NGAO trade study report: LOWFS architecture
KAON 487 (WBS 3.1.2.2.9)

Ralf Flicker (*rflicker@keck.hawaii.edu*)
W.M. Keck Observatory, 65-1120 Mamalahoa Hwy., Kamuela, HI 96743, USA

Viswa Velur (*vnv@astro.caltech.edu*)
Caltech Optical Observatories, CIT, Pasadena, CA 91125, USA

24 May 2007

Abstract

The Shack-Hartmann (SH) and Pyramid (PYR or PWFS) type wavefront sensors (WFS) are evaluated as candidate architectures for the NGAO low-order wavefront sensor (LOWFS). The PWFS is seen to outperform the SH in the majority of scenarios tested in this study, with the notable exception of the single-sub-aperture case in poor image sharpening (250 nm residual high-order error). For a 2×2 sub-division of the pupil, measuring and correcting the five first Zernike modes (Z2-Z6), the pyramid sensor architecture was found to offer significantly superior performance over the Shack-Hartmann over a wide range of conditions. The level of high-order image sharpening of the LOWFS NGS is seen to be a critical parameter that needs to be given priority in the design. A preliminary LOWFS architecture choice is presented, and some initial cost estimates are also presented.

1 Introduction

From the NGAO WBS dictionary definition of 3.1.2.2.9 LOWFS (low-order wavefront sensor) architecture:

> "Consider the cost/benefit and technical maturity of MEMS-based correction within the LOWFS, using MOAO control techniques. Include consideration of additional metrology systems required, if any. Compare with cost/benefit of MCAO system to provide tip/tilt star sharpening. Complete when LOWFS requirements and sky coverage estimates have been documented."

As explained in the work-scope planning sheet, the scope of this trade study has been redefined and de-scoped several times:

> Some modifications to the scope of 3.1.2.2.9 have been made, which tie in with corresponding modifications in 3.1.2.2.10 (cf. that WSPS). Modifications in 3.1.2.2.9 include:
>
> - Include WFS type (SH vs PYR) in performance/cost TS
> - Resulting sky coverage will not be estimated initially in this TS. The WFS architectures are here compared on a 1-on-1 basis, under the assumption that whichever performs the better at a given NGS magnitude will provide the greater sky coverage.

To summarize, what is reported on in this document is a direct comparison of the Shack-Hartmann and Pyramid sensors in a limited set of specific settings, including the number of sub-apertures and corrected modes, and the level of partial high-order correction on the LOWFS. It deserves to be emphasized that this is not an exhaustive system optimization or design quest. The goal being here to distinguish between the competing architectures on equal terms, only a small number of critical parameters were varied and optimized (i.e. the frame rates and modal loop gains), while the majoriry of other parameters assumed some fixed values that were admissible, but almost certainly sub-optimal. This only means that, if studied by themselves, each system configuration could presumably have been optimized somewhat further, and what is presented in this report is not to be taken as the best possible performance achievable in either architecture.

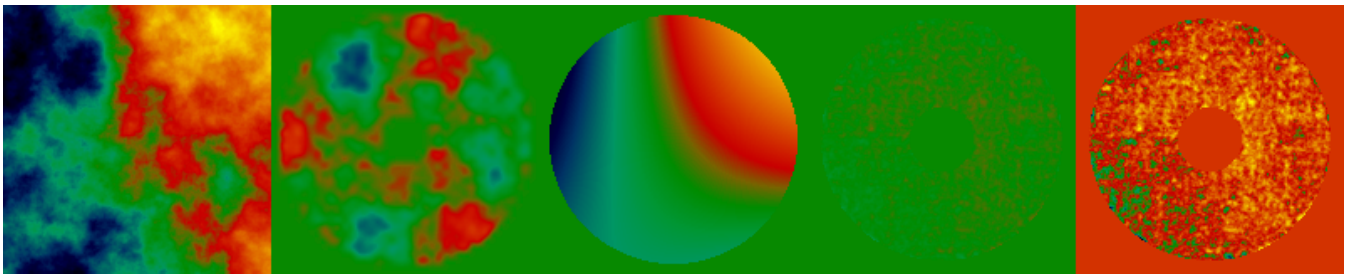

Figure 1: Sample wavefront maps output from the LOWFS simulation tool. From left to right: uncorrected wavefront; high-order DM shape; LOWFS modes shape; the residual wavefront error to scale with the images to its left; the same (res. wfe.) rescaled to the range of the color table. The above simulation used a 2×2 SH LOWFS operating at 143 Hz correcting five Zernike modes (tilt to astigmatism) on a $J = 17$ NGS.

1.1 Methodology

In order to carry out this investigation a new simulation code was written, which constitutes one of the deliverables of this trade study. The code, written in IDL, follows the traditional Monte Carlo type modeling of adaptive optics, but with an emphasis on the LOWFS modeling, which enjoys a number of advanced features that allows the LOWFS to be simulated with a high degree of realism. These features include:

- A high-order WFS and DM delivers a partially corrected wavefront on the LOWFS in closed loop operation.
- multiple WFS types: Pyramid, Shack-Hartmann, linear gradient
- multiple sub-aperture geometries: square, hexagonal, keystone
- detailed detector model
  - photon noise: source (Poisson), sky (Poisson)
  - electronic noise: read-out (Gaussian, digitized), dark current (Poisson)
  - charge diffusion (any additional OTF, photonic or electronic, may be included)
- conditioning of WFS measurements:
  - linearization within the dynamic range
  - constrained outside of the dynamic range
- multi-spectral wavefront sensing for the SH and PYR WFS types, combining any number of a predefined set of photometric wavelength bands (see table 2).
- field stops for spatial wavefront filtering and limiting the sky background

Figures 1 and 2 show some details from running the code. Fig. 1 displays the turbulent wavefront, the high-order and low-order DMs, and the residual wavefront, hile Fig. 2 shows a time history of the low-order modal command signals and residuals (left), and an instantanous detector read out (right). IDL code available upon request from the authors (it requires some non-standard packages).

2 Wavefront sensor modeling

While the low-order subsystem of NGAO will most likely employ multiple low-order wavefront sensors that jointly measures and reconstructs the tomographic null-modes, that composite system is not simulated here. The impact of residual null-modes are addressed elsewhere, and in this stude we examine the properties of the LOWFS in a stand-alone scenario.

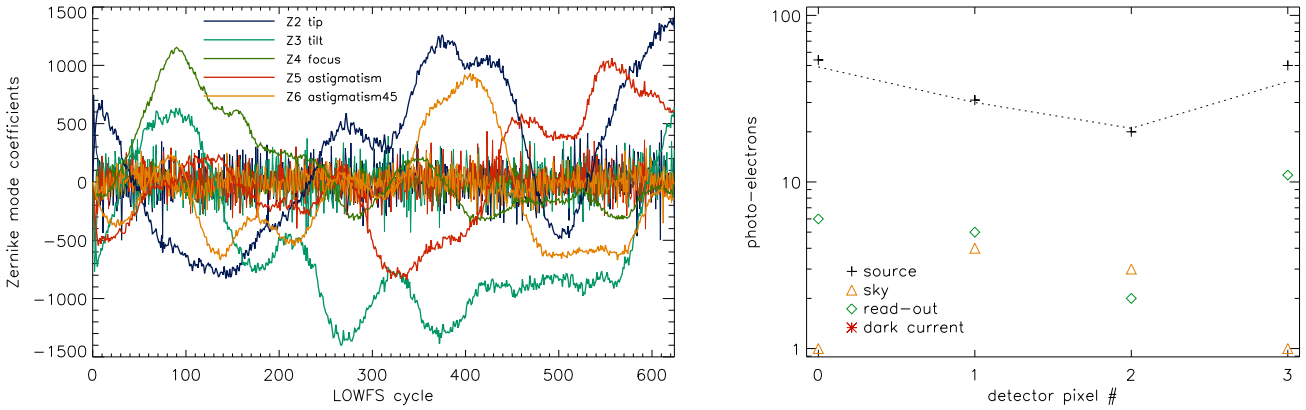

Figure 2: Detail from the LOWFS simulation tool (same configuration as in Fig. 1). Left: history of Zernike coefficients estimated by the LOWFS (thick lines) and the error signal (thin lines). Right: noise details of one sub-aperture, showing the instantaneous levels of source, sky, read-out and dark current photo-electrons. Dotted line is the final detector signal.

2.1 Shack-Hartmann (SH)

The Shack-Hartmann sensor implementation is a straightforward near-field approximation model that extracts the phase over each sub-aperture and propagates it to form a spot on the detector independently of neighboring sub-apertures. This approximation does not correctly model the nonlinear behavior for large spot excursions (large local slopes over the sub-aperture), and for any subdivision of the pupil plane (i.e. whenever there is more than a single sub-aperture), it also fails to account for cross-talk between sub-apertures. Neither effects are deemed to be serious for the current study: by measuring exact transfer curves and constraining the measurement signal, the effects of any spurious nonlinear behavior for large local slopes (or a large noise signal) are minimized. Centroiding is done by a [-3,-1,1,3] type weighting (i.e. projecting the spot onto a tilt mode), which for a quad-cell reduces to the standard formula (cf. e.g. Hardy 1998[6]).

2.2 Pyramid (PYR)

The Pyramid sensor (Ragazzoni 1996[8], Bauman 2003[1], Vérinaud 2004[9], Clare 2004[3]) places a pyramidal prism in the image plane that splits the guide star spot, and propagating to the pupil plane the four re-imaged pupils are collected on a detector as shown in Fig. 3. The pupil plane illumination can be described in a mathematical formalism that is similar to, and shows the close relationship to, a Foucault knife-edge test (see e.g. [9]). The pyramid sensor introduces a number of unique features and new parameters that need to be understood and optimized for a good WFS design, including the issues of modulation and signal reconstruction.

2.2.1 Signal reconstruction

Given the intensity distribution over the four re-imaged pupils, one may employ various pre-processing algorithms to the read out detector signal. This step is equivalent to the centroiding step for a Shack-Hartmann, where the pixel intensities are processed to produce a slope estimate. Three types of signal reconstruction have been suggested: slope estimation, phase reconstruction, and direct signal (i.e. no reconstruction). The direct signal approach simply takes the detector signal, normalizes it to the flux and delivers a $4N$-element (for a 4-sided pyramid) measurement vector of the direct pupil illumination, where N is the number of sub-apertures. The phase reconstruction method applies an inverse Hilbert transform to reconstruct the phase, which turns the pyramid sensor into a direct phase sensor. Finally the slope estimation method is implemented by combining the detector signals as follows:

$$S_x = (I_1 + I_2 - I_3 - I_4)/(I_1 + I_2 + I_3 + I_4) \quad (1)$$

$$S_y = (I_2 + I_3 - I_1 - I_4)/(I_1 + I_2 + I_3 + I_4) \quad (2)$$

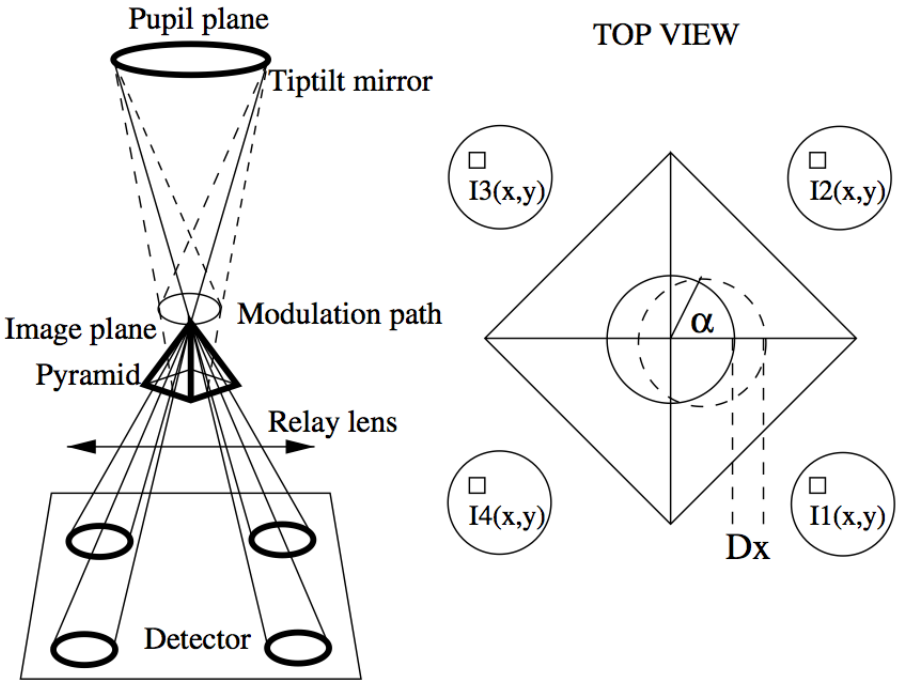

Figure 3: Principles of the pyramid sensor (figure borrowed from Vérinaud 2004).

This is equivalent to the quad-cell centroiding formula used for a sub-aperture in a Shack-Hartmann sensor. Using this formalism, the number of used measurements is halved (to $2N$) with respect to the direct signal approach. Clare 2004 (PhD thesis) shows that the direct signal approach was able to deliver superior performance under certain conditions. The current study was not able to verify this result, and found instead that the slope-based method seemed to be superior in the simulated NGAO LOWFS configuration.

2.2.2 Modulation

Modulation of the focused spot in a circular path around the pyramidal prism serves the function of controlling the sensitivity and linear range of operation of the sensor, in an analogous way to how the centroid gain needs to be adjusted for the spot size in a Shack-Hartmann sensor. For an infinitely sharp spot, the transfer function of a pyramid sensor would become a step function, and even the Airy pattern of a diffraction-limited spot will in most cases have a very limited dynamic range unless some modulation is applied to reduce the sensitivity. It can be shown (see e.g. Clare 2004[3]) that for a 4-sided prism, the optimal modulation path is actually a diamond-shaped pattern that runs parallel to the sides of the prism. For practical reasons, a circular path is usually implemented, and the relative insensitivity to the exact value of the modulation in a realistic environment (see below) suggests that no significant error is induced by doing this. The amplitude of the modulation, i.e. the diameter of the circular path, probably needs to be a dynamically adjusted parameter that adapts to the observing conditions and responds to fluctuations. While a fixed modulation could in principal be used, the sensor would be operating at a sub-optimal setting for most of the time, and this setting may not be suitable for closing the loop initially, etc.

Denoting by α the modulation amplitude in some arbitrary units and by ε a scalar metric derived from the HOWFS closed-loop measurements (e.g. the variance of the HOWFS slope vector properly normalized), one could propose to control the modulation amplitude directly as $\alpha_{k+1} = f(\varepsilon_k)$, where k is the current LOWFS integration cycle and f some scalar function to be defined. If $f = \varepsilon$ we have a direct proportional control, which will be very unstable and sensitive to spikes in ε. One could impose lower and upper limits to α by using a properly scaled and shifted sigmoid function, e.g. $f(\varepsilon) = c_0 + c_1 \tanh(c_2\varepsilon)$, where $\{c_0, c_1, c_2\}$ are fixed scalar coefficients. It is not clear to the author whether the c_1 coefficient should be positive or negative. One may argue that when the residual wavefront error incident on the LOWFS is large, turbulence itself performs part of the function of modulation by blurring the spot, and it could prove to be the best strategy to drive the modulation amplitude in the opposite direction of the

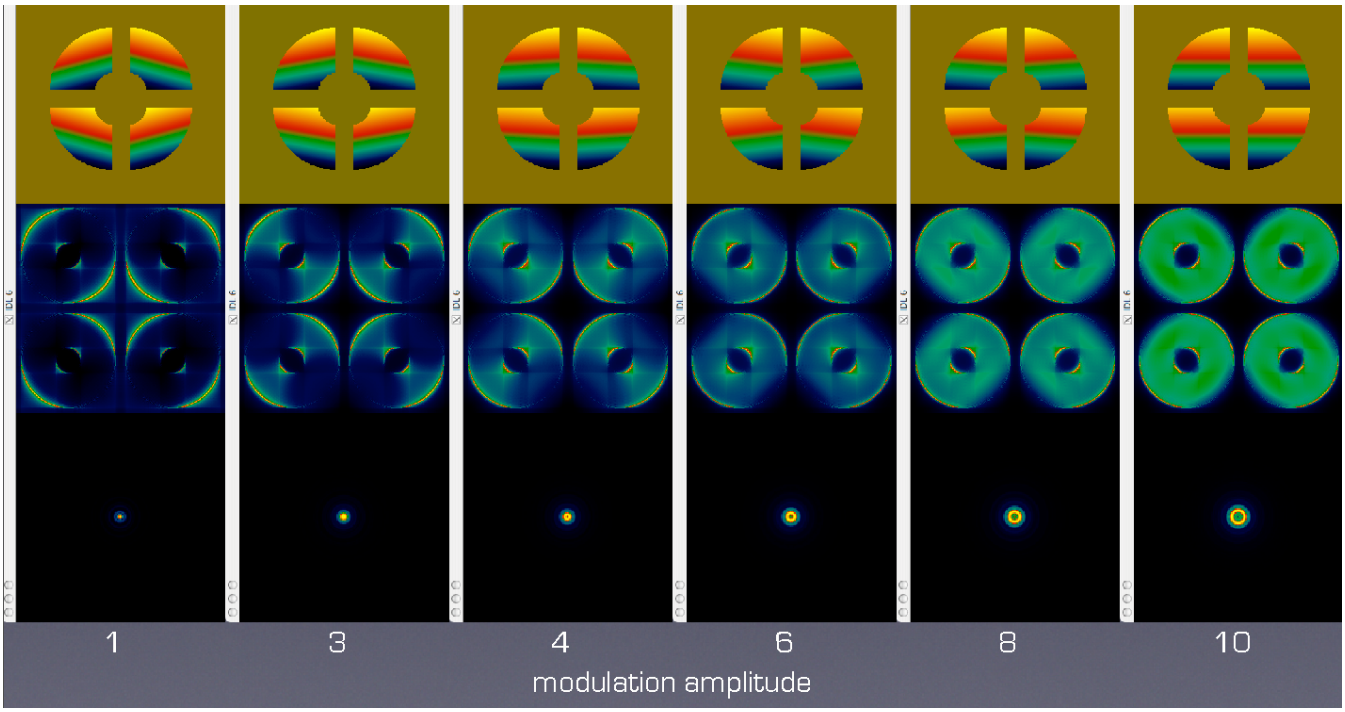

Figure 4: Effect of modulation amplitude on slope reconstruction in a 2×2 Pyramid sensor. Bottom row: modulated spot on pyramid. Middle row: intensity distribution at the re-imaged pupil plane. Top row: artificial visualization of the reconstructed slopes in the four sub-apertures.

HOWFS error metric. This is yet to be investigated in detail. One could also implement a form of leaky integrator to give the controlled variable some inertia and make it less sensitive to fluctuations, which generally tends to improve stability and overall performance. For instance:

$$\alpha_{k+1} = (1 - g_0)\alpha_k + g_0 f(\varepsilon_k), \tag{3}$$

where g_0 is a scalar gain and $1 - g_0$ is the leak factor, which needs to be fairly small in this context. The above algorithm (with $g_0 = 0.5$) is what was implemented in the current LOWFS simulation, although many other types of control laws are of course possible, and need to be explored in more detail.

One final issue concerning the modulation is its relation to the interaction matrix, which in the slope-based mode (presumably also in the direct signal mode) becomes a function of the modulation amplitude. Apart from altering the sensitivity, which could be compensated for if the slope of the transfer function was known (it is), the specific amplitude of the modulation can have a nonlinear side effect on the slope estimation when a sub-division into multiple sub-apertures is used with the pyramid sensor. What can happen at certain settings is that a pure x-tilt is not reconstructed as pure x-tilt in all the sub-apertures, but has a varying contamination of y-tilt, which when reconstructed into Zernike modes can give a contamination from higher-order modes (e.g. focus, astigmatism). This effect is illustrated in the sequence of images in Fig. 4. It is currently not known to the author whether this effect is unique to the slope-based signal reconstruction, or if the phase reconstruction and direct signal methods would suffer from this also. One consequence of this behavior is that it looks like one should allow the reconstruction matrix to be selectable from a set of pre-computed matrices for different modulation settings. This could be done practically without too much trouble, but it was found in simulations that little or no performance gain was achieved by going through this trouble. The main result from experimenting with interaction matrices generated with varying amounts of modulation appears to be that choosing one somewhere in the middle and keeping it fixed works well for a wide range of conditions (specifically, referring to Fig. 4, the matrix corresponding to modulation setting number 5 (not shown) was chosen for the LOWFS simulations in this report, as its sensitivity was suitably in the middle range, and it estimated slopes very cleanly with much less contamination than most other settings, excepting the very large modulation settings, e.g. number 10 in the figure, where instead the sensitivity becomes problematically low).

Turbulence model	7-layer CN-M3	KAON 303
Fractional C_n^2	$[0.471, 0.184, 0.107, 0.085, 0.038, 0.093, 0.023]$	units of μ_0
Altitudes	$[0.0, 2.1, 4.1, 6.5, 9.0, 12.0, 14.8]$	km
Wind speeds	$[6.7, 13.9, 20.8, 29.0, 20.0, 20.0, 20.0]$	m/s
r_0 (at 0.5 μm)	18	cm
d_0 (at 0.5 μm)	4.4	m
τ_0 (at 0.5 μm)	3.8	ms
Outer scale L_0	30	m
Zenith angle	0	
Telescope	circular pupil	73 m^2
Diameter D	10	m
Obscuration	2.5	m
High-order AO	Shack-Hartmann	NGS or LGS
Sub-apertures N	36×36	944 active
Sub-aperture size d	27.8	cm
Piezo-stack array DM	37×37	972 active
Na-layer altitude	90	km
Integration time	1	ms
Delay (additional)	0	
Low-order AO	Shack-Hartmann or Pyramid	NGS
Sub-apertures N	1 or 2×2	
Sub-aperture size d	10 or 5	m
Corrected modes	2 (Z2-Z3) or 5 (Z2-Z6)	Zernikes
Wavefront reconstructor	SVD or MAP	
Wavelength	J+H band	see table 2
Optical transmission	0.728	Nas-det
LOWFS detector	near-IR array	HgCdTe
Quantum efficiency	0.8	
Read noise (Fowler $\times$ 4)	5	e/pixel RMS
Dark current (77K)	0.1	e/pixel/s
Pixels	2×2	quad-cell (SH)
Pixel size	0.5	arc sec (SH)

Table 1: Common and fixed parameters for LOWFS simulations.

2.3 Wavefront reconstruction

Two standard types of wavefront reconstructors were considered for the LOWFS system in this study, the SVD (singular value decomposition) and the MAP (maximum a posteriori, Bayesian) estimator. Anticipating the results to be presented in Sect. 4, we find that the Shack-Hartmann LOWFS system saw little or no benefit from using the MAP estimator, so it employed a SVD throughout, which simplifies to the pseudo-inverse solution in this case since no modes are being filtered. The PWFS employed a MAP estimator for the 2×2 case and the SVD for the 1×1. These wavefront reconstructors have the following form:

$$E_{SVD} = (G^T G)^{-1} G^T, \quad (4)$$

$$E_{MAP} = (G^T G + \gamma C_a^{-1})^{-1} G^T, \quad (5)$$

where G is the interaction ("poke") matrix, and γ is a statistical parameter depending on the seeing and WFS noise level. C_a is the covariance matrix of Zernike coefficients in either the Kolmogorov or von Karman turbulence model (in the Kolmogorov version, this matrix is sometimes called the Noll matrix, see Noll 1976[7]). Using the von Karman model that includes an outer scale, the diagonal elements of this matrix was generated according to Winker 1991[10].

Wavelength Band	central λ (μm)	FWHM (μm)	color correction	zero point flux (10^{10} photons/s)	sky brightness (mag/arcs^2)
U	0.366	0.052	-0.2	19.6	20.90
B	0.438	0.100	-0.6	68.4	22.13
V	0.545	0.083	0.0	40.6	21.99
R	0.641	0.157	0.6	54.1	20.81
I	0.798	0.154	1.3	34.2	20.25
J	1.22	0.260	2.7	26.3	19.60
H	1.63	0.290	3.8	14.2	17.09
K$'$	2.12	0.410	4.9	9.39	16.99
K	2.19	0.320	4.8	8.02	16.78
L	3.45	0.570	6.8	3.67	9.91

Table 2: Photometric system. Based on the Johnson-Cousins-Glass system from Bessel et al. (1998), as compiled and presented in the NGAO_Perf_Budget_Template.xls spread sheet (available at the NGAO Twiki site). The zero point was computed for a collecting area of 79 m^2, and is the photon flux at the Keck Nasmyth from a $V = 0$ star.

3 System and environment parameters

Tables 1 and 3 list the relevant parameters of the AO system components, the simulation environment and LOWFS scenario. The turbulence model is the standard CN-M3. Although the code has the ability to use the segmented Keck pupil, it was not used here as it does not bear critically on the LOWFS characteristics under investigation, and PSFs were not being computed in any case. The near-IR detector specifics were converged upon from consulting multiple sources, but are in general close to the Teledyne Hawaii-2RG characteristics.
(`www.teledyne-si.com/imaging/hawaii2rg.html`)

3.1 Photometric system

The photometric system used for this study is based on a template prepared for NGAO analysis, although not always used consistently by the team. There are some (minor, perhaps) differences between this system and e.g. the one used in the NGAO wavefront error budget Master spread sheet (R. Dekany), but this is not a big worry. While such differences may make comparisons between the LOWFS simulation and spread sheet predictions subject to a magnitude offset, the comparison can still be made unequivocally in terms of photo-detection events (PDEs) rather than magnitudes. The relative performance of the Pyramid and Shack-Hartmann type sensors for a given PDE level will remain unaffected if a new photometry system is adopted, and only the magnitude scale needs to be shifted.

4 Sample numerical results

The Monte Carlo type LOWFS AO simulation was run for 5000 cycles (corresponding to several wind crossing times), for each simulation scenario as specified by the LOWFS NGS magnitude, residual high-order wavefront error, and LOWFS type (SH/PYR and 1x1/2x2 sub-apertures). The level of high-order image sharpening was adjusted to three fixed settings in the following way:

- 80 nm RMS high-order error (85% Strehl Ratio in J band) was produced by making the HOWFS use a bright NGS, emulating the best possible NGAO scenario when the LOWFS NGS is at the center of a narrow-field LGS asterism with very little tomography error.
- 160 nm RMS high-order error (52% Strehl): the HOWFS used a 9th magnitude LGS coinciding with the LOWFS NGS (i.e. no field anisoplanatism, only cone effect). This emulates the situation if the LOWFS had a dedicated LGS to use solely for the purpose of wharpening up its NGS.
- 250 nm RMS high-order error (20% Strehl): the HOWFS used a NGS offset from the LOWFS NGS by 15 arc seconds. This emulates a variety of NGAO scenarios that may approach this level of high-order correction in different ways (wide-field LGS asterism, MCAO, using MOAO algorithms, etc – see e.g. the wide-field plots in KAON 429).

1×1 sub-aperture			2×2 sub-apertures		
J mag	frame rate	PDE	J mag	frame rate	PDE
16	500	294.6	14	1000	232.4
17	250	234.6	15	500	185.6
18	143	163.4	16	250	147.3
19	100	92.9	17	143	102.6
20	83	44,4	18	111	52.5

Table 3: J-band magnitudes, frame rates and resulting photo-detection events (before noise) per sub-aperture and per frame, for a single sub-aperture (left) and a 2×2 WFS (right).

In figure 5 the residual low-order wavefront error (total, from all corrected modes) of the Shack-Hartmann and Pyramid sensors are plotted versus J-band magnitude (but keep in mind sensing is done at J+H). The first thing to note is that the PWFS is seen to outperform the SH in almost every case that was tested with this code so far, although for the tip/tilt-only case the difference is slight. It is notable however that in this configuration the PWFS performs worse than the SH at the poorest image sharpening setting (20% J-band Strehl). It was observed in this scenario that large sporadic errors were induced in the 1×1 PWFS by low-order spot deformations (focus, astigmatism), that were otherwise self-corrected in the 2×2 case, giving much better performance in the latter. Next, it is seen that the 2×2 configuration is more adversely affected at bright NGS by increasing spot size, as the residual high-order wavefront error increses, than is the 1×1 system. One also sees the strong dependence on image sharpening overall, as the performance in every configuration rapidly degrades when going from 160 to 250 nm residual high-order error.

Figures 6 and 7 show a comparison between the current LOWFS simulation code and performance predictions obtained with the NGAO Master spread sheet tool for a similar system configuration. The residual high-order wavefront error in that configuration was 130 nm. The comparison in plotted both versus J magnitude (left) and versus the number of photo-detection events (PDEs) per sub-aperture, as this highlighted some of the intrinsic differences in the two setups. As is seen from the plots versus PDE, the two tools gave their performance estimations in somewhat different parts of the parameter space, with frame rates and PDEs differing significantly. While the comparison is pretty good for the 1×1 case, the 2×2 LOWFS simulation shows a much stronger effect of becoming limited by increasing spot size than the spread sheet tool does.

Around the last data points simulated in each configuration ($J = 20$ for 1×1 and $J = 18$ for 2×2), noise from the three dominant sources was observed to be roughly balanced (by calculating variance of each source). The photon noise on the source NGS at 40-50 PDE (see table 3) was seen to be of a comparable severity as the electronic read-out noise from the detector at 5 electron RMS read noise, with the sky background as limited by the $0.5''$ also being of the same order. Opening up the field stop would make performance background-limited at this end, while a detector with more than 5 electrons RMS read-out noise would make performance read-noise limited instead.

5 Summary & Discussion

Based on the results obtained within this study so far, we can make the following observations, notes and caveats:

- According to the LOWFS simulation, the PYR is seen to outperform the SH in the majority of cases that were tested, although in the tip-tilt-only case the difference is marginal.
- The level of high-order image sharpening of the LOWFS guide stars is seen to have a dramatic impact on LOWFS performance, suggesting that measures be taken to ensure some minimum level of image sharpening.
- The poor performance of both architectures at the 160-250 nm high-order wavefront error settings may be in part due to sub-optimal sensor implementations of various kinds:
  - For the Shack-Hartmann, the detector pixel size ($0.5''$) was not subjected to any rigorous optimization, but a size was chosen that was observed to give good performance in the 80 nm high-order WFE regime, i.e. for a very well-corrected spot on the LOWFS. Performance at 250 nm level may have been improved with larger pixels.

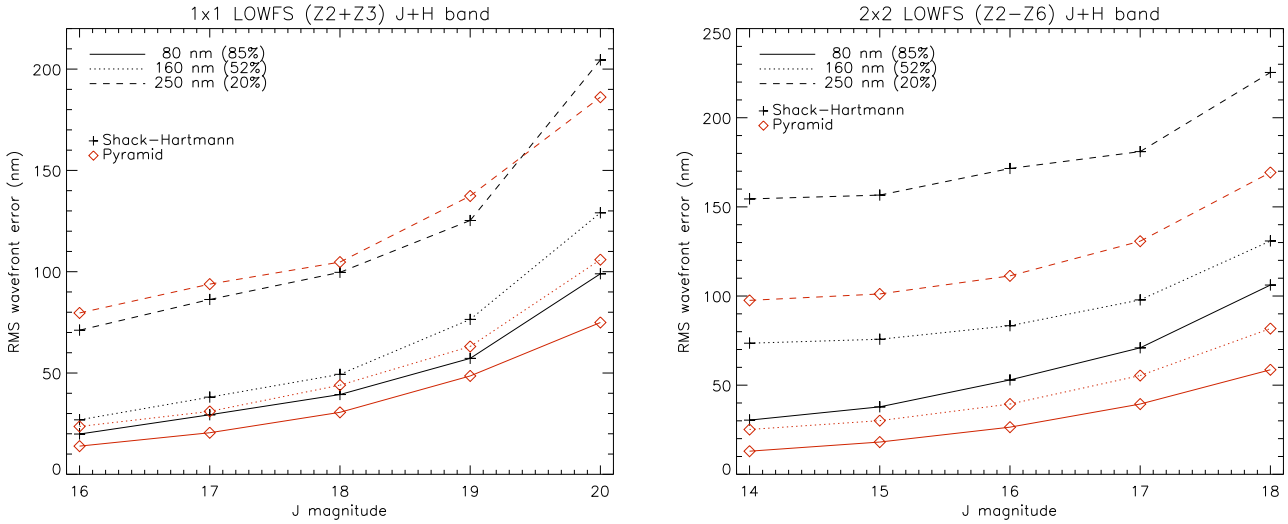

Figure 5: Residual LOWFS RMS wavefront error as a function of J-band magnitude, sensor type (SH – black pluses, PYR – red diamonds) and residual high-order wavefront error shown by different line styles (J-band Strehl ratio within parantheses). Shown on the left is the 1×1 sub-aperture tip/tilt sensor, and on the right the 2×2 sub-aperture tip/tilt+focus+astigmatism sensor.

– For the Pyramid, performance at the 250 nm HO WFE level might have been improved by using a different interaction matrix, as discussed in Sect. 2.2.2. The use of a fixed matrix was based upon performance in the 80 nm WFE regime, and it is possible that a different interaction matrix could have performed better at the 160-250 nm levels.

Based on KAON 470 [4] and the results of the current trade study, we offer the following as a preliminary recommendation for the NGAO LOWFS system architecture:

- 2 near-IR (J+H) tip/tilt sensors with built-in MEMS DM for image sharpening. These sensors may employ e.g. a STRAP type optical design adapted to a near-IR detector.
- 1 near-IR (J+H) pyramid sensor with at least 2×2 sub-apertures, and with a built-in MEMS DM for image sharpening.
- At least 1 LGS dedicated to the LOWFS system for image sharpening.

The STRAP sensor (System for Tip/Tilt Removal with Avalanche Photodiodes, see e.g. Bonaccini 1997[2]) is mentioned here because it is a well documented and routinely used device that have optical properties in common with a pyramid sensor, making it interesting for the current study. It can be regarded as a lenslet-based 1×1 pyramid sensor (Bauman 2003[1]) without modulation. Given the superiority of the PWFS compared to the SH implementation observed in simulations here, even in the 1×1 tip/tilt-only case, we will refer to a near-IR version of a STRAP type sensor as a preliminary option that could make a good candidate for the tip/tilt LOWFS.

With a 2×2 PWFS in slope mode, the configuration listed above gives a total of 12 measurements, which is enough to reliably resolve the first 9 tomography null-modes (and we believe this to be sufficient, compare KAON 470). Although it drives up the cost somewhat, in the cost/performance trade-off it seems to make sense to provide for at least one LGS dedicated to the LOWFS system, in order to be always ensured of decent performance of at least one of the sensors. For each observing scenario, based upon the HOWFS asterism and the geometry of candidate LOWFS NGSs, one would direct the LOWFS LGS to the NGS that would benefit the scenario the most, and have the remaining two low-order sensors drive their MEMS DMs from HOWFS telemetry using MOAO reconstruction algorithms. How such a decision matrix might look is beyond the scope of this study, but to a zeroth-order approximation, although potentially over-simplifying the situation, it might be an acceptable strategy to always use the LOWFS LGS for the 2×2 PWFS in order to ensure a minimum level of performance by this comparatively

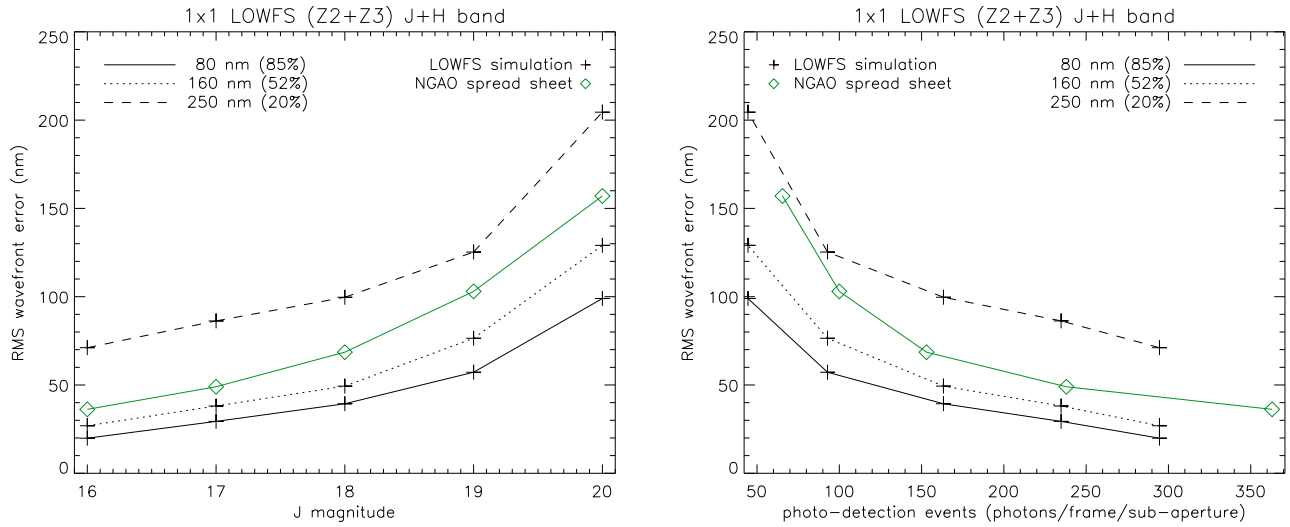

Figure 6: Comparison between the LOWFS simulation (black pluses) and the NGAO Master spread sheet (green diamonds) for the 1×1 case. The spread sheet calculation used a high-order residual RMS wavefront error of 130 nm. Plotted versus J band magnitude on the left, and versus PDEs on the right, which illuminates the differing bandwidth parameters and possible differences in the photometry system.

important low-order sub-system constituent. This strategy is also suggested by the observation that the 2×2 sensor takes a larger hit from spot size on bright stars than does the 1×1 system.

5.1 Technical notes

Field stop. The use of a $0.5''$ field stop throughout all the scenarios reduced the noise from sky background illumination substantially. The field stop could not be made much smaller than this, or the Pyramid sensor started exhibiting problems closing the loop, and in poorer seeing conditions the field stop needs to be opened up anyway. As this study concentrated on comparing the SH and PWFS under otherwiase equivalent conditions, the field stop size was not optimized, but the value $0.5''$ was adopted as a number that worked sufficiently well for the whole range of scenarios simulated.

Dark current. The near-IR dark current that we specified at 0.1 e/pixel/s at 77K was generally seen in these simulations to be a factor $10^2 - 10^3$ from becoming a significant source of noise. While we found no information of the temperature dependence of dark current in HgCdTe arrays, comparing to the behavior of a silicon device (e.g. the e2v CCD39-01 `http://www.e2v.com/download.cfm?type=document&document=914`) it seems likely that dark current would not impact the limiting magnitude even at meat-locker temperatures around -40°C, although this remains unconfirmed.

Modulation. While a significant amount of time was spent during this study on characterizing and understanding the effects of modulation with the PWFS (cf. Sect. 2.2.2), if the spot on the LOWFS has enough residual high-order wavefront error it may not be necessary to apply any modulation at all. The idea that turbulence itself produces an effect similar to modulation had been analyzed by e.g. Costa 2005[5].

5.2 Cost estimate

Based on the architecture suggested in Sect. 5, table 4 drafts an initial cost estimate for the various components and implementation. These are very round numbers listed here only to obtain an order-of-magnitude estimate and identify the cost-driving elements. It was assument that a LOWFS dedicated LGS in this case would be implemented by splitting off light from the HOWFS lasers, rather than purchasing a new unit. In addition to the STRAP-type tip/tilt sensor architecture, an alternative viable architecture (the "large near-IR array" type) was included for comparison. This wide-field method collects any number of available NGS onto a common large detector, which can provide better tip/tilt estimation over a wide field in the event that there are more than two that can be used. In

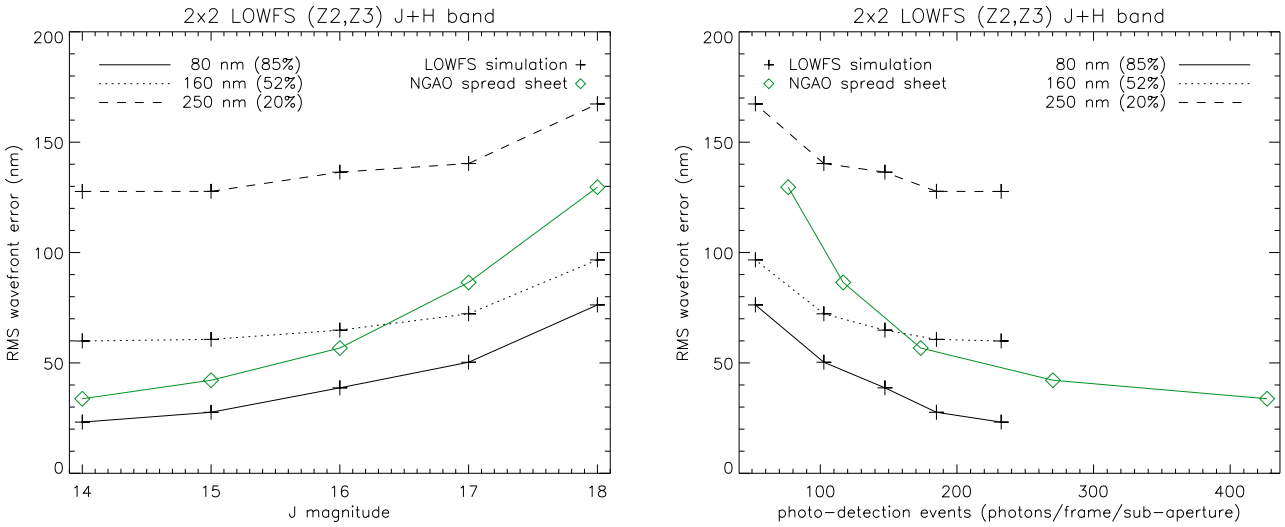

Figure 7: Comparison between the LOWFS simulation (black pluses) and the NGAO Master spread sheet (green diamonds) for the 2×2 case. Only the contribution from tip+tilt are plotted here, and the spread sheet calculation used a high-order residual RMS wavefront error of 130 nm. Plotted versus J band magnitude on the left, and versus PDEs on the right, which illuminates the differing bandwidth parameters and possible differences in the photometry system.

addition, this implementation has the advantage that no beam stearing is needed and the detector may also act as acquisition camera. Some disadvantages of this design include the higher cost of the detector and losing the ability to sharpen the NGS individually using MOAO control and MEMS DMs (the image sharpening is here done over the whole field instead). The performance of this type of sensor needs to be further explored, and the issues regarding plate scale, linearity and dynamic range better documented, but as a rough order of estimate one may expext its performance to be similar to that of the Shack-Hartmann.

References

[1] Brian J Bauman. *Optical Design for Extremely Large Telescope Adaptive Optics Systems*. PhD thesis, University of Arizona, 2003.

[2] D. Bonaccini, D. Gallieni, R. Biasi, F. Gori, M. Ghioni, C. Trottier, and N. N. Hubin. STARP for the VLT instruments [3126-77]. In R. K. Tyson and R. Q. Fugate, editors, *Proc. SPIE Vol. 3126, Adaptive Optics and Applications, Robert K. Tyson and Robert Q. Fugate, Eds., p.580*, volume 3126 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 580–+, 1997.

[3] R. M. Clare. *Wavefront Sensing and Phase Retrieval for Astronomical Imaging*. PhD thesis, University of Canterbury, Christchurch, New Zealand, 2004.

[4] Richard Clare. Keck NGAO sky coverage modeling. Technical report, W.M. Keck Observatory, 2007. KAON 470.

[5] J. B. Costa. Modulation effect of the atmosphere in a pyramid wave-front sensor. *Appl. Opt.*, 44:60–66, January 2005.

[6] J. W. Hardy. *Adaptive optics for Astronomical Telescope*. Oxford University Press, 1998.

[7] R. J. Noll. Zernike polynomials and atmospheric turbulence. *J. Opt. Soc. Am.*, 66(3):207–211, 1976.

[8] R. Ragazzoni. Pupil plane wavefront sensing with an oscillating prism. *J. Modern Opt.*, 43:289–293, 1996.

1x1 TT (large near-IR array)		USD	
Hawaii-2RG		400,000	
MEMS mirror		30,000	
MEMS pupil creating		10,000	
optics to create plate scale		20,000	
2D independent zoom (optional)		20,000	
Dewar+mounting		100,000	
J and H band pass filter		5,000	
	Total:	585,000	
1x1 TT (STRAP style)		USD (ea.)	Qty.
PICNIC array		65,000	2
MEMS mirror		30,000	2
MEMS pupil creating optics		5,000	2
field stops		5,000	2
J&H lenslet-based pyramid		20,000	2
Dewar+mounting		100,000	2
J and H band pass filter		2,000	2
	Total:	227,000	454000
2x2 PWFS		USD	
2x2 lenslet-based Pyramid		20,000	
Re-imaging optics		10,000	
Laser BTO + WFS		200,000	
MEMS corrector		30,000	
MEMS pupil forming optics		10,000	
J and H band pass filter		2,000	
Dewar+mounting		100,000	
PICNIC type detector		65,000	
	Total:	372,000	

Table 4: Tentative cost estimate for a 2×2 PWFS (bottom panel) and two different tip/tilt WFS architectures, the "wide-FoV large array" (top panel) and the narrow-field STRAP type (middle panel).

[9] C. Vérinaud. On the nature of the measurements provided by a pyramid wave-front sensor. *Optics Communications*, 233:27–38, March 2004.

[10] D. M. Winker. Effect of a finite outer scale on the Zernike decomposition of atmospheric optical turbulence. *J. Opt. Soc. Am. A*, 8(10):1568–1573, 1991.