Pyramid versus Shack-Hartmann: Trade Study Results for the NFIRAOS NGS WFS

Jean-Pierre Véran[a], Simone Esposito[b], Paolo Spanò[b], Glen Herriot[a], David Andersen[a]
[a]National Research Council Canada, 5071 W. Saanich Rd., Victoria, BC Canada V9E 2E7
[b]Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, Firenze, Italy

ABSTRACT

NFIRAOS, the first light AO system for the Thirty Meter Telescope, will include a natural guide star (NGS) pyramid wave-front sensor (PWFS). This WFS will have two functions: (i) when there is a bright enough NGS within the science field and the lasers are turned off, the PWFS will act as the fast high order WFS driving the SCAO loop (e.g. for high-contrast imaging); and (ii) when the lasers are in use and the system operates in MCAO mode, the PWFS will act as a slow truth WFS (e.g. to measure drifts in the structure of the sodium layer). The decision to select a PWFS instead of a more conventional Shack-Hartmann WFS (SHWFS) is the outcome of a detailed trade study. In this paper, we summarize the results of this trade study. These include extensive simulation work, which shows that, in the expected operating conditions of NFIRAOS, the PWFS will bring significant performance improvements, including higher Strehl ratio, higher limiting magnitude and lower residual speckle levels for high contrast imaging, even when the system has to correct for significant levels of non-common path aberrations. Our simulation results also provides new insights on the properties of the PWFSs. We also report on opto-mechanical design work, which shows that, with the PWFS, the two functions (i) and (ii) can actually be combined into a single optical path, thus reducing the complexity in terms of number of mechanisms and optical elements. Finally, we discuss the impacts of switching to a PWFS on the other already designed NFIRAOS sub-systems (e.g. the real-time computer), which we have found to be very modest.

Keywords: Adaptive optics, wave-front sensing

1. INTRODUCTION

The Narrow Field IR Adaptive Optics System (NFIRAOS) is the first light facility adaptive optics (AO) system for the Thirty Meter Telescope (TMT) [1]. The NFIRAOS project is currently at the final design stage [2].

NFIRAOS is a Multi-Conjugate AO system, which includes:

- two Deformable Mirrors (DMs) conjugated at 0km (DM0 – 3125 actuators) and 11.2km (DM11.2 – 4548 actuators),
- one Tip/Tilt Stage (TTS) serving as the mount for DM0,
- six Laser Guide Star (LGS) Shack-Hartmann wavefront sensors (WFSs) of order 60x60 sub-apertures
- up to three low order Infrared natural guide star wavefront sensors (OIWFS) within each NFIRAOS instrument,
- one high order visible Natural Guide Star Shack-Hartmann WFS (NGS WFS) of order at least 60x60, which is used for operation without LGS,
- one Truth Wavefront Sensor (TWFS) measuring a natural guide star at low bandwidth, which is used to calibrate for slow-varying biases due to temporal variations in the sodium layer profile in LGS AO mode,
- and a Real-Time Controller (RTC), which processes the inputs from the various WFSs to compute the commands of the deformable mirrors and tip/tilt stage. All the parameters required by the RTC are provided by the Reconstructor Parameter Generator (RPG), which is part of the AO Sequencer. These parameters are optimized and updated as conditions change.

This architecture is summarized in **Figure 1** and **Figure 2**. The science light is going through two AO relays (one for DM11 and one for DM0+TTS), through a science beamsplitter (BS), and off one last off-axis parabola (OAP) and a fold mirror before entering the science instrument, which also contains the OIWFSs.

In LGS mode (**Figure 1**) the laser light follows the same path as the science light up to the science beamsplitter, where it is reflected into the WFS path. The WFS path has another beamsplitter, which reflects the laser light into the LGS WFSs

via a trombone to adjust focus. The visible, non-laser light is transmitted into the truth WFS, via a star selection mechanism. The TWFS measures radial modes at low speed to track changes in the profile of the sodium layer. It is a ~12x12 WFS. It can also track drifts in NCPA compensation, a task for which a higher order of sampling might be desirable.

In NGS mode (**Figure 2**), the science path is the same as in LGS mode, but the LGS WFSs are not used. The visible light is directed into a high-order NGS WFS.
The RTC interfaces with additional telescope and AO sub-systems, including the AO Sequencer, the NFIRAOS Component Controller, the Laser Guide Star Facility System, the NFIRAOS instruments, and the Data Management System (DMS).

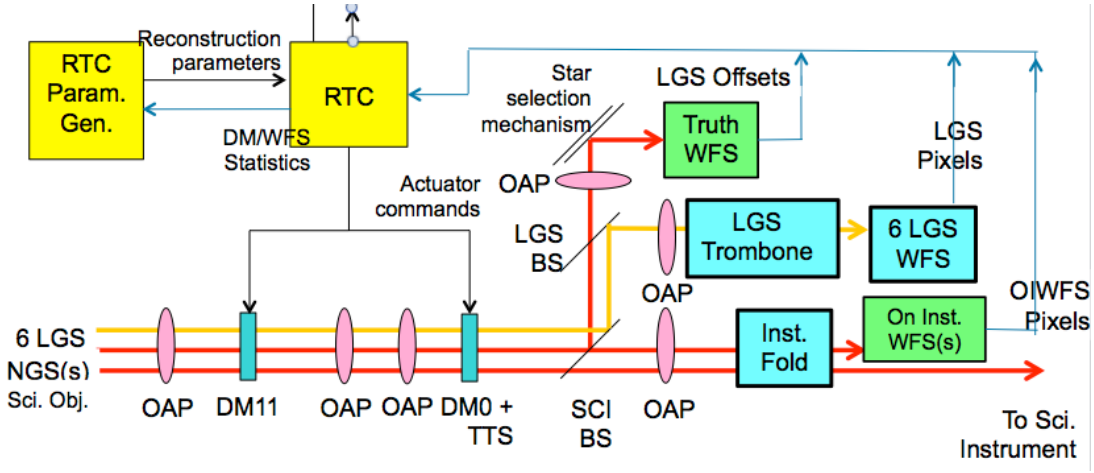

Figure 1: NFIRAOS Simplified Control Diagram
in LGS Mode

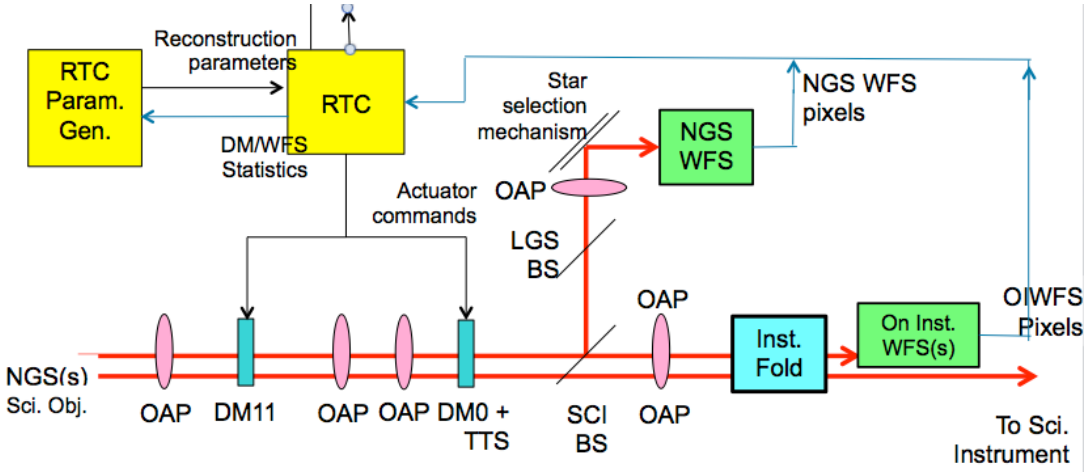

Figure 2 NFIRAOS Simplified Control Diagram
in LGS Mode

In the baseline design, the TWFS and the high order NGS WFS are both Shack-Hartmann WFSs (SHWFSs). They share the same detector, but have two different optical paths, one through a 12x12 lenslet array for the TWFS, one through a 60x60 lenslet array for the high-order NGS WFS. These two SHWFSs could be replaced by a single Pyramid WFS (PWFS), because the sensing order of the PWFS can be adjusted simply by electronically binning pixels onto the detector.

In this paper, we summarize the main results of the trade-study that was carried out, and that led to the decision to change the NFIRAOS baseline design and adopt the PWFS option. We begin, in section 2, by making the case for a PWFS. In section 3, we present the simulation results comparing the expected performance of the SHWFS and the PWFS in the specific case of NFIRAOS, including comparisons in terms of sky coverage and high-contrast imaging. In these simulations, we include a realistic level of non-common path aberrations (NCPAs) between the WFS and the science instrument. Including NCPAs is critical, because so far PWFSs have been used in systems based on adaptive secondary mirrors, where the level of NCPAs is very small. In section 4, we present the opto-mechanical design of the PWFS, and compare it with the baseline design with two SHWFSs, in terms of complexity and cost. Finally, in section 5, we summarize our findings and conclude on the superiority of the PWFS for NFIRAOS.

2. CASE FOR THE PWFS

The Pyramid WFS (PWFS) was originally introduced in 1996 by Ragazzoni, as a generalization of the Foucault optical test [3]. It consists of a pyramid in the focal plane, splitting light into four quadrants. Light from each quadrant is relayed by a lens to a detector in the pupil plane to form four pupil images. These images can be processed to reconstruct the incoming wave-front, at a resolution set by the sampling of the pupil images. The reconstruction process starts by associating pixels at the same location on the four pupil images to form a "virtual" quad-cell. The process is then similar to that of a quad-cell SHWFS, including computing an X and Y signal for each "quad-cell", from which the wave-front can be estimated via a linear operation.

The advantage of the PWFS compared to the SHWFS lies in its potentially much increased sensitivity. In a SHWFS, the wave-front measurement is obtained from the estimation of the spot position in each sub-aperture. The SNR on this estimation is set by the size of the spot, which cannot be smaller than the diffraction limit of a sub-aperture. For NFIRAOS with 60 sub-apertures across the telescope diameter, the minimum spot size is thus $60\lambda/D$, where λ is the WFS wavelength and D=30m. In contrast, the SNR in a PWFS is set by the quality of the spot at the tip of the pyramid. Since there are no sub-apertures, the spot can potentially be as small as λ/D, which is the diffraction limit of the full telescope aperture. Initially of course, the spot is limited by seeing, but as the AO correction kicks in, a spot size of λ/D can potentially be achieved, therefore significantly increasing the sensitivity compared to the SHWFS.

In the conditions described above however, the measurement process is linear only for wave-front of very small amplitude. This can be understood intuitively by realizing that even a very small tilt error of λ/D, would kick most of the light on one or two sides of the pyramid, making two or three of the four pupil images dark. This condition in effect saturates the PWFS, which cannot measure anything except for the direction of the tilt error. In order to preserve linearity of the measurement process on a larger dynamic range, the position of the spot must be modulated, typically in a circle around the tip of the pyramid, during the integration on the detector. This modulation is generally accomplished by a fast tip-tilt mirror located upstream from the pyramid. The amplitude of the modulation is typically $5\lambda/D$ in radius. It can be decreased (e.g. to $3\lambda/D$) if the seeing and/or the AO correction is good (because the amplitude of the wave-front error to be measured is reduced), but must be increased if the seeing and/or the AO correction goes bad (because the amplitude of the wave-front error to be measured becomes larger). The amplitude of the modulation sets an "equivalent" spot size, and this is what drives the real sensitivity of the PWFS. It follows that when larger modulations are required, the gain in sensitivity achieved by the PWFS is reduced.

The sensitivity of the PWFS compared to that of the SHWFS has been analytically studied by Vérinaud in reference [3]. The main results, which are shown in Figure 3, quantify our qualitative description above. We can see three different regimes:

- For low spatial frequencies, the sensitivity of the PWFS increases much faster than that of the SHWFS until a plateau is reached. The smallest the modulation, the fastest the increase, with the height of the plateau remaining the same.
- For medium spatial frequencies, the sensitivity of the PWFS remains constant and larger than that of the SHWFS.
- For high spatial frequencies, the PWFS sensitivity remains constant, while that of the SHWFS keeps increasing. The two sensors have the same sensitivity at the Nyquist frequency.

The linear increase in sensitivity as a function of spatial frequency is the characteristic of a slope (gradient) sensor whereas a constant sensitivity is characteristic of a phase sensor. It follows that the SHWFS is a slope sensor at all frequencies (as we expect), whereas the PWFS is a slope sensor at low spatial frequency but a phase sensor at high

spatial frequency. The amplitude of the modulation sets the cut-off between the two regimes. Because of this property, the PWFS is often referred to as "dual" WFS.

Increased sensitivity directly translates into lower susceptibility to noise. This is why for a given guide star magnitude and given number of degrees of correction, a PWFS-based AO system will normally provide a higher level of correction than a SHWFS-based system. Conversely, a PWFS-based AO system will provide the same level of correction as a SHWFS-based AO system on a fainter star, and therefore will have larger sky coverage.

Another advantage of the PWFS is its increased robustness to aliasing. Measurements of wave-front errors beyond the Nyquist frequency are aliased and are wrongly reconstructed as low order aberrations. This causes aliasing errors. On bright guide stars where the WFS measurement SNR is high, aliasing can become the dominant error after AO correction. The PWFS is less sensitive than the SHWFS to aberrations beyond the Nyquist frequency and therefore has lower aliasing error.

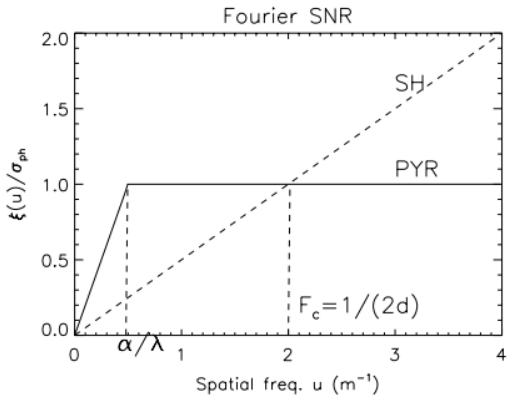

Figure 3: Analytical comparison of the sensitivity of PWFS and that of a SHWFS as a function of the spatial frequency of the wave-front error to be measured. d is the diameter of the SHWFS sub-aperture, Fc is the Nyquist frequency, and α is the amplitude of the modulation.

In section 3.2, we discuss our estimates of the sensitivities of the PWFS and SHWFS, based on our wave-optics simulation. While we were able to broadly confirm the analytical results by Vérinaud, our results also exhibits some interesting differences.

The PWFS has attracted worldwide attention within the AO and astronomical communities after the publication of the results obtained by the FLAO systems, which equip the LBT [5]. These systems achieve unprecedented levels of performance, which is attributed to their use of a PWFS. These results include very high Strehl on bright guide stars (> 80% at H-band in 0.7 arcsec seeing for mangitudes brighter than 10) and high sky coverage (H-band Strehl of 2-10% achieved on 17 magnitude stars, depending on seeing).

3. PERFORMANCE EVALUATION

3.1 Simulation software and parameters

We use the Passata AO simulation software from the Arcteri AO group. At the core of this software is a wave-optics PWFS module that computes a PWFS image from a wave-front map. This computation includes the effect of modulation, by averaging images obtained for a number of positions of the modulating mirror. Photon noise and detector noise can be added to degrade the measurements.

The PWFS module is wrapped up into a conventional closed-loop AO simulation layout, which includes a translating turbulence phase screen, applying DM correction with a lag, measuring the residual wave-front with the PWFS module, applying a control matrix, updating the DM commands via an integrator with an adjustable gain, and computing the AO corrected PSF from the residual wave-front error map.

The AO control is based on modal control. A number of control modes are defined, and the DM is assumed to be able to apply any linear combination of these modes. The interaction matrix is computed by presenting each control mode to the PWFS and recording the corresponding measurement. The control matrix is computed as the pseudo-inverse (least-square reconstructor) of the interaction matrix (the control modes are chosen to exclude any invisible mode).

The Passata software is configured for the NFIRAOS case. The PWFS provides pupil images with exactly 60 pixels in diameter, which would correspond to a Fried configuration with NFIRAOS' 61x61 DM0. The full 30-meter pupil of the TMT is represented on a grid with 600x600 points, which corresponds to a sampling of 1/20 m/pixel. The pupil model is an annulus with the TMT central obscuration, but not the M2 spider, nor any other obstruction. The control modes consist of the 2210 first Karhunen–Loève (KL) modes for a Kolmogorov turbulence. The lowest KL modes are similar to the lowest Zernike polynomials. For higher order modes, the KL modes are more similar to Fourier modes. KL modes can be numbered by radial and azimuthal degrees, like the Zernike polynomials. The radial degree can be loosely related to spatial frequency by considering that radial degree 3 corresponds to a radial frequency of 1/D, and that each sub-sequent order n corresponds to spatial frequency 1/(n-2)D. With this association, the 2210 KL modes span radial orders from 1 (tip and tilt) to 65, and the cut-off frequency of the NFIRAOS DM 1/0.5 m^{-1} is reached at spatial order 63. This justifies that choosing to control 2210 modes to simulate NFIRAOS is reasonable.

To acquire a good interaction matrix, cares must be taken to present modes with proper amplitude. This is especially true for the PWFS, because of its limited dynamic range. We have used an amplitude of A/sqrt(RO) where RO is the radial order of the mode and A=72nm (0.6 rad @ 0.75um) for a 5λ/D modulation and half that for a 3λ/D modulation. We have found that A=72nm was also suitable for the SHWFS. A suitable magnitude is such that the WFS signal is below 0.3, to ensure linearity.

In our simulation, we have used a single turbulence screen, with a Von Karman spectrum with r0=0.186cm and L0=30m. This corresponds to the MK median conditions at zenith. The screen translates at a speed of 8.6m/s, which corresponds to the median MK wind-speed. The screen is generated with the "wave" function from Laurent Jolissaint's PAOLA software on a 4096x4096 grid (on that grid, the TMT pupil is represented by an annulus of outer diameter 600 pixels). It is then rescaled to simulate a zenith angle of 30 degrees. The WFS wavelength is always set to 0.75um.

We have also added several modules to Passata in order to enable the trade-study for NFIRAOS:

- A Shack-Hartmann module with the same interface with the rest of the code as the already existing PWFS module. From a wavefront map, the SHWFS module uses physical optics to compute the spot images in each of the 60x60 sub-apertures. The image is sampled in a quad-cell, with 0.56" pixels.
- Non-common path aberrations (NCPAs): An NCPA map is created by adding 150nm RMS of astigmatism to 70nm RMS of random wave-front error that follows a f^{-2} 2D power spectrum (f here is the spatial frequency). The former corresponds to the expected residual alignment errors and the latter corresponds to the expected polishing errors. Correcting NCPAs requires operating the WFS off-null. So we have modified Passata to allow the user to enter offsets. Offsets can be included either as slope offsets, or (by running the slopes offsets through the control matrix) as wave-front offsets. The latter implementation is chosen in our simulations.
- Pseudo-open loop implementation and minimum variance reconstructor: The Passata software implements a traditional closed loop (CL). In order to support a minimum-variance (MV) reconstructor, we have modified the code to implement pseudo-open loop (POL) AO operation. The original Passata only implements a least-square reconstructor. The wave-front reconstruction matrix R is obtained from the interaction matrix Γ using a pseudo-inverse: $R_{LS} = (\Gamma^T\Gamma)^{-1}\Gamma^T$. The minimum variance reconstructor is implemented as: $R_{MV} = C_a\Gamma^T(\Gamma C_a\Gamma^T+C_n)^{-1}= [\Gamma^T(C_n)^{-1}\Gamma+(C_a)^{-1}]^{-1}\ \Gamma^T(C_n)^{-1}$, where C_a is the covariance of the modes in the atmospheric turbulence and C_n is the covariance matrix of the WFS measurement noise (units of WFS signal square). Since these are KL-modes, C_a is diagonal. The diagonal elements are the variance of the modes, which can be computed as explained in section 3.2. C_n is a diagonal matrix where each element is equal to $(N+4*\sigma^2_e)/N^2$, where N is the average number of photons per sub-aperture per frame and σ^2_e is the variance of the detector read-out noise for each pixel.
- Optical gain estimation: In a quad-cell SHWFS, the WFS signal in each sub-aperture is related to the local wave-front gradient via a centroid gain, which depends on the spot size, which in turn depends on seeing (since the wave-front within a sub-aperture does not undergo any wave-front correction other than global T/T) and other static errors of scales smaller than that of a sub-aperture. Incorrect estimate of the optical gain will result

in loop gain errors, potentially leading to loop instabilities, and in errors in NCPAs corrections, leading to artifacts in the science image. In the NFIRAOS original SHWFS design the centroid gain is tracked by injecting a dither tilt signal d(t) in the WFS path. The dither signal is a pure sinusoid at a known frequency. The signature of this dither signal will appear in the WFS measurements. The amplitude of the measured dither signal m(t) can be estimated by synchronous detection, and compared to the amplitude of the injected dither signal d(t). An error in centroid gain will cause the ratio |d(t)|/|m(t)| to deviate from 1. So the error signal 1-|d(t)|/|m(t)| is fed into an integrator with an adjustable gain to update the current estimate of the centroid gain used to process the WFS measurements. In NFIRAOS, the dither signal can be injected via a tip-tilt mirror (TTM) in front of the WFS. However, because of the finite bandwidth of the TTM, the actual amplitude of the dither signal differs from that of the synthetic sinusoid used to drive the TTM and must be estimated. This is performed using the output signals from the TTM internal sensors, again using synchronous detection. At the same time, the dither signal must be removed from the WFS measurements, so that it does not propagate as noise in the AO loop. In order to do that the phase of the dither signal in the WFS measurements must be estimated as well. This is done through a delay locked loop. For the PWFS, a similar problem exists, in that for a given wave-front to be measured, the WFS measurement depends on the quality of the spot focused on the pyramid: the highest signal is obtained for a diffraction limited spot, and the signal decreases as the quality of the spot decreases. Previous studies by the Arcetri group have shown that this effect can be captured by a global gain factor called optical gain (OG), which is similar to the centroid gain for the SHWFS. As discussed in the introduction, one key element of our study is to evaluate whether this model holds in the case of NFIRAOS where the spot will not be as well corrected as with LBT (because of the lowest density of DM actuators), and whether we can develop a technique to track the OG as conditions change. This development, including its on-sky validation on LBT, is the subject of a separate report in reference [6].

3.2 Fitting error, aliasing and noise propagation

Fitting error:

For the 60x60 NFIRAOS system, the Passata software uses 2210 Karhunen–Loève (KL). These modes are ortho-normal over the NFIRAOS pupil. Note that this pupil is simply a disk with the proper central obscuration, not the real TMT pupil with segment gaps and spiders. Since the KL modes are ortho-normal, phase screens can be projected onto modes using simple dot products. The fitting error and the variance of the modes in the atmospheric turbulence can then be easily estimated by generating random phase screen with proper statistics.

We used 500 independent screens to compute the variance of each mode, and the PSF when all the modes are perfectly corrected (fitting error PSF). For the KL mode basis, the results are shown in Figure 4. As expected, the variance decreases as mode number increases, and correcting all the modes result in a “dark hole”.

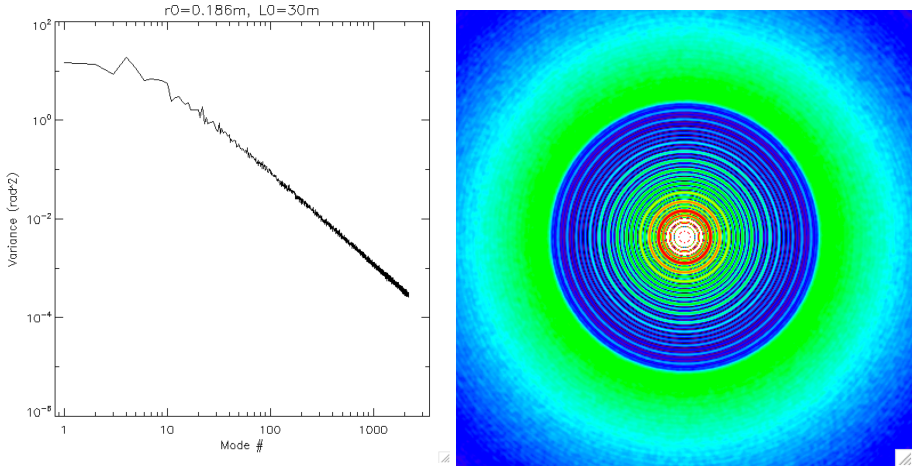

Figure 4: Left: variance of the ON KL modes in an atmospheric turbulence with r0=0.186 and L0=30m. Right: Long exposure PSF when only fitting errors are present.

With the KL-basis, the “dark hole” is a disk. This is because the KL-modes are derived from a circular geometry, without any consideration of the actuator geometry on the DM. In order to verify that the dark hole is indeed square

when the square geometry of the DM actuators is taken into account, we have created an ortho-normal basis that exactly spans the range of the NFIRAOS DM, and where the first 2210 modes are as close as possible to the original KL-basis. This basis is referred to as "DM KL-basis". This basis is created by using a true DM influence function (in our case a bi-cubic spline with a 24% coupling between neighbouring actuators) and deriving an ortho-normal basis that exactly spans the range generated by DM0 (the number of valid actuators for DM0 used for these simulations is 2872). The process is carried out in two steps. The first step uses the Gram-Schmidt method to ortho-normalize the projection of the original 2210 KL modes onto DM0, so that the 2210 first basis modes are as close as possible to the KL modes. The second step is to find the 662 remaining mode and is accomplished by covariance matrix diagonalization. The fitting error is then recomputed with this modal basis and the results are shown in Figure 5.

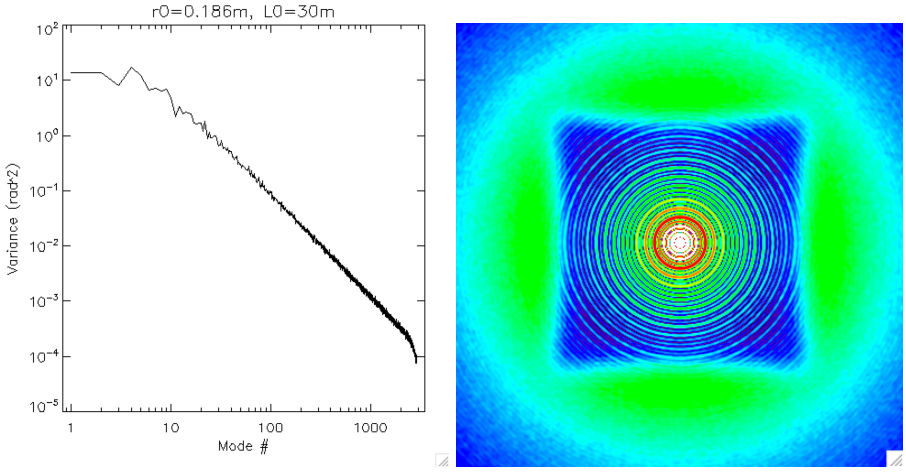

Figure 5: Left: variance of the DM KL-modes in an atmospheric turbulence with r0=0.186 and L0=30m. Right: Long exposure PSF when only fitting errors are present.

We see that the DM KL-basis is able to capture the square geometry of the DM, and produce the "dark square", characteristic of a high contrast image, where the boundaries of the square correspond to the highest spatial frequencies correctable with the DM. The original KL-modes produce a circular PSF, because they do not capture the square DM geometry. For the DM KL-basis, the fitting error is 97.1nm RMS in a turbulence with L_0=30m and r_0=0.186m at zenith (r_0=0.171m at ZD=30 degrees). This corresponds to a phase variance of $0.249*(d/r_0)^{5/3}$, where d=0.5m is the DM actuator pitch projected onto M1. This fitting error is lower than what we can expect from a real AO system. This is because our simulation uses a sampling of 600 pixels for 30m, which corresponds to a sampling of 1/20m/pixel. In 2006 for the NFIRAOS conceptual review, we studied the effect of simulation under-sampling. Based on this study, we estimate that an additional 40nm RMS must be added in quadrature to estimate the true fitting error. This leads to a real fitting error of 105nm RMS, or $0.291*(d/r_0)^{5/3}$. For the original KL-basis, the fitting error is higher: 101.9nm RMS.

Noise propagation:

In the linear regime, the noise propagation on each mode *i* can be directly evaluated from the reconstruction matrix R: $\sigma_i^2=[R\#R^T]_{ii}$. If the modes are orthonormal, such as in the KL-basis, the total noise propagation can be computed as $\sigma_p^2=\Sigma\ \sigma_i^2$, the sum of all the components of n_p. If the reconstruction matrix is computed from an interaction matrix obtained in diffraction limited conditions, then σ_p^2 must be divided by the square of the optical gain OG for the current conditions. The total noise propagation in the AO loop then can be obtained by: $\sigma^2_{wfe}= \sigma_p^2/OG^2 * \sigma^2_{subap}$ where σ^2_{subap} is the measurement error due to noise in each sub-aperture. For both a quad-cell SHWFS and a PWFS with a classical CCD detector, it is given by: $\sigma^2_{subap}=(N+4*\sigma^2_e)/N^2$ where N is the average number of photons detected in a sub-aperture and σ^2_e is the variance of the detector read-out noise for each pixel.

Figure 6 shows the noise propagation coefficient for each mode, for an optical gain equal to 1. The right hand figure plots the average coefficient for each radial order. As expected from theory, the noise propagation for the SHWFS decreases as the inverse square of the radial order. As expected, the noise propagation is lower with the PWFS than with the SHWFS, with the lowest modulation achieving the lowest noise propagation.

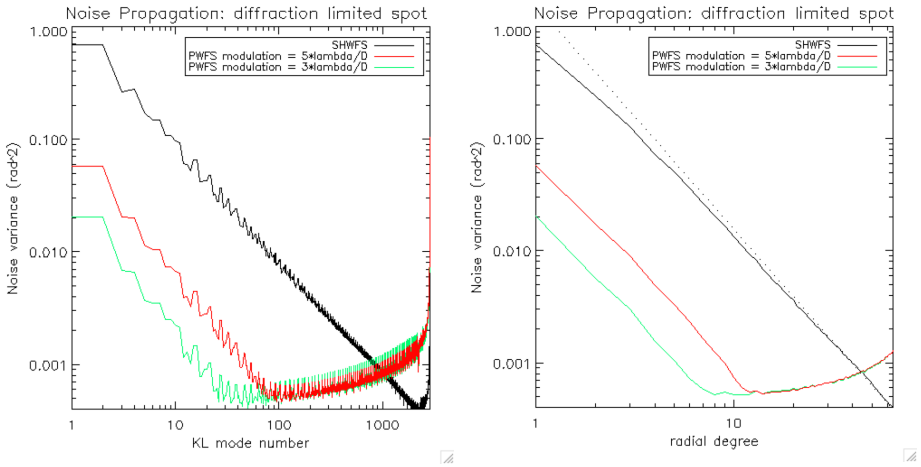

Figure 6: Modal noise propagation coefficient for the SHWFS (black line) and for the PWFS (red line with a 5λ/D modulation and green line for a 3λ/D modulation). On the graph on the right, the average noise propagation per radial degree is plotted. For the SHWFS, the dotted line is a power -2 law

For both the SHWFS and the PWFS, the noise propagation is high for modes higher than the original 2210 KL-modes. Since the variance of these modes in the turbulence is low anyway, these modes are better left uncontrolled, unless the noise is extremely low.

The total noise propagation coefficient σ_p^2 (for a sub-aperture noise variance of 1) is presented in Table 1. Rescaling is obtained by dividing the noise propagation in diffraction limited conditions by the square of the optical gain, which is obtained from the results in Table 4 below. We see that even with the rescaling, the PWFS remains advantageous in terms of noise propagation

	Noise propagation (rad^2) (diffraction)	**Optical gain**	**Noise propagation (rad^2) (rescaled)**
SHWFS	7.1	0.74	13.0
PWFS (3 λ /D)	2.0	0.5	8.0
PWFS (5 λ /D)	2.2	0.5	8.8

Table 1: Total noise propagation coefficient in diffraction limited conditions and total noise propagation rescaled by the square of the optical gain, for the first 2210 modes

The noise propagation is directly related of the sensitivity of the sensor. In the work by Verinaud [4], which we have already referred to in section 2, the sensitivity is defined as the norm of the measurement per photon when one unit of mode is presented. In Figure 7, we show the sensitivities derived from the interaction matrices (left) and compare them with that from the work of Verinaud (right). In our case, the cut-off frequency is Fc=30/D. As discussed before, the radial degrees can be approximately related to spatial frequency by considering that radial degree 3 corresponds to a radial frequency of 1/D and radial degree 63 corresponds to a radial frequency of Fc. As expected, we find that the sensitivity of the SHWFS increases linearly with spatial frequency. For the PWFS, we expect from Verinaud's work that the maximum sensitivity is reached at a radial degree 8 for a 3λ/D modulation and a radial degree 13 for a 5λ/D modulation, which is indeed the case. However, we see that after the sensitivity peaks, it starts decreasing instead of remaining constant as predicted by Verinaud. The reason for this drop in sensitivity remains to be studied. However, we believe that the reason lies in the diffraction due to the edge of the pyramid, which blurs the high spatial frequencies and that is neglected in Verinaud's calculations. Finally, we note that this drop in sensitivity helps with aliasing.

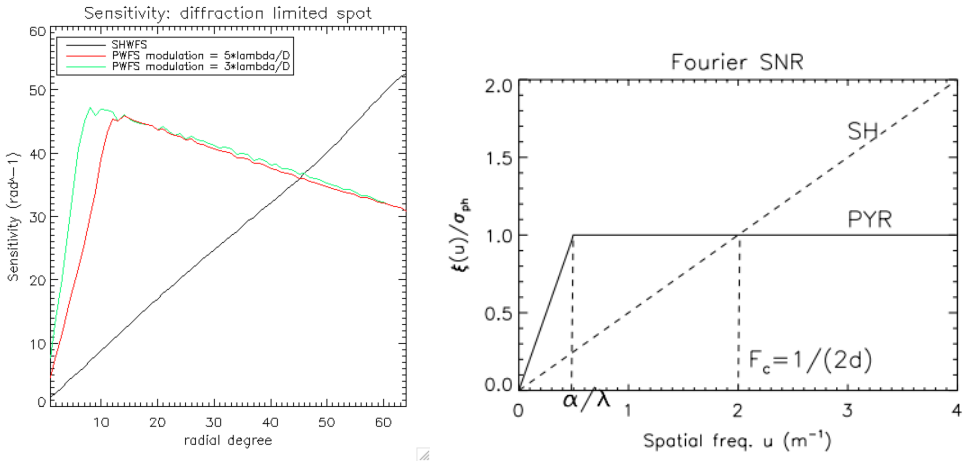

Figure 7: Left: sensitivity as a function of the radial degree, as computed from the interaction matrices. Right: sensitivity as a function of spatial frequency found by Vérinaud [4].

For the low order modes (order 1 or 2), the sensitivity of the PWFS is ~4 times higher than that of the SHWFS. This gain is comparable to the difference between the spot size of a SHWFS (60λ/D) and the "equivalent" spot size of a PWFS, which is obtained by applying a modulation of radius 5λ/D to a spot of size λ/D. As expected, the sensitivity increases when the modulation is reduced.

Robustness to aliasing

Besides having lower noise propagation, the PWFS is expected to be more robust to aliasing. This is because the sensitivity of the SHWFS keeps increasing beyond the system cut-off frequency Fc, whereas the sensitivity of the PWFS is expected to remain constant. We see this behavior in Figure 7, which suggests that the actual sensitivity of the PWFS is even decreasing at high spatial frequencies. Since spatial aliasing arises from spatial frequencies beyond the system cut-off frequency Fc, its effect can be evaluated using the same sort of Monte-Carlo simulation as for the fitting error, except that the un-correctable wave-front error maps are measured by the WFS and reconstructed. We also used 500 independent realizations and found that the aliasing errors shown in Table 2. We see that the aliasing for the PWFS is significantly lower than that of the SHWFS. We expect that if we take into account the under-modelling error of our simulation (due to under-sampling the phase screens, see the paragraph on fitting error above), the difference will be even higher, since the sensitivity to the highest order modes not included in our simulation will be higher for the SHWFS.

	SHWFS	PWFS 5λ/D mod	PWFS 3λ/D mod
Aliasing error (nm RMS)	64.5	22.1	23.0
Aliasing error / Fitting error	63%	22%	23%

Table 2: Aliasing error for different WFS. r0=0.186m and L0=30m

For each mode the variance of the aliasing error is shown in Figure 8. It is interesting to note that for all WFSs, tip and tilt carry most of the variance.

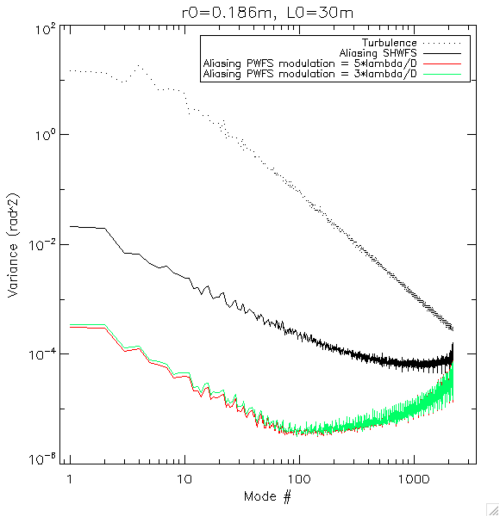

Figure 8: Variance for each of the ON KL-modes. Dotted line: in the uncorrected atmospheric turbulence; black for the SHWFS; red and green for the PWFS for different modulations.

3.3 End-to-end simulation results on bright guide stars

Our simulation work has focused first on bright star simulations (up to magnitude 12). The main simulation parameters are summarized in Table 3:

Atmospheric turbulence	One phase screen r0=0.186m x ZD=30 degrees L0=30m, translating at V=8.6m/s
Telescope	Disk D=30m, 10% central obscuration. No other obscuration for telescope
AO loop	Sampling frequency = 800Hz Servo-lag = 2 frames Controlled modes: 2210 original KL-modes Integrator gain = 0.5
All WFSs	Sensing wavelength = 750nm Flux = 2.06629e+06 photo-detection per frame for a magnitude 8 stars CCD Read-out noise = 2.11e- @ 800Hz per pixel per frame Photon noise
PWFS	Modulation: radius=lambda/5 simulated in 32 steps Pupil image diameter = 60 pixels Distance between adjacent pupil images = 72 pixels Dithering signal: synthetic 80Hz sinusoid with amplitude 6nm RMS of astigmatism. OG tracking loop gain = 0.02
SHWFS	Number of sub-apertures = 60 in diameter Configuration = 2x2 quad-cell Pixel size = 0.56 arcsec No dithering
NCPAs	150nm RMS of astigmatism + 70nm RMS @ f^{-2}
Imaging	λ=1.6um (H-band), x4 oversampling from Nyquist Exposure time = 2s

Table 3: Main simulation parameters

Note that for the SHWFS, we did not implement dithering. This is because for the SHWFS, dithering in the common path would require too high an amplitude. Instead we adjusted the optical gain by ensuring that the DC DM component had exactly the right amount of NCPA correction. This makes the SHWFS results somewhat optimistic. The PWFS results include the propagated dithering signal, which is only 9.5nm RMS.

Residual wave-front errors

Least-square reconstructor

The residual wavefront errors and optical gain tracking results using the least-square reconstructor are presented in Table 4. The RMS is the average total wave-front RMS, whereas RMS TT is the average RMS of the tip and tilt modes only. All RMS values are in nm. NCPA Val is the average value of the NCPA component in the DM shape (should be 1 for perfect NCPA correction), and gammaOpt is the average optical gains. Averages are taken over the 2 seconds (1600 iterations) exposure, which starts 0.5s (400 iterations) after the loop is closed.

	PWFS	SHWFS
No Noise	RMS = 117.3 115.7 114.4 RMS TT = 13.0 10.5 11.8 NCPA Val = 1.036 1.045 1.020 gammaOpt = 0.489 0.499 0.498	RMS = 138.2 135.0 132.9 RMS TT = 31.6 27.1 17.6 NCPA Val = 1.004 0.998 1.001 gammaOpt = 0.736 0.753 0.739
Mag 8	RMS = 117.5 115.9 114.6 RMS TT = 13.2 11.1 12.5 NCPA Val = 1.035 1.046 1.016 gammaOpt = 0.488 0.498 0.496	RMS = 138.5 135.4 133.3 RMS TT = 31.7 27.5 18.3 NCPA Val = 1.004 0.998 1.001 gammaOpt = 0.735 0.753 0.739
Mag 10	RMS = 119.0 117.1 115.7 RMS TT = 14.2 10.6 11.0 NCPA Val = 1.040 1.042 1.013 gammaOpt = 0.487 0.495 0.491	RMS = 140.3 137.7 135.3 RMS TT = 32.0 31.1 22.0 NCPA Val = 1.003 0.997 1.001 gammaOpt = 0.740 0.760 0.744
Mag 12	RMS = 145.4 143.8 140.2 RMS TT = 47.8 49.5 40.9 NCPA Val = 1.047 0.999 1.004 gammaOpt = 0.442 0.431 0.438	RMS = 163.6 160.8 159.8 RMS TT = 41.0 43.4 34.4 NCPA Val = 1.003 0.999 1.001 gammaOpt = 0.761 0.783 0.763

Table 4: Results of end-to-end simulation for SHWFS and PWFS for difference guide star magnitudes. gammaOpt is the average optical gain over the last 2 seconds of the simulation

- As expected, the RMS always increases when magnitude increases, but noise effects only start to kick in above magnitude 10.
- The optical gain tracking loop is always stable, with NCPA values always close to 1. The optical gain for PWFS is lower than for SHWFS. This is consistent with the SHWFS spot being seeing limited, and therefore less affected by residual wave-front errors.
- For the SHWFS, gammaOpt is roughly constant regardless of magnitude, whereas for PWFS, gammaOpt decreases by 10% from Mag 10 to Mag 12. This is consistent with spot quality degrading on the pyramid due to higher residual wave-front RMS.
- For bright NGSs, PWFS has a lower residual T/T, consistent with lower aliasing as noise is negligible. For Mag 12 however, RMS TT for PWFS is even higher than for SHWFS. This is not consistent with noise propagation coefficient. We believe that this higher level of residual TT is due to increased non-linearities due to increased wave-front errors, converting into TT errors. It might be possible to improve the results by reducing the number of controlled modes.
- In the "No-noise" case, we would expect the performance of the PWFS and of the SHWFS to differ only by the difference in aliasing error reported in Table 2. The difference is actually somewhat larger (70nm RMS instead of 61nm RMS), suggesting perhaps some non-linear behavior for the high order modes in the SHWFS.

Minimum variance reconstructor

We have run simulations with the minimum variance reconstructor (see section 3.1) to see if it can make a difference, especially for the SHWFS. The simulation was run for a magnitude 10 stars: N=115.8 photo-detections per sub-aperture and per frame @ 800Hz), RON=2.11e- at 800Hz.

The results are shown in Table 5. The results are in line with that reported in Table 4. They show that using the minimum variance reconstructor does not significantly help the SHWFS.

	Total WFE RMS (nm)	Tip-tilt WFE only (nm)

PWFS Least-square	114	12
SHWFS Least square	138	27
SHWFS minimum variance	137	27

Table 5: Simulation results for a magnitude 10 star using different WFS and different reconstructors

Note in the PWFS results the "bump" at radial degree 2, which corresponds to the astigmatism we inject in the dithering process.

PSF structure

The SHWFS and PWFS PSFs for a 2s second exposure are compared in Figure 9. This is for magnitude 8. Only the 1.1"x1.1" central region of the PSF is presented, which corresponds to the controlled area, as for Figure 4. Compared to Figure 4, we see that the PSF area within the controlled region contains residual speckles, and that the SHWFS image contains significantly more speckles than the PWFS. These speckles are called by high-contrast imaging specialists "pinned" speckles. They are pinned to the diffraction rings and evolve very slowly: they only change when the current atmospheric turbulence pattern clears the telescope aperture, which typically takes 5-6 seconds. These quasi-static speckles mask potential companions and are the most difficult to subtract: faster speckles average out, whereas static speckles can be subtracted out, e.g. by angular differential imaging (which is why we have not included the uncorrected static contributions from the telescope, such as segment gaps and spider and from NFIRAOS optical elements).

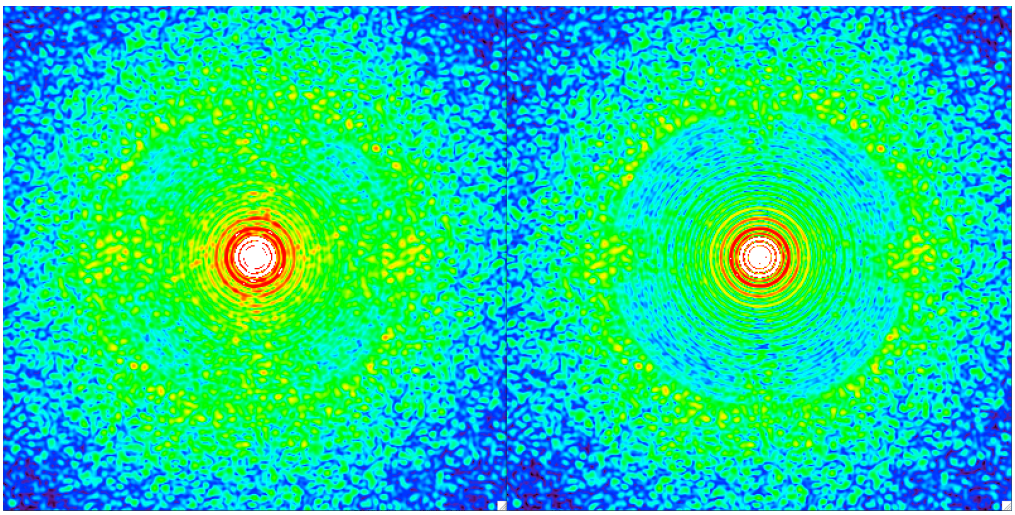

Figure 9: Central 1.1x1.1 arcsec region of a 2s-exposure PSF for a magnitude 8 guide-star, for a SHWFS (left) and for a PWFS (right). The PSF are presented in log scale, normalized to their maximum intensity and with a 10^{-3} upper threshold applied.

High contrast imaging

In this section, we try to characterize the AO corrected PSFs using metrics that are relevant to the field of high-contrast imaging

Contrast profile

Figure 10 shows the radial profile of the PSF contrast. The PSFs are in H band and are obtained for two cases: NGSs of magnitude 8 and of magnitude 12. The curves in green and white are for PWFS and SHWFS respectively. The contrast values are calculated as averages over rings of size 10 pixels, with respect to a maximum value of 1. Considering that the PSF is oversampled by a factor 4 the contrast is computed for a focal plane patch of 2.5λ/D.

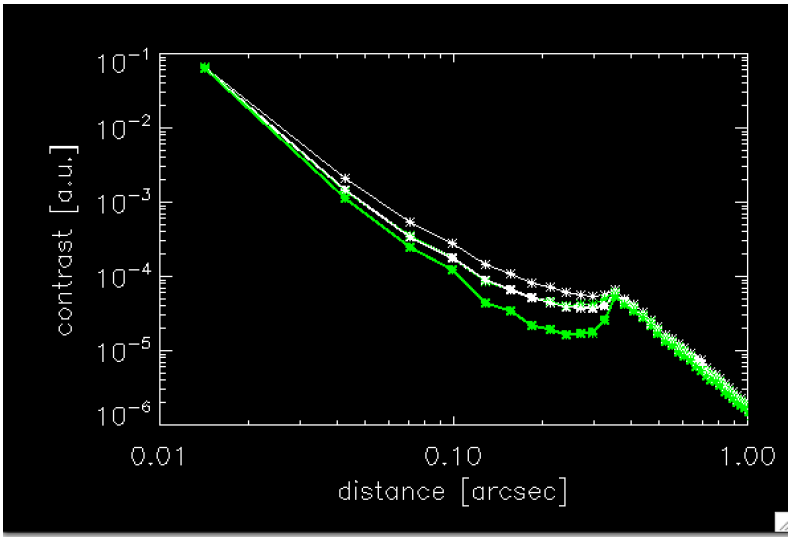

Figure 10: PSF contrast radial profile for mag 8 and for mag 12. The no noise case gives the same results of mag 8. Green and white curves refer to PWFS and SHWFS respectively.

We see that the gain in contrast for PWFS compared to the SHWFS is a factor 2-3 in the range 0.15" to 0.3"

Speckle subtraction using ADI

In this section, we attempt to evaluate the noise due to the changes in the speckle pattern in the AO corrected PSF when using Angular Differential Imaging (ADI). The ADI technique is based on computing differences between images taken as the science field rotates. Therefore, we have computed the difference of two uncorrelated AO corrected images (normalized to a maximum of 1), computed the speckle noise level as the standard deviation on rings 2.5λ/D-wide, and averaged over three image pairs. The result, as a function of the center distance, is plotted in Figure 11 below. In particular we see that the flux variation due to speckle noise is 2-3 times lower in the range 0.1"-0.3" for the PWFS.

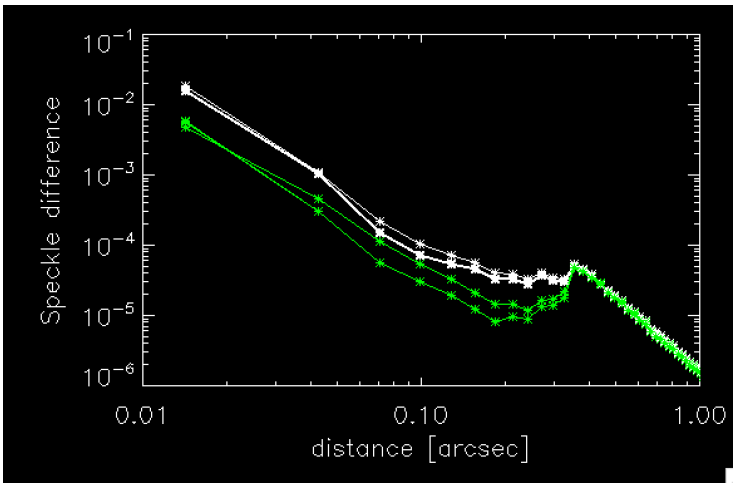

Figure 11: Average difference of the speckle pattern flux in two different AO corrected images, normalized to a maximum value of 1.

Again as above green and white curves are for PWFS and SHWFS, while the two curves of the same color represents reference stars of mag 8 and 12.

SNR radial profile estimate

Using the PSF Contrast Profile from Figure 10 and the Difference Profile from Figure 11, we can try to estimate the Signal-to-Noise Ratio (SNR) to detect a planet around a central star. In our case we assume that the flux of the planet is collected in a square box of size 2.5x2.5λ/D. The exposure time of the single image is set to 2s and the total number of

images N_{exp}=2000 (total exposure time of 70min). The star is assumed to have a magnitude 6 and the planet a magnitude 16. The SNR as a function of the distance from the reference star is shown in Figure 12. It is computed as follows:

- Signal: S = (Planet Flux) * 0.8 * (Strehl Ratio),
- Var Noise = 4* [(Difference Profile * Star Flux)^2 + (Contrast Profile * Star Flux)]
- Noise = sqrt(Var Noise) / sqrt(N_{exp})
- SNR = Signal / Noise

0.8 is the estimated throughput of the science path, and the factor 4 accounts for the fact that on average, the planet flux will be split between 2x2 pixels.

Figure 12 shows that in an annular region between 0.1 and 0.3 arcsec the SNR achieved by the PWFS is improved by a factor 2-5 compared to the SHWFS. Because the SNR scales with the square root of the exposure time gains in exposure time by a factor 4-25 can be expected with the PWFS to achieve a given SNR.

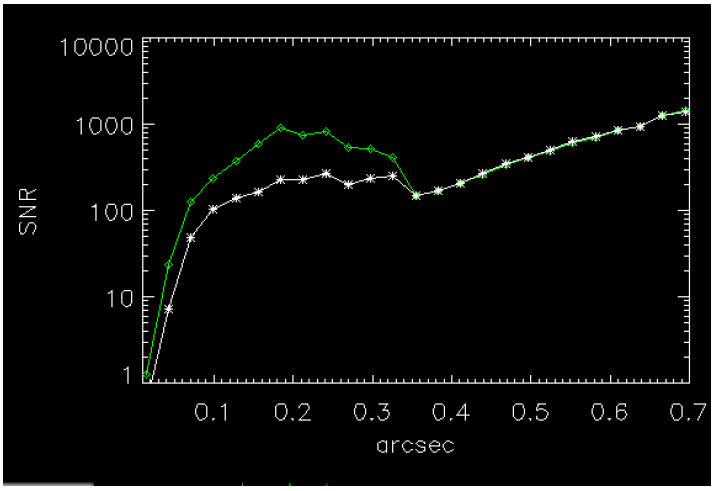

Figure 12: SNR for the detection of a 16mag planet around a 6mag star used as a reference for the AO system. The SNR is plotted as a function of the center distance. Green and white curves are for PWFS and SHS respectively.

Other effects

This study of high contrast imaging is fairly simplistic and by no means complete. Many effects that will reduce the SNR are not included. While purely static aberrations and diffraction effects can be subtracted out, slow drifts, e.g. caused by beam wander on optics, will cause more residual speckles. These however, could be attenuated with techniques such as apodization, coronography, higher quality optics and better image processing techniques. However, pinned speckles will always remain as ultimate limiting errors, and therefore reducing their remains critical.

3.4 Sky coverage

Performance as a function of magnitude

In order to evaluate the end-to-end performance of NFIRAOS as a function of guide star magnitude, we need to add the following contributions, which are not accounted for in our simulations:

- Under-modeling errors of 40nm RMS due to using a sampling of only 1/20m, as discussed in section 3.2
- Residual uncorrectable errors. As per the PDU performance budget, these are Opto-mechanical implementation errors (75nm RMS) and AO component errors and higher-order effects (60 nm RMS), which sums in quadrature to 96nm RMS

With these allocations, the performance on a magnitude 8 star becomes: 158 nm RMS for the PWFS and 174nm RMS for the SHWFS

In order to find the performance for guide star fainter than magnitude 12, we rely on earlier simulations performed by Lianqi Wang at TMT for the SHWFS. The problem is that these simulations are not performed by the same simulation software (MAOS for the SHWFS, Passata for the PWFS) and not with exactly the same observing parameters. These results are shown in Table 6.

	8	9	10	11	12	13	14	15	16
Fs (Hz PWFS)	800	800	800	800	800	400	200	100	50
NModes (PWFS)	2210	2210	2210	2210	2210	2210	1500	800	400
RMS (nm PWFS)	154.1	154.6	155.4	158.3	170.7	185.2	232.5	326.4	425.2
RMS (nm SHWFS)	161.0	162.7	166.8	179.9	194.3	227.3	283.0	345.2	517.0
Quad diff (nm)	46.6	50.7	60.6	85.5	92.8	131.8	161.3	112.4	294.1

Table 6: Simulation results from previous work. The last row is the quadrature difference between the second and third last rows.

If we assume that the change in RMS as a function of guide star magnitude reasonably captures the effect of increased noise, we are able to derive the performance shown in Figure 13. It shows that the 50% H-band Strehl level is reached at magnitude 12.2 for a SHWFS, but only at magnitude 13.5 for the PWFS. For this level of correction (which is the one assumed in the TMT science cases for NGS AO), the limiting magnitude is therefore 1.3 higher for the PWFS. At magnitude 12.2 where the SHWFS achieves 50% of Strehl, the PWFS achieves 61%. This results in 34% saving in exposure time, which is quite significant.

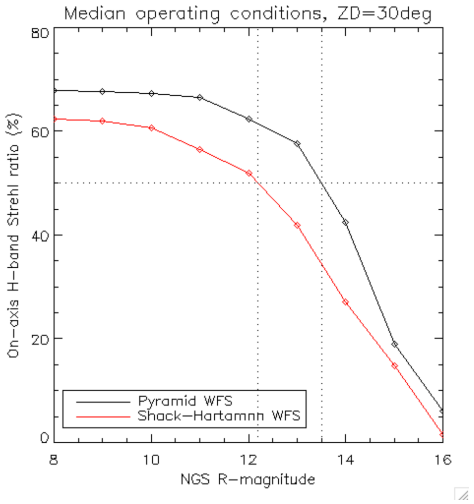

Figure 13: H-band Strehl ratio as a function of guide star magnitude for the SHWFS and the PWFS

Sky coverage evaluation

The increased limiting magnitude results in increased sky coverage. Sample sky coverage simulations using the NOMAD catalogue hosted by Vizier in the galactic plane at 60 degrees of longitude show an increase in sky coverage from 0.5% for the SHWFS to 1.5% for the PWFS (assuming 20" spaxels). These numbers remain low because of the required Strehl ratio is 50% in H-band, which translates into bright stars. However, the difference is significant and can be appreciated in Figure 14: stars that are bluish (about 9-10 of them) can be observed with the SHWFS, but stars bluer than green (more than 30) can be observed with the PWFS.

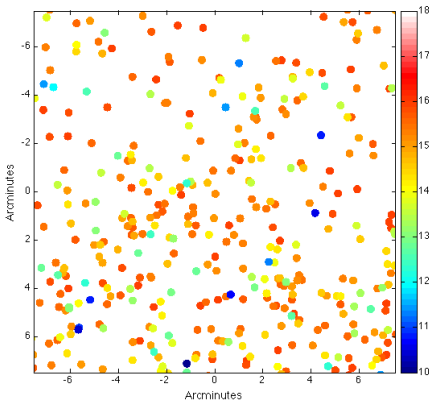

Figure 14: Sample sky coverage illustration. The dots represent a 20" diameter region of the galactic plane (60 degrees longitude) where a star brighter than a given magnitude can be found.

3.5 Truth wave-front sensing

In LGS MCAO mode, the TWFS must measure wave-front errors resulting from changes in the sodium layers. Signatures of these changes appear mostly in radial modes (especially Z11), but other modes may carry information on higher order drifts, e.g. created by print-through, etc.

A PWFS can provide to the TWFS the same advantages as for the high order NGS WFS: reduced sensitivity to noise and to spatial aliasing. This can potentially allow increasing the frame rate for better tracking capabilities. The PWFS is also advantageous in that the spatial sampling can be adjusted through binning to the SNR of the guide star. That way, higher order measurements can be extracted from bright guide stars, without compromising sky coverage.

It is worth noting that the TWFS is expected to operate anywhere in the 2' NFIRAOS FOV. This means that a) the NGS might not be well very well corrected, and b) the TWFS must have sufficient dynamic range to make its measurements on top of the uncorrected field dependent static aberrations. At the edge of the NFIRAOS FOV, the amplitude of these aberrations is expected to be 200nm RMS, mostly coming from beam shearing on M3 (~150nm RMS) and on the NFIRAOS optical elements (~90nm RMS). In order to provide the required dynamic range, more modulation is likely to be required. This, plus the poorer image quality of the spot, will reduce some of the advantages of the PWFS. At this point, we have not attempted to simulate the TWFS process. However, simulation and experiments using a PWFS on an extended source suggest that the PWFS remains more sensitive than the SHWFS even for sources extended by up to 1.6", as reported in [7]. Also, in November 2014, the Arcetri team reported that they were able to successfully use the FLAO PWFS to provide Truth sensing for the ARGOS system. The ARGOS system uses Rayleigh beacons to provide GLAO correction, so the spot on the FLAO pyramid would not be well corrected in that case.

4. NFIRAOS VNW bench design with a PWFS

4.1 PWFS optical design

The design of the NFIRAOS has been modified to accommodate a PWFS instead of a SHWFS. The layout of the WFS bench is shown in Figure 15.

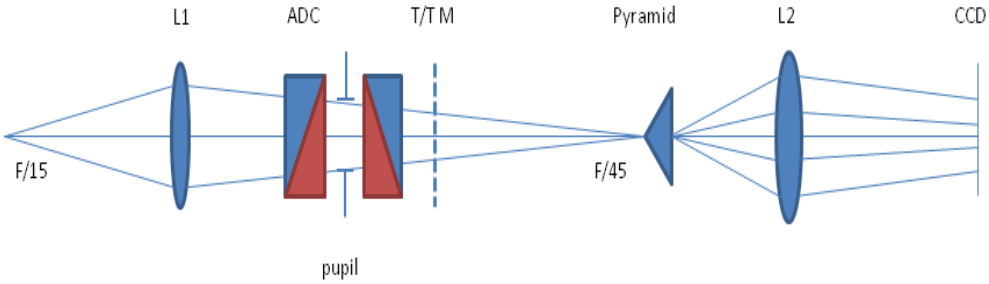

Figure 15: Optical layout of the NFIRAOS PWFS

The F/15 entrance beam is relayed to a F/45 beam via lens L1, which produces a focal image at the tip of the pyramid. Behind the pyramid, a relay lens L2 produces the four pupil images onto the CCD. This CCD is the 256x256 21um pixel CCID-74 from MIT/LL, which is the same as the one baselined for the Shack-Hartmann design. Two counter-rotating double prisms provide ADC correction in the F/45 beam, which also includes a fast tip-tilt mirror that provides the modulation and dithering capabilities. Note that the TTM is not exactly in the pupil plane. This produces a small blurring of 0.1pixel RMS on the pupil images during a full modulation cycle of radius 5λ/D. This will lead to a slight reduction of the PWFS sensitivity to high spatial frequencies, comparable to the blurring effect due to charge diffusion between pixels on the CCD (~0.3pixel FWHM blur for the CCID-74)

With the SHWFS, we are pretty much limited to 60x60 sub-apertures, due to the need of having guard pixels around each quad-cells. For the PWFS however, we can increase the sampling resolution. In the current PWFS design, each pupil image covers a disk of 96 pixels in diameter and the centers of the pupil images are separated by 120 pixels. The higher sampling allows reducing the aliasing when the GS is bright enough, whereas on-chip binning allows reducing the number of pixels to 96, 48, 32, 24, 16, 12, etc as needed, when the GS is fainter (note: it is unclear at this time whether binning by more than 4 will be supported – still even if binning is limited, we have more versatility than with the SHWFS).

4.2 Mechanical packaging

The packaging of the PWFS and of the SHWFS are shown in Figure 16, with the same scale. The SHWFS layout must include two paths, one for laser mode where a 12x12 lenslet array is used, and one for the NGS mode, where a 60x60 lenslet array is used. No lenslet array and only one path is required with the PWFS, since the sampling resolution is adjusted through on-chip binning. We see that both designs fit the assigned 800mm x 500mm space envelope for the VNW bench (outer rectangle).

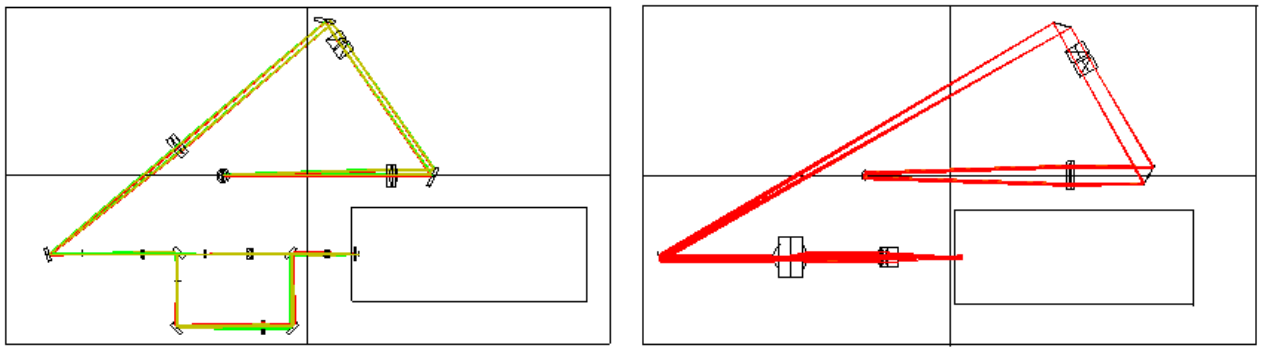

Figure 16: Left: packaging of the SHWFS. Right: packaging of the PWFS. The black rectangle represents the CCD camera housing.

4.3 Impact on complexity, cost and schedule

The PWFS design has 2 lenses (one doublet and one triplet), 4 fold mirrors and 1 pyramid, plus an ADC. The SHWFS has 6 lenses (all doublets), 8 fold mirrors and 2 lenslet arrays, plus an ADC. Fewer optical elements result in a significant cost saving in the fabrication and the assembly of the WFS bench.

The SHWFS requires a mechanism to deploy the two fold-mirrors that allow switching to the TWFS. This mechanism requires at least one motor, which must operate at -30C.

4.4 Impact on other NFIRAOS sub-systems

As part of the trade study, we have also evaluated the impact of switching to a PWFS on the other, already designed, NFIRAOS sub-systems. We have found no impact on the RTC, on the AO Executive Software and on the design of the visible camera recording the WFS signal. The NFIRAOS Component Controller requires a new modulation board, in order to produce the spot modulation on the pyramid within the integration time on the detector. In terms of alignment complexity, we have found that the alignment tolerances within the PWFS are slightly tighter, because the magnification of the pupil images on the CCD needs to be tightly controlled. However, the alignment tolerances between the PWFS and NFIRAOS are slightly looser, because there is no need to provide the ability to clock the PWFS.

5. CONCLUSION

Our trade study has clearly demonstrated the advantages of the PWFS for NFIRAOS. These include higher Strehl ratios, higher sky coverage and better high contrast imaging capabilities in NGS mode, even in presence of significant NCPAs. It will therefore increase NFIRAOS ability to work without lasers. In LGS modes, the PWFS provides better TWFS capability, especially with the ability to adjust the spatial sampling as a function of the signal level. We have found that adopting the PWFS will reduce final cost and risk of NFIRAOS because of simpler opto-mechanics (fewer motors operating in the -30C NFIRAOS enclosure and fewer optical elements).

The downsides of the pyramid lie mostly in the limited experience in this more novel technology, potentially lengthening the integration and test phase, as well as the risk in procuring a pyramid, which is much less commonly available than a lenslet array.

Overall, we have found that the pluses outweigh the minuses, and our recommendation was to adopt the PWFS. This recommendation has now been accepted by the TMT Project Office, and NFIRAOS will proceed with a PWFS.

ACKNOWLEDGMENT

We are grateful to the TMT AO group for their support, and specifically for Lianqi Wang for providing some simulation results to anchor ours. We are also grateful to Fernando Quiros-Pacheco, now at GMT, for his help in using Passata.

REFERENCES

[1] Sanders, G. H., "Thirty Meter telescope project update," In Ground-based and Airborne Telescopes V Proc. of SPIE, Vol. 9145, (2014).
[2] Herriot, G., et al., “NFIRAOS: first facility AO system for the Thirty Meter telescope,” In Adaptive Optics Systems IV Proc. of SPIE Vol. 9148, (2014).
[3] Ragazzoni, R., “Pupil plane wavefront sensing with an oscillating prism”, Journal of Modern Optics, vol. 43, Issue 2, p.289-293, (1996)
[4] Vérinaud, C., On the nature of the measurements provided by a pyramid wave-front sensor, Optics Communications, Volume 233, Issue 1-3, p. 27-38, (2004).
[5] Esposito, S. et al., “First light AO (FLAO) system for LBT: final integration, acceptance test in Europe and preliminary on-sky commissioning results”; Proc. SPIE 7736 (2010).
[6] Esposito, S. ,et al., “Non common path aberration correction with non linear WFSs”, this conference
[7] Pinna, E., Design and numerical simulations of the GMT Natural Guide star WFS, Proc. SPIE 9148 (2014).