage@onera.fr

2 Non-Common Path Aberrations

2.1 Classical case: non coronagraphic imaging

The non-common path aberrations are the main limitation of high dynamic XAO systems. As illustrated on the Figure1, these aberrations are the quasi-static aberrations introduced all the way on the imaging path, from the beam splitter down to the imaging camera. These aberrations are therefore unseen, and remain uncorrected by the AO loop.

A procedure, called Pseudo-Closed Loop and detailed in[5], allows the measurement and compensation of these aberrations. The measurement procedure relies on the phase diversity focal plane Wave Front Sensor to estimate the aberrations introduced all the way down to the imaging camera. The compensation of these aberrations is done by the AO system itself, via a modification of the reference slopes. Closing the loop on modified reference slopes allows one to precompensate the NCPA aberrations.

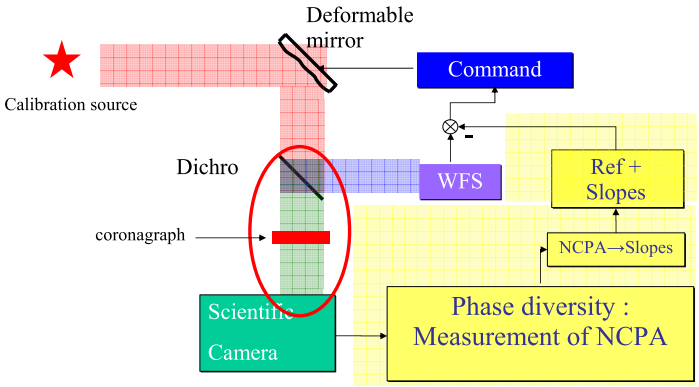

Fig. 1. Scheme of an AO system. The main component are indicated. The NCPA are introduced on the imaging path, all the way down to the scientific camera.

2.2 Coronagraphic imaging

The case of coronagraphic imaging complicates slightly the procedure. Firstly, the classical phase diversity algorithm is based on a classical image formation model, that does not account for a coronagraphic mask. Secondly, among the NCPA aberrations, the ones introduced upstream of the coronagraphic mask are known to be the main source of light residual in the imaging focal plane. On the Figure1, the coronagraph is the limitation between the upstream and downstream aberrations.

3 SPHERE baseline solution for NCPA calibration

The SPHERE baseline solution for NCPA calibration is described in [1]. In order to estimate the quasi-static aberrations upstream of the coronagraphic plane, a process in two steps is done. All the calibration process only uses classical phase diversity, and therefore needs to remove the coronagraphic mask from the optical path. The different focal planes are illustrated on Figure2.

- Firstly the term ϕ_{IRDIS} is estimated by using the Pseudo-Closed Loop procedure in a classical way. The calibration source is at FP1, the phase of diversity is produced by the deformable mirror, and IRDIS images are used,
- then the sole downstream aberrations ϕ_{DOWN} are estimated by a one-shot phase diversity measurement. This measurement uses a calibration source positionned at FP3, the phase of diversity is produced by the source motion along optical axis (equivalent to a defocus) and uses IRDIS images,
- the upstream aberrations ϕ_{UP} are deduced from the difference $\phi_{IRDIS} - \phi_{DOWN}$.

The PCL procedure used to perform the calibration of ϕ_{IRDIS} aberration consists in an iterative procedure summarized in the following points:

- Close the loop on initial reference slopes,
- acquisition focused and defocused image,
- processing images with classical phase diversity,
- modification of reference slopes to pre-compensate residual phase

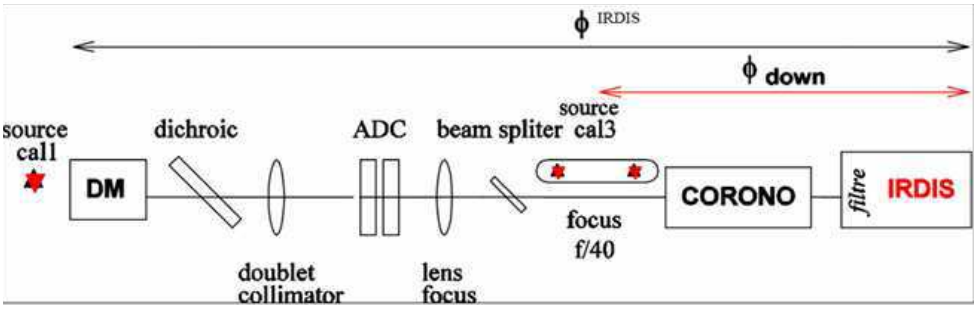

Fig. 2. Detailed scheme of the optical path, from entrance calibration source to the imaging camera (IRDIS). The source cal1 is used for calibration of ϕ_{IRDIS} aberrations. The source cal3 is used for the calibration of aberrations downstream of the coronagraph ϕ_{DOWN}.

4 Experimental validation of NCPA compensation on SAXO

The capacity of PCL procedure to measure and compensate for NCPA has been validated during the AIT of the XAO system of SPHERE. As indicated in the previous section, the baseline solution implies non-coronagraphic imaging, and therefore implies classical phase diversity. As a first result, only the whole aberrations (ϕ_{IRDIS}) have been measured and compensated for, using the PCL procedure.

In the absence of the IRDIS camera during SAXO AIT, the images have been acquired on the DTTS (Differential Tip-Tilt Sensor). This IR detector performs imaging at 1.6 microns, in a focal plane close to the coronagraphic focal plane. Images of size 128x128 pixels have been used, at a high SNR in order to ensure an accurate phase diversity estimation. The phase diversity measurement is done in the pixel basis, and performed on a 32x32 pixels array.

The Figure3 shows the RMS value (in nm) of the residual phase estimated by phase diversity after NCPA compensation. The first point (iteration 0) corresponds to the phase residual without any compensation, which is 32nm RMS. This means to an internal Strehl Ratio of 89.9% in the visible, or 98.5% at in H-band. This SR corresponds to the bench internal quality in closed loop, without any NCPA compensation. Then the iterative PCL procedure allows to decrease the residual WFE down to 4 nm in 4 iterations. This means a Strehl ratio of 99.8% in the visible, and 99.9% in the H-band.

The estimation of Strehl ratio, which quantifies the image quality is a hard task. The value given here are computed from the phase estimation given by phase diversity, following the Marechal criterion:

$$SR = e^{-\sigma^2_{phi}}. \tag{1}$$

The results presented here validate the capacity of PCL procedure to estimate correctly the quasi-static aberrations of the whole imaging path, without any coronagraphic mask. Still, the procedure presents several identified limitations:

- The need for removing the coronagraphic mask from the system during the calibration is a hard system constraint, as the mask has to be replaced at the same place with a high precision.
- The differential measurement increases the estimation noise by a factor $\sqrt{2}$.
- The accurate estimation of aberrations needs an iterative procedure, as implemented in the PCL. The sole ϕ_{IRDIS} aberrations are measured with an iterative process, the ϕ_{DOWN} aberrations show a reduced accuracy precision.

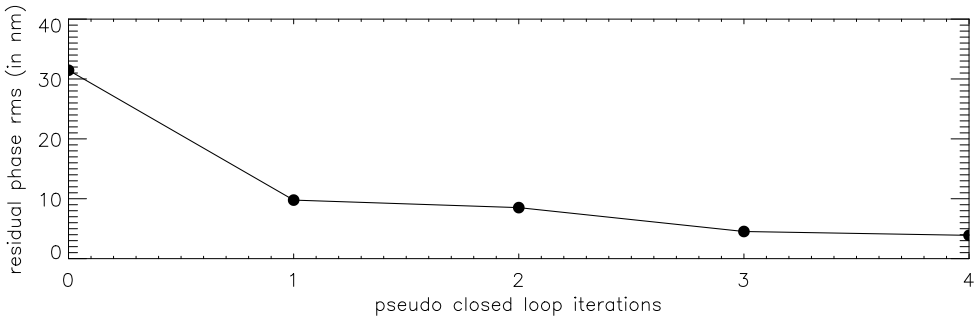

Fig. 3. Residual WFE wrt iteration of PCL. 4 iterations are shown. The first point (iteration 0) corresponds to the WFE without any compensation.

5 Coronagraphic phase diversity (COFFEE): a new approach

A new approach to the baseline solution is proposed here. This approach is based on the extension of the phase diversity wave-front sensor to the coronagraphic imaging. The extension of phase diversity to coronagraphic images is of prime interest. Firstly, the light distribution in a coronagraphic image is mainly driven by the aberrations upstream of the coronagraphic mask. The estimation of aberration from coronagraphic images therefore directly gives information about upstream aberrations. Secondly, the coronagraph does not need to be removed, and the calibration no more needs the second calibration source at FP3, which releases the constraint on the system complexity.

5.1 Modification of classical phase diversity imaging model

The extension of phase diversity to coronagraphic imaging is done by modifying the image formation model assumed in the phase diversity measurement. A simplified model is proposed here, based on the propagation of light in the optical system presented in Fig. 4. The optical coronagraphic imaging system is composed of an AO-capable telescope, a coronagraph, and a detector plane. All aberrations are assumed to be introduced in a pupil plane. We distinguish between the upstream aberrations ϕ_{up} and the downstream ones ϕ_{down}.

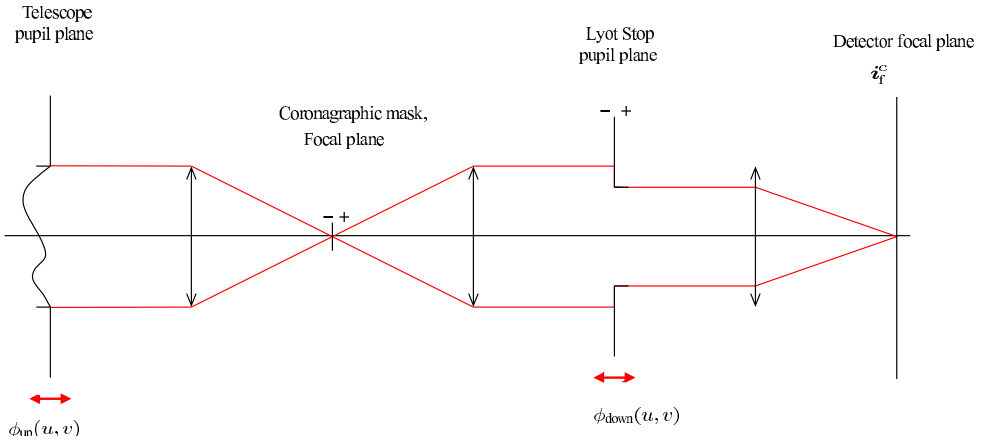

Fig. 4. Optical scheme of a coronagraphic imager. The upstream and downstream static aberrations, as well as the adopted notations are denoted ϕ_{up} and ϕ_{down} respectively.

As in conventional (i.e. non-coronagraphic) phase diversity, we shall assume that we use the focal plane camera to record at least two images that differ only from a known aberration, or *phase diversity* ϕ_{div}, which can be introduced conveniently upstream of the coronagraph by the DM of the AO subsystem, as in the non-coronagraphic case. We shall denote by

All aberrations, denoted by φ_x where x = "up" or "down" or "div", are each expanded on a basis $\{b_x^k\}_k$, which is typically either Zernike polynomials or the pixel indicator functions in the corresponding pupil plane:

$$\varphi_x(u,v) = \sum_k \phi_x^k \, b_x^k(u,v), \tag{2}$$

where the summation is, in practice, limited to the number of coefficients considered sufficient to correctly describe the aberrations. We shall denote by $\boldsymbol{\phi}_x$ the vector concatenating the set of unknown aberration coefficients ϕ_x^k.

Assuming hereafter that the observed object is a point-source, a discrete model of these images is:

$$\boldsymbol{i}_{\mathrm{f}}^c = f \, . \, \boldsymbol{h}_{\mathrm{d}} \star \boldsymbol{h}_{\mathrm{c}}(\boldsymbol{\phi}_{\mathrm{up}}, \boldsymbol{\phi}_{\mathrm{down}}) + \boldsymbol{n} \tag{3}$$

$$\boldsymbol{i}_{\mathrm{d}}^c = f \, . \, \boldsymbol{h}_{\mathrm{d}} \star \boldsymbol{h}_{\mathrm{c}}(\boldsymbol{\phi}_{\mathrm{up}} + \boldsymbol{\phi}_{\mathrm{div}}, \boldsymbol{\phi}_{\mathrm{down}}) + \boldsymbol{n}' \tag{4}$$

where f is the recorded flux, $\boldsymbol{h}_{\mathrm{d}}$ the known detector PSF, $\star$ denotes a discrete convolution, $\boldsymbol{n}$ and $\boldsymbol{n}'$ represent the measurement noises.

$\boldsymbol{h}_{\mathrm{c}}(\boldsymbol{\phi}_{\mathrm{up}}, \boldsymbol{\phi}_{\mathrm{down}})$ is our model for the coronagraphic point source image, a.k.a. "coronagraphic PSF" although it is of course field-dependent. In this paper it is taken from [6] and depends on $\boldsymbol{\phi}_{\mathrm{up}}$ and $\boldsymbol{\phi}_{\mathrm{down}}$. This model assumes that the coronagraphic PSF is the one of a perfect coronagraph. Note that this model could also be used for long-exposure images, in which case it has an additional parameter, namely the residual phase structure function.

The measurement noise $\boldsymbol{n}$ (and similarly $\boldsymbol{n}'$) comprises both photon and detector noises. Because calibration is performed with high photon levels, we adopt a non-stationary white Gaussian model, which is a good approximation of a mix of photon and detector noises. Its variance is the sum of the photon and detector noise variances: $\sigma_n^2(k,l) = \sigma_{\mathrm{ph}}^2(k,l) + \sigma_{\mathrm{det}}^2$. The former can be estimated as the image itself thresholded to positive values, and the latter can be calibrated prior to the observations.

It is to be noted that the phase estimation for the coronagraphic phase diversity is performed in the same way as in the classical phase diversity. This means by using the minimisation of a maximum a posteriori criterion, as detailed in[5]. The parameters estimated during the minimisation are the upstream aberrations $\boldsymbol{\phi}_{\mathrm{up}}$, and the downstream aberrations $\boldsymbol{\phi}_{\mathrm{down}}$. The flux f is analyticaly estimated.

5.2 Optimisation of the phase of diversity

The choice of the diversity phase is crucial as it has at least three impacts. Indeed, depending on the chosen diversity:

1. sign ambiguities on the sought phase can be removed—or not;
2. the noise propagation from the images to the estimated phase can vary;
3. local minima in the criterion can be removed, or at least pushed away from the global minimum—or not.

In order to address item (1), we shall, as for conventional phase diversity, use a diversity that is an even phase map (see, e.g., [4]). Note that in this whole section, we only consider upstream aberrations, and the downstream aberrations are null.

In order to address item (2), we plot the total intensity in the diverse image as a function of the chosen diversity, the latter ranging from the traditional defocus Z_4 to spherical aberration Z_{11}. For any of these curves, the maximum sensitivity (to the same Zernike as the chosen diversity at least) is obtained when the slope is maximum; it is obtained for RMS values of the diversity typically between 0.7 rad RMS and 0.8 rad RMS.

In order to adress item (3), we noticed that using a phase of diversity composed of a mix of defocus and astigmatism is a suitable one: $a_4^{\mathrm{div}} = a_5^{\mathrm{div}} = 0.8$ rad RMS. This is an empirical choice, and needs furthermore study. Particularly the question of optimal phase of diversity is still pending.

5.3 Validation and performance estimation by simulations

The Figure 5 shows coronagraphic PSF simulated for various types of coronagraphic mask. PSF on the left is simulated by using the perfect coronagraph model, which is the model assumed by COFFEE. The image is sampled at 5 x shannon at $\lambda = 0.55\,\mu$m, with 30 nm RMS of upstream aberrations. For the proof of concept presented herein, these aberrations are limited to two tens of Zernike polynomials (up to Z_{23}).

The validation of COFFEE is done by calibrating interaction matrix. This matrix is the concatenation of COFFEE measurement $\hat{\phi}_{up,k}$ obtained with the following procedure:

- simulation of coronagraphic images $\boldsymbol{i}^c_{\mathrm{f}}$ and $\boldsymbol{i}^c_{\mathrm{d}}$, assuming $\phi_{up} = a_k Z_k$ and $a_k = 30nm$, $phi_{do} = 0$, no noise, perfect coronagraph model
- estimation of $\hat{\phi}_{up,k}$ and $\hat{\phi_{do,k}}$ with COFFEE

The column k of interaction matrix is therefore composed of the estimated upstream aberrations, decomposed on Zernike modes, when the simulated upstream phase is composed of the sole zernike mode Z_k, with amplitude 30nm RMS. If the estimation is correct, the interaction matrix is diagonal. If the estimation is not correct, terms will appear out of the diagonal. This interaction matrix quantifies the behavior of COFFEE with respect to Zernike modes.

The Figure 6 shows on the left the corresponding interaction matrix. The estimation by COFFEE is quite good as the interaction matrix is highly diagonal. The estimation error is mainly driven by numerical aspects, and equals to $10^{-7}nm$.

The perfect coronagraph imaging model is a simplified model. We thus also validated COFFEE robustness to different coronagraph models. The 3 other PSF on Figure 5 show coronagraphic PSF simulated with different coronagraph models. Roddier and Roddier Phase Mask [RRPM], Four Quadrant Phase Mask [F4PM], and Lyot coronagraph. The same upstream and downstream aberrations are considered in all cases.

The same interaction matrix is calibrated for each of the considered coronagraph model, following the procedure:

- simulation of coronagraphic images $\boldsymbol{i}^c_{\mathrm{f}}$ and $\boldsymbol{i}^c_{\mathrm{d}}$, assuming $\phi_{up} = a_k Z_k$ and $a_k = 30nm$, $phi_{do} = 0$, no noise, coronagraph model X (X=RRPM, F4PM or Lyot)
- estimation of $\hat{\phi}_{up,k}$ and $\hat{\phi}_{do,k}$ with COFFEE

The corresponding interaction matrices are shown on Figure 6. The shape of these three interaction matrices clearly show that COFFEE, implemented with a perfect coronagraph model is robust to the RRPM coronagraph model. Indeed, the RRPM interaction matrix is the only one to be diagonal.

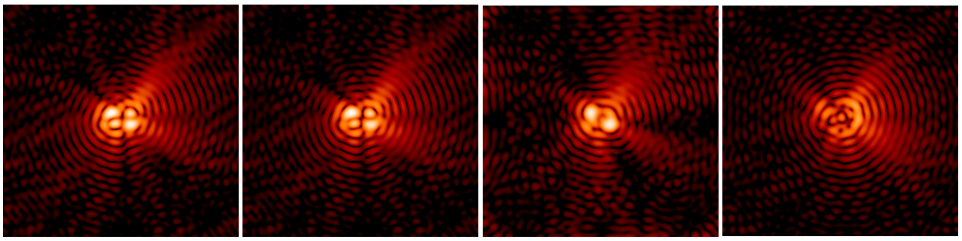

Fig. 5. Example of simulated coronagraphic PSF for various coronagraph model. Perfect coronagraph, RRPM, F4PM and Lyot coronagraph are considered.

5.4 Experimental validation, on a Tip ramp estimation

Some very first experimental validations have been performed on an ONERA test bench, coupled to an apodized RRPM coronagraph [7]. The bench design is the following one:

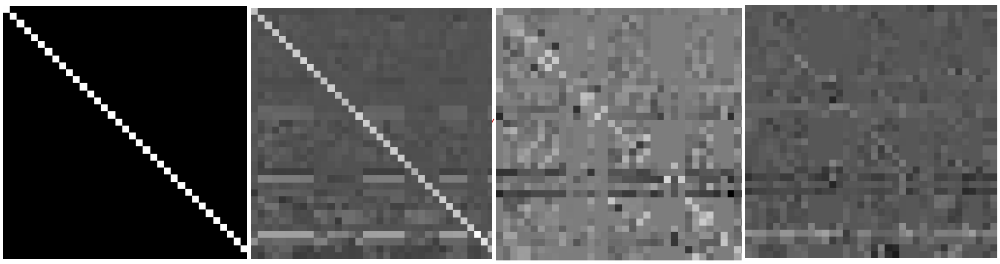

Fig. 6. Interaction matrix for the different coronagraph considered in Fig.5.

- Reference fibered source in the visible, $\lambda = 675nm$
- optical relay, magnification 1, allowing a translation of the PSF on the coronagraphic mask
- RRPM adapted to $\lambda = 677nm$, without apodisation
- Lyot stop of 95% of entrance pupil
- imaging camera, sampled at 5xshannon

The COFFEE WFS has been validated on experimental data acquired on BOA. The purpose of this validation is to estimate a Tip ramp introduced upstream of the coronagraph. This Tip ramp is introduced by translating the source along the X direction, therefore translating the PSF with respect to the RRPM mask.

This preliminary validation is performed on a bench without AO capacity. For sake of simplicity, the phase of diversity is chosen to a simple defocus, introduced by translating the source along the Z axis, so as to introduce a diversity $\phi_{div} = 0.8radZ_4$.

Pairs of focused and defocused images are acquired for various X positions of the source, going from $a_2 = -1.5radians$ to $a_2 = 1.5radians$, which means a translation of $-\lambda/D$ to λ/D. One estimation is performed for each X position by COFFEE. In order to illustrate the bench coronagraphic images, the Figure 7 shows on [left, top] side one coronagraphic PSF obtained on the bench. The X position of the source is voluntarily set to a non-zero value, to emphasize the effect on the PSF shape. On the [left, bottom] side of the Figure 7, a simulated coronagraphic PSF is computed using the following items:

- a perfect coronagraph model
- the upstream aberrations estimated by COFFEE on the experimental data of Figure 7, [left top].

The Figure 7 shows the result of upstream Tip ramp estimation. The X axis shows the introduce Tilt, varying from -1.5 radians to 1.5 radians. The Y axis shows the Tip value estimated by COFFEE on experimental images. The estimations are very close to simulation result performed on similar simulated data. The simulation accounts for the bench limitations, which are:

- A RRPM coronagraph, non adapted to wavelength,
- no pupil apodization.

The non linear behavior is mainly driven by the limitations of the bench, and are to be resolved in the coming weeks with a RRPM adapted to the wavelength of the source, and a pupil apodisation.

6 Conclusion

The baseline solution for NCPA compensation has been shown. The procedure performs a measurement and pre-compensation of NCPA at the coronagraphic level, and not in the detector plane, thanks to a differential measurement. A first step has been demonstrated during SAXO AIT, with the compensation of the NCPA down to 4nm RMS at the detector level. A novel solution, based on an extension of phase diversity WFS to the coronagraphic case has been developped, and validated by means of simulation. The so-called COFFEE estimator is based on a perfect coronagraph model, and allows the estimation of ϕ_{up} as well as ϕ_{do}. A good compliance to RRPM coronagraph has been demonstrated in simulation. The very first experimental results show a good compliance with simulation, for the case of a Tilt ramp estimation. Some full experimental validation has to be done on high orders as well as low orders upstream aberrations. Simulations have to be performed to quantify the capacity of the novel solution on the SPHERE system.

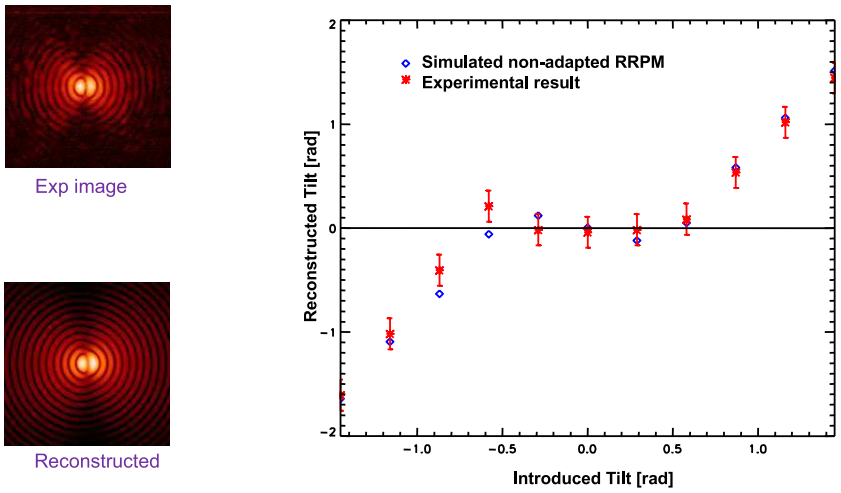

Fig. 7. [left,top] Experimental coronagraphic image acquire on ONERA test bench. The source is X translated in order to introduce 0.7 radians of upstream Tip. [left,bottom], the simulated PSF computed from the COFFEE estimation of Tip of left image. [right] COFFEE estimation of an upstream Tip ramp. The result of simulation shows coherence with the experimental result. Simulation accounts for the bench limitation (RRPM not adapted to wavelength, no apodisation).

References

1. T. Fusco, G. Rousset, J.-F. Sauvage, C. Petit, J.-L. Beuzit, K. Dohlen, D. Mouillet, J. Charton, M. Nicolle, M. Kasper and P. Puget, *High order Adaptive Optics requirements for direct detection of Extra-solar planets. Application to the SPHERE instrument.*, Opt. Express, **14** (17), pp. 7515–7534 (2006).
2. B. A. Macintosh, J. R. Graham, D. W. Palmer, R. Doyon, J. Dunn, D. T. Gavel, J. Larkin, B. Oppenheimer, L. Saddlemyer, A. Sivaramakrishnan, J. K. Wallace, B. Bauman, D. A. Erickson, C. Marois, L. A. Poyneer and R. Soummer, *The Gemini Planet Imager: from science to design to construction*, Dans *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. **7015** de *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* (juillet 2008).
3. R. Racine, G. A. Walker, D. Nadeau and C. Marois, *Speckle Noise and the Detection of Faint Companions'*, Pub. Astron. Soc. Pacific, **112**, p. 587 (1999).
4. L. M. Mugnier, A. Blanc and J. Idier, *Phase Diversity: a Technique for Wave-Front Sensing and for Diffraction-Limited Imaging*, Dans *Advances in Imaging and Electron Physics*, sous la direction de P. Hawkes, vol. **141**, chap. 1, pp. 1–76. Elsevier (2006).
5. J.-F. Sauvage, T. Fusco, G. Rousset and C. Petit, *Calibration and Pre-Compensation of Non-Common Path Aberrations for eXtreme Adaptive Optics*, J. Opt. Soc. Am. A, **24** (8), pp. 2334–2346 (août 2007).
6. J.-F. Sauvage, L. M. Mugnier, G. Rousset and T. Fusco, *Analytical expression of long-exposure adaptive-optics-corrected coronagraphic image. First application to exoplanet detection*, J. Opt. Soc. Am. A, **27** (11), pp. A157–A170 (novembre 2010).
7. M. N'diaye, K. Dohlen, S. Cuevas, P. Lanzoni, F. Chemla, C. Chaumont, R. Soummer and E. T. Griffiths, *Experimental results with a second-generation Roddier Roddier phase mask coronagraph*, , **509**, p. A8 (janvier 2010).