

Comparative Performance of a 3-Sided and 4-Sided Pyramid Wavefront Sensor

Johanan L. Codona³, Michael Hart^{1,2}, Lauren H. Schatz², and Mala Mateen³

¹ HartSCI LLC, 2555 N. Coyote Dr. #114, Tucson, AZ 85745

² College of Optical Sciences, University of Arizona, 1630 E University Blvd, Tucson, AZ 85719

³ Air Force Research Laboratory, Kirtland AFB, NM 87117

ABSTRACT

We investigated the relative performance of three-sided and four-sided pyramid wavefront sensors (PWFS3 and PWFS4). Three-sided pyramids are easier to manufacture and are reasonably expected to be less sensitive to sensor shot and read noise, simply because there are 3/4 as many pixels contributing noise for the same plate scale. For this study, we assumed high Strehl ratios, allowing us to analyze the performance without modulation. We assumed a linear reconstructor method for processing the WFS images. The background-subtracted and flat-fielded images were normalized by the total image count to make the response independent of guide star flux. We then subtracted an ideal flat-wavefront image, leaving us with a delta-image that goes to zero when the wavefront is flat. The pupil image plate scale is selected to allow adequate sampling of the intensity compared with the DM resolution. This minimizes the number of pixels for the highest desired resolution, minimizing the amount of read and shot noise entering the reconstructor. We simulated an AO-equipped 1.5 m telescope with a 12×12 actuator DM and 36 sensor pixels across a pupil image. The delta-image (image minus ideal image) pixels were unfolded into a vector that is multiplied by a reconstructor matrix to give the DM actuator updates. We used 2500 weak Kolmogorov phase screens to estimate both the PWFS3 and PWFS4 reconstructors using SVD and tested on an independent set of wavefronts. The optimum number of SVD modes was selected by minimizing the single-update residuals for both the PWFS3 and PWFS4. We found that the single-update suppression was nearly the same for both cases, the PWFS3 wavefront variance being reduced to 8% while the PWFS4 was 7.5%. However, as expected, the PWFS3 was less sensitive to read noise due to fewer image pixels required for a given spatial resolution. The measured WFE residual due to noise was consistent with the square root of the ratio of the pupil image areas, 87%.

1. INTRODUCTION

The pyramid wavefront sensor (PWFS) [1] is known to be fundamentally more sensitive to wavefront error (WFE) than other types of WFS conventionally used for ground-based space situational awareness (SSA) observations because it exploits the coherence of light across the full telescope aperture. The objective of the research was to establish the performance of the three-sided pyramid wavefront sensor (PWFS3) and compare it to the performance of the more conventional four-sided pyramid (PWFS4). Numerical modeling and analytic assessment were carried out in two independently configured models to compare the sensitivity of the two varieties of PWFS to photon noise inherent in the beacon source as well as read noise expected from a realistic detector. Open-loop and closed-loop performance has been quantified in terms of total rms residual WFE as well as the spatial spectrum of the error.

At the most basic level, the PWFS is a generalization of the Foucault knife-edge test [2]. A concave optic under test is illuminated from the center of curvature. A knife edge is inserted in the reflected beam near the image plane. If the rays are not all converging on the same point, some will be blocked while others will pass around the knife edge, generating a characteristic illumination pattern. It is a geometric test that highlights those areas of the optic with slope errors that prevent the reflected light rays from being directed to the correct location in the image. Fundamentally, the Foucault test acts as a WFS that is sensitive to the local wavefront slope.

The PWFS establishes a practical sensor built on the Foucault concept. The beam is split in both directions simultaneously by the facets of a pyramid in the focal plane so that both directions of the wavefront gradient can be recovered. An image of the pupil is formed from the light reflecting off each facet of the pyramid. Rather than throwing away half the light by blocking it, the portions of the beam are all separately imaged so no information is lost.

1.0 Simulation and Analysis

We studied the relative performance of the 3 and 4-sided PWFS using a numerical simulation of an AO system, simplified to consider the single-update suppression of a typical residual wavefront and the relative effect of photon and read noise. The model is capable of simulating AO performance with pyramid modulation, but it was not used in this study. The wavefront reconstruction used the simple linearized assumption that, in high Strehl ratio conditions, the deviation in the intensity in and around the PWFS pupil images from the perfectly-corrected versions is related to the wavefront aberrations by a linear functional. Since both types of PWFS have nonlinear intensity responses for larger aberrations, any deviation from the linear model will appear as noise, limiting how much correction can be achieved in a single update. We did not evaluate this nonlinear noise contribution here. As a discrete representation of the linear relationship between wavefront and changed intensity, we related the change in intensity measured with the WFS camera pixels to the deviations from flat at the actuator positions using a wavefront reconstructor matrix. This matrix was estimated by using a large collection of Kolmogorov phase screens, and using SVD [3] to find a best least-squares fit to the data. It is through testing the two pyramid WFS reconstructors that we are able to derive their relative performance and sensitivity to noise.

1.1 Numerical simulation model

We developed a numerical model using our wave optics package, AOSim3 (Fig. 1). The model includes a 1.5 m telescope with a 20% central obstruction. The telescope includes an 83.3x angular magnification, reducing the beam to 18 mm where the deformable mirror (DM) is applied and the wavefront sensing is performed. The DM is modeled as a square grid of actuators using a 2-D cubic spline interpolation to find the resulting optical path length (OPL) map. The DM is placed in the reduced pupil with a 12×12 grid of actuators across the beam. The modeled DM actually had a 16×16 actuator array with the actuators whose segments were less than 50% illuminated, tied to the nearest illuminated actuator. This was the same for both the PWFS3 and PWFS4, with 116 actuators being controlled. The pyramidal prism was modeled as a thin phase screen imprinted with a pyramid-shaped OPL map and a programmable dispersion relationship. Since our analysis only used monochromatic light, the dispersion feature was not exercised.

The model includes support for pyramid modulation, but it was disabled for this analysis. Modulation is used to increase the linear range of the PWFS in situations where the uncorrected wavefront slopes are so great that light is

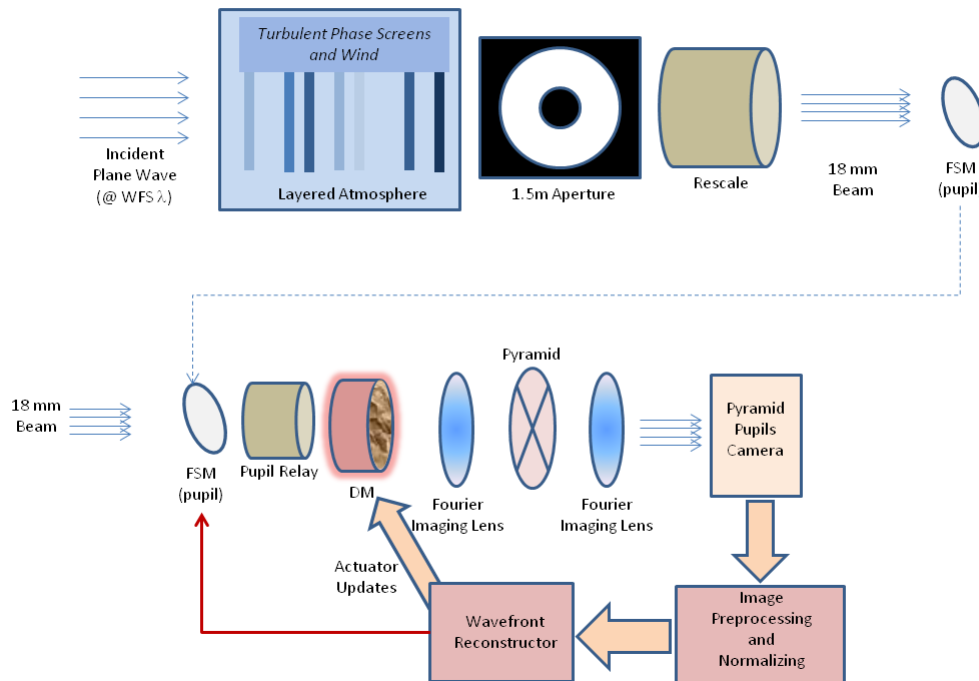


Figure 1: Block diagram of the AOSim3 pyramid simulation.

often found predominantly in a single pupil image, saturating the measurement above or below the narrow high-sensitivity dynamic range. By rapidly translating the image around the pyramid during a single exposure, the magnitude of the angle over which light from a steeply tilted patch of the wavefront will fall into more than one of the pyramid facets is increased. This allows slope measurements to be made over a larger dynamic range, although with a lower accuracy. Modulation is particularly useful when bootstrapping a pyramid WFS-based AO system. In the limit of best correction, however, the modulation should be less than λ/D , or stopped altogether. Therefore, when modeling the best performance of a pyramid WFS, modulation should be turned off. This also helps accelerate the simulation since a single WFS frame can be computed with only one pass through the system.

The model includes an extended atmosphere, comprised of multiple turbulent phase screens and wind layers. For the present study, the analysis primarily used a single phase screen located at the entrance pupil. This simplification allowed the model to run substantially faster and sacrificed no fidelity over the small fields-of-view used to estimate AO system performance.

For a particular time step in the model, the phase screens are adjusted to their respective wind-translated positions. A monochromatic plane wave incident on the topmost atmosphere phase screen is interpolated to the field grid positions and propagated through the atmosphere one turbulent layer at a time until it reaches the telescope's entrance pupil (Fig. 1). There it is reduced to the exit pupil size, phase-shifted by the DM's computed OPL map, and imaged into the focal plane by forward Fourier transforming the field and setting the sample spacing to correspond to a desired imaging lens. The amplitude of the field is also scaled to conserve energy. The imaging lens was assumed to be achromatic and have a 10 cm focal length. In this intermediate focal plane, the wave is passed through the pyramid optic with an appropriate translation depending on the optional use of pyramid modulation. Finally, the field is once again forward-Fourier transformed to the now multi-pupil image plane where the magnitude-squared field is computed.

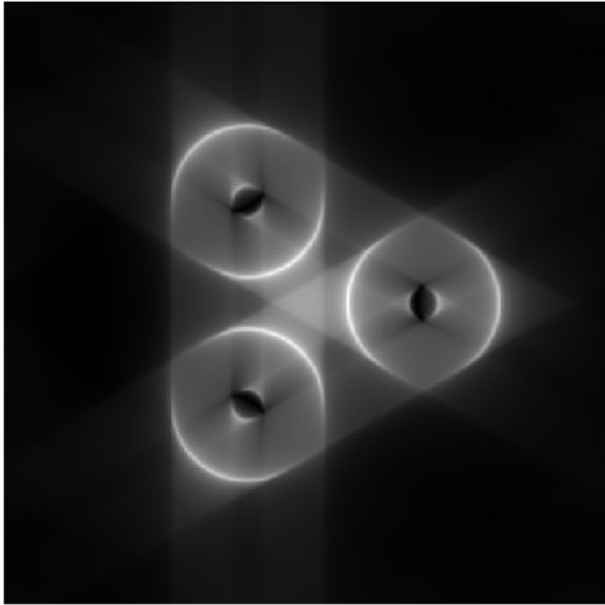
Photon noise in the resulting image is modeled by normalizing the irradiance image to unit volume, multiplying by the expected number of photons, and computing a realization of a Poisson process for each pixel in the image. Gaussian read noise is also included, an independent realization for each pixel of each modeled image.

1.2 Model outputs

To find the intensity changes from the ideal flat wavefront, we first collect a set of images with a plane wave input (Fig. 2). Since the relative pixel values of the operational images and the calibration images will depend on wavefront beacon brightness, atmospheric transparency, etc., we must first process the images to make them compatible. First, the images should be background-subtracted and flat-fielded, with any bad or noisy pixels replaced by the median of its neighbors. Then the image is normalized by its sum, resulting in an estimate of the intensity that is independent of incident flux. The same procedure is performed on images resulting from an aberrated wavefront, as shown in Fig. 3. This figure shows the response of the PWFS to the same aberrated wavefront, in this case an uncorrected Kolmogorov phase screen with a Fried length r_0 of 15cm at a wavelength of 500 nm, with an rms WFE of 526 nm. This is a much deeper aberration than we assume for our high-Strehl analysis, and it exhibits larger intensity deviates than we expect to encounter, other than during the initial bootstrapping cycles after the AO closed-loop has been activated. The images are pre-processed and normalized the same as the ideal calibration images and "delta-image" differences are shown in Fig. 4.

We only need to consider a set of regions-of-interest (ROI) that include the separate pupil images; it is not obvious that the light in between the pupil images contains substantial new information that is not also captured by the light inside and immediately near the pupils. Also, the simulation computes these patterns at a higher spatial resolution than would actually be captured by the WFS camera. Figure 5 illustrates the case of the PWFS3 used for reconstruction. In this figure, the three separate pupil images for an aligned, unaberrated plane wave have been cut out from the simulation's full frame, computed at high resolution. They are then binned down to the actual 32×32 resolution appropriate for $r_0 = 5$ cm and a 1.5 m entrance pupil. For concise visual representation, the three pupil images are stacked as the RGB planes of a full-color image. Including an extra pixel on either side of each nominal pupil image (i.e. 34 pixels across each pupil), results in $3(1+32+1)^2 = 3468$ pixels for the PWFS3, and $4(1+32+1)^2 = 4624$ pixels for the PWFS4. The delta-image pixels for each of the pupils were taken as the components of a Hilbert space vector, to be multiplied by the reconstructor matrix to find the 116 actuator updates.

PWFS3 Ideal Intensity, $\gamma=1/2$



PWFS4 Ideal Intensity, $\gamma=1/2$

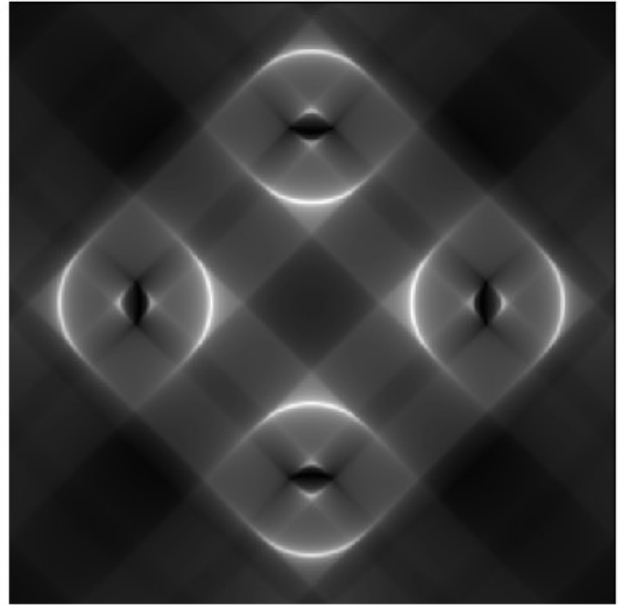
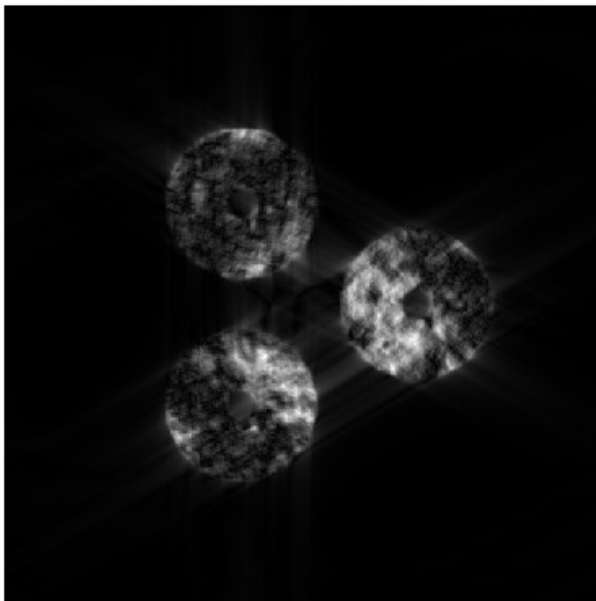


Figure 2: The ideal intensity patterns in the final imaged pupil plane for PWFS3 (left) and PWFS4 (right). The intensities are shown on a square root scale. These patterns are subtracted from those resulting from aberrated wavefronts, defining the goal for a zero-seeking AO servo.

PWFS3 Aber_15cm Intensity, $\gamma=1/2$



PWFS4 Aber_15cm Intensity, $\gamma=1/2$

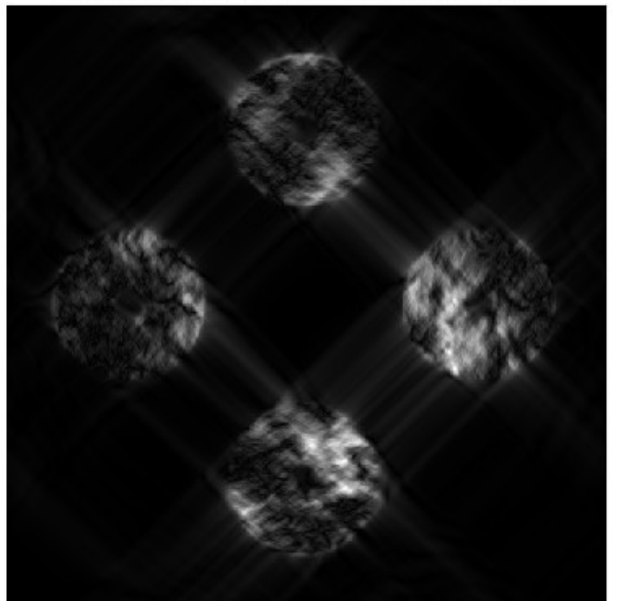


Figure 3: The PWFS3 and PWFS4 intensity patterns with an aberrated wavefront with a Fried length of 15 cm. The test aberration is the same in both cases.

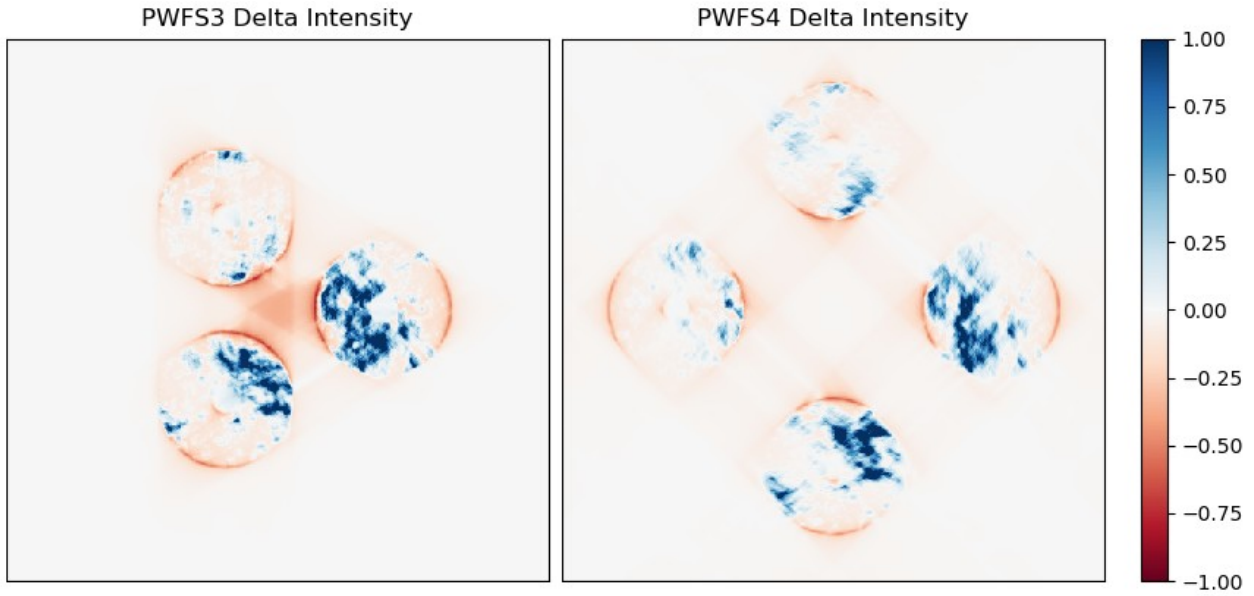


Figure 4: The corresponding PWFS3 and PWFS4 delta-intensity patterns. The same aberrated wavefront was used in both cases.

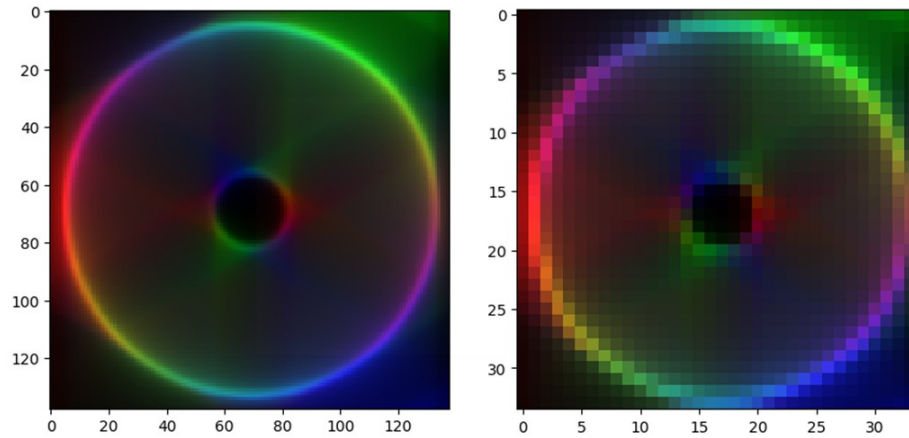


Figure 5. The PWFS3 pupil images plotted as RGB color planes at full simulation resolution (left) and reduced to camera pixel resolution (right). The axis labels are pixels.

1.3 Building the reconstructor

Our assumption that the wavefront aberrations are linearly related to the change in intensity from the ideal light pattern can be written

$$\alpha = \mathbf{R} \delta \mathbf{I}, \quad (1)$$

where α is the vector of actuator offsets that best represent the wavefront residual, $\delta \mathbf{I}$ is the vector of normalized delta-image pixel values, and \mathbf{R} is the reconstructor matrix that remains to be determined. If the linear model is reasonably correct, then Eq. 1 is simultaneously true for any small aberration and can be written for a set of actuator test vectors α , and the corresponding PWFS delta-image pixel vectors $\delta \mathbf{I}$

$$\alpha = \mathbf{R} \delta \mathbf{I}. \quad (2)$$

We can use singular value decomposition, SVD [3], to factor the delta-intensity data matrix into a pair of unitary matrices, U and V , and a diagonal matrix of singular values, S .

$$USV' = \delta I \quad (3)$$

The column vectors of U are the “SVD modes” of the delta intensity. The SVD matrices are used to find the pseudo-inverse [3] as

$$\delta I^{-1} \approx V \hat{S}^{-1} U' \quad (4)$$

where \hat{S}^{-1} is a diagonal matrix of the inverses of the first to the N -th singular values, the rest being set to zero. This can now be used to multiply both sides of Eq. 1 from the right, giving us a least-squares approximation to the reconstructor, parameterized by the number of SVD modes used in the delta-image pseudo-inverse.

$$R \approx \alpha V \hat{S}^{-1} U'. \quad (5)$$

The actuator modes corresponding to the delta-intensity modes are found by multiplying them with the reconstructor, or

$$U_{\text{actuators}} = \alpha V \hat{S}^{-1} U'_{\delta I} U_{\delta I}. \quad (6)$$

We used 2500 Kolmogorov phase screens to estimate the SVD modes and build the reconstructors, with 500 more used for testing the results. The same set of phase screens was used with the PWFS3 and PWFS4. The phase screens were scaled to have an rms WFE of 100 nm, placing them in the range of hoped-for AO performance. This gives us roughly the same amount of nonlinear noise that we might expect in practice. The DM was not used to correct the wavefront in the open-loop tests, but the displacements of the wavefront at the actuator positions were used for the actuator test vectors, while the full-resolution phase screen was used to compute the delta-intensity images and corresponding WFS data vectors. This was repeated for each of the 2500 phase screens. Figure 6 shows the first 50 SVD modes for the PWFS3 delta-intensity, while Fig. 7 shows the corresponding first 50 DM modes. Likewise, Fig. 8 shows the first 50 SVD modes for the PWFS4 delta-intensity, while Fig. 9 shows the corresponding first 50 DM modes.

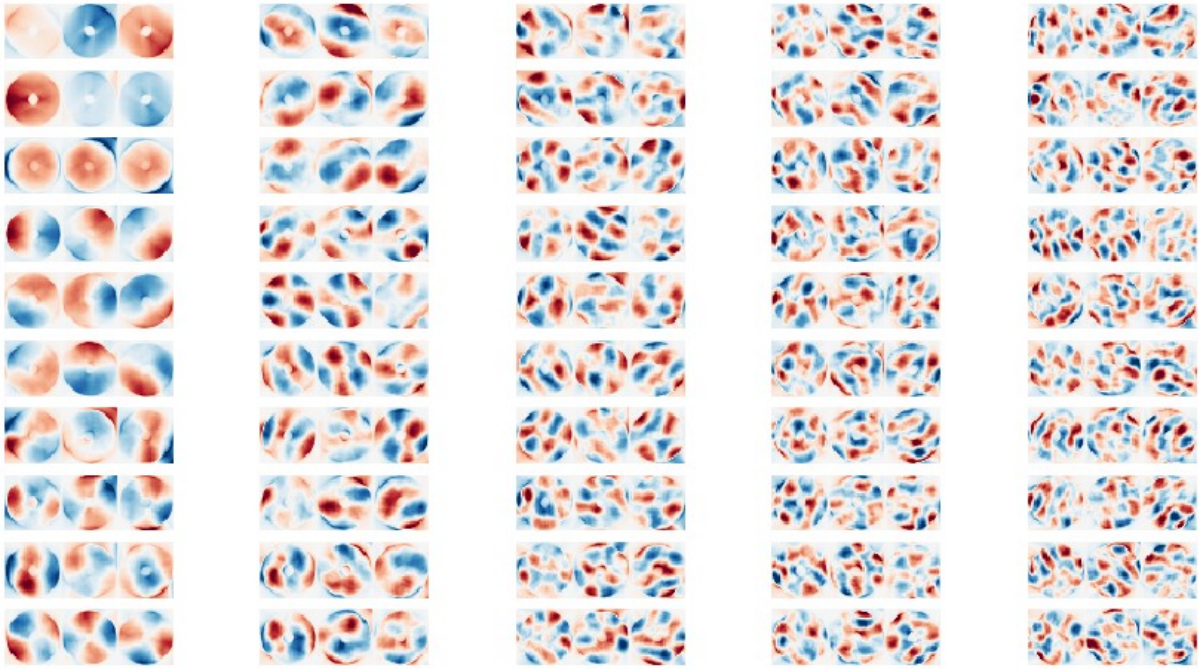


Figure 6: The first 50 delta-intensity SVD modes for the PWFS3 (decreasing singular values run top-to-bottom and left-to-right). Red and blue are opposite sign intensity displacements.

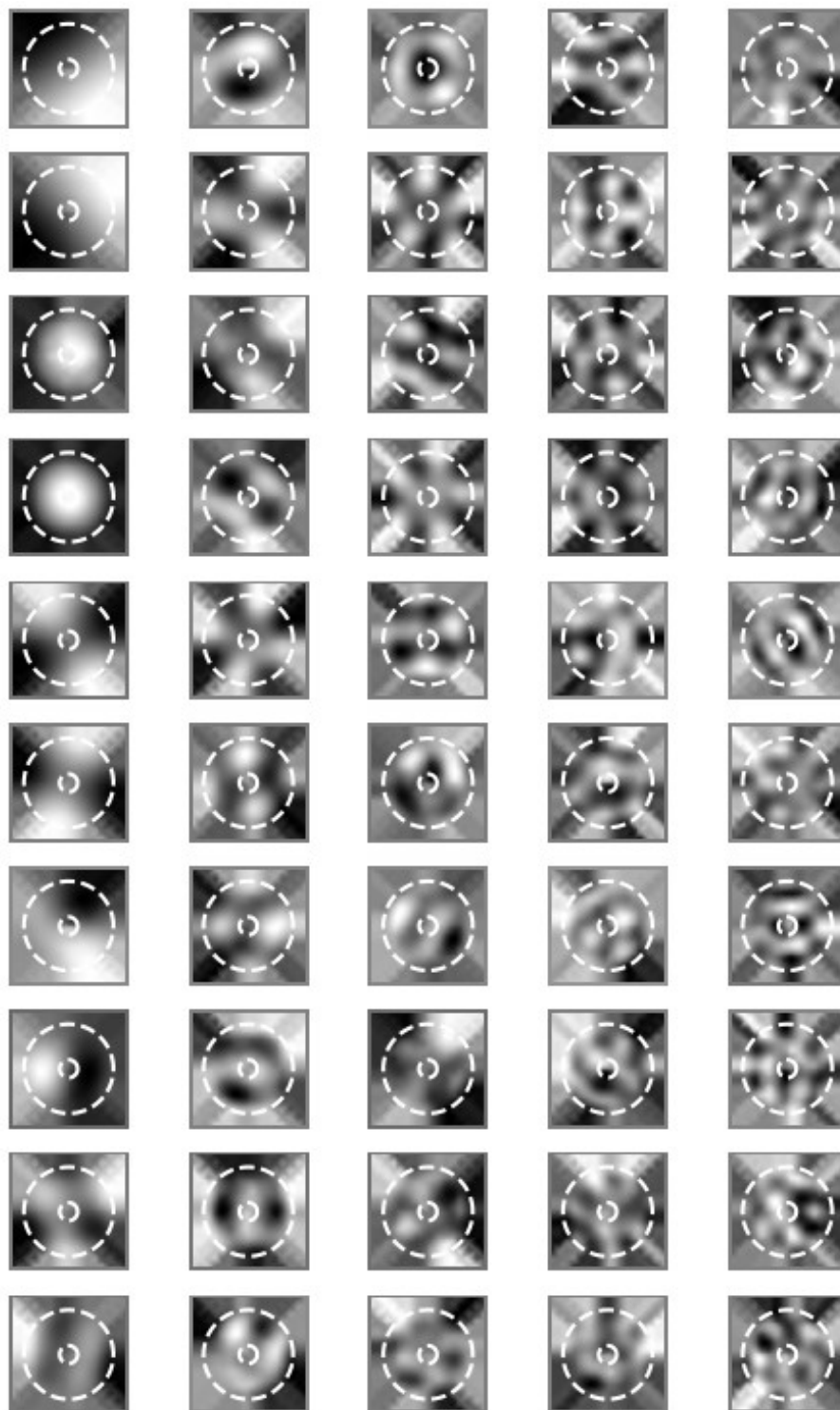


Figure 7: The corresponding first 50 DM modes for the PWFS3.

1.4 Selecting the number of SVD modes

To find the best number of SVD modes to include in each of the reconstructors, 100 additional test phase screens were used following the same procedures described above. For each number of modes, a reconstructor was built and used to estimate the actuator values for each of the test cases. The estimated actuator values were then subtracted from the values determined directly from the phase screen, and the rms WFE was computed for each case

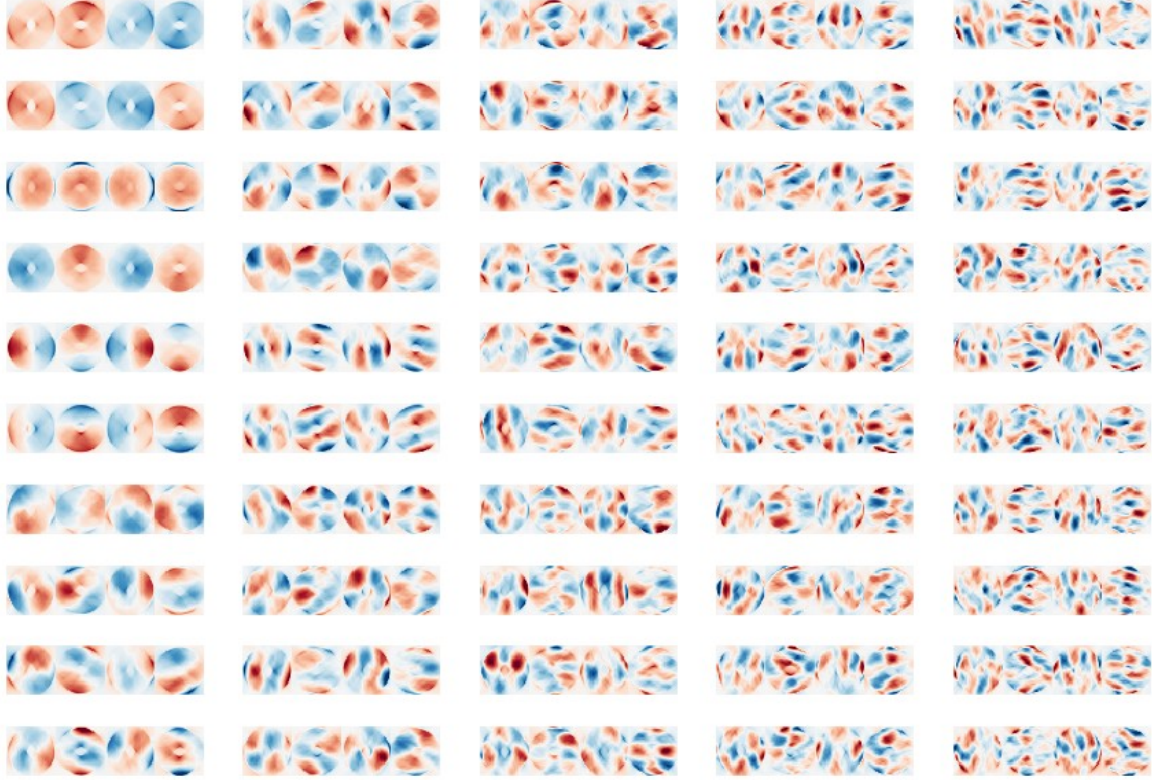


Figure 8: The first 50 delta-intensity modes for the PWFS4.

and averaged over all of test phase screens. The ratio of starting WFE to the residual is a measure of how well the PWFS would perform in an open loop, single update correction to a random aberrated wavefront. The resulting after/before rms WFE for PWFS3 and PWFS4 is shown in Fig. 10 as a function of number of SVD modes included in the reconstructor. Over a wide range of included modes, the PWFS4 performs slightly, yet consistently better than PWFS3, by about 0.5%. The best PWFS4 residual in this case was about 8.5% while the PWFS3 residual was approximately 9%. Note that if too many modes are included in the reconstructor matrices, more than about 1000, the residual WFE begins to get worse. This is because the total number of phase patterns used to construct the influence matrices was 2500, and as the number of SVD modes retained approaches this number, the reconstructor begins to fit random details in the pupil image data instead of more general trends.

While the finding is consistent, a half-percent difference in the open loop WFE residual is insignificant to the expected performance in a closed-loop AO system with other sources of noise, such as fitting error and photon noise. The operational difference between the two systems in this noise-free case would not be noticeable and would deliver essentially the same performance.

1.5 Noise propagation with a linear reconstructor

Once implemented in a live AO system, the wavefront correction accuracy is degraded by a number of noise sources. Each of these results in a random error in the wavefront estimation, resulting in a worse correction. Critically important is the WFE caused by read noise in the WFS camera. This will depend on the camera, the plate scale of the sensor, and other details like down-sampling, but we can estimate the relative importance of photon and read noise between our two PWFS.

In the limit of near-perfect correction, the multi-pupil intensity pattern approaches that of the ideal profile. Since the ideal value is determined at very high SNR during calibration, it will be relatively noise-free. Thus, the noise in the delta-intensity pattern will be almost entirely from the PWFS camera frame. We can use the reconstructor \mathbf{R} as an error propagator to find the resulting actuator position variance. If the processed and binned image has a detected noise error of ϵ_{ij} in the (i,j) pixel, written as a vector of pixel values ϵ_l for photon noise, and ϵ_n for sensor noise.

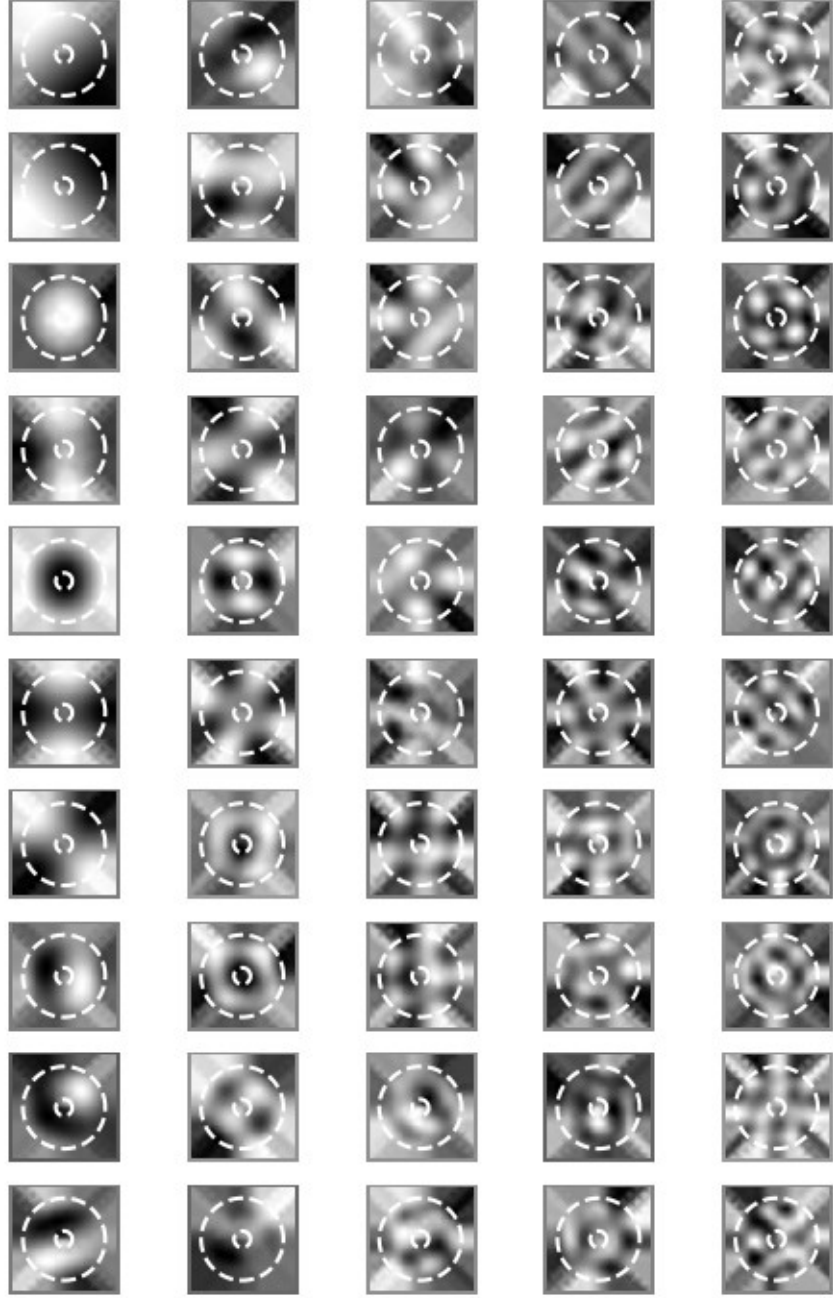


Figure 9: The first 50 DM modes for the PWFS4.

The vector version of the average intensity is \mathbf{I} , and the normalized ideal reference vector is \mathbf{v} . We write a realization of the delta intensity vector as

$$\delta \mathbf{I} = (\mathbf{I} + \boldsymbol{\epsilon}_I + \boldsymbol{\epsilon}_n) / N_{total} - \mathbf{v} \quad (7)$$

where N_{total} is the total flux in the image. Assuming the noise types are uncorrelated, the vector of delta intensity variance is

$$\sigma_{\delta I}^2 = (\langle \boldsymbol{\epsilon}_I^2 \rangle + \langle \boldsymbol{\epsilon}_n^2 \rangle) / N_{total}^2. \quad (8)$$

A realization of the image with random noise errors leads to a realization of reconstructed actuator positions \mathbf{z} as

$$\mathbf{z} + \boldsymbol{\epsilon}_{zI} + \boldsymbol{\epsilon}_{zn} = \mathbf{R}[(\mathbf{I} + \boldsymbol{\epsilon}_I + \boldsymbol{\epsilon}_n) / N_{total} - \mathbf{v}] \quad (9)$$

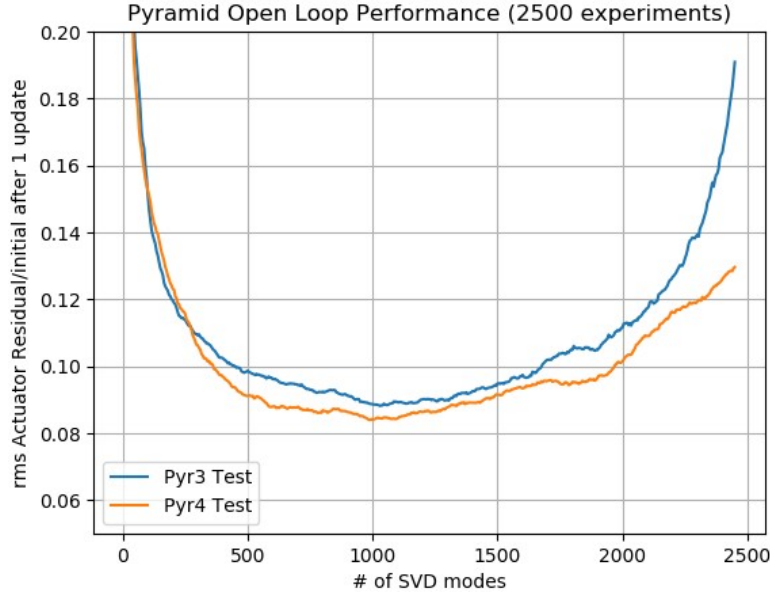


Figure 10 The open loop residual rms WFE relative to the input aberration for PWFS3 and PWFS4 as a function of the number of SVD modes used in the reconstructor.

where the realization of wavefront error due to photon noise is

$$\epsilon_{zI} = \mathbf{R} \epsilon_I / N_{total} \quad (10)$$

and due to sensor noise is

$$\epsilon_{zn} = \mathbf{R} \epsilon_n / N_{total}. \quad (11)$$

Taking the average of the outer product of the computed wavefront error with itself gives the actuator error covariance as

$$\langle \epsilon_z \epsilon_z^T \rangle = \langle [\mathbf{R} \epsilon_I / N_{total}] [\mathbf{R} \epsilon_I / N_{total}]^T \rangle. \quad (12)$$

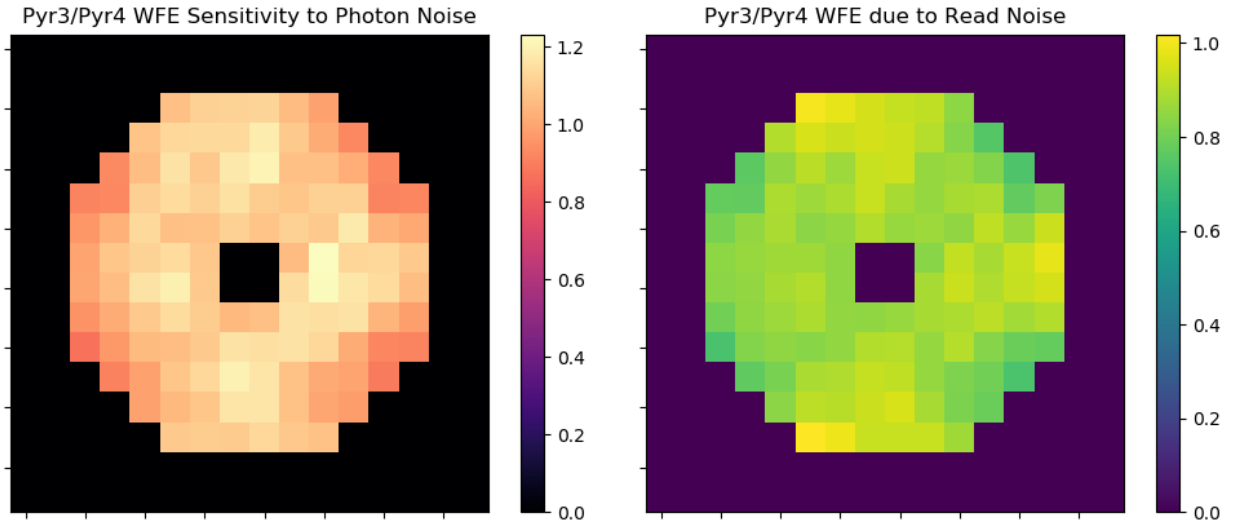


Figure 11: Plots of the PWFS3/PWFS4 ratios of rms WFE due to photon noise (left) and read noise (right). The photon noise ratio follows the relative open-loop suppression of the two reconstructors, while the read noise sensitivity is consistent with there being fewer noisy pixels in the PWFS3 relative to PWFS4. Due to the difference in scaling with guide star brightness, the read noise advantage becomes the more important effect with fainter targets.

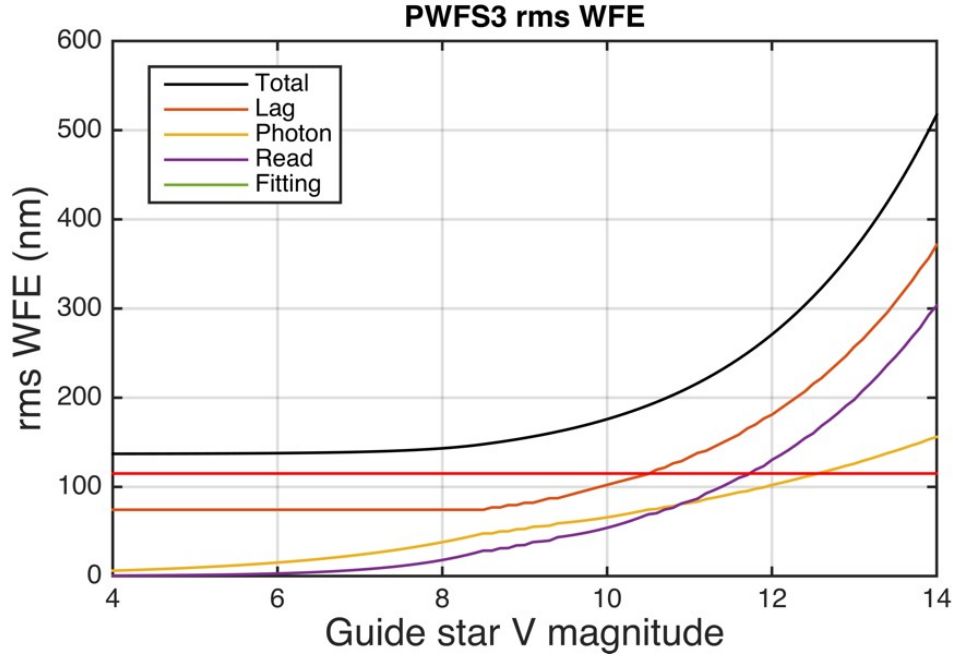


Figure 12: Sources of WFE for the PWFS3 vs. guide star magnitude.

This can be simplified as

$$\langle \epsilon_z \epsilon_z^T \rangle = \mathbf{R} \langle \epsilon_n \epsilon_n^T \rangle \mathbf{R}^T / N_{total}^2 . \quad (13)$$

Assuming the noise is statistically independent between pixels, the noise covariance collapses to a diagonal matrix of the noise variance, giving

$$\langle \epsilon_z \epsilon_z^T \rangle = \mathbf{R} \text{diag}(\sigma_i^2) \mathbf{R}^T / N_{total}^2 . \quad (14)$$

Eq. 14 is appropriate when the noise depends on position, like photon noise. If, the noise is statistically the same in each pixel (i.e. homoscedastic) with variance σ^2 (e.g. read noise) then

$$\langle \epsilon_{zn} \epsilon_{zn}^T \rangle = \sigma_n^2 \mathbf{R} \mathbf{R}^T / N_{total}^2 . \quad (15)$$

In the case of photon noise, if there are enough photons in each pixel to justify using the central limit theorem, the variance in the pixel value depends on position and is

$$\langle \epsilon_l^2 \rangle = I . \quad (16)$$

This, plus photon noise being uncorrelated between pixels, gives the wavefront covariance due to photon noise as

$$\langle \epsilon_{zl} \epsilon_{zl}^T \rangle = \mathbf{R} \mathbf{v} \mathbf{R}^T / N_{total} . \quad (17)$$

Just computing the diagonal terms gives us a map of the actuator variance based only on the detector performance, the reconstructor, and the total number of photons detected in the sensor camera. We performed this calculation on both pyramid sensors using the best open loop reconstructors found above. The ratios of the WFE due to photon noise and read noise are shown in Figure 11. In both cases, the actual noise sigmas drop out, giving a direct comparison of the noise sensitivity for the two pyramid geometries. On the average, the sensitivity to photon noise follows the open-loop suppression ratio, leaving the 4-sided pyramid slightly ahead for our reconstructors. The WFE sensitivity to read noise is about 13% better in the 3-sided pyramid's case, consistent with the square-root of the ratio of the number of noisy pixels involved in the measurement. Since the scaling with beacon brightness differs between photon and read noise, the relative read noise advantage becomes the more important effect for fainter beacons, giving a performance advantage for PWFS3 with faint guide stars.

1.6 Relative performance in closed-loop

The remaining trade off is to set the WFS exposure time to minimize the total WFE. For this analysis we assumed a value of τ_0 of 1.6 ms at 500 nm, along with a maximum WFS frame rate of 1600 fps (the limit of a Zyla camera),

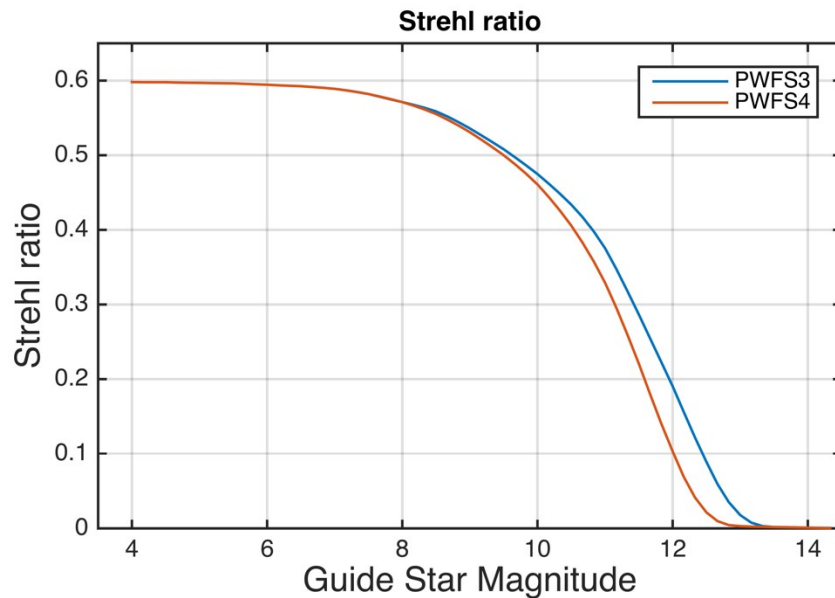


Figure 13: J-band Strehl ratios as a function of guide star magnitude.

and an estimate of processing lag overhead of 850 μ s, giving a minimum lag error of 75 nm. With a relatively bright guide star and our 12 \times 12 DM, fitting error dominates, with the minimum exposure time and processing lag causing the second largest error. With fainter sources, the photon noise and read noise contributions become stronger, the latter being affected by the image normalization. At some point, achieving the minimum WFE requires that we start to increase the exposure time. This causes the lag error to increase, which eventually grows to become more significant than all other sources of error combined. The various contributions to the WFE are plotted in Fig. 12 for the PWFS3 as a function of the guide beacon's V magnitude. The Strehl ratio corresponding to the total WFE are calculated using the Maréchal approximation at an assumed J band science wavelength, shown in Fig. 13 for both PWFS3 and PWFS4. The PWFS3 maintains its Strehl ratio to about half a magnitude fainter than PWFS4. The Strehl ratio plot shows that for bright guide stars, there should be no discernible difference between the performance of the two sensors. Of course, performance in that limit would improve further by use of a faster WFS frame rate or finer sampling in the pupil. We note that the sensitivity improvement is largely due to the inclusion of less sensor read noise, giving the 0.5 magnitude advantage even when using a rather good camera detector with a read noise of just 0.9 e- rms. If we were to use a detector with 3e- rms read noise instead, both sensors would deliver degraded performance, but the gap in performance would also widen to approximately 1 magnitude.

2.0 CONCLUSIONS

The comparative PWFS simulation and analysis in the *noise-free case* indicates that at least for non-iterated open loop wavefront aberration suppression, the performance of the 4-sided and 3-sided pyramids is nearly the same, with the PWFS4 suppressing about 0.5% better. However, it is our experience that such a small difference in open loop suppression would not translate to performance gains in closed loop.

The more important performance difference is in noise sensitivity, where for faint sources the PWFS3 showed a reduction of 13% compared to the PWFS4. This is attributable to the fewer pixels used by the PWFS3. The importance of this reduction in the faint guide star limit where read noise begins to dominate is close to twice that value in terms of residual WFE because of the knock-on effects of balancing other error sources. The closed-loop advantage of PWFS3 becomes even more pronounced with noisier detectors.

The regime where the PWFS3 excels is exactly where GEO satellites are found. Because these objects are of an angular size substantially smaller than the seeing limit, an AO system based around a PWFS3 will be able to exploit the coherence of the light and will be optimally suited to correcting the image to the highest possible Strehl ratio. The contrast ratio in the vicinity of the GEO will also be maximized with a PWFS3, allowing enhanced detection of faint closely-spaced objects (CSO) compared to AO systems using either a Shack-Hartmann WFS or a PWFS4. The PWFS3 is therefore the right choice for SSA application to GEO objects.

3.0 ACKNOWLEDGMENT

This work has been supported by AFRL under contract FA9451-17-P-0515. The opinions expressed are those of the authors and do not necessarily reflect those of the United States Air Force.

4.0 REFERENCES

1. Ragazzoni, R., “Pupil plane wavefront sensing with an oscillating prism,” *J. Modern Optics* **43**, 289–293 1996.
2. Malacara, Daniel, ed. *Optical shop testing*. John Wiley & Sons, 2007.
3. Barrett, Harrison H., and Kyle J. Myers. *Foundations of image science*. John Wiley & Sons, 2013.