Pyramid WFS Tolerance Study for NFIRAOS NGS AO

Lianqi Wang[1]*, Jean-Pierre Veran[2], and Brent Ellerbroek[3]

[1]TMT International Corporation, 100 West Walnut St STE 300, Pasadena, CA, 91124, USA
[2]National Research Council Canada, 5071 West Saanich Rd, Victoria BC. V9E 2E7
[3]Retired

ABSTRACT

The pyramid wavefront sensor (WFS) is becoming a popular choice for natural guide star (NGS) based adaptive optics (AO) systems on account of its improved sensitivity and reduced spatial aliasing effect compared to the Shack-Hartmann WFS. However, AO system performance degradations due to implementation errors regarding the Pyramid WFS are still not very well understood. These include the quality of the pyramid (sharpness of the edges and vertex, orthogonality of the faces), image quality of the relay optics, pyramid-to-detector misregistration, etc.. Understanding these effects is crucial to setting proper tolerances, deciding which degree of freedom should be adjustable and optimizing the performance to cost ratio. In this paper, we present our simulations methods and results on the pyramid WFS and these implementation errors for the NGS AO mode of the Thirty Meter Telescope (TMT) facility Narrow Field Infrared AO System (NFIRAOS). We have found that the AO system performance is mildly sensitive to certain effects, such as pupil image distortion or drift, but more sensitive to other effects, such as pyramid imperfections, etc.. For the TMT pupil geometry, we find that it is also important to update the control matrix as the telescope pupil rotates on account of the pupil segmentation due to the M2 support struts.

1. INTRODUCTION

Pyramid WFS (PWFS)[2] is increasingly popular in modern NGS based AO systems due to its increased sensitivity compared to Shack-Hartmann WFS. In this paper, we describe the modeling and tolerance study results of PWFS for TMT NFIRAOS. Section 2 describes our numerical modeling approach for the PWFS. Section 3 describes the optical gain optimization. Section 4 describes the implementation error sources and tolerance. Finally Section 5 shows the importance of correct pupil function in the interaction matrix for optimal performance.

2. NUMERICAL MODELING

2.1 Pyramid

Inverse Fourier Transform (FFT) on the complex point spread function (PSF) gives the pupil function. In order to simulate the effect of the Pyramid, we add a tip/tilt OPD error onto each quadrant of the complex PSF, such that the pupil function from each quadrant of the PSF is shifted on the detector . Let's take the first quadrant for example. What we want to shift is exactly $W/4$ along both x and y axis, where W is the size of of the FFT grid, which is twice or more the diameter of the pupil to avoid aliasing and cross talk. In the following equations, we determine how much tilt should be introduced by the pyramid.

The complex PSF Φ is obtained from FFT of the input complex pupil function ψ

$$\psi(x,y) = A(x,y)e^{\frac{2\pi i}{\lambda}\varphi(x,y)} \tag{1}$$

$$\Phi(k_x, k_y) = \hat{\psi}(\frac{k_x}{\lambda}, \frac{k_y}{\lambda}) = \int \psi(x,y)e^{-\frac{2\pi i}{\lambda}(xk_x+yk_y)}dxdy \tag{2}$$

*Send correspondence to lianqiw@tmt.org. AO4ELT5, Tenerife, Spain. June 2017

where A is the amplitude function and φ is the OPD, all defined at points (x, y). The inverse FFT on the PSF give the complex pupil function:

$$\psi(x, y) = \int \Phi(k_x, k_y) e^{\frac{2\pi i}{\lambda}(xk_x + yk_y)} \frac{dk_x dk_y}{\lambda^2} \tag{3}$$

In order to shift $\psi(x, y)$ by (d_x, d_y) amount

$$\psi'(x, y) = \psi(x - d_x, y - d_y) = \int \Phi(k_x, k_y) e^{\frac{2\pi i}{\lambda}((x - d_x)k_x + (y - d_y)k_y)} \frac{dk_x dk_y}{\lambda^2} \tag{4}$$

$$= \int \Phi(k_x, k_y) e^{-\frac{2\pi i}{\lambda}(d_x k_x + d_y k_y)} e^{\frac{2\pi i}{\lambda}(xk_x + yk_y)} \frac{dk_x dk_y}{\lambda^2} \tag{5}$$

all we have to do is to apply tilt of $T(k_x, k_y) = e^{-\frac{2\pi i}{\lambda}(d_x k_x + d_y k_y)}$ to the complex PSF. Set d_x and d_y to $W/4$ or $N\Delta x/4$ where N is FFT array size and Δx is sampling of the pupil. Given that $k_x = \frac{\lambda p}{N\Delta x}$ and $k_y = \frac{\lambda q}{N\Delta x}$, the term is simply

$$T(p, q) = e^{-\frac{\pi i}{2}(p + q)} \tag{6}$$

where p, q are the index in the PSF points, with $(0, 0)$ at the PSF center. The term for other quadrants can be easily replicated by multiplying a factor of 1 or -1 to p or q depending on the quadrant. Modulation of the pyramid can be modeled by samply laterally shifting $T(p, q)$.

For real optical systems, the spatial frequency (k_x, k_y) are related to the focal plane location (x, y) by the focal length f. The tip/tilt functions can be rewritten as

$$T(k_x, k_y) = e^{-\frac{2\pi i}{\lambda}(\frac{d_x}{f} fk_x + \frac{d_y}{f} fk_y)} \tag{7}$$

or

$$T'(x, y) = e^{-\frac{2\pi i}{\lambda}(d_x x/f + d_y y/f)} \tag{8}$$

with $x = fk_x, y = fk_y$, and $d_x = d_y = D/2$, where D_{optics} is the diameter of the pupil. The slope of the pyramid would be $1/[2(n-1)f/\#]$ where n is the refractive index of the pyramid, and $f/\# = f/D$ is the f-number of the optics preceding the pyramid. A slightly larger angle can be used to further separate the pupil. For a $f/45$ system, the angle is only 1.3 degrees if $n = 1.5$. A double pyramid is often used to make each angle larger but has to compensate for chromatic aberration in the mean time.[1]

2.2 Gradients

After obtaining pupil images on the four quadrants, the gradients can be computed simply using quad cell algorithm with four values from corresponding pupil locations in each quadrant. let $I_{ij}[k]$ be the intensity value at (i, j) at "subaperture" k, with i, j range from 0 to 1. The gradient at subaperture k is

$$g_x[k] = \alpha \frac{I_{10}[k] + I_{11}[k] - I_{00}[k] - I_{01}[k]}{\Sigma_{ij} I_{ij}[k]} \tag{9}$$

$$g_x[k] = \alpha \frac{I_{01}[k] + I_{11}[k] - I_{00}[k] - I_{10}[k]}{\Sigma_{ij} I_{ij}[k]} \tag{10}$$

where α is the gain, determined by applying a known mode to the OPD and scaling the gradients to match the input mode. To the first order, the subaperture total intensity is independent of the wavefront. We can use this property to linearize the quad cell algorithm to make it more noise resilient. Let the averaged subaperture total intensity be $I_0 = \left\langle \Sigma_{ij} I_{ij}[k] \right\rangle_k$, and define area of subaperture k as the normalized subaperture intensity:

$$A[k] = \frac{\Sigma_{ij} I_{ij}[k]}{I_0} \tag{11}$$

We can rewrite the gradient computation as

$$g_x[k] = \frac{\alpha}{I_0 A[k]} \left(I_{10}[k] + I_{11}[k] - I_{00}[k] - I_{01}[k] \right). \tag{12}$$

For instantaneous measurements with noise, I_0 is determined in real time, while $A[k]$ is determined during calibration.

2.3 Noise Covariance Matrix

From Eq 12 we can determine the gradient measurement noise. Assume the noise in each pixel is not correlated and ignore the noise in the average subaperture intensity, we have:

$$\langle g_x[k] g_x[k] \rangle = \left(\frac{\alpha}{I A[k]} \right)^2 4\sigma_{I,k}^2 \tag{13}$$

where $\sigma_{I,k}^2$ is the noise in each pixel of subaperture k due to photon and read out noise.

2.4 Interaction Matrix

The deformable mirror (DM) to WFS interaction matrix from can be computed by poke an actuator each time and compute the gradients from the resulting optical path different (OPD).

The interaction matrix is dependent on the optical gain of the system. If we compute the DM actuator to WFS interaction matrix G_A without turbulence. For a certain DM shape a, let the measurement be g,

$$g = G_A a \tag{14}$$

and computed DM error vector be $\hat{a}$, using least square reconstructor,

$$\hat{a} = G_A^\dagger g \tag{15}$$

When turbulence is present, and assume the real optical gain is decreased by a factor of α, while the old gain value is still used, the gradient measurement will be increased by the same factor

$$g' = \alpha g \tag{16}$$

the computed DM error vector is now

$$\hat{a}' = G_A^\dagger g' = \alpha \hat{a} \tag{17}$$

To fix the problem, we use dithering mechanism to inject a sinusoidal tip/tilt onto the FSM, and measure the response of the WFS to determine the new optical gain. Assume we can determine the gain decrease factor α, we can adjust the optical factor using inside pyramid WFS gradient algorithm by a factor of $1/\alpha$ to restore the correct gradient measurement

$$g'' = \alpha g / \alpha = g. \tag{18}$$

There is no need to modify the reconstructor, as it is using the correct actuator to subaperture gradient information, which is independent on optical gain.

3. OPTICAL GAIN

Non-Common Path Aberration (NCPA) refers to differential aberrations along WFS optics path and science optics path usually caused by optics following WFS beam splitter. We need to apply an offset on the DM so that the aberrations in science path is corrected optimally. The DM OPD offset plus any aberration in the WFS path creates a net OPD offset in front of the Pyramid WFS, which consequently measures an average gradient offset. By subtracting this as a reference gradient, we can have the pyramid WFS measure zero average gradient and therefore maintain the DM OPD offset.

The following measurement model describes the PWFS process:

$$s = g\nabla\phi_{atm} + g'\nabla\phi_{ncpa} + g''\nabla\phi_{dither} \tag{19}$$

Here, various $\nabla\phi$ represent the gradient of wavefront measured by PWFS without external turbulence affect and the various g describes the gain variation when measuring turbulence or NCPA aberrations. Although not spelled out explicitly, g can also a function of the wavefront aberration strength, especially when RMS wavefront error is large. The last component $\nabla\phi_{dither}$ is the gradient of the dither signal, which may be small amount of tip/tilt or astigmatism.

Due to non-linear effect, the optical gain of the PWFS depends on the 1) Zernike order of the aberration, 2) strength of the aberration, and 3) present of any external aberration. We take three different modes and study their interaction in PWFS measurement: 1) 70 nm of astigmatism, 2) 150 nm of polishing error with -2 power law, and 3) one instance of closed loop turbulence, all of which are defined on TMT pupil.

Figure1 shows the results for different amount of those modes. The left panel shows the measurement of astigmatism mode. The blue curve is for just various amounts of astigmatism, without any external aberration. The response is linear with gain of 1. The red curve shows the results when an constant amount of closed loop turbulence is added as external turbulence. The response is again linear, but the optical gain is now close to 0.5. The green curve shows the result when proportional percentage of polishing error is added as external aberration. The gain starts at around 1, but rolls off after about 40% and is close to 0.6 at 100% aberration. The light blue curve is with both proportional percentage of polishing error and constant turbulence. The gain starts at around 0.5, as with the red curve, but again roll off due to increasing polishing error. The last orange curve is with proportional percent of both polishing error and turbulence. The gain starts at around 1, but rolls off rapidly. From above, we can conclude that the PWFS response to astigmatism mode is linear when no or constant external turbulence is present, but with gain decreasing with increasing strength of external turbulence.

The middle panel shows the measurement of polishing error. For the standalone (blue curve) measurement, the gain starts at 1, but rolls off after 40% aberration. When turbulence is added, the gain decreases again. However, when astigmatism is added, the response does not change, as shown in the green curve which is overlaps the blue curve, and the red curve which overlaps with the cyan curve.

The right panel shows the measurement of closed loop turbulence. For the standalone (blue curve) measurement, the gain starts at 1, but rolls off after 40% aberration. When astigmatism is added, the response does not change (curve not shown), but when polish error is added, the gain is further reduced.
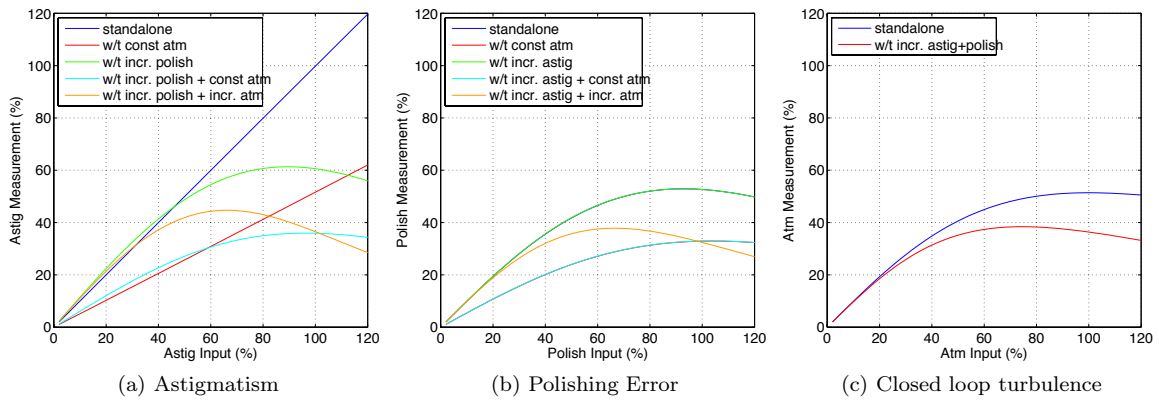

(a) Astigmatism   (b) Polishing Error   (c) Closed loop turbulence

Figure 1: Optical Gain Linearity with $5\lambda/D$ radius modulation

3.1 Determine Optical Gain

From 19, the gradient of turbulence $\nabla\phi_{atm}$ is related to the measurement s as

$$\nabla\phi_{atm} = \frac{1}{g}(s - g'\nabla\phi_{ncpa} - g''\nabla\phi_{dither}). \tag{20}$$

By alternating the dither signal sinusoidally and making use of synchronous detection, we can determine the optical gain of the dither signal at present condition:

$$g'' = \frac{\left|\langle s(k)e^{2\pi ik/N}\rangle\right|}{\left|\langle\nabla\phi_{dither}e^{2\pi ik/N}\rangle\right|}, \tag{21}$$

where N is the number of points in each dither cycle. The contribution of dither single in 20 can be ignored if it is very small, as in the case of common path dithering.

There is no easy way to estimate g, but notice that it (as shown in red curve of Figure1(b)) is fairly close to the gain of g'' (as shown in orange curve in Figure 1(a)), we can substitute it by g''. The measurement is now simply as

$$\nabla\phi_{atm} = \frac{s}{g''} - \frac{g'}{g''}\nabla\phi_{ncpa}. \tag{22}$$

Using the Pyramid WFS model or lab measurement, we can readily determine the gradient of the NCPA offset $\nabla\phi_{ncpa}$, however the gain g' still needs to be determined.

For low Zernike order NCPA aberrations (like astigmatism), the PWFS has large linear dynamic range, we can safely assume $\frac{g'}{g''} = 1$. The measurement is simplified as

$$\nabla\phi_{atm} = \frac{s}{g''} - \nabla\phi_{ncpa}. \tag{23}$$

This also applies to weak high order NCPA aberrations (≤ 70 nm -2 power law) when the response is still linear.

3.2 Simulations Results

Figure 2 shows the RMS wavefront error versus the amount of NCPA along science path. Fitting error is removed when computing WFE. The blue curves show dithering with tip/tilt on the PWFS modulation mirror (non-common path dithering). The red and green curves shows dithering using an astigmatism mode of 10 nm RMS on the DM with 4 (green) or 20 (red) points in a sinusoidal scaling curve. In general, the performance penalty due to NCPA is negligible for polishing error $<$ 70 nm or low order Zernike mode (e.g., astigmatism) $<$ 200 nm. Beyond this, the performance starts to gradually degrade.

Figure 3 plots the variation of optical gain as a function of the Fried parameter r_0. The optical gain decreases with stronger turbulence (smaller r_0) as well as in the presence of NCPA. Larger modulation ($10\lambda/D$) tends to result in higher optical gain and less sensitivity to NCPA due to its increased linearity.

Finally, Figure 4 shows the performance variations against guide start brightness at median seeing conditions ($r0 = 0.18m$). NCPA of 70 nm astigmatism and 150 nm of polishing error is including in the science path. The Pyramid WFS gains around 1 magnitude over SHWFS under similar conditions.

In summary, with such significant variation of the optical gain, it is critical to determine the optical gain during observation to achieve optimal performance.

4. IMPLEMENTATION ERRORS

In this section we describe various implementation errors of the Pyramid WFS, the method to model them in simulation, and the effects in wavefront error.
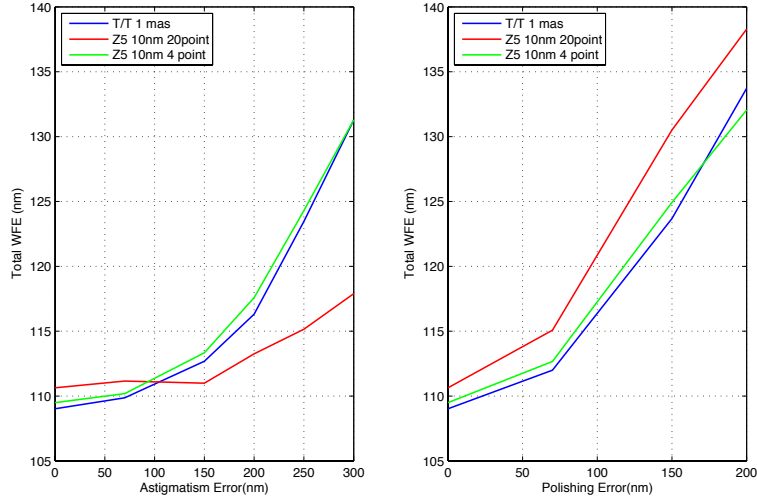
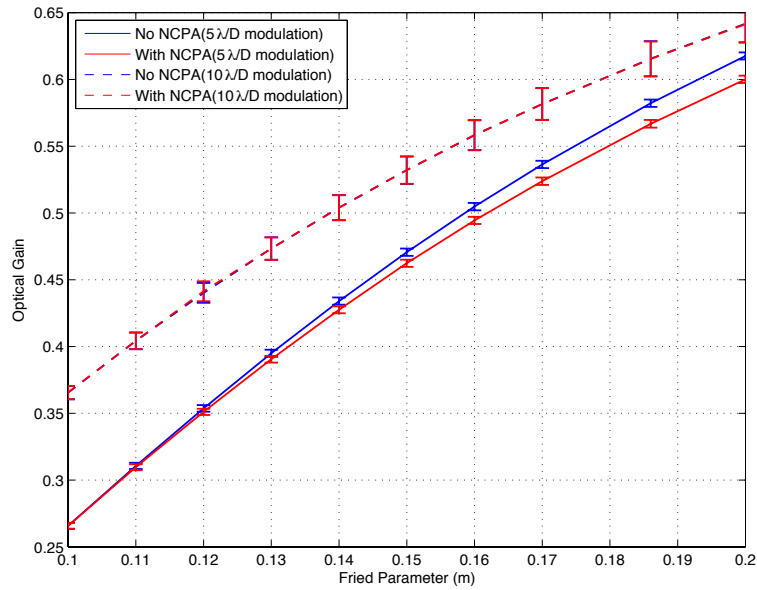Figure 2: RMS WFE versus amounts of NCPA along science path. Fitting error is removed when computing WFE.

Figure 3: Optical Gain versus r_0.

4.1 Implementation Effects

1. Pupil image location and size

   (a) error in pyramid vertex angle affects distance between sub-pupil,

   (b) mis-registration error between PWFS and detector,

   (c) magnification error.

2. Pupil drift

3. Common pupil distortion caused by NFIRAOS re-imaging optics

4. Differential pupil distortion due to radial elongation caused by optics aberration.
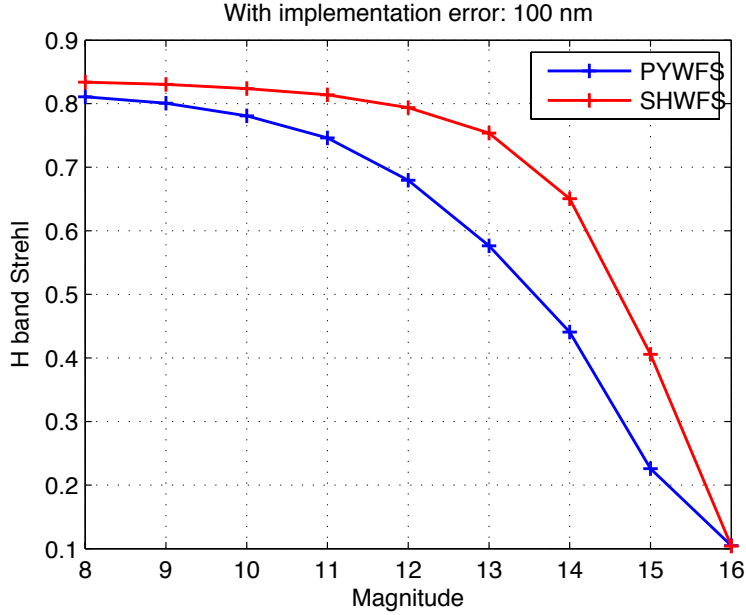
5. Pupil blur

Figure 4: H band Strehl comparison of Shack-Hartmann WFS and Pyramid WFS at median seeing.

Implementation Error Terms	Tolerance	Impact to WFE
Common pupil distortion	0.2 pix RMS	0
Differential pupil distortion	0.4 pix p/v	16
Pupil image location and size	A few pixels	12
Round edge and vertex	22 micron flat edge/vertex	16
Pupil image quality (blur)	0.6 pix FWHM	15
Charge diffusion (blur)	0.7 pix FWHM	17
Modulation error	<2 uas	0
Pupil drift	<0.1 pixel	0
Total		35

Table 1: Implementation error terms, tolerance and impact to RMS WFE.

   (a) relay optics blurring

   (b) optics chromatic aberration

   (c) detector charge diffusion.

   (d) modulation mirror smearing (mirror not on pupil)

6. Pyramid round edge and vertex

7. Focal plane effects

   (a) modulation error

   (b) symmetric and asymmetric extended source

   (c) ADC residual

Table 1 shows the implementation errors, the tolerance and impact to RMS WFS.
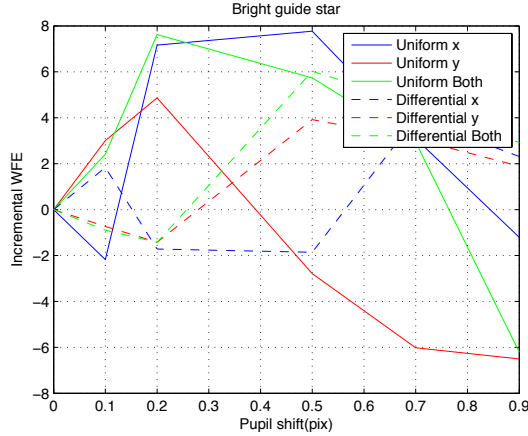
Figure 5: Incremental WFE due to shift between pupil and pixels

4.2 Modeling Approach and Results

4.2.1 Pupil Image Location

We model these effects by shifting the pixels when sampling the pupil onto detectors. Shift larger than half a pixel are subtracted by the rounded integer to simulate redefining of the pupil. The interaction matrix is measured during simulation and therefore compensates automatically for the error. The only effect is possible reduced sensitivity.

Figure 5 shows the incremental wavefront under different amount of pupil shifts. The solid curves are for uniform pupil shift along x, y or both axis for the indicated amount in the x axis. The dash curves are for differential pupil shifts where different sub-pupil are shifted to different directions. Overall the impact to wavefront error is limited to 8 nm maximum. We didn't simulate magnification error of the pupil. But because we oversampled the sub-pupil (96x96, whereas the DM is only order 60x60), the effect is expected to be small.

We allocated 12 nm for these effects in the error budget, and set the pupil position tolerance to a few pixels.

4.2.2 Pupil drift

We rely on PWFS to measure the pupil location and feed back to the telescope control system to reposition the pupil, which requires PWFS pupil stability to be much less than a percent. The effect on gradient measurement is therefore negligible.

4.2.3 Common pupil distortion

The common pupil distortion effects all sub-pupil in the same way and will therefore not impact gradients measurements. The distortion is accounted for in the interaction matrix and will have negligible effect in wavefront error. We set the tolerance to 0.2 pixel RMS.

4.2.4 Differential pupil distortion

We modeled the pupil elongation effect by compressing the detector pixel coordinate along the radial direction. Table 2 shows the incremental wavefront error caused by various amount of pupil elongation. Here, a value of 1 is equivalent to pupil being elongated by $2\sqrt{2}$ pixels along the radial direction. The factor of 2 is for both sides, and $\sqrt{2}$ is due to diagonal effect. We set the tolerance to 1.5 pixel p/v, and the corresponding impact to wavefront error is around 18 nm.

Maximum Offset (pixel)	Incremental WFE (nm)
0.5	-2.3
0.7	5.2
0.8	9.5
1	7.3
1.5	18.3
2	47.1

Table 2: Incremental WFE due to pupil elongation

RMS blurring (pixel)	Incremental WFE
0.1	2.1
0.2	4.0
0.3	7.6
0.4	9.7
0.5	12.0
0.6	14.7
0.7	17.0
0.8	19.9
0.9	22.7
1.0	25.4

Table 3: Incremental WFE due to blurring.

4.2.5 Pupil Blurring

The various blurring effects are modeled as by convolving the image with a Gaussian blurring function with certain FWHM, just like charge diffusion. Table 3 shows the incremental WFE due to increasing amount of blurring effect. We specify the effect of detector charge diffusion to 0.7 pixel FWHM and everything else to 0.6 pixel FWHM RMS.

4.2.6 Round pyramid edge and vertex

We modeled this effect by modifying the OPD of the pyramid along the vertex and edge of the pyramid. The PSF in the simulation has a sampling of 1.7 milli-arcseconds (mas) per pixel. We simulated a pyramid with 3.4 mas flat surface along the edge and tip, corresponding to 22 micron[2]. The incremental wavefront error obtained during simulation is 16 nm. The effect for the smaller imperfection will be even smaller.

4.2.7 Focal plane effects

The modulation error is caused by asymmetry of the modulation circle. Due to the tight requirements of creating dither pattern of tip/tilt, the tolerance has been set to $0.003\lambda/D$, which results in negligible tip/tilt error. The extended source and ADC residual can be modeled as larger modulations in filled circle. Figure 6 shows the effect of symmetric extended object on performance of classic AO mode, with FWHM of the object shown in the legend (arcsec). A size of 2 arcsec results in 2-magnitude penalty for the same performance level.

Asymmetric objects may result in higher order effects after tip/tilt error has been corrected because non-linear response of different subapertures. The effect will be studied and reported in a future paper.

5. INTERACTION MATRIX
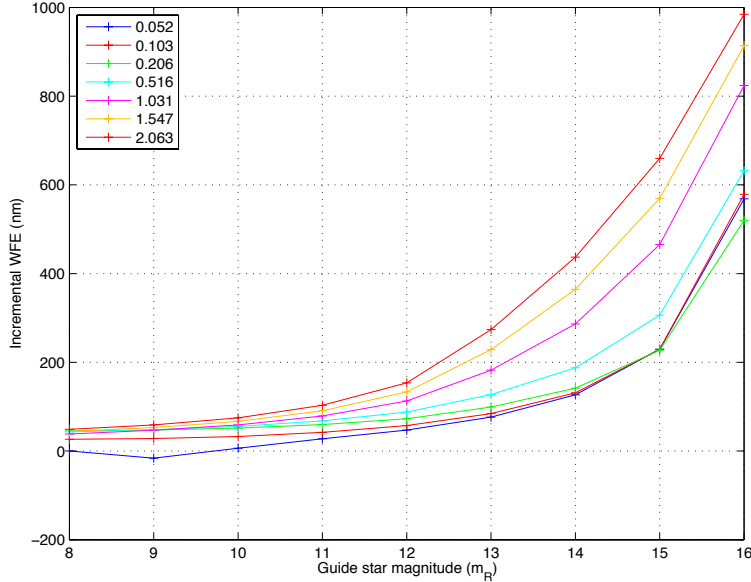
---

[2]The focal ratio at the pyramid is f/45.

Figure 6: Impact on classic AO performance with symmetric extended object (FWHM in arcsec shown in legend).

Pupil in calibration	Incremental WFE (nm)
TMT_10%	14.2
TMT_20%	19.4
Circular	30.9
Circular_10%	35.2
Circular_20%	36.6

Table 4: Incremental WFE caused by pupil illumination mismatch

5.1 Pupil Function

If the PWFS interaction matrix is calibrated using NFIRAOS calibration sources, there will be two important mismatch between the pupil illumination during calibration versus on sky observation:

1. There is no internal TMT pupil mask within NFIRAOS. The pupil will therefore be circular.

2. The illumination on the pupil (DM0) with NFIRAOS calibration sources are not uniform, with 10% to 20% drop on the edge compared to the center.

We simulated the following cases:

1. baseline: use TMT pupil mask with uniform illumination during calibration.

2. TMT_10%(20%): use TMT pupil mask with 10% (20%) drop in illumination on the edge during calibration.

3. Circular: use circular mask with uniform illumination during calibration.

4. Circular_10%(20%): use circular mask with 10% (20%) drop in illumination on the edge during calibration.

Table 4 shows the incremental wavefront error of those effects. It is apparent that mismatch between TMT pupil and the circular pupil gives more error than pupil illumination uniformity. Our solution is to derive pupil distortion from the measured interaction matrix, and incorporate it into numerically computed the interaction for use on sky.

Pupil angle difference (degree)	Incremental WFE (nm)
0.4	5.8
0.5	9.1
0.6	10.6
0.7	170.5
0.8	174.3
0.9	178.2
1	183.3

Table 5: Incremental WFE due to pupil angle mismatch in interaction matrix

5.2 Pupil Rotation

With the elevation/azimuthal mount and irregular shape of the primary mirror, the pupil illumination rotates with telescope zenith angle. If the PWFS interaction matrix is calibrated using a certain pupil rotation angle and used at another angle, there will be a mismatch of illumination for certain subapertures near the edge or covered by secondary mirror support struts. The secondary mirror support strut is 25 cm in width and causes severe measurement bias for the partially covered subapertures. It causes significant illumination pattern change in less than 1 degree of pupil rotation.

We carried out simulations to quantify the incremental wavefront error due to using interaction matrix measured/computed at a slightly different pupil rotation angle. Table 5 shows the results. We can see that the incremental wavefront jumps from negligible to unacceptable after 0.7 degrees of pupil rotation. It turns out there is a significant piston difference between sections of the pupil separated by the secondary mirror support struct.

With the telescope zenith angle range of 0 to 65 degrees, the TMT pupil can rotate from 0 to ~65 degrees in NFIRAOS input. With 0.5 degree update separation, we have 130 measurements to make. The worse case input pupil rotates 15 arcseconds/sec, which is 120 seconds for 0.5 degrees.

6. CONCLUSION

In this paper, we summarized the simulations result on PWFS implementation error sources, tolerances, and error budgets. We also discussed the importance of correct pupil function in the interaction matrix.

REFERENCES

[1] E. Pinna S. Esposito A. Tozzi, P. Stefanini. The double pyramid wavefront sensor for lbt, 2008.
[2] Roberto Ragazzoni. Pupil plane wavefront sensing with an oscillating prism. *Journal of Modern Optics*, 43(2):289–293, 1996.