Calibration of quasi-static aberrations in high-contrast astronomical adaptive optics with a pupil-modulated point-diffraction interferometer

NICOLÁS DUBOST,* NAZIM ALI BHARMAL, AND RICHARD M. MYERS

CfAI, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
**nicolas.s.dubost@durham.ac.uk*

Abstract: The direct detection and imaging of exoplanets requires the use of high-contrast adaptive optics (AO). In these systems quasi-static aberrations need to be highly corrected and calibrated. In order to achieve this, a high-sensitivity wavefront sensor, the pupil-modulated point-diffraction interferometer (m-PDI), is presented. This sensor modulates and retrieves both the phase and the amplitude of an incoming electric field. The theory behind the wavefront reconstruction, the visibility of fringes, chromatic effects and noise propagation are developed. Results show this interferometer has a wide chromatic bandwidth. For a bandwidth of $\Delta\lambda = 50\%$ in units of central wavelength, the visibility of fringes and the response of the WFS to low and high-order aberrations are almost unaffected with respect to the monochromatic case. The WFS is, in contrast, very sensitive to variations in the size of its pinhole. The size of the pinhole is shown to affect the sensor's linearity, the dynamic range and the amount of noise. Larger pinholes make the sensor less sensitive to low-order aberrations, but in turn also decrease the effects of misalignments.

OCIS codes: (120.0120) Instrumentation, measurement, and metrology; (120.3180) Interferometry; (220.1080) Active or adaptive optics; (010.7350) Wave-front sensing.

References and links

1. T. Fusco, G. Rousset, J.-F. Sauvage, C. Petit, J.-L. Beuzit, K. Dohlen, D. Mouillet, J. Charton, M. Nicolle, M. Kasper, P. Baudoz, and P. Puget, "High-order adaptive optics requirements for direct detection of extrasolar planets: Application to the SPHERE instrument,," Opt. Express **14**, 7515–7534 (2006).
2. C. Marois, D. Lafreniere, R. Doyon, B. Macintosh, and D. Nadeau, "Angular Differential Imaging: A Powerful High-Contrast Imaging Technique," Astrophys. J. **641**, 556–564 (2006).
3. J.-L. Beuzit, M. Feldt, K. Dohlen, D. Mouillet, P. Puget, F. Wildi, L. Abe, J. Antichi, A. Baruffolo, P. Baudoz, A. Boccaletti, M. Carbillet, J. Charton, R. Claudi, M. Downing, C. Fabron, P. Feautrier, E. Fedrigo, T. Fusco, J.-L. Gach, R. Gratton, T. Henning, N. Hubin, F. Joos, M. Kasper, M. Langlois, R. Lenzen, C. Moutou, A. Pavlov, C. Petit, J. Pragt, P. Rabou, F. Rigal, R. Roelfsema, G. Rousset, M. Saisse, H.-M. Schmid, E. Stadler, C. Thalmann, M. Turatto, S. Udry, F. Vakili, and R. Waters, "SPHERE: a Planet Finder Instrument for the VLT," Proc. SPIE **7014**, 29–34 (2008).
4. O. Guyon, F. Martinache, V. Garrel, F. Vogt, K. Yokochi, and T. Yoshikawa, "The Subaru Coronagraphic Extreme AO (SCExAO) System: Wavefront Control and Detection of Exoplanets with Coherent Light Modulation in the Focal Plane," Proceedings of SPIE **7736**, 773624 (2010).
5. J.-F. Sauvage, T. Fusco, G. Rousset, and C. Petit, "Calibration and precompensation of noncommon path aberrations for extreme adaptive optics," JOSA A **24**, 2334–2346 (2007).
6. F. Martinache, O. Guyon, N. Jovanovic, C. Clergeon, G. Singh, T. Kudo, T. Currie, C. Thalmann, M. McElwain, and M. Tamura, "On-sky Speckle Nulling Demonstration at Small Angular Aeparation with SCExAO," Publications of the Astronomical Society of the Pacific **126**, 565–572 (2014).
7. F. Zernike, "Diffraction Theory of the Knife-Edge Test and its Improved Form, The Phase-Contrast Method," Monthly Notices of the Royal Astronomical Society **94**, 377–384 (1934).
8. R. N. Smartt and W. Steel, "Theory and Application of Point-Diffraction Interferometers," Japan. J. Appl. Phys. **14**, 351–356 (1975).

9. M. Loupias, M. Langlois, E. Thiébaut, M. Tallon, and J. Leger, "Status of an extreme adaptive optics testbench using a self-referenced Mach-Zehnder wavefront sensor," Proc. SPIE **9909**, 99094F (2016).
10. M. N'Diaye, A. Vigan, K. Dohlen, J.-F. Sauvage, A. Caillat, A. Costille, J. H. V. Girard, J.-L. Beuzit, T. Fusco, P. Blanchard, J. Le Merrer, D. Le Mignant, F. Madec, G. Moreaux, D. Mouillet, P. Puget, and G. Zins, "ZELDA, a Zernike wavefront sensor for the fine measurement of quasi-static aberrations in coronagraphic systems: concept studies and results with VLT/SPHERE," Proc. SPIE **9909**, 99096S (2016).
11. N. A. Bharmal, R. M. Myers, A. G. Basden, and A. P. Reeves, "An interferometric wavefront sensor for high-sensitivity low-amplitude measurements," Proc. SPIE **8447**, 84472J (2012).
12. H. Medecki, E. Tejnil, K. A. Goldberg, and J. Bokor, "Phase-shifting point diffraction interferometer," Opt. Lett. **21**, 1526 (1996).
13. J. K. Wallace, S. Rao, R. M. Jensen-Clem, and G. Serabyn, "Phase-shifting Zernike interferometer wavefront sensor," Proc. SPIE **8126**, 81260F (2011).
14. M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," Journal of the Optical Society of America **72**, 156 (1982).
15. M. N'Diaye, K. Dohlen, T. Fusco, and B. Paul, "Calibration of quasi-static aberrations in exoplanet direct-imaging instruments with a Zernike phase-mask sensor," Astronomy & Astrophysics **555**, A94 (2013).
16. 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," SPIE Proc. **7015**, 701518 (2008).
17. J.-F. Sauvage, T. Fusco, C. Petit, S. Meimon, E. Fedrigo, M. Suarez Valles, M. Kasper, N. Hubin, J.-L. Beuzit, J. Charton, A. Costille, P. Rabou ., D. Mouillet, P. Baudoz, T. Buey, A. Sevin, F. Wildi, and K. Dohlen, "SAXO, the eXtreme Adaptive Optics System of SPHERE: overview and calibration procedure," SPIE Proc. **7736**, 13 (2010).
18. B. Macintosh, J. R. Graham, P. Ingraham, Q. Konopacky, C. Marois, M. Perrin, L. Poyneer, B. Bauman, T. Barman, A. S. Burrows, A. Cardwell, J. Chilcote, R. J. De Rosa, D. Dillon, R. Doyon, J. Dunn, D. Erikson, M. P. Fitzgerald, D. Gavel, S. Goodsell, M. Hartung, P. Hibon, P. Kalas, J. Larkin, J. Maire, F. Marchis, M. S. Marley, J. McBride, M. Millar-Blanchaer, K. Morzinski, A. Norton, B. R. Oppenheimer, D. Palmer, J. Patience, L. Pueyo, F. Rantakyro, N. Sadakuni, L. Saddlemyer, D. Savransky, A. Serio, R. Soummer, A. Sivaramakrishnan, I. Song, S. Thomas, J. K. Wallace, S. Wiktorowicz, and S. Wolff, "First light of the Gemini Planet Imager," Proceedings of the National Academy of Sciences **111**, 12661–12666 (2014).

1. Introduction

In astronomical adaptive optics (AO) the application of high-contrast techniques, including extreme adaptive optics (XAO), has a particular focus towards the direct detection and imaging of exoplanets. It has been estimated that, in systems such as the Gemini Planet Imager (GPI) [1], the non-common path aberrations (NCPA) between the imaging arm and the WFS would have to be corrected to less than 10 nm rms, within the first 100 modes in order to achieve their targets. Other slowly changing aberrations are produced by telescope flexures, temperature differences, or the movement of elements such as the atmospheric dispersion corrector or the derotator. These aberrations are at the origin of quasi-static speckles which vary in the order of minutes, hours, or even days depending on their origin. For example, at Gemini with Altair NIRI, it was reported the PSF evolves on timescales of 10-60 minutes [2]. Because the aberrations are long lived, they do not average out over time and are more difficult to deal with.

In their first approach, systems like VLT-SPHERE [3] and Subaru-SCExAO [4] relied respectively on methods such as phase-diversity [5] and speckle-nulling [6]. For the latter, it has been estimated the converging time needs to be under 15-20 minutes to leave time for the science acquisition (this is only a requirement and converging times can be much lower). Another successful technique is Angular Differential Imaging [2] (ADI), which relies on the differential motion between the sky and the pupil to distinguish speckles from astrophysical signals. This technique is reliant on a stable PSF, being able to increase its noise attenuation by a factor of 3 when dealing with PSFs that evolve in the order of 60 minutes, rather than when they evolve in the order of 10 minutes.

The problem of measuring small quasi-static speckles in real-time can be addressed with the use of interferometers. Since the light of a star is not temporally coherent, such an approach would have to produce a self-generated reference beam to exactly match the piston and the polarisation of the test beam. If this is not done, but instead the test beam is interfered with a

beam produced by a local source, no interference would be possible. By introducing a portion of the input light into a small pinhole, point-diffraction interferometers (PDI) are able to produce nearly flat reference beams that do not require piston calibration. The first version of a PDI, the Zernike phase-contrast test was introduced by Zernike [7]. The theory was further generalized and the term coined by Smartt and Steel [8]. The difference between the Zernike phase-contrast test and Smartt's PDI is the first introduces a phase change in the pinhole, while the second uses an opaque film to reduce in different amounts the amplitude of the light transmitted through and around the pinhole. When a sensor is based on the Zernike phase-contrast test it is also dubbed a Zernike sensor. These interferometers can be seen as common-path configurations of Mach-Zehnder interferometers with one output. Mach-Zehnders are also a type of PDI and have been integrated into XAO systems [9]. Although all PDIs are susceptible to residual tip-tilts, which is discussed later in this paper, those that are common-path have fewer parts, are more compact and more robust against external perturbations such as vibrations and air turbulence in the optics.

ZELDA, a Zernike sensor, has successfully been demonstrated in SPHERE [10]. As standard PDIs do, Zernike sensors encode wavefront phase directly into light intensity. The advantage of this is, under small aberrations, the measured light intensity is proportional to the sine of the phase, allowing for quick wavefront reconstruction. But this kind of PDI relies on the assumption that the illumination is homogeneous across the pupil. Therefore, heterogeneous variations of intensity in the pupil will lead to an erroneous estimate of the phase. For example, in single-conjugate AO (SCAO) systems, uncorrected scintillation at the pupil level could register as phase. Current XAO systems are not multi-conjugate and cannot deal with scintillation.

This paper further develops the theory for a solution first presented by Bharmal [11]. This approach is an improved version of the phase-shifting PDI (PS-PDI) presented by Medecki [12]. This technique is not to be confused with the phase-shifting PDI later introduced by Wallace [13]. The later is similar to a Zernike sensor in that it adds a piston phase to the core of the point-spread function (PSF) in order to produce the reference beam. In Medecki's approach, the phase-shift is introduced as a tilt at the pupil with a small-angle beam-splitter. Because the tilt produces a modulation of the electric field in the pupil, we have decided instead to refer to our method as the pupil-modulated PDI (m-PDI). An introduction to the working principle of the modified m-PDI is then followed by the mathematical formalism. The mathematical model is used to derive the propagation of sources of error, such as photon noise, read-out noise and chromatic effect. The linearity and the dynamic range is then tested against variations in multiple parameters.

In Section 2 we present the concept and mathematical formalism of the improved m-PDI, later deriving from it the propagation of photon-shot noise and read-out noise in Section 3. The effects of Strehl and of chromatic bandwidth on the visibility of fringes is later studied in Section 4. A sensitivity analysis of the sensor's linearity and dynamic range are carried out in Section 5. Finally, Section 6, the sensor is tested in a simulation of a real system and the sensitivity is presented as the exposure time required to achieve a given signal-to-noise ratio.

2. The pupil-modulated point-diffraction interferometer

2.1. Principle

The scheme of the m-PDI is given in Fig. 1. At the entrance pupil, an aberrated wavefront goes through a small-angle beam splitter, which in this case is a grating. The beam is then split into modes, some of which will later be interfered at the exit pupil. The central mode (i.e. mode 0) goes through a narrow pinhole in the focal plane filter mask. This filters out higher frequencies leaving a flat beam which will be used as the reference beam. As shown in Section 5, the size of the pinhole and, thus, the frequencies left in the reference beam have an important effect on the sensor's linearity, sensitivity and dynamic range. Mode +1 goes through a larger aperture which filters enough frequencies to avoid aliasing. This is important to produce an unambiguous test

beam, as will be explained in Section 2.2. Both beams later interfere at the exit pupil producing fringes. The interference fringes are an image of the line-pairs in the grating, which, as it will be shown later, are modulated by both the phase of the electric field and its amplitude.

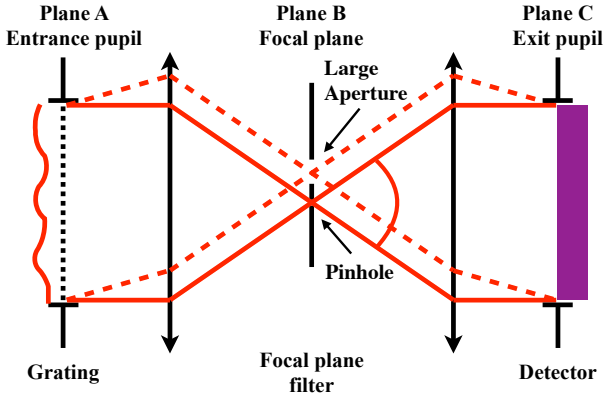


Fig. 1. Layout of the pupil-modulated point-diffraction interferometer. The grating splits the beam into modes. Mode 0, shown in a solid red line, goes through a point-diffraction pinhole in the focal plane. Mode +1 goes through a larger aperture.

In Medecki's PDI the apertures in the focal plane filter mask are swapped. Mode 0 goes through the large aperture while mode +1 goes through the pinhole. In this configuration, since mode +1 carries less light than the central mode and is being filtered, the reference beam has less light than the test beam. This reduces the visibility of the fringes. Another disadvantage is the pinhole filters light by wavelength, letting through only a narrow bandwidth. By letting mode +1 go through a larger aperture, more wavelengths can go through. This is a necessary feature in astronomical AO because of the limited amount of light available.

2.2. *Formalism*

For an aberrated wavefront, the electric field

$$\Psi_0 = Pe^{i\varphi} = P_0(1-\epsilon)e^{i\varphi}, \tag{1}$$

where P is the amplitude of the electric field, P_0 is its average across the pupil, ϵ is a zero-mean function describing the local scintillation and φ is the phase. The function P is considered to be 0 outside of the telescope pupil, hence defining the telescope's aperture shape. In this case, the small-angle beam splitter will be considered to be a grating. The grating is located on Plane A, as defined in Fig. 1, and is described by a square wave function G_T of period T. This function alternates between the value 0 representing an obstruction and 1 representing full transmission. For simplicity, the function will be considered to be even, so its description in the Fourier domain is also even and real. The wavefront after the grating is described by

$$\Psi_A = \Psi_0 \cdot G_T. \tag{2}$$

Just before the focal plane filter mask on Plane B, the wavefront is described by

$$\widehat{\Psi}_A = \widehat{\Psi}_0 \otimes \widehat{G}_T, \tag{3}$$

where $\widehat{A}$ is the Fourier transform of a $\mathcal{F}[A]$ and the $\otimes$ symbol represents the convolution operation. The hat notation and $\mathcal{F}$ will also include the Fourier optics scaling factor $1/\lambda f$, where λ is the

wavelength and f is the lens' focal length. With these considerations taken into account, the Fourier transform of G_T is

$$\widehat{G}_T(k,\zeta) = \frac{1}{2}\delta(k,\zeta) + \frac{1}{\pi}\sum_{m=1,2,\ldots}^{\infty}\frac{1}{2m-1}\left[\delta\left(k - \frac{\lambda f(2m-1)}{T},\zeta\right) + \delta\left(k + \frac{\lambda f(2m-1)}{T},\zeta\right)\right], \tag{4}$$

where $\delta(k,\zeta)$ is a Dirac delta function, k and ζ are Cartesian position coordinates in the focal plane in meters and the indices $m = 1, 2, \ldots$ represent the modes produced by the grating. Since the dispersion modes produced by the grating spread parallel to the axis described by k, ζ will not be included in further equations unless necessary.

At Plane B the wavefront is multiplied by the focal plane filter mask

$$M = M_0 + M_{+1}, \tag{5}$$

where M_0 is a circular top-hat functions equal to 1 for $|(k,\zeta)| < D_{B,0}/2$, M_{+1} is a square top-hat function equal to 1 for $|k - \lambda_0 f/T| < D_{B,+1}/2 \cap |\zeta| < D_{B,+1}/2$, 0 elsewhere, $D_{B,0}$ and $D_{B,+1}$ denote their respective diameter and side, and λ_0 is the central wavelength around which the instrument is designed. The apertures' geometry is presented in Fig. 2. The central wavelength

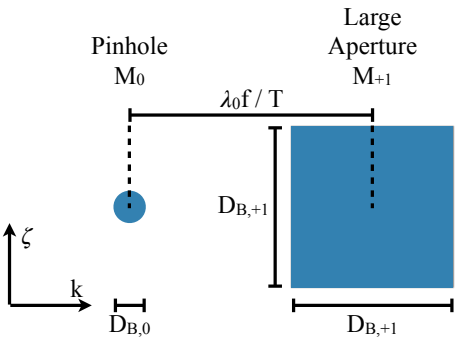


Fig. 2. Focal plane filter mask, in Plane B and with a square large aperture M_{+1}.

λ_0 is the wavelength that goes right through the center of the large aperture M_{+1}. When the instrument is fed with polychromatic light, λ_0 may not be the same as λ_c, which is the wavelength at the center of the light's spectrum. Contrary to M_0, M_{+1} can be either a circle, a rectangle or any other shape. The square aperture configuration lets through the same spatial frequencies in both the X and the Y direction. The axes k and ζ in the focal plane are respectively parallel to X and Y in the pupil planes.

After the mask, the electric field is

$$\Psi_B = \widehat{\Psi}_A M = \frac{1}{2}\widehat{\Psi}_0(k)M_0 + \frac{1}{\pi}\widehat{\Psi}_0\left(k - \frac{\lambda f}{T}\right)M_{+1}. \tag{6}$$

When $D_{B,0}$ is in the vicinity of $\lambda f/D_A$ or smaller, where D_A is the diameter of the entrance pupil, then the first term of Eq. (6) can be considered a point-diffraction source

$$\Psi_{B,0}(k) = \frac{1}{2}\widehat{\Psi}_0(k)M_0 \simeq b\delta(k) \tag{7}$$

of amplitude b. On the other hand, in the second term of the expression, M_{+1} acts as a low-pass filter on $\widehat{\Psi}_0$, leaving Eq. (6) as

$$\Psi_B \simeq \Psi_{B,0}(k) + \frac{1}{\pi}\widehat{\Psi}_{LP}\left(k - \frac{\lambda f}{T}\right), \tag{8}$$

where Ψ_{LP} has been low-pass filtered by M_{+1}.

The light is then propagated into Plane C, where the detector reads the intensity function

$$I_C = \Psi_C \Psi_C^* = \widehat{\Psi}_B \left(\widehat{\Psi}_B\right)^*, \tag{9}$$

where $*$ notes the complex conjugate of a function. In order to retrieve the phase of the original wavefront a Fourier transform is applied to I_C, leaving

$$\widehat{I_C} = \widehat{\widehat{\Psi}_B \widehat{\Psi}_B^*} = \Psi_B(k) \otimes \Psi_B(-k) \tag{10}$$

$$= \left(\Psi_{B,0}(k) + \frac{1}{\pi}\widehat{\Psi}_{LP}\left(k - \frac{\lambda f}{T}\right)\right) \otimes \left(\Psi_{B,0}(-k) + \frac{1}{\pi}\widehat{\Psi}_{LP}\left(-k - \frac{\lambda f}{T}\right)\right) \tag{11}$$

All the terms in Eq. (11) developed and graphically represented in Fig. 3. In the figure, the lateral

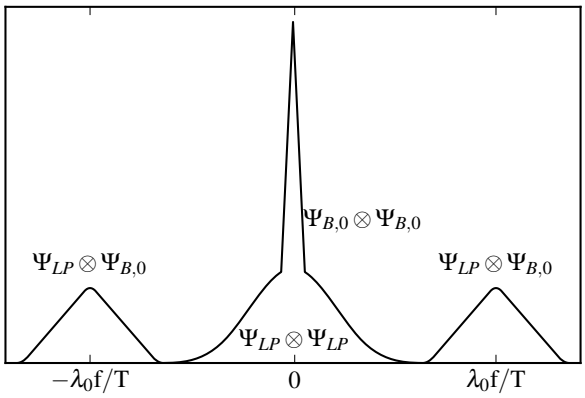


Fig. 3. Fourier transform of the wavefront intensity in Plane C, where Ψ_{LP} is Ψ_0 low-pass filtered by mask aperture M_{+1} and P is the point-diffraction source produced by mask aperture M_0.

sidebands contain the frequency modulated spectrum of the original phase, convolved with the point-diffraction source $\Psi_{B,0}$. As a consequence, the smaller the point-diffraction source, the higher the fidelity between the sideband and the original wavefront. It is also important to notice that, since Ψ_{LP} has a finite bandwidth, the bandwidth of the central signal $\Psi_{LP} \otimes \Psi_{LP}$ is twice as large as the those of the lateral ones. Therefore, in order to avoid aliasing

$$\frac{3}{2} D_{B,+1} < \frac{\lambda_0 f}{T}, \tag{12}$$

since the bandwidth of Ψ_{LP} is set by the aperture M_{+1} of diameter $D_{B,+1}$ in the focal plane filter mask. Note that this means the highest spatial frequency that can be sampled in the X direction is defined by the period of the grating T as

$$k_{max} < 1/3T. \tag{13}$$

As a comparison, the period of the grating sets the maximum spatial frequency in the X direction in a similar way the size of subapertures do in a SH. For simplicity in this case M_{+1} has been set to be a square so that k_{max} equals the maximum frequency in the Y direction ζ_{max}. But in principle, for an arbitrary size M_{+1} in the ζ direction, ζ_{max} is independent of the grating's size and only depends on the size μ_p of the pixels sampling I_C. In this case the relationship would be

$$\zeta_{max} < 1/2\mu_p. \tag{14}$$

After calculating the Fourier transform of I_C, the next steps to demodulate the wavefront are to filter out the undesirable terms multiplying by M_{+1}, scaling by the amplitude of the point-diffraction source $\Psi_{B,0}$, shifting by $\lambda_0 f/T$ and inverting the Fourier transform. Applying these operations gives

$$\begin{aligned}
&\mathcal{F}^{-1}\left[\frac{\pi}{b}\left(\widehat{I}_C M_{+1}\right)\otimes\delta\left(k+\frac{\lambda_0 f}{T}\right)\right]\\
&=\mathcal{F}^{-1}\left[\frac{\pi}{b}\left(\Psi_{B,0}\otimes\frac{1}{\pi}\widehat{\Psi}_{LP}\left(k-\frac{\lambda_0 f}{T}\right)\right)\otimes\delta\left(k+\frac{\lambda_0 f}{T}\right)\right]\\
&=\mathcal{F}^{-1}\left[\frac{\pi}{b}\Psi_{B,0}\otimes\frac{1}{\pi}\widehat{\Psi}_{LP}(k)\right]\simeq\mathcal{F}^{-1}\left[\widehat{\Psi}_{LP}(k)\right]=\widetilde{\Psi}_0,
\end{aligned}\tag{15}$$

where $\widetilde{\Psi}_0$ is the estimate of Ψ_0. Since the electric field Ψ_0 was defined in Eq. (1) as a function of amplitude and phase, it is in principle possible to retrieve both from the estimate $\widetilde{\Psi}_0$. In this case, the estimated phase is not a function of the amplitude as it is in standard PDIs and in Zernike sensors. Because of this, a non-homogeneous illumination of the pupil does not lead to errors in the estimation of the phase. This algorithm of phase retrieval that consists of using the Fourier transform of the interferogram to first retrieve the electric-field was introduced by Takeda [14]. The focus of this paper will be the estimation of the wavefront's phase, rather than the estimation of the amplitude.

3. Noise propagation

Introducing the simplification in Eq. (7) into Eq. (11), the intensity in the exit pupil becomes

$$I_C = b^2+\frac{1}{\pi^2}P^2+\frac{2b}{\pi}P\cos\left(\frac{2\pi}{T}x-\varphi_{LP}\right),\tag{16}$$

where φ_{LP} is φ low-pass filtered by M_{+1}. On one hand this simplified expression will prove useful to study the effects of Strehl and chromaticity in Section 4. On the other hand, it is the starting point for the analytical derivation of noise propagation. The details of the derivation can be found in Appendix A. It is emphasized that, as the main purpose of the m-PDI is to act as a WFS, only the effects of noise on the retrieved phase are studied.

If the illumination of the pupil is considered to be homogeneous so that $P(x, y) = P_0$, where P_0 is a constant value, and the intensity at the entrance pupil is $I_0 = P_0^2$, the error contribution of the read-out noise and the photon-shot noise is

$$\sigma_R^2+\sigma_P^2=\frac{2\pi^2}{9N_P^2b^2I_0}\delta S_R^2+\frac{2\pi^2}{9N_P^2b^2I_0}\left(b^2+\frac{1}{\pi^2}I_0\right),\tag{17}$$

where N_P is the number of pixels sampling a line-pair and δS_R^2 is the power of the read-out noise.

For the case of maximum fringe visibility, i.e. when $b = P_0/\pi$, and for a small line-pair sampling with $N_P = 4$, the noise is

$$\sigma_R^2+\sigma_P^2\simeq\frac{1.4}{I_0^2}\delta S_R^2+\frac{0.3}{I_0}.\tag{18}$$

The noise propagation is flat across the spatial frequency spectrum, different to a Shack-Hartmann WFS, and can be improved by increasing the number N_P of pixels per line-pair.

Figure 4 shows a comparison between simulation results and the model described by Eq. (17). The simulation is done with a pinhole diameter $D_{B,0} = \lambda_0/D_A$, which translates into $b \simeq 0.22$ Simulation results confirm noise propagation is flat across the frequency spectrum and show

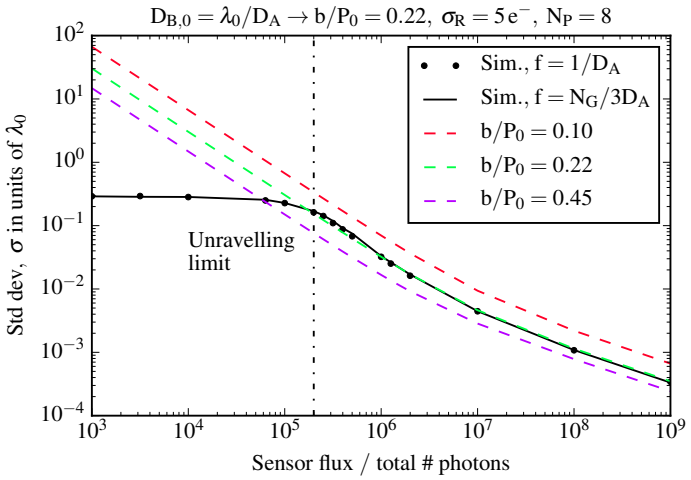


Fig. 4. Propagation of read-out and photon-shot noise as a function of sensor flux available before the entrance pupil. Simulated data is presented for two sinusoidal wavefronts fed into the WFS, with frequencies $f = 1/D_A$ and $f = N_G/3D_A$. Simulations are compared to estimations using Eq. (17) for 3 values of b normalized by P_0.

good agreement with the theory until the unraveling limit. The simplest form of wavefront reconstruction for this sensor consists in reconstructing the electric field and then retrieving the phase, as stated in Section 2.2. Because the phase produced by noise can not be unraveled, the noise reaches a limit. After this limit reconstructed signals are meaningless.

For design purposes, it is interesting to study the dependency of read-out and photon-shot noise to the reference beam's amplitude b. Figure 5 shows the propagation of noise as a function of the reference beam's normalized amplitude b/P_0, for a given sensor flux and as described by Eq. (17). Since the noise decreases as b/P_0 increases, there is an incentive to maximize the

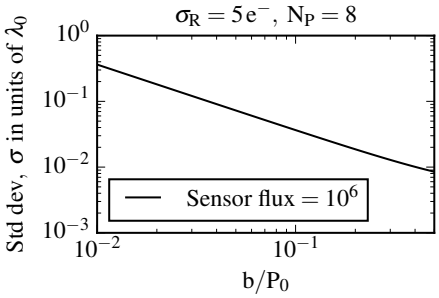


Fig. 5. Propagation of read-out and photon-shot noise as a function of normalized reference beam amplitude, for a given sensor flux.

intensity of the reference beam by having the largest possible pinhole. But as it has already been stated in this section, there is a pinhole size which maximizes visibility. As it will later be shown in Section 5 a larger pinhole leads to a reduction in dynamic range and in accuracy to low-order aberrations.

4. Visibility loss from chromatic bandwidth and a decrease in Strehl

A simple way of including the effects of Strehl into the analytical model is to consider $b = b_0\sqrt{S}$ where b_0 is the highest possible value of b produced with a flat non aberrated wavefront and S is Strehl.

For a rectangular chromatic bandwidth $\Delta\lambda$ around a central wavelength λ_0 and $P = P_0$, the intensity in the exit pupil is

$$\begin{aligned} I_C(x, y, \Delta\lambda, \lambda_0) &= \frac{1}{\Delta\lambda} \int_{\lambda_0-\Delta\lambda/2}^{\lambda_0+\Delta\lambda/2} I_C(x, y, \lambda) d\lambda \\ &\simeq S b_0^2 + \frac{P_0^2}{\pi^2} + \frac{2\sqrt{S} b_0 P_0}{\pi} \operatorname{sinc}\left(\frac{\Delta\lambda \varphi_{0,LP}}{2\lambda_0}\right) \cos\left(\frac{2\pi}{T} x - \varphi_{0,LP}\right), \end{aligned} \tag{19}$$

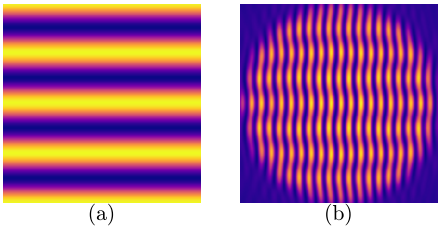
where $\operatorname{sinc}(x) = \sin(x)/x$ and $\varphi_{0,LP}$ is the low-pass filtered phase for λ_0. The first thing to notice about the equation is the modulation of the fringes only depends on the phase of the central wavelength. The effects of the chromatic bandwidth are all wrapped inside the sinc function. Combined, the effects of the Strehl and of the chromatic bandwidth are given by

$$V(x, y) = \sqrt{S} \operatorname{sinc}\left(\frac{\Delta\lambda \varphi_{0,LP}(x, y)}{2\lambda_0}\right), \tag{20}$$

where $0 < V < 1$ will be considered a proxy of visibility.

It is important to note that V depends both in general on S and locally on $\varphi_{0,LP}(x, y)$. In other words, on the one hand a decrease in Strehl produces a general visibility loss across the pupil, and on the other aberrations also introduce local visibility losses proportional to the bandwidth. But local effects are so small they can be neglected. For example, for $\Delta\lambda = 0.2$ and $\varphi_{0,LP} = 0.25$ both in units of λ_0, then $\operatorname{sinc}\left(\Delta\lambda\varphi_{0,LP}(x, y)/2\lambda_0\right) = 0.996$. In contrast, for a sinusoidal aberration with an amplitude equal to the previous value, then the general term is $\sqrt{S} = 0.11$, which produces a strong loss of visibility.

To test Eq. (20), a sinusoidal aberration perpendicular to the fringes is put across the input pupil as shown in Fig. 6. This input allows the measurement the visibility V for known values of $\varphi_{0,LP}$.



(a) (b)

Fig. 6. (a) Input wavefront at entrance pupil A. The sinusoidal input is perpendicular to the grating's line-pairs and to the resulting fringes on the exit pupil. (b) Intensity at exit pupil C

Figure 7 shows a comparison between the theoretical visibility loss in the monochromatic case against three polychromatic simulated cases. As expected from the model, the effect of the chromatic bandwidth is negligible when contained under 50%. The biggest mismatch between the simulation and the analytical model is close to 15% and takes place in the low Strehl regime around $S \simeq 0.16$, as shown in Fig. 7. This disagreement lies outside of most of the scientific cases for this WFS.

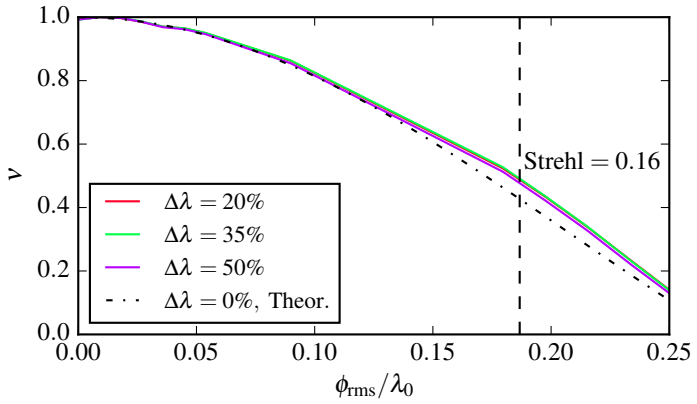


Fig. 7. Loss of visibility as a function of wavefront phase RMS in units of wavelength, for a sinusoidal aberration, perpendicular to the grating's line-pairs. Solid lines represent simulated polychromatic cases and the dashed line represents the theoretical monochromatic prediction. Visibility is measured in the place of maximum aberration $\varphi_{0,LP}$. The vertical line marks the largest disagreement between the theoretical model and the simulation.

5. Accuracy and dynamic range

The design of the m-PDI involves several parameters of which are here considered the number of line-pairs or grooves in the pupil N_G, the number of pixels sampling the interference fringes N_P, the chromatic bandwidth $\Delta\lambda$ in units of the central wavelength λ_0 and the focal plane pinhole's diameter $D_{B,0}$. The values these parameters take will determine the accuracy, linearity and dynamic range of the resulting WFS. Figure 8 shows how the response of the WFS concept changes when these parameters take on different values. Results for different central wavelengths are not presented since all plots remain the same as long as the axes are in units of λ_0.

In the initial configuration, as well as in most configurations, our sensor remains linear for twice the range of the Zernike sensor [15]. The figure shows the linearity and the dynamic range are very insensitive to changes in the number of pixels per interference fringe N_P and to the chromatic bandwidth. As a consequence the m-PDI has a large chromatic bandwidth which compensates for the loss of light on the grating and on the spatial filter. The sensor's throughput, given perfect optics, is given by

$$\eta = \frac{b^2}{P_0^2} + \frac{1}{\pi^2}, \tag{21}$$

which for $D_{B,0} = \lambda/D_A$ gives $\eta = 0.15$. This means that with a bandwidth of $\Delta\lambda = 50\,\%$, the sensor takes in as much light as another sensor with a throughput of 1 and a bandwidth $\Delta\lambda = 7.5\,\%$

In a middle ground, changes to the number of line-pairs across the pupil N_G produce a small reduction of 16% in the response to spherical aberrations. Since the maximum spatial resolution is proportional to N_G and that both N_G and N_P can easily be changed without much effect on the linearity and dynamic range, the number of pixels can be kept low, all while having a high resolution.

Finally, the size of the pinhole is the variable with the greatest effect on the accuracy. Indeed, in Fig. 8 panel (v-a) and for $D_{B,0} = 2\lambda_0/D_A$ the response to tip (or tilt) is reduced by 50% as most of the PSF's core is well inside the pinhole. This is balanced by an increase in the amplitude of the reference beam b and a subsequent decrease in photon-shot and read-out noise as shown in Fig.5. Doubling the size of the pinhole increases b by a factor of 1.9 up to 0.41, which in

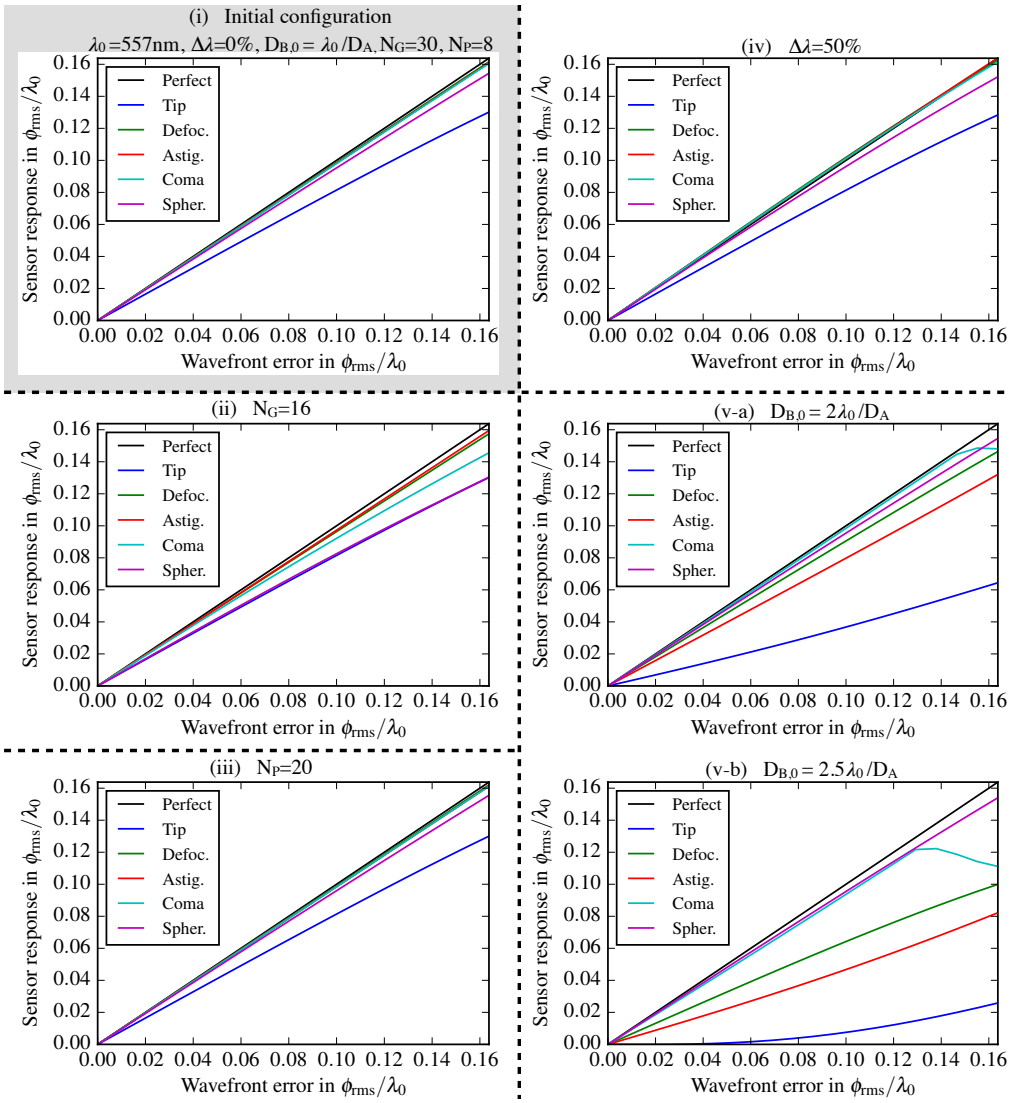


Fig. 8. Response of the m-PDI to wavefront errors for different low-order aberrations and in different configurations. Each configuration is produced by changing one parameter from (i) the initial configuration. The 4 parameters that are changed are: (ii) the number of line-pairs in the pupil N_G, (iii) the number of pixels sampling an interference fringe N_P, (iv) the chromatic bandwidth $\Delta\lambda$ in units of λ_0 and (v) the focal plane pinhole's diameter $D_{B,0}$.

turn decreases noise by 42%. This configuration could be desirable in cases where the WFS does not have to measure tip and tilt but the greatest possible value for b is required to minimize photon-shot and read-out noise. As the pinhole's diameter becomes greater than $2.5\lambda_0/D_A$, one by one and almost from the lowest to the highest order, the responses to different aberrations start to decrease. This can be seen in Fig. 8 panel (v-b). Here $b = 0.42$, leading to a reduction of 43% on the noise compared with the initial configuration. It is also possible to notice there is sharp decline in accuracy to coma, abruptly restricting the dynamic range for this mode. The origin of this phenomenon as well as its observability in the fringe pattern is to be the subject of further study. Larger pinholes are not desirable as all the accuracy is be lost with no noise reduction. For

increasingly higher pinhole diameters, b converges to 0.45, which at most produces a decrease of 46% with respect to the initial configuration.

The effects of spatial filtering on the WFS's accuracy to different spatial modes are better explored in Fig. 9. The figure shows the transfer function relative to spatial frequency for different

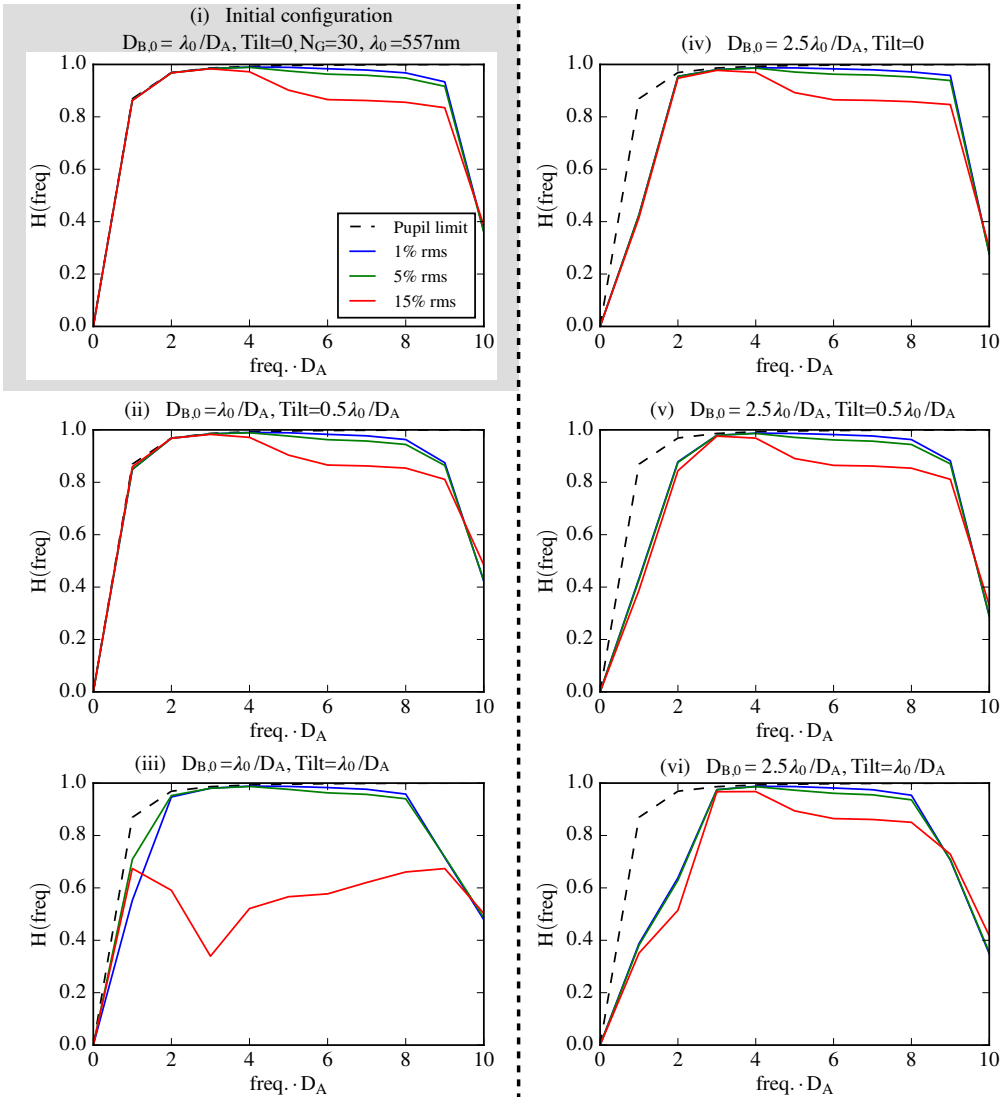


Fig. 9. Transfer function with respect to spatial frequency for different aberration amplitude and in different configurations. Columns have different pinhole sizes and rows have different preexisting tilts.

pinhole sizes and different pre-existent tilts. Tilts can be caused by pointing errors or other sources of misalignments. The transfer function spans all the spatial frequencies that can be measured with $N_G = 30$, as described by Eq. (13). As can be observed in all the panels, the transfer functions suffers from a sharp decline as it approaches the maximum frequency k_{max}. It is also important to notice that the maximum sensitivity to low order aberrations is given by the shape of the pupil. In this case, a circular pupil has a poor transfer function for frequencies with a spatial scale in the vicinity of its diameter.

The first thing to notice is the difference between the left panels, which all have a small pinhole size $D_{B,0} = \lambda_0/D_A$, and the right panels, with a larger pinhole size $D_{B,0} = 2.5\lambda_0/D_A$. With a bigger pinhole size the WFS looses sensitivity to small aberrations. This happens because all aberrations that are not filtered by the pinhole are present in the test beam and therefore become invisible to the sensor.

The disadvantage of a small pinhole though are aberrations can more easily drain it of the light necessary to produce the test beam. Figure 9 shows, from top to bottom, how the transfer function changes with an increasingly larger preexisting tilt. At first, the effects go unnoticed, but in panel (iii) there is clear loss of dynamic range. This doesn't happen in the panel to the right, panel (vi), where the same tilt is applied but the pinhole is larger. The advantage of the larger pinhole is to keep more light into the reference beam despite aberrations.

Finally, the transfer function acts as a good indicator of the shape of the signal-to-noise ratio (SNR) as a function of spatial frequency. This is the case since, as Fig. 4 shows, noise is flat across the spectrum. The sensitivity of the m-PDI in terms of SNR is better discussed in the following section.

6. Application to a real system

Here, the application of the sensor to a real system is considered. The sensor will get a 5% share of light from a beam-splitter, so other operations can be performed in parallel (like science observations). It is worth noticing the sensor could also be fed the entirety of an unused chromatic band by a dichroic. This alternative is not presented here. The system will be similar to other XAO systems that have been developed [16, 17], but with a lower read-out noise to consider newer detector technology. These systems are optimized to work around $\lambda = 1.6\,\mu$m. In the case of GPI, for its first light the system was able to close the AO loop for stars $I < 8$ mag, but was expected to operate down to $I \simeq 10$ mag under better seeing conditions [18]. Table 1 shows the parameters of the simulation.

Table 1. Parameters used for simulating the exposure time.

Parameters	Values
Central wavelength λ_0	1.625 μm
Bandwidth $\Delta\lambda$	20%, 50%
Zero mag. flux density	1080 Jy
Apparent star magnitude	10
Telescope diameter	8 m
Telescope transmission T_{tel}	40%
Beamsplitter transmission T_{BS}	5%
Line-pairs across pupil N_G	60, 185
Pixels per line-pair N_P	4
Read-out noise	1 e^-

Figure 10 shows the sensor's sensitivity to the order and amplitude of aberrations, represented as the exposure time needed to achieve SNR = 1. The simulation is performed for two different bandwidths, $\Delta\lambda = 20\,\%$ and $\Delta\lambda = 50\,\%$. In the results, the exposure time is inversely proportional to the chromatic bandwidth. This expands upon previous results showing the sensor's achromaticity for all orders.

As mentioned, the transfer function shapes the curves of exposure time. The exposure time is also inversely proportional to the transfer function. This means higher amplitude aberrations will present curves that are less flat around the center than lower amplitude ones.

The simulation was also performed for a low-order case with $N_G = 60$, and a high-order

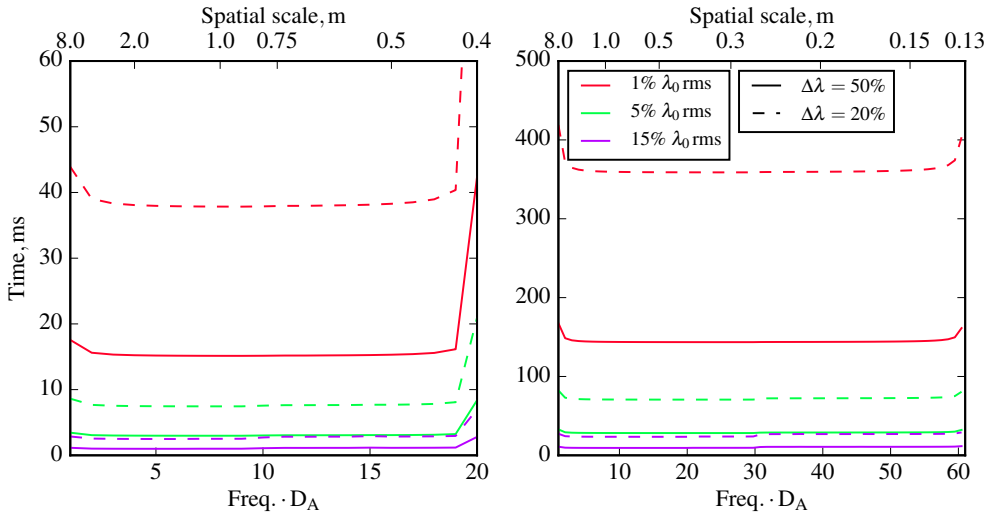


Fig. 10. Exposure time required to achieve SNR = 1 as a function of aberration spatial scale. (left) $N_G = 60$, (right) $N_G = 180$. The exposure time is presented for different amplitudes of aberration and for different chromatic bandwidths. The parameters of the simulation can be found in Table 1.

case, with $N_G = 185$. In the low-order case, the exposure times required are in the order of milliseconds and tens of milliseconds, meaning the sensor can operate at frequencies of up to hundreds of Hz. This is not fast enough to act as the main sensor of an XAO loop operating at frequencies that exceed 1 kHz, but would be enough for real-time quasi-static aberrations measurements. In the high-order case the sensor operate at frequencies of a few Hz. This is still enough to measure quasi-static aberrations down to the deformable mirror's scale, with actuators with a 13 cm separation.

7. Conclusions

The principle for a m-PDI has been presented. Later an analytical model was developed. The model allows for electric field reconstruction, i.e. to retrieve both the phase and the amplitude of an incoming electric field. The propagation of photon-shot noise and detector read-out were derived. The model was also extended to quantify chromatic effects and the influence of Strehl on the visibility of fringes. Both sets of predictions, on the propagation of noise and on the loss of visibility were tested against simulations built on first principles. The analytical models show good agreement with the simulations, showing the WFS is well understood.

An important result shows this interferometer has a wide chromatic bandwidth. Not only is the visibility of fringes unaffected by a wider bandwidth, but neither is the accuracy of the WFS to low nor high order aberrations. When $\Delta\lambda = 50\%$ the high accuracy of the monochromatic case is conserved. A wide chromatic bandwidth allows to compensate the loss of light at the grating and at the spatial filter. The trade-off of a wider chromatic bandwidth is a reduced range of measurable aberrations.

The WFS model is also shown to be sensitive to the size of the pinhole. While doubling the pinhole's diameter from λ_0/D_A to $2\lambda_0/D_A$ reduces the sensitivity to tip and tilt by 50%, it also decreases photon-shot and read-out noise by 42%. Increasing the size of the pinhole beyond this point is ill-advised as the noise can only decrease a further 4% while loosing all the sensitivity to low-order aberrations.

Finally, the exposure times required are in the order of milliseconds in a low-order case, meaning in this system the sensor can operate at frequencies of up to hundreds of Hz. This is not fast enough to act as the main sensor of an XAO loop operating at frequencies that exceed 1 kHz, but would be enough for real-time quasi-static aberrations measurements.

The exposure time required to achieve a given SNR was shown to be mostly flat across the spatial spectrum. This time is inversely proportional to the chromatic bandwidth and the aberration's amplitude, and is in the order of milliseconds for low-order aberrations, and tenths of a second for high-order aberrations. Such timescales are reasonable for the real-time measurement of quasi-static aberrations.

Appendix A: Analytical derivation for the propagation of photon-shot and read-out noise

As presented in Section 3, the intensity in the exit pupil can be described as

$$I_C = b^2 + \frac{1}{\pi^2}P^2 + \frac{2b}{\pi}P\cos\left(\frac{2\pi}{T}x - \varphi_{LP}\right), \tag{22}$$

where φ_{LP} is φ low-pass filtered by M_{+1}. For a small φ_{LP} the expression can be approximated as

$$I_C = b^2 + \frac{1}{\pi^2}P^2 + \frac{2b}{\pi}P\cos\left(\frac{2\pi}{T}x\right) + \frac{2b}{\pi}P\sin\left(\frac{2\pi}{T}x\right)\varphi_{LP}. \tag{23}$$

Back in the Fourier domain and considering $P \simeq P_0$, the demodulated phase is

$$\widehat{\varphi_{LP}} = \frac{\pi}{2bP_0}\left(\widehat{I_C} \otimes \left[-i\delta\left(k - \frac{\lambda_0 f}{T}\right) + i\delta\left(k + \frac{\lambda_0 f}{T}\right)\right]\right) M_{+1}\left(k + \frac{\lambda_0 f}{T}\right), \tag{24}$$

where the mask M_{+1} has been centered around the origin.

When white noise n_w of mean value $\bar{n}_w = 0$ and power δS_w^2 is added to the detected signal I_C and put through the demodulation the result is

$$\frac{\pi}{2bP_0} i\left(-\widehat{n}_w\left(k - \frac{\lambda_0 f}{T}\right) + \widehat{n}_w\left(k + \frac{\lambda_0 f}{T}\right)\right) M_{+1}\left(k + \frac{\lambda_0 f}{T}\right), \tag{25}$$

where $\widehat{n}_w\left(k - \frac{\lambda_0 f}{T}\right)$ and $\widehat{n}_w\left(k + \frac{\lambda_0 f}{T}\right)$ are uncorrelated. This means the sum of both results in a white noise $n_{w'}$ where $\delta S_{w'}^2 = 2\delta S_w^2$. The resulting noise is then filtered by M_{+1}. Using Parseval's theorem the power of the filtered noise can be computed as

$$\delta S_{w'}^2 = \left\langle \int_{-\lambda_0 f/2\mu_p}^{\lambda_0 f/2\mu_p} \int_{-\lambda_0 f/2\mu_p}^{\lambda_0 f/2\mu_p} |\widehat{n}_{w'}|^2 \, dk d\zeta \right\rangle = \int_{-\lambda_0 f/2\mu_p}^{\lambda_0 f/2\mu_p} \int_{-\lambda_0 f/2\mu_p}^{\lambda_0 f/2\mu_p} \langle |\widehat{n}_{w'}|^2 \rangle dk d\zeta, \tag{26}$$

For a square filter

$$M_{+1}\left(k + \frac{\lambda_0 f}{T}\right) = \mathrm{rect}\left(\frac{3Tk}{2\lambda_0 f}\right), \tag{27}$$

the power of the filtered noise is

$$\begin{aligned}\delta S_{w',M_{+1}}^2 &= \int_{-\lambda_0 f/2\mu_p}^{\lambda_0 f/2\mu_p} \int_{-\lambda_0 f/2\mu_p}^{\lambda_0 f/2\mu_p} \langle |\widehat{n}_{w'}|^2 \rangle rect\left(\frac{3Tk}{2\lambda_0 f}\right) dk d\zeta \\ &= \int_{-\lambda_0 f/3T}^{\lambda_0 f/3T} \int_{-\lambda_0 f/3T}^{\lambda_0 f/3T} \langle |\widehat{n}_{w'}|^2 \rangle dk d\zeta = \left(\frac{2\mu_p}{3T}\right)^2 \delta S_{w'}^2.\end{aligned} \tag{28}$$

Let the number of pixels sampling one line-pair of period T be $N_P = T/\mu_p$, then the noise's variance induced by white noise is

$$\sigma_w^2 = \left(\frac{\pi}{2bP_0}\right)^2 \delta S^2_{w',M_{+1}} = \left(\frac{\pi}{2bP_0}\right)^2 \frac{4}{9N_P^2}\delta S^2_{w'} = \frac{2\pi^2}{9N_P^2 b^2 P_0^2}\delta S^2_w. \tag{29}$$

This equation can be directly applied to read-out noise n_R. With respect to photon-shot noise n_P it is important first to notice the noise between two pixels is uncorrelated. Formally, this is

$$\mathrm{Cov}(n_P(x), n_P(x+\xi)) = \begin{cases} I_C(x), & \xi = 0 \\ 0, & \forall \xi \neq 0 \end{cases} \tag{30}$$

where $Cov()$ is the covariance between to functions and ξ is the distance between any pair of pixels of coordinates x and $x + \xi$. For such a correlation the power spectral density is flat and equal to the average power of the noise

$$\overline{\delta S_P^2} = \frac{1}{N_G T} \int_0^{N_G T} I_C dx, \tag{31}$$

where N_G is the number of line-pairs in the pupil. If the average phase $\bar{\varphi} = 0$ across the pupil, then the above can be approximated to

$$\overline{\delta S_P^2} = b^2 + \frac{1}{\pi^2} P_0^2. \tag{32}$$

Then, if the intensity at the entrance pupil is $I_0 = P_0^2$, the error contribution of the read-out noise and the photon-shot noise is

$$\sigma_R^2 + \sigma_P^2 = \frac{2\pi^2}{9N_P^2 b^2 I_0}\delta S_R^2 + \frac{2\pi^2}{9N_P^2 b^2 I_0}\left(b^2 + \frac{1}{\pi^2} I_0\right). \tag{33}$$

Funding

Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) (72160371); Science and Technology Facilities Council (STFC) (ST/L002213/1,ST/L00075X/1,ST/P000541/1).