Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a Zernike phase-mask sensor
II. Concept validation with ZELDA on VLT/SPHERE


1 Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
e-mail: mamadou@stsci.edu
2 Aix Marseille Université, CNRS, LAM (Laboratoire d’Astrophysique de Marseille) UMR 7326, 13388 Marseille, France
3 European Southern Observatory, Alonso de Cordova 3107, Vitacura, Santiago, Chile
4 ONERA, The French Aerospace Lab, BP72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France
5 CNRS, IPAG (Institut de Planétologie et d’Astrophysique de Grenoble), UMR 5274, BP 53, 38041 Grenoble Cedex 9, France
6 Université Grenoble Alpes, IPAG, 38000 Grenoble, France

Received 1 April 2016 / Accepted 3 June 2016

ABSTRACT

Warm or massive gas giant planets, brown dwarfs, and debris disks around nearby stars are now routinely observed by dedicated high-contrast imaging instruments that are mounted on large, ground-based observatories. These facilities include extreme adaptive optics (ExAO) and state-of-the-art coronagraphy to achieve unprecedented sensitivities for exoplanet detection and their spectral characterization. However, low spatial frequency differential aberrations between the ExAO sensing path and the science path represent critical limitations 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 our previous work, we proposed a wavefront sensor based on Zernike phase-contrast methods to circumvent this problem and measure these quasi-static aberrations at a nanometric level. We present the design, manufacturing, and testing of ZELDA, a prototype that was installed on VLT/SPHERE during its reintegration in Chile. Using the internal light source of the instrument, we first performed measurements in the presence of Zernike or Fourier modes introduced with the deformable mirror. Our experimental results are consistent with the results in simulations, confirming the ability of our sensor to measure small aberrations ($<50$ nm rms) with nanometric accuracy. Following these results, we corrected the long-lived non-common path aberrations in SPHERE based on ZELDA measurements and estimated a contrast gain of 10 in the coronagraphic image at 0.2″, reaching the raw contrast limit set by the coronagraph in the instrument. In addition to this encouraging result, the simplicity of the design and its phase reconstruction algorithm makes ZELDA an excellent candidate for the online measurements of quasi-static aberrations during the observations. The implementation of a ZELDA-based sensing path on the current and future facilities (ELTs, future space missions) could facilitate the observation of cold gaseous or massive rocky planets around nearby stars.

Key words: instrumentation: high angular resolution – instrumentation: adaptive optics – techniques: high angular resolution – telescopes – methods: data analysis

1. Introduction

Circumstellar disks and planetary companions around nearby stars are routinely observed on the ground by several facilities with exoplanet direct-imaging capabilities (e.g. Beuzit et al. 2008; Macintosh et al. 2008; Guyon et al. 2010; Hinkley et al. 2011; Skemer et al. 2012; Close et al. 2014). 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 (Macintosh et al. 2014; Vigan et al. 2015). 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 (Galicher et al. 2014; Chilcote et al. 2015; Vigan et al. 2016; Maire et al. 2016; Zurlo et al. 2016; Bonnefoy et al. 2016). Similar to the recent discovery of 51 Eri b (Macintosh et al. 2015), 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 of the formation and evolution of planetary systems.

To achieve direct imaging and spectroscopy of companions orbiting nearby stars, these ground-based instruments rely on a combination of extreme adaptive optics (ExAO) system for the fine control of the wavefront errors that are due to atmospheric turbulence and optic imperfections, coronagraphy for starlight suppression, and dedicated observational strategies and post-processing methods to retrieve the signal of the substellar mass companions. With their near-infrared capabilities, these instruments can observe faint planetary-mass companions in thermal emission and study young or massive gaseous planets with contrast ratios down to $10^{-3}−10^{-5}$ at 0.2–0.3″, corresponding to solar system scales for stars within 100 pc.
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 (AO) instruments. Their importance was well known (e.g. Fusco et al. 2006) 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 close to the coronagraph as an image position sensor, to ensure that any differential image movement is compensated for. Thermally induced differential defocus was minimized by design, and on-coronagraph focus is optimized at the start of observations. Higher-order aberrations, responsible for the residual quasi-static speckles that limit the high-contrast performance of the instrument, were first of all minimized by tightly specifying all optics in the differential path, and second 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 (Sauvage et al. 2007), 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 (e.g. Macintosh et al. 2005; Martinez et al. 2012, 2013). 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. Other unforeseen limiting effects have also been experienced with these new instruments of unprecedented performance. In particular, the low-wind effect (LWE) where piston patterns appear across the spiders in monolithic-pupil telescopes, has turned out to be devastating as it systematically occurs when the outdoor wind speed (i.e., 30 m below the VLT platform) is below 1 to 3 m/s (Sauvage et al. 2016). Around 20% of observing time is affected by this effect, causing strong speckle in the first Airy ring and even splitting of the main peak into up to four peaks of varying intensity. The cause of this effect has been identified as being energetic transfer between spiders that are radiatively cooled by the night sky, and the slowly passing dome air. While strategies such as repainting spiders using a low-emissivity paint are being studied to control the root cause of this particular case of dome seeing, real-time phase measurements may be needed to correct for its effect.

Over the past few years, several methods have been proposed to circumvent the NCPA problem (e.g. Gonsalves 1982; Wallace et al. 2010a; Paul et al. 2013). We have proposed the use of a Zernike phase mask sensor to calibrate the NCPA seen by the coronagraph in exoplanet direct imagers (N’Diaye et al. 2012, 2014; Dohlen et al. 2013). However, and to the best of our knowledge, no experimental validation has been performed on a real instrument for the measurement of coronagraphic aberrations.

In this paper, we propose to validate the Zernike phase-mask sensor for the measurement of differential aberrations in a real exoplanet direct-imaging instrument, SPHERE at the VLT. We designed a sensor called ZELDA (which stands for Zernike sensor for Extremely Low-level Differential Aberrations) and installed a manufactured prototype in SPHERE during its reintegration at the ESO Paranal observatory in 2014. We present the first experimental results, using the internal light source of the instrument. After recalling the principle and formalism of the Zernike sensor, we present the design, properties, and manufacturing of ZELDA. We then evaluate the performance of our wavefront sensor in the presence of different aberration modes and analyze its sensitivity to spectral bandpass. We finally present the first results of ZELDA-based wavefront correction on VLT/SPHERE for the long-lived speckles and derive the contrast gain achieved with this calibration.

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 Paper 1.

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:

$$I_C = P^2 + 2b^2 \lambda^2 (1 - \cos \theta)^2 + 2pb \sin \varphi \sin \theta - \cos \varphi (1 - \cos \theta),$$

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

\[ I_C = P^2 + 2b^2(1 - \cos \theta) + 2Pb \left[ \varphi \sin \theta - (1 - \varphi^2/2)(1 - \cos \theta) \right]. \]  
(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 Paper I). Assuming a normalized amplitude in the entrance pupil \( P = 1 \), we obtain the solution

\[ \varphi = -1 + \sqrt{3 - 2b - (1 - I_C)/b}. \]  
(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 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. Design and properties

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). As an illustration, we assume a normalized entrance pupil plane amplitude and measurements are performed at \( J = \lambda_0 \) and then, the phase error at any given point in the pupil is then reduced to

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

Figure 3 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) location of that 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, \( dI_C/d\varphi = 0 \). In our mask design, the dynamic of the sensor ranges between \(-0.140 \lambda_0 \) and \( 0.360 \lambda_0 \), as illustrated with the vertical lines in Fig. 3.

2.3. Prototype

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 (N’Diaye et al. 2010, 2011) and
Zernike wavefront sensors (Dohlen et al. 2006; N’Diaye et al. 2014), has been found superior to the more classical additive process where SiO$_2$ is deposited onto a fused silica substrate (Guyon et al. 1999). 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 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. A picture of our prototype is shown in Fig. 4.

The excellent shape of masks made by this procedure is evident from the optical profilometry made using a Wyko interference microscope and presented in Fig. 5. Table 1 gives the measured depth and diameter for this prototype, showing good agreement with our specifications: a relative error better than 1% is achieved for both dimensions. The root mean square (rms) roughness within the machined area is 0.9 nm, which proves identical to that of the substrate outside of the mask. While a slight rounding of the edges can be seen, the transition zone is less than 1 µm wide, corresponding to about 1% of the mask diameter, which is within the specifications and the accuracy range of the manufacturing process. The slight offset of the mask shape does not impact the ZELDA performance as the wavefront reconstruction algorithm is adjusted with the characteristics of the manufactured phase mask to achieve optimal phase error measurements.

![Fig. 4. Picture of the ZELDA prototype in its assembly before installation in the SPHERE coronagraph wheel in April 2014.](image)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Specification</th>
<th>Laboratory measurement</th>
<th>Relative difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter $d$</td>
<td>69.9 µm</td>
<td>70.7 µm</td>
<td>1.1%</td>
</tr>
<tr>
<td>Depth $z$</td>
<td>814.1 nm</td>
<td>814.6 nm</td>
<td>0.1%</td>
</tr>
</tbody>
</table>

3. ZELDA performance

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 SPHERE, and we compared the ZELDA measurements with the theoretical modes to estimate the performance. 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. In this section we describe the data acquisition and analysis (Sect. 3.1) before reviewing the results obtained with different modes (Sect. 3.2) and the sensitivity to spectral bandpass (Sect. 3.2.2).

3.1. Data acquisition and processing

3.1.1. Calibration of the aberrations: sensitivity factor

For optimal performance, the modes were introduced in the system in closed-loop mode using an offset to the reference slopes of the Shack-Hartmann (SH) wavefront sensor (WFS) of SPHERE. This approach allowed us to benefit from an extremely high-quality point-spread function (PSF, Strehl ratio > 90% at $\lambda_0$) at the level of the ZELDA phase mask, since the system works fully in closed-loop mode. We used the inverse of the influence matrix of the system to convert the phase that we wished to introduce at the surface of the DM into a voltage offset at the level of the individual actuators. This voltage vector was then converted into reference slope offsets using the interaction matrix of the system, which is calibrated daily as part of the calibration plan of the instrument.

To be accurate, this procedure requires a calibration step to precisely estimate the amount of aberrations that are introduced.
for a given offset on the reference slopes, which from now on we refer to as the sensitivity factor. In other words, we wish to determine the real amount of aberrations that are introduced when we request 1 nm RMS of aberrations. This amplitude calibration was performed by introducing a known aberration of a given amplitude onto the DM and measuring the loss of Strehl ratio induced at the level of the PSF in the focal plane. For small amounts of aberrations, Maréchal’s approximation tells that the Strehl ratio can be expressed as $S = e^{-\sigma^2}$, with $\sigma^2$ the variance of the wavefront error. In practice, we introduced a ramp of focus instead of a single value to make the measurement more robust to noise. Then an inverse exponential function with three free parameters was fit to the Strehl values measured for each focus value. The free parameters correspond to the position of the maximum, the sensitivity factor that we searched for, and a scaling factor that is necessary to take the static aberrations of the system into account.

When performing this calibration, the AO system is operated in particular conditions. First, the light source is an internal calibration source, hence allowing an operation with very high flux that would correspond to a star of negative R-band magnitude. All potential problems related to noise propagation are therefore completely avoided. Second, the aberrations to be compensated for are considered static, or with characteristic evolution time so small with respect to the AO-loop frame-rate (1.2 kHz) that any temporal filtering by the loop can be considered as negligible. To project the aberration phase map onto the reference slopes, the calibrated interaction matrix of the AO loop is used. This matrix represents the sensitivity of the SH wavefront sensor to all the modes controlled by the system, it therefore accounts for the system response as accurately as possible. As a consequence of this procedure, the sensitivity factor is valid for any other mode, although it has been calibrated only on the focus.

### 3.1.2. Data acquisition in SPHERE

The data were acquired in November and December 2015 during daytime technical time. All measurements were made internally using the light sources available in the calibration unit of the instrument (Wildi et al. 2009). Images are taken with the infrared dual-band imager and spectrograph (IRDIS; Dohlen et al. 2008), one of the two scientific subsystems of SPHERE dedicated to the detection and characterization of giant planets in the near-infrared (Vigan et al. 2008, 2010). An exposure time of 10 s is used, providing a signal-to-noise ratio high enough for the analysis. The acquired data are summarized in Table 2.

At the beginning of each data acquisition session, the calibration of the sensitivity factor (see Sect. 3.1.1) was performed to ensure an accurate knowledge of the amount of aberrations introduced in the ZELDA tests. For each test, the PSF was manually centered on the ZELDA phase mask by looking at the intensity image on the IRDIS detector. Since ZELDA transforms phase errors into intensity variations, performing an accurate centering visually is relatively easy, especially since the overall amount of aberrations in SPHERE is low (Vigan et al. 2016). The PSF is centered by changing the reference slopes of the near-IR differential tip-tilt sensor (DTTS) of the instrument, the purpose of which is to maintain the PSF centered and stable on the coronagraph. As a result, the tip-tilt cannot be controlled directly with the reference slopes of the SH WFS because the DTTS will ensure that the PSF is always centered on the same location. For the tip-tilt tests with ZELDA, the DTTS loop was opened to disable this control of the centering and allow the PSF to move freely.

For each test, three types of data were acquired: ZELDA data with the PSF centered on the phase mask, a clear pupil reference with the PSF outside of the phase mask, and instrumental backgrounds with the same integration time as the ZELDA and clear pupil data. The second type of data gives us the contribution of the amplitude aberrations. All data were acquired in the Fe II near-infrared narrow-band filter centered on $\lambda = 1652$ nm ($\Delta \lambda = 24$ nm), for which the geometry of the ZELDA phase mask was originally designed.

The ramp in amplitude was chosen to sample the range within which $\varphi$ is monotonic, which is from $-250$ to $+250$ nm peak-to-valley (PtV), and it was extended to $+600$ nm PtV to study the nonlinearity observed beyond this range. The range was sampled by 35 measurement points, with a coarse sampling at the extremes and a finer sampling close to zero. All modes were tested with the same ramp. Additional tests were performed with other filters than the Fe II, see Sect. 3.2.2.

The Zernike modes were calculated on an annular geometry (Mahajan 1981) to take into account the fact that the central actuators of the SPHERE DM are not seen by the WFS and are not controlled in closed-loop mode by the system. The Fourier modes were tested up to ten cycles/pupil in both horizontal and vertical directions (x and y). At nine cycles/pupil some instabilities of the system started to be observed (high-order AO loop opening, probably due to aliasing effects) for the largest amplitudes (more than 200 nm), therefore the ramp was limited to a smaller extent for these modes.

#### 3.1.3. Processing

After acquisition, the data of each sequence were processed uniformly. The images were first background subtracted, and the bad pixels were corrected for using a sigma-clipping procedure. Then we normalized the ZELDA pupil image by the clear pupil image itself for each pixel, and finally, for each aberration and amplitude, we calculated the phase from the normalized ZELDA images following Eq. (4).

Figure 6 shows all the annular Zernike modes measured with ZELDA in the system when 400 nm PtV of aberrations are introduced. The central actuators, not seen by the WFS and controlled by the system, have been numerically masked out. Overall, the shape of the Zernike modes is well reproduced and clearly recognizable. The regular pattern of the DM actuators is clearly visible, as is the effect of individual dead or stuck and known actuators.

#### 3.2. Measurements and comparison with theory

We analyzed the performance of ZELDA over several spatial frequency regimes by using different Zernike and Fourier modes. For any given mode and amplitude of aberration, we compared the wavefront error measured by ZELDA with the theory based on our numerical model. The consistency between the two was used to validate our concept.

In our tests on SPHERE, the introduced aberration was measured by ZELDA for different error sources: the quasi-static aberrations in our system, dead actuators in the DM, the photon noise and detector readout noise. Since these sources of errors alter the estimate of the introduced aberration, we adopted a differential data analysis strategy to minimize their effect on the ZELDA measurement. For each mode, we first measured the quasi-static errors in the absence of added aberrations on the DM and subtracted this reference from all the other measurements.
amplitude given by ZELDA. We finally compared the theoretical sampling as the experimental map to determine the aberration dead actuators. The theoretical measurement map has the same data, we excluded the points that are measured at the location of amplitude on DM in the absence of noise. As for experimental modeled the theoretical ZELDA measurement for the introduced sensor.

Errors, the photon noise and detector noise) and derived the aberration to mitigate the remaining error sources (residual errors, the photon noise and detector noise). These last where an aberration is introduced, leaving mostly the introduced aberration, the photon noise and the detector noise. These last two contributors are small thanks to the high signal-to-noise (S/N) ratio of the original images. The dead actuators on the DM generate large amplitude errors that go beyond the dynamic range of ZELDA, making the aberrations at the corresponding pupil locations challenging to estimate. We ruled out the range of ZELDA, making the aberrations at the correspond-

Table 2. Summary of acquired data.

<table>
<thead>
<tr>
<th>Mode</th>
<th>Date</th>
<th>Amplitude range (nm PrV)</th>
<th>Steps</th>
<th>Filter</th>
<th>Wavelength (nm)</th>
<th>Bandwidth (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zernike modes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tilt, Z1</td>
<td>2015-11-30</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Tilt, Z2</td>
<td>2015-11-30</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Focus, Z3</td>
<td>2015-11-30</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Astig x, Z4</td>
<td>2015-11-30</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Astig y, Z5</td>
<td>2015-11-30</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Coma y, Z6</td>
<td>2015-12-16</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Coma y, Z7</td>
<td>2015-12-16</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Trefoil x, Z8</td>
<td>2015-12-17</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Trefoil y, Z9</td>
<td>2015-12-17</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Spherical, Z10</td>
<td>2015-12-17</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier modes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fourier, x, 1 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 2 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 3 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 4 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 5 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 6 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 7 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 8 cyc./pup.</td>
<td>2015-12-01</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 9 cyc./pup.</td>
<td>2015-12-01</td>
<td>−200 ... +400</td>
<td>31</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 10 cyc./pup.</td>
<td>2015-12-01</td>
<td>−135 ... +135</td>
<td>21</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 1 cyc./pup.</td>
<td>2015-12-17</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 2 cyc./pup.</td>
<td>2015-12-17</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 3 cyc./pup.</td>
<td>2015-12-17</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 4 cyc./pup.</td>
<td>2015-12-18</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 5 cyc./pup.</td>
<td>2015-12-18</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 6 cyc./pup.</td>
<td>2015-12-18</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 7 cyc./pup.</td>
<td>2015-12-19</td>
<td>−250 ... +600</td>
<td>35</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 8 cyc./pup.</td>
<td>2015-12-19</td>
<td>−200 ... +200</td>
<td>27</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 9 cyc./pup.</td>
<td>2015-12-19</td>
<td>−200 ... +200</td>
<td>27</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, y, 10 cyc./pup.</td>
<td>2015-12-19</td>
<td>−135 ... +135</td>
<td>21</td>
<td>Fe II</td>
<td>1642</td>
<td>24</td>
</tr>
<tr>
<td>Fourier, x, 5 cyc./pup.</td>
<td>2015-12-16</td>
<td>−250 ... +600</td>
<td>35</td>
<td>CntH</td>
<td>1573</td>
<td>23</td>
</tr>
<tr>
<td>Fourier, x, 5 cyc./pup.</td>
<td>2015-12-16</td>
<td>−250 ... +600</td>
<td>35</td>
<td>BH</td>
<td>1625</td>
<td>290</td>
</tr>
<tr>
<td>Fourier, x, 5 cyc./pup.</td>
<td>2015-12-19</td>
<td>−250 ... +600</td>
<td>35</td>
<td>H$_2$</td>
<td>2124</td>
<td>31</td>
</tr>
</tbody>
</table>

where an aberration is introduced, leaving mostly the introduced aberration, the photon noise and the detector noise. These last two contributors are small thanks to the high signal-to-noise (S/N) ratio of the original images. The dead actuators on the DM generate large amplitude errors that go beyond the dynamic range of ZELDA, making the aberrations at the corresponding pupil locations challenging to estimate. We ruled out the range of ZELDA, making the aberrations at the correspond-

3.2.1. Zernike modes

Figure 7 displays the theoretical and experimental measurements of the Zernike sensor for tip-tilt and defocus in the top and bottom panels, showing good agreement between theory and experience. The same general behavior has been observed with other Zernike modes and Fourier modes, as reported in Appendices A and B. The discrepancies between theory and experiment are thought to be related to the accurate calibration of the sensitivity factor. We note that the value of the sensitivity factor varies by up to 10% from one day to the next. No systematic measurements were made to quantify the variations at higher temporal frequency, but it is reasonable to assume that some variations can be expected over the course of 5–6 h (typical length of our data acquisition sessions), leading to the small differences observed with respect to the theory.

The Zernike sensor shows a linear response around the zero point, enabling a simple reconstruction of the small
coronagraphic aberrations and fast convergence with close-loop compensation toward the zero point in one or two iterations. In this regime, the slope of the response curve is equal to one for all the Zernike modes except for the tip-tilt errors. As underlined in Paper I, this behavior is believed to be due to the modification of the light distribution going through the decentered mask with respect to the star image. With a careful calibration, this effect will have no impact on the ability of the Zernike sensor to measure tip-tilt errors.

For the aberrations outside the linear range, the Zernike sensor still proves efficient by considering a close-loop operation between with the measurement and correction. The measurements converge towards the linear regime of the wavefront sensor after a few iterations, as shown by Vigan et al. (2011). Since ZELDA is assumed to operate in the presence of small aberrations, our data analysis is based on a quadratic relation between the measured intensity and the wavefront errors (see Sect. 2). Investigating a more accurate reconstruction expression could increase the measurement accuracy for large phase errors but this is beyond the scope of the paper.

### 3.2.2. Sensitivity to spectral bandpass

The above results have been given for an internal light source passing through a Fe II narrow-band filter of central wavelength $\lambda_c = 1642$ nm and bandwidth $\Delta\lambda = 24$ nm, for which the ZELDA sensor has been originally designed. In Paper I we numerically studied the sensitivity of the Zernike sensor to chromatic effects, showing the reliability of our concept for a light source with well-known central wavelength. We now probe the performance of our sensor experimentally by measuring aberrations in light source filtered with different chromatic parameters on SPHERE.

We analyzed the ZELDA response curve for a Fourier aberration mode with five cycles per pupil in three different filters: Fe II ($\lambda_c = 1642$ nm, $\Delta\lambda = 24$ nm), BH ($\lambda_c = 1625$ nm, $\Delta\lambda = 290$ nm), and H$_2$ ($\lambda_c = 2124$ nm, $\Delta\lambda = 31$ nm). Based on the nature of the considered filters, we estimated the introduced aberration with a data analysis based on the terms $\theta$ and $b$ that were re-evaluated at $\lambda = \lambda_c$, as shown in Fig. 2. The ZELDA measurements for these three filters are given in Fig. 8.

For each filter, the experimental measurements agree well with the numerical values. The match in BH and H$_2$ filters seems slightly better than in Fe II filter, and we see two reasons that might explain this result: first of all, a better calibration of the sensitivity factor when the test with BH and H$_2$ filters were performed, and second, the smaller amount of aberrations with respect to the wavelength at redder filters, enabling a more accurate measurement of the small errors by ZELDA, provided that the sensor has enough dynamic range to operate.

More interestingly, the second and third panels underline the ability of ZELDA to perform accurate aberration measurements using filters with a wide spectral band ($\Delta\lambda = 290$ nm) or centered on a wavelength largely shifted from the wavelength of design ($\lambda_c = 2124$ nm instead of $\lambda_0 = 1642$ nm, i.e., a shift of 482 nm), as long as the data analysis accounts for the filter characteristics in its reduction. This interesting property of our sensor could prove beneficial for the estimation of chromatic aberrations by operating multichromatic aberration measurements with ZELDA at several narrow-band filters, ideally in parallel. Such measurements could possibly lead to contrast gain in post-processing methods based on spectral deconvolution (Sparks & Ford 2002; Vigan et al. 2008) and hence, improvement in the extraction of the spectral information for the faintest planetary companions.

### 4. Results with ZELDA-based wavefront correction

In this section, we now use ZELDA for the purpose it was originally designed for: measuring and compensating for the NCPA of SPHERE at the level of the coronagraphic mask to improve the quality of the focal-plane images. First, the amount of aberrations introduced when changing the reference slopes was calibrated. Then a ZELDA measurement was acquired and analyzed to produce an optical path difference (OPD) map, which is presented in the top left of Fig. 9. 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. 9, middle row). We used a $25 \lambda/D$ window because it provides a better correction than a $20 \lambda/D$. The reason is that because the Hann window falls exactly at zero at the window edge, spatial frequencies that could be corrected with the DM are too strongly attenuated or even completely filtered out in the process. In our limited time with the instrument, we did not have the opportunity to investigate further than testing different window sizes. The use of other filtering windows may provide better results.

The filtered OPD map was then projected onto the reference slopes of the WFS, and a new ZELDA image was acquired and analysed, providing the OPD map showed in Fig. 9, 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

Fig. 7. Top plot: response curves of the ZELDA sensor for tip-tilt (top) and defocus errors (bottom) in simulations (black line) and during the experiment on VLT/SPHERE (colored crosses). Bottom plot: difference between the experiment and the theory.

Fig. 8. Top plot: response curves of the ZELDA sensor for the Fourier mode with 5 cycles per pupil along the x-axis direction in simulations (black line) and during the experiment on VLT/SPHERE (blue crosses), using Fe II (top), BH (middle), and H₂ (bottom) filters. The characteristics of the filter are summarized in Table 2. Bottom plot: difference between the experiment and the theory.
show that the NCPA of the system apparently has indeed been corrected. A slight residual tip-tilt in the OPD map after correction remains, particularly visible in the Hann-filtered version of the OPD map. 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 (see Sect. 3.1.2), for which the reference slopes are changed manually at the beginning of the test to center the PSF on the ZELDA mask. The optimal approach would be to correct the high orders by modifying the reference slopes of the SH WFS and to correct the tip-tilt by modifying the reference slopes of the DTTS. We were unable to investigate this approach for the tests presented here, but it will represent a necessary improvement in future tests.

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. A neutral density filter was used to acquire the reference PSF without saturating the peak. Corresponding instrumental backgrounds were also acquired. The resulting coronagraphic images are showed in the bottom row of Fig. 9. 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 $\sim 10 \lambda/D$. After the NCPA correction, the whole AO-corrected region appears much cleaner: the speckles close to the axis have almost disappeared to reveal a very regular annulus at the edge of the coronagraphic mask, similar to what is expected from the theoretical design of the SPHERE APLC (Guerri et al. 2011). In the $4 - 8 \lambda/D$ range, the static speckles are also strongly attenuated. In the remaining AO-corrected region, the speckles are also attenuated, which is particularly visible along the horizontal and vertical directions, where a strong static pattern of speckles was previously visible.

To quantitatively assess the performance gain after NCPA compensation, we plot in Fig. 10 the azimuthal standard deviation of the coronagraphic images as a function of separation, normalized to the peak flux 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 Paper I, 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.

5. Conclusion

In Paper I, we proposed a Zernike phase-mask sensor for the measurement of coronagraphic aberrations in exoplanet direct-imaging instruments. In this paper, we have presented the first experimental results and wavefront error correction with our concept in a real instrument, allowing us to improve the image quality in the coronagraphic images for exoplanet observation.

We designed and manufactured a prototype called ZELDA, which was installed inside VLT/SPHERE during its reintegration in Paranal in 2014. We validated the concept experimentally with tests performed on the internal point source where we manually introduced Zernike and Fourier modes on the DM. Following these encouraging results, we measured and compensated for the long-lived NCPA on SPHERE using the measurements from ZELDA. A contrast gain of up to one order of magnitude was reached at an angular separation of 5 \( \lambda/D \) from the axis in the coronagraphic images, providing very encouraging results for the observation of the giant gaseous planets down to the raw contrast limit set by the coronagraph. These performance results are at least on the same order as the contrast gain achieved by Paul et al. (2014) using the coronagraphic phase-diversity method named COFFEE. Additional studies will be performed to fairly compare the NCPA measurements and the contrast performance obtained by ZELDA and COFFEE.

While our first results are very encouraging, we have identified ways to improve the NCPA correction and thus the contrast performance. Our NCPA correction here relied on a ZELDA OPD map that was filtered in Fourier space with a Hann window as a first attempt. The quality of our wavefront error compensation is expected to be even more improved by investigating different sizes and natures of filter windows for the Zernike sensor map. Time constraints made our approach for the ZELDA-based wavefront calibration unsatisfactory for the correction of tip-tilt and higher-order modes because we did not account for all the specificities of SPHERE. Additional tests will enable us to take advantage of all the control loops in the instrument and achieve an optimal correction of the NCPA. We will also investigate combinations of ZELDA with IFS or polarimetric imager to calibrate chromatic or polarimetric aberrations.

Our experiment was performed with an internal source during our run on SPHERE, potentially representing a first step toward implementing a ZELDA sensing path on the instrument to improve its high-contrast observations on-sky. First phase-error maps have been obtained on sky with our concept to diagnose LWE observed on the coronagraphic images at wind speed below 1 m/s (Sauvage et al. 2016). Additional thorough studies will lead us to operate on-sky observations with ZELDA-based wavefront calibration and reach of the ultimate contrast limits of the instrument.

From a practical point of view, the NCPA correction with ZELDA in SPHERE could be implemented following two approaches. The off-line approach, in which the NCPA would be measured at the beginning of the night, and the same correction would be applied to all observations during that night. This assumes that most of the NCPA do not vary significantly over the course of a few hours. This is similar to the original calibration scheme foreseen for the instrument, where the NCPA could be calibrated once per day using phase diversity techniques (Sauvage et al. 2007). The online approach is more complex to implement since it requires replacing the entire near-infrared DTTS with a ZELDA-based sensor. The main gain of the online implementation is that it allows sensing not only tip-tilt variations but also higher-order NCPA in real-time during the observations, providing the equivalent of the calibration wavefront sensor in the Gemini/GPI instrument (Wallace et al. 2010b). This solution is currently under study for future evolutions of SPHERE.

The design simplicity and fast algorithm speed of our Zernike sensor for phase reconstruction makes its use very appealing for online measurements of the quasi-static coronagraphic aberrations during observations in various instruments. Given its properties, such a concept has already been adopted as a low-order wavefront sensor for the WFIRST mission and its coronagraphic module (Zhao 2014; Spergel et al. 2015). Current high-contrast imaging facilities and future exoplanet direct imagers on ELTs or envisioned post-JWST space observatories (Dalcanton et al. 2015) might thus benefit from the ZELDA-based wavefront correction to increase the signal-to-noise ratio of the planetary companions in the coronagraphic images and hence expand the discovery space of observable exoplanets by reaching deeper contrasts at small angular separations.

Acknowledgements. The authors are very grateful to the referee, Frans Snik, for his insightful suggestions and helpful comments to improve the quality of the original manuscript. This work is partially supported by the National Aeronautics and Space Administration under Grants NNX12AG05G and NNX14AD3G. The authors also acknowledge the support from CNRS (France), MPIA (Germany), INAF (Italy), FINES (Switzerland), MPIA (Heidelberg, Germany), LAM (Marseille, France), LESIA (Paris, France), Laboratoire Lagrange (Nice, France), INAF – Osservatorio di Padova (Italy), Observatoire de Genève (Switzerland), ETH Zurich (Switzerland), NOVA (Netherlands), ONERA (France) and ASTRON (Netherlands) in collaboration with ESO. SPHERE was funded by ESO, with additional contributions from CNRS (France), MPIA (Germany), INAF (Italy), FINES (Switzerland) and NOVA (Netherlands). SPHERE 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 numbers RI3-CT-2004-001566 for FP6 (2004–2008), grant number 226604 for FP7 (2009–2012) and grant number 312430 for FP7 (2013–2016).

References

Bloomhof, E. E., & Wallace, J. K. 2004, Optics Express, 12, 6240

A79, page 10 of 14
Appendix A: Zernike modes

During our tests on SPHERE, we measured not only the response curves of ZELDA for tip-tilt and defocus modes (see Sect. 3.2.1), but also for high-order modes: astigmatism, coma, trefoil, and spherical aberration. Figure A.1 display the results for these modes, showing good consistency between the simulated and experimental measurements. The discrepancies observed between theory and experiment for the tip-tilt error and defocus are possibly due to the imperfect calibration of the sensitivity factor. Our sensor enables the measurement of Zernike modes with nanometric accuracy, which is interesting for realignment purposes and online compensation of aberrations that are due to thermal or optomechanical drifts of the optics for an exoplanet direct-imaging instrument.

Appendix B: Fourier modes

We have also analyzed the response of ZELDA to Fourier modes with different spatial frequencies to assess the ability of our concept to capture speckles at different angular separations from the star position in the coronagraphic images. Section 3.2.1 shows the results obtained by ZELDA for the Fourier mode with a spatial frequency of five cycles per pupil in different filters. In addition to these tests, we also perform systematic ZELDA measurements of the Fourier modes with spatial frequencies ranging from two to ten cycles per pupil in both $x$- and $y$-axis directions, but using only the Fe II filter.

Figure B.1 presents the response curves for ZELDA with Fourier modes, showing again a very good match between the theoretical and experimental values. As for the Zernike modes, the discrepancies between theory and experiment for the Fourier modes are believed to be related to the errors in the calibration of the sensitivity factor. Our concept allows accurately measuring the aberrations corresponding to the bright speckles at the smallest separations from the star position in the coronagraphic images, thus enabling the observation of the closest and faintest planetary companions around the host star down to the raw contrast limit set by the coronagraph.
Fig. A.1. **Top plot**: theoretical and experimental aberration measurements for different Zernike modes as a function of the wavefront error in the entrance pupil plane. **Bottom plot**: difference between the experiment and the theory.
Fig. B.1. **Top plot:** theoretical and experimental aberration measurements for different Fourier modes in x- and y-axis directions as a function of the wavefront error in the entrance pupil plane. **Bottom plot:** difference between the experiment and the theory.