ZELDA, a Zernike sensor for accurate calibration of aberrations in coronagraphic instruments: validation in VLT/SPHERE and preliminary on-sky results


Aix Marseille Univ, CNRS, LAM, Laboratoire d’Astrophysique de Marseille, Marseille, France
Université Côte d’Azur, Observatoire de la Côte d’Azur, CNRS, Laboratoire Lagrange, UMR 7293, CS 34229, 06304 Nice Cedex 4
ONERA, The French Aerospace Lab BP72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France
European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile
Université Grenoble Alpes, IPAG, F-38000 Grenoble, France
CNRS, IPAG, F-38000 Grenoble, France
European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany

ABSTRACT
Gas giant planets, brown dwarfs, and debris disks around nearby stars are now routinely observed by dedicated high-contrast imaging instruments on large, ground-based observatories. These facilities combine extreme adaptive optics and coronagraphy to achieve unprecedented sensitivities for exoplanet detection and spectral characterization. However, residual quasi-static phase aberrations in these coronagraphic systems represent a critical limitation for the detection of giant planets with a contrast lower than a few $10^{-6}$ at very small separations ($<0.3''$) from their host star. In 2013 we proposed ZELDA, a Zernike wavefront sensor to measure these phase aberrations at a nanometric level. Relying on a prototype installed in VLT/SPHERE in 2014, we performed multiple tests in December 2015 and demonstrated the capability of our concept to accurately measure well-known aberrations on internal source. We also showed a contrast improvement by a factor of 10 at 0.2'' for this instrument by calibrating its residual aberrations with our device. After a brief review of this experimental validation, we present preliminary on-sky results on SPHERE obtained in March 2017. This new data set relies on ZELDA estimate of quasi-static aberrations on the internal source. We finally discuss perspectives for future ELT instrumentation with ZELDA.

Keywords: High-contrast imaging, non-common path aberrations, wavefront sensing

1. INTRODUCTION
Circumstellar disks and planetary companions around nearby stars are routinely observed on the ground by several facilities with exoplanet direct-imaging capabilities. Of these facilities, the instruments VLT/SPHERE and Gemini Planet Imager (GPI) have recently seen first light in 2013-2014, providing unprecedented sensitivity and inner working angle for exoplanet observations. Since their commissioning, they have shed light on known or newly detected planetary companions with insights on their physical characteristics (orbit and mass) and atmospheric chemical features through spectral characterization and photometric and astrometric information. Similar to the recent discovery of 51 Eri b or HIP 65426 b, large surveys of nearby stars with these instruments are expected to unveil more gas giant planets, providing clues for comparative exoplanetology and enabling a better understanding on the formation and evolution of planetary systems.

Send correspondence to A. Vigan: arthur.vigan@lam.fr
To achieve direct imaging and spectroscopy of companions, these ground-based instruments rely on a combination of extreme adaptive optics (ExAO) system, coronagraphy, dedicated observational strategies and post-processing methods. Differential aberrations between the ExAO sensing path and the science path, so-called non-common path aberrations (NCPA), have been identified as setting high-contrast performance limits for adaptive optics instruments. Their importance was well known at the start of the development of the recently commissioned planet imagers, GPI and SPHERE, and various strategies were implemented to minimize them. In particular, SPHERE has a differential tip-tilt correction system, using a camera just before the coronagraph as an image position sensor, to ensure a compensation for any differential image movement.

Responsible for the residual quasi-static speckles that limit contrast performance, higher-order aberrations were minimized first of all by tightly specifying the surface errors of all optics in the differential path and secondly by implementing AO calibration strategies to minimize residual aberrations. For SPHERE the differential optics worked better than specified, and consequently, the adopted calibration strategy, which is based on phase diversity techniques, was not found to improve the final image quality and was finally discarded. Still, the remaining NCPA are on the order of a few tens of nanometers, preventing coronagraphs from achieving their ultimate performance. These wavefront errors can be split into two contributions: the long-timescale aberrations that are due to the optical surface errors or misalignments in the instrument optical train and the slowly varying instrumental aberrations that are caused by thermal or opto-mechanical deformations as well as moving optics such as atmospheric dispersion correctors. They lead to static and quasi-static speckles in the coronagraphic images, which represent critical limitations for the detection and observation of older or lighter gaseous planets at smaller separations. More precise measurement strategies are required to measure and correct for these small errors with accuracy and achieve deeper contrast (down to $10^{-7}$, representing the ultimate contrast limit of these instruments) for the observation of the faintest companions.

Over the past few years, several methods have been proposed to circumvent the NCPA problem. We have proposed the use of a Zernike phase mask sensor to calibrate the NCPA seen by the coronagraph in exoplanet direct imagers. This phase-contrast method uses a phase-shift mask to modulate the phase differential aberrations into intensity variations in the pupil plane. Since differential aberrations in exoplanet imagers are small, a linear or quadratic relation between the wavefront errors and the pupil intensity enables reconstructing the differential aberrations at nanometric accuracy with a simple, fast algorithm, making calibration in real time possible. Laboratory demonstration of the concept have been carried out in this context. We have also performed preliminary tests and obtained encouraging results of the Zernike sensor on the coronagraphic testbed in Marseille. In 2014, we implemented a prototype called ZELDA (Zernike sensor for Extremely Low-level Differential Aberrations) in the SPHERE instrument at the VLT.

The first validation of ZELDA was presented in N’Diaye et al. (2016). In this proceeding, we briefly present the principle and formalism of the ZELDA sensor, then we summarize the results of the concept validation in SPHERE, before presenting new results that were recently obtained on-sky. Finally, we discuss perspectives for ELT instrumentation with ZELDA.

### 2. ZELDA SENSOR

#### 2.1 Principle and formalism

The ZELDA sensor is based on phase-contrast techniques that were proposed by Zernike (1934) to measure NCPA in high-contrast imaging instruments with nanometric accuracy. We briefly recall the principle and formalism of this Zernike sensor that were detailed in a previous paper.

This sensor uses a focal plane phase mask to produce interference between a reference wave created by the mask and the phase errors present in the system (Fig. 1). As a result, this sensor converts the aberrations in the entrance pupil into intensity variations in the exit pupil. This phase-to-intensity conversion depends on the mask characteristics, that is, the diameter $d$ and the depth that is related to the sensor phase delay $\theta$. In the following, $\lambda$ and $D$ denote the wavelength of observation and the telescope aperture diameter. We recall the expression of the ZELDA signal $I_C$ as a function of the phase error $\varphi$ for a given pixel in the entrance pupil:

$$I_C = P^2 + 2b^2(1 - \cos \theta) + 2Pb [\sin \varphi \sin \theta - \cos \varphi (1 - \cos \theta)] ,$$
where $P$ and $b$ denote the amplitude pupil function and the amplitude diffracted by the focal plane phase mask of diameter $d$. For small phase errors, we can consider a quadratic case where we only maintain first- and second-order terms of $\varphi$ in the Taylor expansion term. Then we have

$$I_C = P^2 + 2b^2(1 - \cos \theta) + 2Pb\left[\varphi \sin \theta - \left(1 - \varphi^2/2\right)\left(1 - \cos \theta\right)\right] .$$  \hspace{1cm} (2)

The phase can be recovered from the intensity by solving this second-order equation. For a phase mask with depth $\theta = \pi/2$ and angular diameter $1.06 \lambda_0/D$, where $\lambda_0$ denotes the wavelength of design, $b$ ranges between 0.4 and 0.6 (see Fig. 3 of N'Diaye et al. 2013). Assuming a normalized amplitude in the entrance pupil $P = 1$, we obtain the solution

$$\varphi = -1 + \sqrt{3} - 2b - \frac{1 - I_C}{b} .$$  \hspace{1cm} (3)

The formalism of the Zernike phase-mask sensor is valid for any aperture geometry, including primary mirror segmentation, central obstruction, or spider struts, since the intensity measurement and the phase reconstruction is performed inside the geometric pupil at any given point. This property makes our sensor particularly interesting for measurements in the context of a deformable mirror (DM) with dead actuators as in SPHERE, since their respective points do not alter the measurements of the other points inside the geometric pupil.

### 2.2 ZELDA design for VLT/SPHERE

Our design for VLT/SPHERE corresponds to the case where $d = 1.076 \lambda_0/D$ and $\theta = 0.440\pi \lambda_0/\lambda$ with $\lambda_0 = 1.642 \mu$m. With such a mask diameter, $b$ has a chromatically dependant profile similar to that of an Airy pattern twice the size of the pupil (Fig. 2, left). As an illustration, we assume a normalized entrance pupil plane amplitude and measurements are performed at $\lambda = \lambda_0$ and then, the phase error at any given point in the pupil is then reduced to

$$\varphi = -1.208 + 1.230\sqrt{2.590 - 1.626b - 0.813(1 - I_C)/b} .$$  \hspace{1cm} (4)

Figure 2 (right) shows the ZELDA signal at $\lambda = \lambda_0$ as a function of the wavefront error for a given pixel with $b = 0.5$. The intensity received by a pixel depends on the wavefront error (WFE) at the location of that

![Figure 1. Principle of the ZELDA analysis with wavefront errors in the entrance pupil plane A to estimate, a phase mask centered on the stellar signal in the following focal plane B, and the intensity measurement in the re-imaged pupil plane C. A linear or quadratic reconstruction of the aberrations is performed from the recorded intensity with nanometric accuracy.](image-url)
Figure 2. *Left:* Radial profile of the amplitude $b$ diffracted by a mask of size $1.076 \lambda_0/D$ and phase shift $\theta = 0.444\pi \lambda_0/\lambda$ at the central wavelength of the filters that are used during our tests with SPHERE. The dashed line defines the entrance pupil function $P_0$. *Right:* ZELDA pupil plane intensity as a function of phase aberration for a given pixel in the pupil, assuming a mask-diffracted wave amplitude $b = 0.5$. The dynamic range of our sensor is represented with dashed vertical lines. The linear and second-order phase reconstruction are displayed with dot and dashed lines.

pixel on a sinusoidal function. However, the sinusoid is not symmetric about zero aberration, giving rise to an asymmetric dynamic range defined by the monotonic range around zero. The limits of the dynamic range are given by the changes of gradient sign of $I_C$, that is, $dI_C/d\phi = 0$. In our mask design, the dynamic of the sensor ranges between -0.14 $\lambda_0$ and 0.36 $\lambda_0$, as illustrated with the vertical lines in Fig. 2 (right).

The phase mask consists of a circular shape machined into the front face of a fused silica substrate by the aid of photolithographic reactive ion etching. This subtractive process, which has been experimented with and optimized in the context of the Roddier & Roddier coronagraph $^{31,32}$ and Zernike wavefront sensors $^{27,33}$ has been found superior to the more classical additive process where SiO$_2$ is deposited onto a fused silica substrate $^{34}$. While the ion-etching process offers extremely steep edges and precisely defined phase steps, it is also monolithic, avoiding any interfaces between materials that might give rise to spurious interference effects. The phase mask prototype was manufactured by the SILIOS company in two steps. First, a circular hole was generated and transferred into photoresist by UV photolithographic projection, leaving the surface to be machined naked. Then, reactive ion etching was applied by exposing the surface to SF$_6$ gas. The shape of the manufactured mask, measured with a microscope, was within 1% of the specifications $^{29}$.

### 3. ZELDA PERFORMANCE FOR NCPA CORRECTION

#### 3.1 Absolute performance of the sensor

In 2015 we performed multiple validation tests of the ZELDA wavefront sensor in SPHERE. This Zernike sensor was installed in the infrared coronagraphic wheel of the instrument during its reintegration in Paranal in 2014. For our test purposes, the coronagraphic mask is replaced with ZELDA and the system is set up in pupil-imaging mode to perform our phase aberration measurements. During these tests, we introduced Zernike and Fourier modes on the high-order DM of SPHEREx, and we compared the ZELDA measurements with the theoretical modes to estimate the performance (see Fig.3). The modes were chosen to cover low-order aberrations and spatial frequencies, which are the main components of the NCPA that we wish to correct to improve the performance of an ExAO coronagraphic system. They were introduced at various amplitudes to explore the dynamic range of the sensor. These tests demonstrated the ability to measure aberrations at sub-nanometric accuracy in the linear range of the sensor. All the details of the tests, data acquisition and data processing are presented in N’Diaye et al. (2016).$^{29}$
3.2 Results with ZELDA-based wavefront correction

The ZELDA was used to measure and compensate for the NCPA of SPHERE at the level of the coronagraphic mask to improve the quality of the focal-plane images. The ZELDA measurement was acquired and analyzed to produce an OPD map, which is presented in the top left of Fig. 4. In addition to the static pattern of DM actuators, the OPD map clearly shows low spatial frequency aberrations at the level of a few dozen nanometers (RMS), corresponding to uncorrected NCPA. This map cannot be projected directly onto the reference slopes of the WFS. The sampling of the pupil on the IRDIS detector is such that ZELDA is sensitive to spatial frequencies higher than 190 cycles/pupil, while the SPHERE high-order DM has only $40 \times 40$ actuators, which in theory allows corrections of up to $20 \lambda/D$. To avoid any spatial aliasing, the ZELDA OPD map was first filtered in Fourier space using a Hann window of size $25 \lambda/D$ (Fig. 4, middle row).

The filtered OPD map was then projected onto the reference slopes of the WFS, and a new ZELDA image was acquired and analyzed, providing the OPD map shown in Fig. 4, top right panel. In this new OPD map, the low spatial frequencies that were clearly visible before the correction have now completely disappeared, and the map appears much flatter. The OPD maps show that the NCPA of the system apparently has indeed been corrected. A slight residual tip-tilt in the OPD map after correction remains. This is not the result of an imperfect measure, but simply arises because the tip-tilt is controlled in close-loop mode by the DTTS, for which the reference slopes are changed manually at the beginning of the test to center the PSF on the ZELDA mask. A new implementation of the procedure now enables to remove this residual tip-tilt.

To verify the quality of the NCPA compensation, we acquired coronagraphic images with IRDIS at 1593 nm (H2 filter, see, e.g., Vigan et al. 2010), using the Apodized Pupil Lyot Coronagraph (APLC; Soummer 2005) optimized for the $H$ band. Two data sets were acquired: one using the default reference slopes of the WFS (before correction), and one using the reference slopes updated using the ZELDA measurement (after correction). For each data set, we acquired a 2 min coronagraphic image and 2 min reference PSF image where the PSF was moved out of the coronagraphic mask. The resulting coronagraphic images are showed in the bottom row of Fig. 4. The visual difference between the two images is striking inside the AO-corrected area. The image before correction is dominated by speckles close to the axis, up to $7-8 \lambda/D$. It also shows a strong horizontal and vertical pattern of speckles that extend from the edge of the corrected region down to $\sim 10 \lambda/D$.

To quantitatively assess the performance gain after NCPA compensation, we plot in Fig. 5 the azimuthal standard deviation of the coronagraphic image intensity as a function of separation, normalized to the peak flux.
Figure 4. Illustration of the correction of the SPHERE NCPA using ZELDA measurements with IRDIS. The top row shows the OPD maps measured with ZELDA before (left) and after (right) the correction, presented at the same color scale. The middle row shows the same maps filtered in Fourier space using a Hann window of size 25 $\lambda/D$. Low spatial frequency aberrations are clearly visible on the left, while they have disappeared after the correction, except for a small amount of residual tip-tilt (see text). The bottom row shows the equivalent focal plane coronagraphic images before (left) and after (right) the compensation of the NCPA, measured at 1593 nm (IRDIS H2 filter) and presented at the same color scale. The gain is obvious close to the center, but also noticeable farther out where the intensity of the speckles in the corrected area has decreased significantly.
Figure 5. Normalized azimuthal standard deviation profiles before (plain line) and after (dashed line) correction of the NCPA using ZELDA, as a function of angular separation. The contrast gain is plotted in the bottom panel. The dotted line corresponds to the edge of the coronagraphic mask (90 mas). The measurements correspond to the coronagraphic images presented in Fig. 4. They are compared to the theoretical performance of the SPHERE APLC (red, dash-dotted line).

of the reference off-axis PSF. The bottom panel of the figure shows the gain in contrast between the two curves. Within 2-16 $\lambda/D$, there is a gain in contrast of a factor more than 2, with even a peak at more than 10 around 5 $\lambda/D$. This agrees very well with estimates from N’Diaye et al. (2013), where we estimated a possible gain over a factor of 10, and it is a strong confirmation of the potential of ZELDA to compensate for NCPA. We also plot the simulated theoretical performance of the SPHERE APLC in presence of the amplitude aberrations directly measured in the instrument and using an image of the Lyot stop, but without any phase aberrations. Within 5 $\lambda/D$, we reach this theoretical performance, which means that the NCPA at low spatial frequencies are almost entirely corrected for. We note that on the internal source the instrument pupil is purely circular, with no central obscuration or spiders, but the Lyot stop still includes elements to mask the diffraction of the central obscuration, the spiders, and the bad actuators of the DM. This will result in a slightly better performance of the coronagraph on sky than on the internal source.

4. PRELIMINARY RESULTS ON-SKY

The first on-sky results were obtained in March 2017 at Paranal in very average observing conditions. The calibration procedure was to first determine and compensate for the NCPA using ZELDA on the internal source, and then apply the exact same correction on-sky. On the internal source (Fig. 6, left), the compensation for the NCPA results in a visible improvement of both the images and the raw contrast. However, we note that the gain is slightly smaller than what was obtained in 2015, but also the starting point is at a much higher contrast level than in 2015. For instance, the raw contrast without NCPA compensation at $0.2'' \times 10^{-5}$ in 2015, while it was closer to $7-8 \times 10^{-5}$ in 2017. The origin of this difference is not entirely clear and still needs to be investigated.

On-sky, there is no difference between the non-compensated and compensated case, i.e. compensating for the NCPA did not result in any visible improvement of the images or raw contrast curves. There are several possible explanations. The first one is that the level of turbulent residuals completely covers the static aberrations of the system. Although it is certainly the case in some areas of the image, the quasi-static speckles are still visible in the
images within the AO-correction region. Another possibility is that there is a difference in the instrumental setup between the internal source and the on-sky configurations. On sky, the derotator and atmospheric dispersion correctors (ADCs) are in variable positions that depend on the pointing of the telescope, which certainly induces some differences. In addition, the spectral beam-shift, which result in different parts of the optics being seen by different wavelengths, also has a significant influence that would need to be quantified. Finally, the presence of amplitude aberrations due to the telescope optics also certainly has an effect. These different hypotheses and their respective contribution will be quantified with further tests in 2018.

An important problem to keep in mind is that the calibration of the NCPA is currently done on the internal source, which certainly does not take into account all the effects that are seen when observing a real star. In this context, it will be extremely important to obtain ZELDA measurements on sky, and compensate for the NCPA directly from these measurements. Many questions still remain in this case:

- what will be the absolute accuracy of the sensor?
- how long will we need to integrate to average the turbulent residuals?
- will we even be able to measure the NCPA in such a configuration?

5. CONCLUSION

We have demonstrated that the prototype of Zernike wavefront sensor proposed in N’Diaye et al. (2013), ZELDA, is fully functional in VLT/SPHERE to measure and compensate for the non-common path aberrations. We are able to reach sub-nanometric accuracy, which results in a gain of a factor $\sim 10$ in contrast at very small angular separations. For on-sky compensation, the first results do not show any contrast improvement, however there are still many open questions. The fact that the instrumental configuration changes significantly between the internal source and on-sky observations is certainly an important limitation, which calls for direct ZELDA measurements of the NCPA on sky.

In the context of the Extremely Large Telescopes, designing wavefront sensors that are able to measure small wavefront errors is a critical issue to address for high-contrast imaging observations. The first-light instruments MICADO and HARMONI both include in their science case the characterization of the giant exoplanets that are detected by SPHERE and GPI. In the case of HARMONI, a dedicated high-contrast mode is envisioned.
(Carlotti et al. in prep.), and it includes an on-line ZELDA wavefront sensor for measuring and compensating for the NCPA in real time. Having such a sensor will be essential to reach the technical specification on the contrast of $10^{-6}$ at 200 mas.

Finally, for future extremely high-contrast instrumentation on ELTs, ZELDA can be considered to extremely fine cophasing of the telescope segments. We have already conducted initial simulations that demonstrate the performance of ZELDA for this task, and laboratory validations are currently investigated in Marseille and Nice.

**ACKNOWLEDGMENTS**

This work is partially supported by the National Aeronautics and Space Administration under Grants NNX12AG05G and NNX14AD33G issued through the Astrophysics Research and Analysis (APRA) program (PI: R. Soummer). MN would like to thank Rmi Soummer and Laurent Pueyo for their support. MN would also like to acknowledge the ESO Chile Visiting Scientist program. Additionally, MN warmly acknowledges Frantz Martinache for his support through the KERNEL project (ERC grant #683029). Finally, AV would like to thank the Paranal staff, and particularly G. Zins, for their patience and support when performing the tests presented in this work.

SPHERE is an instrument designed and built by a consortium consisting of IPAG (Grenoble, France), MPIA (Heidelberg, Germany), LAM (Marseille, France), LESIA (Paris, France), Laboratoire Lagrange (Nice, France), INAF - Osservatorio di Padova (Italy), Observatoire de Genève (Switzerland), ESO Chile Visiting Scientist program. SPHERE was also received funding from the European Commission Sixth and Seventh Framework Programmes as part of the Optical Infrared Coordination Network for Astronomy (OPTICON) under grant number RI3-Ct-2004-001566 for FP6 (2004-2008), grant number 226604 for FP7 (2009-2012) and grant number 312430 for FP7 (2013-2016).

**REFERENCES**


