

Generalised optical differentiation wavefront sensor: a sensitive high dynamic range wavefront sensor

S. Y. HAFFERT

Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands

*haffert@strw.leidenuniv.nl

Abstract: Current wavefront sensors for high resolution imaging have either a large dynamic range or a high sensitivity. A new kind of wavefront sensor is developed which can have both: the Generalised Optical Differentiation wavefront sensor. This new wavefront sensor is based on the principles of optical differentiation by amplitude filters. We have extended the theory behind linear optical differentiation and generalised it to nonlinear filters. We used numerical simulations and laboratory experiments to investigate the properties of the generalised wavefront sensor. With this we created a new filter that can decouple the dynamic range from the sensitivity. These properties make it suitable for adaptive optic systems where a large range of phase aberrations have to be measured with high precision.

© 2016 Optical Society of America

OCIS codes: (010.1080) Active or adaptive optics; (010.1285) Atmospheric correction; (010.7350) Wave-front sensing; (350.1260) Astronomical optics; (120.5050) Phase measurement.

References and links

1. F. Roddier "The effects of atmospheric turbulence in optical astronomy," *Prog. Opt.* **19** 281–376 (1981).
2. M. Schwertner, M. J. Booth, M. A. A. Neil, and T. Wilson, "Measurement of specimen-induced aberrations of biological samples using phase stepping interferometry," *J. Microsc.* **213**, 11–19 (2004).
3. M. Schwertner, M. J. Booth, and T. Wilson, "Specimen-induced distortions in light microscopy," *J. Microsc.* **228**, 97–102 (2007).
4. J. Davis and W. Tango, "Measurement of the atmospheric coherence time," *Publ. Astron. Soc. Pac.* **108**, 456 (1996).
5. R. V. Shack and B. C. Platt, "Production and use of a lenticular Hartmann screen," *J. Opt. Soc. Am.* **61**, 656 (1971).
6. O. Guyon, "Limits of adaptive optics for high-contrast imaging," *Astrophys. J.* **629**, 592–614 (2005).
7. R. Ragazzoni, "Pupil plane wavefront sensing with an oscillating prism," *J. Mod. Opt.* **43**, 289–293 (1996).
8. R. Ragazzoni and J. Farinato, "Sensitivity of a pyramidal wave front sensor in closed loop adaptive optics," *Astron. Astrophys.* **350**, L23–L26 (1999).
9. R. Ragazzoni, E. Diolaiti, and E. Vernet, "A pyramid wavefront sensor with no dynamic modulation," *Opt. Commun.* **208**, 51–60 (2002).
10. A. Burvall, E. Daly, S. R. Chamot, and C. Dainty, "Linearity of the pyramid wavefront sensor," *Opt. Express* **14**, 11925–11934 (2006).
11. R. A. Sprague and B. J. Thompson, "Quantitative visualization of large variation phase objects," *Appl. Opt.* **11**, 1469 (1972).
12. J.C. Bortz, "Wave-front sensing by optical phase differentiation," *J. Opt. Soc. Am. A* **1**, 35–39 (1984).
13. B. A. Horwitz, "New pupil-plane wavefront gradient sensor," *Proc. SPIE* **2201**, 496 (May 31, 1994).
14. J. E. Oti, V. F. Canales, and M. P. Cagigal, "Analysis of the signal-to-noise ratio in the optical differentiation wavefront sensor," *Opt. Express* **11**, 2783–2790 (2003).
15. F. Hénault, "Wavefront sensor based on varying transmission filters: theory and expected performance," *J. Mod. Opt.* **52**, 1917–1931 (2005).
16. J.E. Oti, V. F. Canales, and M. P. Cagigal, "Improvements on the optical differentiation wavefront sensor," *Mon. Not. R. Astron. Soc.* **360**, 1448–1454 (2005).
17. D. Greggio, D. Magrin, J. Farinato, R. Ragazzoni, M. Bergomi, M. Dima, L. Marafatto, and V. Viotto, "Avoiding to trade sensitivity for linearity in a real world WFS," *Proceedings of the Third AO4ELT conference*, 36 (2013).
18. C. Vérinaud, "On the nature of the measurements provided by a pyramid wave-front sensor," *Opt. Commun.* **233** (2004).
19. B. Welsh, "A Fourier series based atmospheric phase screen generator for simulating nonisoplanatic geometries and temporal evolution," *Proc. SPIE* **3125**, 327 (1997).
20. G. I. Taylor, "The spectrum of turbulence," (*ISO 4*) *Proc. R. Soc. A* **164**, 476–490 (1938).
21. T. von Karman, "Progress in the statistical theory of turbulence," *PNAS* **34**, 530–539 (1948).

22. W. Dali Ali, A. Ziad, A. Berdja, J. Maire, J. Borgnino, M. Sarazin, G. Lombardi, J. Navarrete, H. Vazquez Ramio, M. Reyes, J. M. Delgado, J. J. Fuensalida, A. Tokovinin and E. Bustos, "Multi-instrument measurement campaign at Paranal in 2007. Characterization of the outer scale and the seeing of the surface layer," *Astron. Astrophys.* **524**, A73 (2010).
23. R. C. Cannon, "Optimal bases for wave-front simulation and reconstruction on annular apertures," *J. Opt. Soc. Am. A* **13**, 862–867 (1996).
24. W. H. Southwell, "Wave-front estimation from wave-front slope measurements," *J. Opt. Soc. Am.* **70**, 998–1006 (1980).
25. D. Peter, M. Feldt, B. Dorner, T. Henning, S. Hippler, and J. Aceituno, "PYRAMIR: Calibration and Operation of a Pyramid Near-Infrared Wavefront Sensor," *Pacific* **120**, 872–886 (2008).
26. S. Esposito, A. Riccardi, L. Fini, A. T. Puglisi, E. Pinna, M. Xompero, R. Briguglio, F. Quirós-Pacheco, P. Stefanini, J. C. Guerra, L. Busoni, A. Tozzi, F. Pieralli, G. Agapito, G. Brusa-Zappellini, R. Demers, J. Brynnel, C. Arcidiacono, and P. Salinari, "First light AO (FLAO) system for LBT: final integration, acceptance test in Europe, and preliminary on-sky commissioning results," *SPIE* **7736**, 773609 (2010).
27. L. A. Poyneer and B. Macintosh, "Spatially filtered wave-front sensor for high-order adaptive optics," *J. Opt. Soc. A* **21**, 810–819 (2004).
28. N. O 'Mahony, "RoboDIMM - The ING's new seeing monitor," *The newsletter of the Isaac Newton Group of telescopes* **7**, 22–24 (2003).
29. C. Plantet, S. Meimon, J. M. Conan, and T. Fusco, "Revisiting the comparison between the Shack-Hartmann and the pyramid wavefront sensors via the Fisher information matrix," *Opt. Express* **23**, 28619–28633 (2015)
30. E. Gendron, M. Brangier, G. Chenegros, F. Vidal, Z. Hubert, G. Rousset, and F. Pouplard, "A new sensor for laser tomography on ELTs," *Adaptive Optics for Extremely Large Telescopes Conference*, 05003 (2010).
31. R.K. Komanduri, K. F. Lawler, and M. J. Escuti, "Multi-twist retarders: broadband retardation control using self-aligning reactive liquid crystal layers," *Opt. Express* **21**, 404–420 (2013).
32. M. N. Miskiewicz and M. J. Escuti, "Direct-writing of complex liquid crystal patterns," *Opt. Express* **22**, 12691–12706 (2014).

1. Introduction

Adaptive optics is used in many fields to correct aberrations created by turbulent media such as the atmosphere or intracellular fluids [1–3]. We do not have control over these turbulent processes, and their effects change on very short time scales and create completely different environments [4]. Because of this it is important to be able to measure the wavefront aberrations in a large range of conditions. These measurements have to be done with sufficient precision and accuracy so that the applied correction compensates the aberrated wavefront. These two requirements of dynamic range and sensitivity are coupled in most of the current wavefront sensors. If the dynamic range of the wavefront sensor is increased the sensitivity will suffer and vice versa which is shown in Table 1.

The most commonly used wavefront sensor is the Shack-Hartmann wavefront sensor [5](SHWFS), which uses a lenslet array to sample the wavefront. From the array of spots the wavefront gradient can be reconstructed. But the SHWFS is not optimal for high resolution phase measurements [6] because the number of phase measurements is directly coupled with the dynamic range and sensitivity.

The pyramid wavefront sensor [7–9] (PWFS) uses a pyramidal prism to split the focal plane in four quadrants. For the PWFS the number of phase measurements is only dependent on the amount of pixels that are used to sample each pupil. So the sensitivity and dynamic range are independent from the number of phase measurements. This is already a clear advantage for the PWFS. The downside of the PWFS is the limited dynamic range. The dynamic range of the PWFS can be increased by dynamic modulation [7, 10]. By increasing the modulation radius the dynamic range becomes larger, but the sensitivity decreases.

The Optical Differentiation wavefront sensor (ODWFS) [11–16] works on the basis of filtering light with a linear amplitude filter in a focal plane filtering setup. This can be thought of as a continuous Foucault knife edge test instead of the normal discrete knife edge test. This wavefront sensor has the opposite characteristics of the PWFS. The dynamic range of the ODWFS is very high, but the sensitivity is low.

In Table 1 there is an overview of the characteristics of the WFSs that are discussed in this

Table 1. Overview of different wavefront sensors. The dynamic range is the maximal local tilt that can be measured and the sensitivity is the variance in the wavefront measurements due to photon noise. The variables stand for; D the entrance aperture diameter, λ the wavelength, N is the number of photons per phase measurement, N_{subap} is the amount of sub-apertures of the SHWFS, and N_a is the number of Airy rings that fit within the field of view of the WFS. For the ODWFS the field of view is the size of the filter, for the MPWFS it is the modulation radius and for the SHWFS it is the lenslet field of view. These relations were adapted from [6, 7, 14]

Wavefront sensor	Dynamic range	Sensitivity
SHWFS	$N_a N_{\text{subap}} \lambda / D$	$N_a N_{\text{subap}} \lambda / \sqrt{2ND}$
PWFS	λ / D	$\lambda / \sqrt{2ND}$
MPWFS	$N_a \lambda / D$	$N_a \lambda / \sqrt{2ND}$
ODWFS	$N_a \lambda / D$	$N_a \lambda / \sqrt{2ND}$

paper. The expression for the dynamic range of the PWFS, MPWFS and ODWFS are very similar. This is because their working principle is the same. The PWFS is the same as an ODWFS but with a very small amplitude filter, and the modulation of the MPWFS acts as creating an effectively linear amplitude filter. The MPWFS is originally also based on the work by phase visualization technique of Bortz with amplitude filters [12]. The dynamic ranges of these WFSs are limited by the edge where the transmission becomes one. For an infinite steep curve such as the PWFS, the dynamic range is limited by the size of the Airy disk. This is also the case for the other two, but usually N_a is much larger than 1 so the finite size of the spot can be ignored. The expression for the dynamic range of the SHWFS is different because its measurement are done in the pupil plane instead of the focal plane as the others do. The dynamic range is limited to the field of view of a single lenslet, which is twice the NA of the lenslet.

What can be read from the sensitivity column is that all of them are proportional to λ / D . Increasing the size of the optical system thus always leads to a two fold advantage. The size of the spot becomes smaller and there are more photons. But increasing the aperture diameter also decreases the dynamic range. As can be seen in this table there is a linear trade off between dynamic range and sensitivity. If the dynamic range is increased by a factor of two, the sensitivity will go down by a factor of two. This holds for all the WFS's that are discussed here. For the SHWFS this is only true if the number of detector pixels is fixed, which is often a real constrain in optical design.

The trade off can be partially bypassed by using two wavefront sensors, one for the large low order modes, and one for the higher order modes. This is not the best solution, because this will introduce non-common path errors (NCP) between the WFS's due to misalignment or non perfect optics. The NCPs can be minimized by placing the WFS's as close together as possible, which is used in the Very Linear Wavefront sensor [17]. Next to the NCPs errors that can be introduced due to technical difficulties, the amount of photons will also be split up. With fewer photons the measurement noise will increase.

The solution proposed here is the combination of both a PWFS and the ODWFS in a single wavefront sensor. This is the Generalised Optical Differentiation wavefront sensor (G-ODWFS). By generalizing optical differentiation to nonlinear filters a new shape for the amplitude filter could be created, which has both a high sensitivity and a high dynamic range. By combining them in one WFS the NCP errors are avoided and all the photons are used in one measurement, and this increases the signal to noise.

In Section 2 the principles of generalised optical differentiation are shown together with the discussion of a new filter shape. In Section 3 numerical simulations are used to compare the G-ODWFS to the PWFS and ODWFS. The 4th Section shows the experimental results in the

lab with a prototype. The 5th Section discussed a possible design for high speed broadband operation.

2. Principles of generalised optical differentiation

2.1. Focal plane filtering

The complex field of a coherent source at the entrance aperture of an optical system is

$$U_{\text{in}} = A(\vec{r}) \exp(i\phi(\vec{r})) \quad (1)$$

where A is the amplitude of incoming light and ϕ is the phase. For an aberration free point source both $A(\vec{r})$ and $\phi(\vec{r})$ are constant over the entrance aperture. The effects of Fourier filtering of this complex field can be described by,

$$U_{\text{T}} = \mathcal{F} \{ T \cdot \mathcal{F} \{ U_{\text{in}} \} \} \quad (2)$$

,where U_{T} is the final complex field, U_{in} is the input field and T is the spatial filter. The effects of a continuous spatial filter can be easily evaluated analytically by using a Taylor expansion of the spatial filter,

$$T(k_x, k_y) = \sum_{n=0}^{\infty} t_n \left(\frac{k_x}{k_m} \right)^n, \quad (3)$$

where T is the filter function, t_n is the n th order coefficient, k_x the spatial frequency in the x direction and k_m the maximal spatial frequency which is set by the size of the filter. This normalization reduces the spatial frequency coordinates to $k_x/k_m \in [-1, 1]$. Spatial frequencies are used here because the filter is in the focal plane. The expansion can be substituted in Eq. (2).

$$U_{\text{T}} = \mathcal{F} \left\{ \sum_{n=0}^{\infty} t_n \left(\frac{k_x}{k_m} \right)^n \cdot \mathcal{F} \{ U_{\text{in}} \} \right\}. \quad (4)$$

The Fourier transform is linear so the order of the summation and the transform can be exchanged,

$$U_{\text{T}} = \sum_{n=0}^{\infty} t_n \mathcal{F} \left\{ \left(\frac{k_x}{k_m} \right)^n \cdot \mathcal{F} \{ U_{\text{in}} \} \right\}. \quad (5)$$

The intensity of the pupil image consists of two terms; one with the power of the modes themselves, and one with all the cross terms. With $U_n = \mathcal{F} \left\{ \left(\frac{k_x}{k_m} \right)^n \cdot \mathcal{F} \{ U_{\text{in}} \} \right\}$, which is the electric field response due to the n th filter mode, the intensity is then given by

$$I_{\text{T}} = \sum_{n=0}^{\infty} t_n^2 |U_n|^2 + 2 \sum_{m \neq n} t_n t_m \Re \{ U_n U_m^* \}. \quad (6)$$

2.2. Wavefront sensing by spatial filtering

It is possible to retrieve the phase of the incoming field with the results of the previous section. Two mirrored amplitude filters are used for the retrieval. Mirroring a function has two effects, the even part of the function stays the same but the odd part switches sign. In the case of a basis expansion the coefficients of the even modes will stay the same but the coefficients of the odd modes will switch sign. For the mirrored filters, denoted T and R , this results in

$$T(k_x, k_y) = R(-k_x, k_y), \quad (7)$$

$$t_n = (-1)^n r_n, \quad (8)$$

$$t_n^2 = r_n^2. \quad (9)$$

Here t_n and r_n are the n th basis coefficients of respectively the T and R filter. Both filters will make a pupil plane image given by Eq. (6). The difference between the two pupil images is

$$I_R - I_T = \sum_{m \neq n, n+m=\text{odd}} 4t_n t_m \Re\{U_n U_m^*\}. \quad (10)$$

The sum is over all modes where n and m are different and the sum of the two are odd. The first such a combination is 0, 1. These two modes are used to sense the phase gradient. The zeroth order mode is a constant, so $U_0 = U_{\text{in}}$. Each higher order Taylor term multiplies the electric field by a power law and the result of filtering by a power law is

$$\mathcal{F}\{k_x^n \cdot \mathcal{F}\{U(\vec{r})\}\} = (-i)^n \frac{\partial^n U(\vec{r})}{\partial x^n}. \quad (11)$$

The wavefront gradient is encoded by the first order terms of the expansion. So the filter coefficients should have the property $t_n \ll t_1, t_0$ for $n > 1$. If terms higher than the first order are negligible and if U_0 and U_1 are substituted, the intensity difference is,

$$I_R - I_T = 4 \frac{t_0 t_1}{k_m} \phi_x A^2. \quad (12)$$

Here ϕ_x is phase derivative in the x direction. This is still proportional to the amplitude of the incoming electric field. This can be solved by dividing through the sum of the two pupil images that can be shown to be,

$$I_R + I_T = 2 \sum_{n=0}^{\infty} t_n^2 |U_n|^2 + 4 \sum_{n \neq m, n+m=\text{even}} t_n t_m \Re\{U_n U_m^*\}. \quad (13)$$

Here the sum is over all the power in the modes themselves and over the combinations where m and n are not equal. This time the sum of the two should be even. The difference in outcome can be thought of as the equivalent between retrieving the odd and even part of a function. If the terms to first order are retained the sum of pupil intensities become,

$$I_T + I_R = 2t_0^2 A^2 \left(1 + \left[\frac{t_1}{k_m t_0} \right]^2 \left\{ \left(\frac{A_x}{A} \right)^2 + \phi_x^2 \right\} \right). \quad (14)$$

Here A_x is the derivative of the amplitude in the x direction. In this sum the two terms on the right can influence the normalisation. The first one is due to amplitude changes. The most extreme example is the edge of the telescope, where the transmission changes from 1 to 0. When reconstructing the phase from the pupil images the edges should be treated very carefully. A solution for this is to ignore the edge pixels [15]. Other amplitude aberrations could come from the intrinsic amplitude changes in the coherent source. These are no problems as long as the amplitude changes are small compared to the absolute amplitude. The second normalisation bias is due to the phase aberration itself. In most cases, and certainly in closed loop AO operation, $(t_1/t_0 \phi_x)^2$ is much smaller than 1. So for atmospheric turbulence this can be neglected.

The normalised difference then becomes

$$\frac{I_R - I_T}{I_R + I_T} \approx \frac{2t_1}{k_m t_0} \phi_x. \quad (15)$$

Conceptually this can be seen as response of the phase derivative to the normalised derivative of the filter. For an incoherent source each contributing source point can be added in intensity, which gives the following normalised difference

$$\frac{I_R - I_T}{I_R + I_T} \approx \frac{2t_1}{k_m t_0} \frac{\sum_{i=0}^{N_i} A_i^2 \phi_{x,i}}{\sum_{i=0}^{N_i} A_i^2} = \frac{2t_1}{k_m t_0} \langle \phi_x \rangle_I. \quad (16)$$

So for an incoherent source the intensity-averaged phase gradient is measured. The incoherent source should be smaller than the whole filter for the previous equation to hold.

For broadband conditions the filter size scales with wavelength, as k_m is a function of wavelength. For a broadband source with a bandwidth of $\Delta\lambda$ centred around λ_0 the following equation holds,

$$\frac{I_R - I_T}{I_R + I_T} \approx \frac{2t_1}{t_0} \frac{\int_{\lambda_0 - \Delta\lambda/2}^{\lambda_0 + \Delta\lambda/2} A^2 \phi_x k_m^{-1} d\lambda}{\int_{\lambda_0 - \Delta\lambda/2}^{\lambda_0 + \Delta\lambda/2} A^2 d\lambda} = \frac{2t_1}{t_0} \langle \phi_x k_m^{-1} \rangle_\lambda. \quad (17)$$

This is the spectrally weighted response. From this expression it can also be shown that the wavefront sensor will measure the wavefront gradient achromatically. The product of ϕ_x and k_m^{-1} is,

$$\phi_x k_m^{-1} = k_m^{-1} k \frac{\partial W}{\partial x}. \quad (18)$$

Here W is the wavefront to be measured. The wavelength dependency of k_m and ϕ_x cancel each other in the product, because k_m is proportional to the wave number k .

Without assuming a certain shape for the filter a few properties of these optical derivative filters can be determined. First of all if a linear filter is used this formulation is exact, no higher order terms are measured. This means that when the phase has to be sensed without influences of higher order derivatives, the filter should be linear. The shape of the linear filter can be seen in Fig. 1.

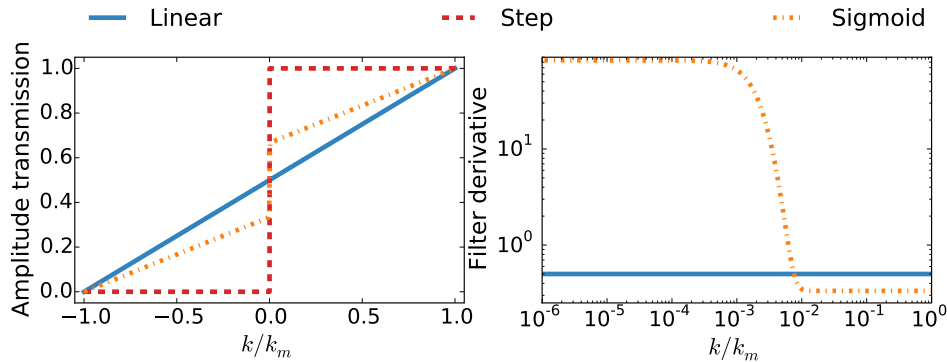


Fig. 1. The three filters used in the simulations: linear, step and sigmoid. The figure on the right shows the derivative of the filters. The derivative of the step filter is not shown since that is an infinitely high delta function.

The second property is the sensitivity. The sensitivity is how well a change in phase can be sensed with a given number of photons. This can also be seen as how efficient a WFS is at retrieving phase information from a single photon. The sensitivity is inversely proportional to the slope of the filter,

$$\text{Var}[\phi_x] = \left(\frac{k_m t_0}{2t_1} \right)^2 \text{Var} \left[\frac{I_R - I_T}{I_R + I_T} \right]. \quad (19)$$

The variance of the normalized differences can be estimated using propagation of errors. Because gradients are measured only half of the available photons can be used per phase gradient measurement. The final phase measurement noise per sub-aperture with read noise included is then

$$\text{Var}[\phi_x] \leq \left(\frac{k_m t_0}{2t_1} \right)^2 \left(\frac{2}{N} + \frac{8\sigma_{\text{read}}^2}{N^2} \right). \quad (20)$$

Here N is the total amount of photons for both directions and σ_{read}^2 is the variance of the read noise. The variance in the equation above is only an upper limit and changes when the light is unevenly distributed between two pupils.

The dynamic range is set by the bounds where the filter reaches 0 and 1 in transmission. A larger dynamic range requires a larger physical filter, but this creates a shallower slope. A shallower slope results in a lower sensitivity as can be seen from Eq. (20). If the filter becomes larger then k_m also becomes larger, and the sensitivity of the phase measurements is proportional with k_m . So here we see essentially the same trade-off that the SHWFS and the modulated PWFS have. There is a linear trade off between the dynamic range and sensitivity. The only method to go around this trade off is by tweaking the filter profile in such a way that t_1 becomes larger without decreasing the size of the filter. The change in t_1 by considering other filter shapes, like a square root profile, is small and is on the order of unity. This means that global changes to the filter will not lead to a better trade off. A non linear trade off is possible by making local changes to the filter shape as will be shown in the next section.

2.3. Design of a novel, hybrid filter

Considering the previous section there are two extreme filter shapes. The first is a linear function, which has the highest dynamic range for a given filter size. The second is a step filter, which has the highest sensitivity because of the infinitely steep slope. Both filters can be seen in Fig. 1. Both filters have been used before in WFS's. The linear filter is the ODWFS, and the step filter is a PWFS.

Step filters have the lowest noise propagation, due to the steepness, and are the most sensitive to phase changes. But the dynamic range is very small, due to the zero width of the step. Another effect of the step function is the discrete change in amplitude. This creates a lot of higher order modes in the basis expansion. These higher order effects diffract the light away from the modes that are used for wavefront sensing.

By combining the two extreme filters, we can trade off the sensitivity and the dynamic range by tweaking their relative power and create a filter that is optimal made for the problem at hand. The linear combination of the two extreme filters lead to a linear filter with a step in the middle. By adding the step the noise propagation will be better, but the step adds a discontinuity. The discontinuity creates strong diffraction effects by inclusion of higher order modes, which decrease the sensitivity. To reduce the diffraction the step should be a smooth function for which the width and height of the step can be easily controlled. This is realised with the following Fermi function like profile

$$T(k_x, k_y) = \frac{(1-\beta)}{2} \left(1 + \frac{k_x}{k_m}\right) + \frac{\beta}{1 + \exp\left(-\frac{k_x}{\sigma k_m}\right)}. \quad (21)$$

This profile is the sum of a linear function with a sigmoid function. The t_0 and t_1 parameters are both 1/2 for a linear filter which is included in the pre factor. The parameter β determines the relative size of the step and σ the width of the step. An example of this filter is shown in Fig. 1. To create a linear filter $\beta \rightarrow 0$, and to create a step $[\beta \rightarrow 1, \sigma \rightarrow 0]$. With this filter shape it is possible to interpolate between the two extremes. To determine the useful parameter space of the filter we consider the case when β is very small and the case where σ is very large. If β is very small then the filter will still act like the linear filter and nothing is gained over the linear one. By increasing σ the width of the step is increased, and at a certain point the width of the step is larger than the filter size. In that case $\exp(-k/\sigma k_m) \approx 1$ and the function is well described by

$$T(k_x, k_y) = \frac{1-\beta}{2} \left(1 + \frac{k_x}{k_m}\right) + \frac{\beta}{2}, \quad (22)$$

a linear filter with decreased response.

Therefore if the G-ODWFS wavefront sensor has to be sensitive, the step width should be small. By making the step small in width, it will at a certain point become smaller than the Airy pattern of the optical system. This will increase the amount of power that is diffracted away, like in a PWFS. In the case that the step width becomes much smaller than the airy disk, the sigmoid profile can be regarded as a step function. In this case the gain can not be calculated by using Eq. (15). It is also not possible give an explicit expression for the gain for large wavefront aberrations, due to the step. The proportionality constant has to be calculated numerically. For small wavefront aberrations the response can be derived analytically. It is the direct sum of a linear filter and a step filter weighted by their relative importance. So the normalised difference is,

$$\frac{I_R - I_T}{I_R + I_T} \approx (1 - \beta) \frac{2t_1}{k_m t_0} \phi' + \beta \frac{D}{\pi} \phi'. \quad (23)$$

with D the aperture diameter. The response for the step is taken the same as that of a PWFS [18]. The step height is independent from the filter size for this profile, and this decouples the sensitivity from the dynamic range. Now that the sensitivity no longer is coupled to the dynamic range, it will be possible to create a filter where the sensitivity and dynamic range can be independently chosen for the science at hand. There is still a small coupling between the two parameters but the trade off becomes highly non linear. Only in the extreme case when β approaches 1 will the coupling increase.

The size of the filter can be expressed in the number of Airy rings it encompasses. The maximal spatial frequency then becomes $k_m = 2\pi N_A/D$, with N_A the number of airy rings. The relative gain in sensitivity of the G-ODWFS compared to that of the pyramid is then

$$\frac{g_H}{g_{PWFS}} = \beta + \frac{1 - \beta}{2N_A}. \quad (24)$$

Here we can see that the relative gain for small wavefront aberrations is almost independent of the size of the filter. The relative sensitivity of the wavefront sensor is thus β . If the step size is 1/3 we expect that the WFS will be 3 less sensitive than the PWFS. If the filter becomes smaller in size the gain will converge to the gain of a PWFS. So here we see the non-linear trade off between sensitivity and dynamic range. The dynamic range can be arbitrarily increased by creating a larger filter. This creates a lower sensitivity for large aberrations but the sensitivity for small wavefront aberrations will not be influenced.

This is a convenient property for closed loop AO systems. Because as the AO system feedback loop is closed it will start to correct for wavefront errors, and at the moment the wavefront errors are small enough the WFS will have an increased sensitivity. The benefit of this over a pure step filter is that there is still some response for large aberrations. This increases the locking speed of the AO system, which is help full in dynamic conditions where the wavefront errors can change rapidly. The benefit of this over a linear filter is the increased sensitivity for small aberrations.

The direct combination of the two profiles has an advantage over the method with two separate WFSs. Suppose there is a way to split the photons up between the two WFSs with a ratio of β . Because the photons are split up before detection the photons on one sensor will not help to reduce the photon noise on the other. While this is the case when you directly combine the two. If a the step is included in the linear filter then for large aberrations there is already a larger difference between the pupil images intensities. This increases the SNR of the normalized differences compared to a standard ODWFS.

3. Simulations

3.1. Simulation set-up

A simple closed-loop system with a time-variable phase screen is implemented to test the G-ODWFS. The dynamic phase screens simulate the effects of imaging through the atmosphere when looking at a point source. The atmospheric turbulence is modelled by using a Fourier series based expansion [19]. The effects of the complete atmosphere are concentrated in several layers. The strength and height of each layer is taken from Guyon [6]. The temporal evolution of a single phase screen is assumed to follow Taylor's Frozen flow [20]. The wind speed is set to 10 m/s and the direction is randomly chosen. The current implementation uses a modified Von-Karman power spectrum [21] with a 100 m outer scale [22] and a 1 mm inner scale, with 150 by 150 samples in the Fourier domain and 256 by 256 samples in the real domain.

The deformable mirror is modelled as perfect because only the effects of the wavefront sensor are of interest. The shape of the deformable mirror is given by a superposition of a modal basis $\phi = \sum_{n=1}^N a_n M_n$, where a_n is the n th modal coefficient and M_n is the n th mode. The Karhunen-Loeve basis with a Kolmogorov spectrum is used, because there is no analytical expression for the covariance function of the modified Von-Karman spectrum. The basis is created by direct decomposition of the covariance matrix of the power spectrum over the telescope aperture [23]. For all simulations a 100-mode deformable mirror is used.

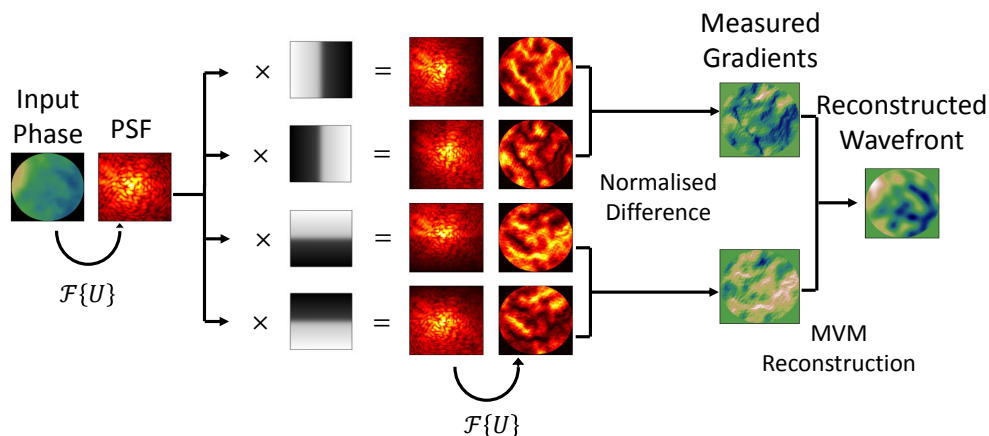


Fig. 2. In this figure a single measurement loop is shown. The measurements start with a phase distortion caused by either the atmosphere or a non flat shape of the deformable mirror. This creates an aberrated psf, which is then split up in four copies. Each copy is filtered by a differently oriented amplitude filter. The amplitude filter in this example is the hybrid filter, with $\beta = 1/3$. Another propagation creates four pupil images. The pupil images are then combined to retrieve the wavefront gradients. From the wavefront gradients the input wavefront is reconstructed with the matrix-vector multiplication method.

The wavefront sensor is simulated as shown in Fig. 2. The input wavefront is Fourier transformed to create the PSF, which is then multiplied by each filter creating four new PSFs. The total amount of filtered PFSs is four because we need two for each direction. The final PSFs are then Fourier transformed again to create the pupils. The linear filter is simulated in the same way as the hybrid, only the filter shape is changed. The pyramid is simulated by multiply the PSF with a transparent quadrant. Each quadrant generates a different pupil by Fourier transforming the filtered PSF. The interference effects between different pupils is neglected in these simulates. This is the same as placing the pupils far away from each other.

The final pupils are all sampled on 32 by 32 pixel grids. This sampled pupil is also used

for the wavefront reconstruction. The wavefront model coefficients are reconstructed by the matrix-vector multiplication method [24].

The effects of photon noise and read noise are added to the final pupil images. First the total intensity of the PSF before the filters is normalized to one and then multiplying by the total amount of photons. With this the intrinsic loss for each filter is taken into account. For the ODWFS and the G-ODWFS the effect of splitting the light into two for x and y is taken into account by dividing the intensity by two. Poisson noise is then generated by using the pixel value as the expectation value for Poisson random variables. Most wavefront sensors for high contrast imaging use EM-CCDs with low read noise. The read noise can be much smaller than 1.0 photon-electron per pixel. To include some effects of read noise, it is approximated with a Poisson distribution with $\lambda = 1.0$ photon-electron.

The closed loop simulation begins with the creation of an input wavefront. The phase on the deformable mirror is then subtracted from the input phase. This creates the residual phase pattern, which is used to measure the image quality and perform the wavefront sensing. This is an iterative process the deformable mirror starts with a flat shape. For all simulations a time step of 1 ms is used, together with a source flux of $1.25 \cdot 10^8$ photons per second. The simulation uses a 4.2 meter telescope, so the photon flux corresponds to an 8th magnitude star. Each sub-aperture receives roughly 155 photons per frame. Each normalised difference will have half of the amount of photons because these have to be split up for the two gradient directions. The simulations are all monochromatic at $0.5\mu\text{m}$, so the wavefront sensing and the imaging is done at the same wavelength. The performance estimates are done before correcting the current frame. This introduces a single frame of lag.

3.2. Calibration of the AO system

To calibrate the relation between the wavefront sensor and the deformable mirror the influence matrix has to be determined. We have,

$$\vec{s} = A\vec{a}, \quad (25)$$

where \vec{s} is the measurement vector, A the influence matrix and \vec{a} the input signal for the DM. The columns of the influence matrix are the slopes of the response to the input modes. This model is linear while the gain of the new G-ODWFS is non-linear. From Eq. (23) it can be seen that there is still a linear relation for small wavefront aberrations. So this will give the correct relation for small wavefront errors. For larger wavefront errors the gain is reduced. So the compensation for small wavefront errors will be correct, while the large wavefront errors are underestimated. This does not matter in a closed-loop AO system, the large scale aberrations are then corrected in a few iterations instead of one.

The influence matrix calibrates the gain of the filter. To determine the slope of a single mode, two measurements have to be done. First the mode is applied to the DM with a positive modal coefficient and then a second measurement with the same mode but with a negative modal coefficient. This removes any bias. The two input vectors create two measurement vectors that are combined as followed,

$$A_i = \frac{\vec{s}_+ - \vec{s}_-}{a_+ - a_-}, \quad (26)$$

where A_i is the i th column of the influence matrix, \vec{s}_\pm is the measurement corresponding to the a_\pm applied mode coefficient. There are two constraints when choosing the input mode coefficients for the calibration. The wavefront sensor will not measure a response if the input coefficient creates a response that is buried in the noise. Secondly if the input is too large either the wavefront sensor response will saturate, or the DM will reach its maximum stroke. The input can therefore not be chosen too small or too large. This is especially relevant for the PWFS, because its dynamic range is very small. A small input coefficient will also lead to large errors in the influence matrix due to the division. The influence matrix of the pyramid wavefront sensor will therefore be

noisier than other wavefront sensors that can be calibrated with larger amplitudes. The difficulty of calibrating a pyramid wavefront sensor has already been noted several times [25, 26].

The actuation matrix is the matrix that transforms the measurement vector back to mirror mode coefficients. If the influence matrix is a square matrix, the actuation matrix is the inverse of the influence matrix. Often the influence matrix is rectangular so an inverse matrix does not exist. To create the actuation matrix the pseudo inverse of A is calculated by using singular value decomposition (SVD). With an SVD the number of reconstructed modes can be easily regulated by setting the unwanted modes' singular values to zero during the inversion. For all simulations the full inversion is used, so no singular values are set to zero.

For the simulations the calibration amplitude was set to 0.01 wave as not to saturate the PWFS. The linear and sigmoid were calibrated with the same amplitude. The calibrations are done without noise, as AO systems in principle can be calibrated with an internal light source so that the amount of photons can arbitrarily be increased.

3.3. Simulation results

The filters that are used in the simulations are shown in Fig. 1. The sigmoid filter is compared to the PWFS and the ODWFS which are the extreme filters and also have been simulated and tested in previous work. The sigmoid filter has a step size $\beta = 1/3$ and a width $\sigma = 0.001$. The dynamic range of tip/tilt aberration for the three filters can be seen in Fig. 3. The linear filter has the largest dynamic range and the pyramid the smallest. The sigmoid filter has the same dynamic range as the linear filter, because the response increases as the input aberration is increased. The slope of the response is very shallow for the larger inputs. If the response is needed at those points, the measurement error should be very small. The pyramid wavefront sensor is limited to a range of 1.0λ .

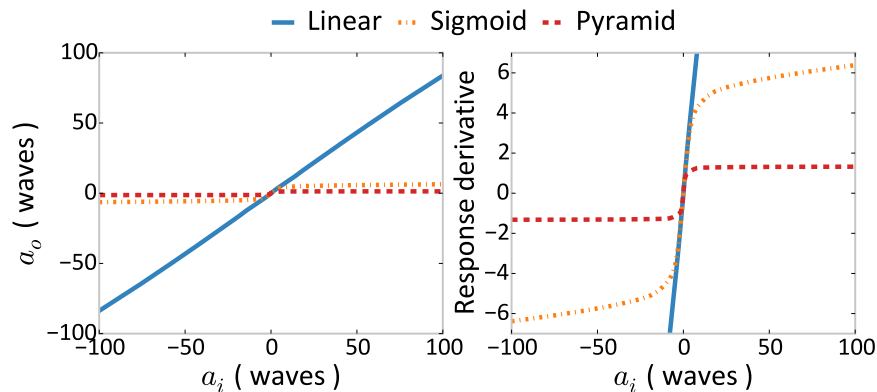


Fig. 3. These two figures show the dynamic range of the filters. The figure on the right is a zoomed version of the left figure. The input tilt coefficients on the x axis are plotted against the measured tilt coefficients.

The metric used for comparison between wavefront sensors in the close loop simulations is the Strehl ratio, S . Two different static cases are looked at. In the first case the phase errors are created by a random superposition of the deformable mirror modes. All modes that are in the phase screen can be corrected, so every deviation from perfect is then due to photon noise. The second case uses a static atmospheric phase screen, which has more modes than the deformable mirror can correct. This will show the errors due to the influence of uncorrected modes and reconstruction errors. The results are shown in Fig. 4. In the case of a phase error

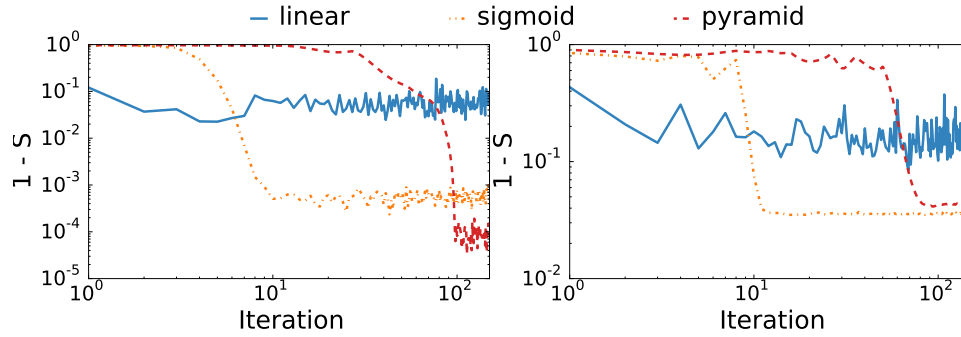


Fig. 4. In these figures the deviation from unity of the Strehl ratio is plotted against time. The left figure shows the change in strehl when the phase screen is created by taking a superposition of deformable mirror modes. The figure on the right shows the increase in Strehl when a static atmospheric phase screen is used. All wavefront sensors start with the same phase screen.

caused by the deformable mirror, the pyramid wavefront sensor reaches the highest Strehl ratio. The sigmoid follows the pyramid very closely. The linear filter has the lowest Strehl ratio as is expected from Eq. (19). The correction time scales also show the effect of the dynamic range. If the dynamic range of a WFS is high it will measure large aberrations correctly, and the AO system will be able to correct the aberration within a single iteration. If the dynamic range is not high enough the aberration amplitude will be underestimated. Therefore several iterations are necessary to compensate the full aberration. Thus a higher dynamic range results in a larger temporal bandwidth, resulting in faster corrections which are important when there are dynamical changes.

This simulation shows the limiting Strehl that can be achieved by the different wavefront sensors with the given amount of photons, because all modes can be corrected. If the WFS could measure the aberrations more precise, the compensation would be better and the Strehl ratio would be higher. So this simulation effectively shows the sensitivity of the WFS.

In the high Strehl regime the Strehl ratio can be approximated as $S = \exp(-\sigma_{\phi,rms}^2)$, and $S = 1 - \sigma_{\phi,rms}^2$ when the Strehl is very close to unity. The plotted $1 - S$ corresponds in that case to the square of the wavefront rms in radians. The final wavefront rms's in units of wavelengths are $0.039 \pm 0.007\lambda$, $0.0039 \pm 0.0005\lambda$ and $0.0015 \pm 0.0003\lambda$ for the linear filter, the sigmoid filter and the pyramid respectively. The ratio between the PWFS and the sigmoid filter is 2.6 ± 0.6 , and the relative gain from Eq. (23) is 2.95. This shows that the hybrid gain can predict the sensitivity of the new WFS. The sigmoid only loses a factor of 2.6 in sensitivity compared to the pyramid wavefront sensor, while achieving a dynamic range which is 10 times higher.

The right plot of Fig. 4 shows the correction of a simulated static atmospheric phasescreen. In this case there are many more modes than the deformable mirror can correct. All limiting Strehl ratios are lower than in the first case due to the increase in uncorrected modes. Especially the difference between the sigmoid filter and the pyramid sensor is interesting. The $1 - S$ value of the pyramid is larger than the sigmoid filter, which implies that the sigmoid filter reaches a higher Strehl ratio. The final wavefront errors are $0.064 \pm 0.010\lambda$, $0.0300 \pm 0.0002\lambda$ and $0.0330 \pm 0.0002\lambda$ respectively for the linear filter, the sigmoid filter and the pyramid. The very small spread of the final wavefront error for the sigmoid and pyramid imply a high stability. Multiple static wavefront maps have been tested and in all cases the sigmoid filter performed slightly better than the PWFS.

The final wavefront error is determined by two components, the measurement noise and the reconstruction error. The measurement noise is caused by the propagation of photon noise, as given by Eq. (20). The pyramid wavefront sensor should be most robust against photon noise, which is also shown in the left graph of Fig. 4. The second error is due to the phase reconstruction. The matrix vector multiplication method is a Least Squares fitting procedure. With the SVD inversion the gradient measurements are fitted to the input mode coefficients. So fitting errors can decrease the quality of the reconstruction. The singular values of the interaction matrix can be seen in Fig. 5. We can see that the pyramid is the most sensitive because its singular values are the highest. And the sigmoid is in between the PWFS and the ODWFS.

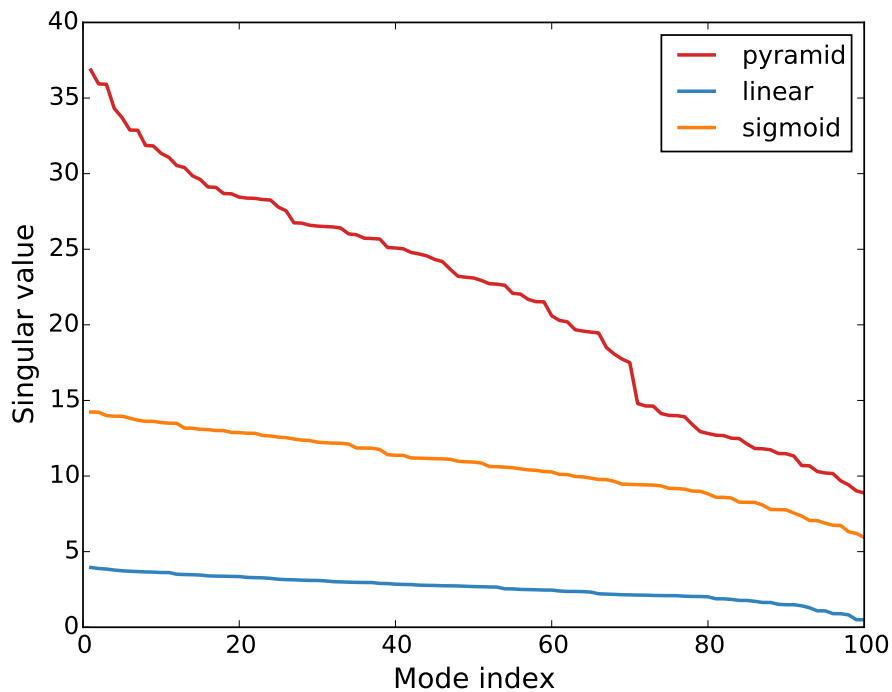


Fig. 5. The singular value is plotted against mode number. The pyramid has the highest singular values. The sigmoid is in between the pyramid and the linear.

There are four different kind of modes that can influence the performance of an AO system. The first are modes that can be controlled and sensed, these are the modes that create the control matrix. The second are modes that cannot be controlled but can be sensed, this happens when a WFS has more degrees of freedom in the measurements than the number of actuators on a DM. The third are modes that can be controlled but not sensed, like DM waffle patterns. And last of all are the modes that can not be controlled and sensed, these are high order modes cause aliasing in the WFS measurements if the WFS is not spatially filtered [27]

The difference in the second simulation is not due to the first kind of modes because the first simulation shows that the pyramid performs better. It is also not due to modes like waffle patterns because these are not included in the simulation. So the only possibility is either aliasing or additional fitting errors due to the second type of modes. Even though the SVD creates an orthogonal basis it can still lead to fitting errors for calibrated modes. For the DM control orthogonal modes usually are used. While these modes can be orthogonal their gradients

generally are not. This causes fitting errors which can be significant. This can be solved in the same way as aliasing is treated by adding a spatial filter. The effect of a higher order mode can be found by using that mode as an input wavefront aberration and use the reconstructor for the controllable modes. This returns a phase reconstruction in controllable mode space, if this vector is not zero then there is an effect from the higher order modes on the controllable modes.

The amount of influence a higher order mode has on the controllable modes is defined as the mean squared amplitude per higher order mode and per controllable mode. The PWFS has a coupling of 4 percent and is the highest of the tested WFSs. This has also been compared to a rooftop WFS which had a coupling of 2.2. This leads to the suspicion that it is caused by the diffraction effects off the edges. The ODWFS has a coupling of 2.3 percent and the G-ODWFS has a coupling of 2.4 percent. These cross coupling coefficients do not quantitatively explain the difference in Strehl ratio, because the cross coupling has to be weighed by the atmospheric power spectrum. The coefficients also depend very strongly on the sampling of the final pupil images and the sampling of the initial telescope pupil. But in all cases the PWFS has the highest cross coupling and the ODWFS the lowest, with the G-ODWFS in the middle. This can explain the difference in the second simulation where there are modes in the atmosphere which the DM can not correct, but will influence the controllable mode measurements.

Figure 4 shows that all three wavefront sensors are sensitive enough for most adaptive optics systems. The limiting factor is usually not the wavefront sensor but the deformable mirror. Only extreme ao systems reach high enough Strehl ratios to reach the limit of wavefront sensing. The largest gain is shown in dynamical simulations where the atmosphere evolves in time while the system is operating in closed loop. In Fig. 6 two simulations are shown for different atmospheric conditions. In the left figure the atmospheric conditions correspond to an input $D/r_0 = 84$, and for the right figure $D/r_0 = 21$. These two values of r_0 are the left and right sided 5 percent limits taken from the RoboDIMM data of Observatorium Roque de los Muchachos at La Palma. 90 percent of the time the atmospheric conditions will be between these two limits [28]. Each simulation uses the same phase screens for the wavefront sensors, so that they are correcting the same atmosphere at each point in time. Ten simulations are performed with the same atmospheric conditions. The mean of each simulation is plotted together with the standard deviation at every point in time.

From Fig. 6 it is clear that the sigmoid filter outperforms both wavefront sensors if there is strong turbulence. For the weaker conditions as in Fig. 6 on the right, the performance is also better than both. The difference between the sigmoid and the pyramid is very small but on average it reaches a higher strehl. The effect of the increased dynamic range shows itself in the temporal bandwidth. Due to the high dynamic range the linear filter is already at its sensitivity limit in ten iterations. The sigmoid filter needs 20 iterations while the pyramid needs 50 iterations. In Fig. 7 a representative snapshot of the dynamical simulation with the better atmospheric conditions is shown. The snapshot makes the difference in correction quality even more apparent. The sigmoid filter is able to correct the PSF up to the 4th Airy ring. The pyramid only reaches the 2nd Airy ring with visible distortion in the Airy rings. The linear filter can't even stabilize the first Airy ring stable. The sigmoid filter is able to reach the static case limiting Strehl ratio, indicating an excellent temporal correction.

4. Experimental work

4.1. Experimental setup

A prototype of the G-ODWFS sensor was developed in the optical lab. The schematic setup is shown in Fig. 8. The input source of the setup is a CPS240 diode laser operating at 635 nm. The laser is focused by a lens with a focal length of 250.0 mm, and a 25.0 mm diameter. At the focal point the laser is spatially filtered by a 10 μm pinhole. The filtered laser is collimated with a second lens onto the deformable mirror. The collimation lens has a focal length of 140.0 mm and

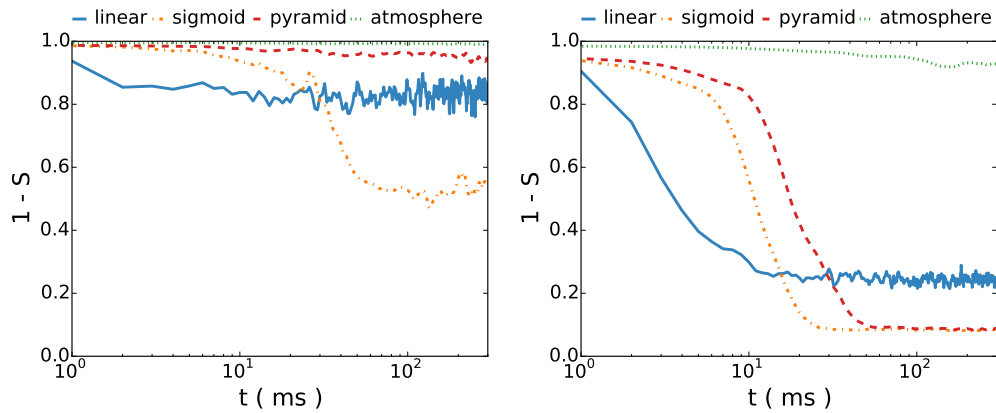


Fig. 6. Two different dynamical closed loop simulations. In the left figure the turbulence strength is $D/r_0 = 84$, and for the right figure it corresponds to median conditions in the visible at $D/r_0 = 21$. All wavefront sensors are looking at the same sky so that they correct the same phase screen at every point in time. For both atmospheric conditions 10 simulations were done. The mean $1 - S$ of these simulations are plotted with solid lines. The standard deviation of each wavefront sensor is given by the coloured region. $1 - S$ is plotted for the different wavefront sensors and the natural effects of the atmosphere. In both cases the sigmoid filter out-performs both wavefront sensors.

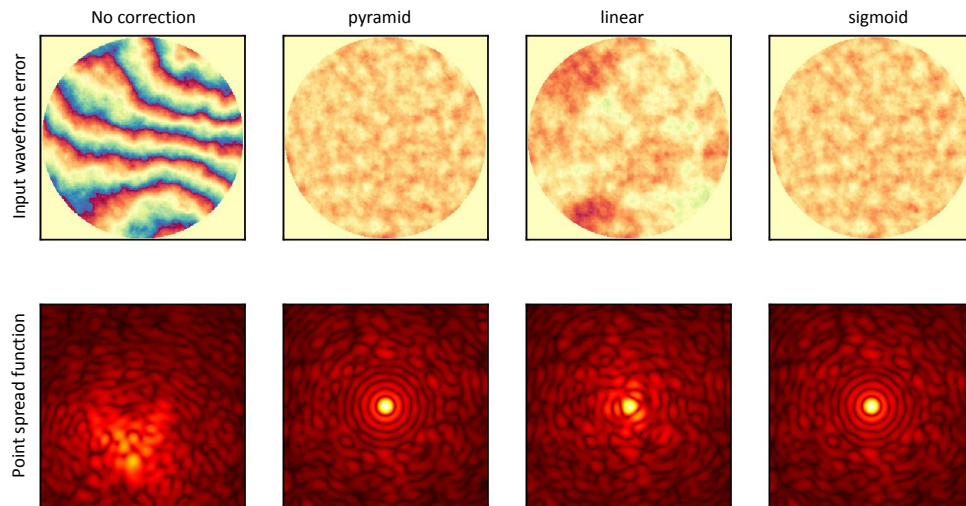


Fig. 7. Snapshot of a single dynamical simulation with $D/r_0 = 21$. The top row shows the residual phase distortion due to the atmosphere. Red is a positive π phase difference and blue is a negative π phase difference. Up to four Airy rings can be clearly seen in the corrected PSF of the pyramid and the sigmoid. The linear filter cannot compensate enough to stabilize the first Airy ring. The PSFs are normalised with respect to the maximum of the diffraction limited PSF.

a diameter of 25.0 mm. The DM is an Alpao DM 97-15, which is a continuous facesheet DM with 97 actuators and a clear aperture of 13.5 mm. This DM is used for creating and correcting pupil plane aberrations. After being reflected from the DM, the laser is reimaged by an identical lens. The DM is positioned at the focal plane of both lenses. The reimaged laser is focused through a non-polarising beam splitter and half goes to the focal plane filtering wavefront sensor and half to a Shack-Hartmann wavefront sensor.

The input to the Shack-Hartmann wavefront sensor is spatially filtered by a square mask, to eliminate the cross talk between subapertures due to large spot offsets. The output of the spatial filter is collimated with a 60 mm focal length achromatic lens. The collimated beam is then sampled by a microlens array. The microlens array has a pitch of 500 μm and a 32.82 mm focal length. The pupil is sampled by 11 microlenses across its diameter. After passing through the microlens array the array of spots is reimaged by a pair of lenses onto the wavefront sensor camera. The lenses are achromatic lenses with respectively a focal length of 60.0 mm and 35.0 mm. The wavefront sensor camera is an Andor iXon3 EM-CCD, with 128 by 128 pixels running at a speed of 513 Hz with full frame readout. Each subaperture is sampled by 8x8 pixels.

For the focal plane filtering wavefront sensor, different amplitude filter shapes need to be tested. With a Holoeye LC2002 spatial light modulator placed between two polarisers it can be used as a programmable amplitude filter. The transmission will depend on the applied pixel signal and the wavelength, because the phase retardation is wavelength dependent. To exclude chromatic effects of the spatial light modulator a 635 nm laser is used. The Holoeye SLM consists of 800 by 600 pixels, with a pitch of 32 μm and a 55% fill factor. The filters are cropped to 600 by 600 pixels to make them symmetric. All pixels that are outside this region have the lowest possible transmission. The linear polarisers used for the filtering are, a LPVISA050 and a LPVISB050 from Thorlabs. The orientation of the polarisers relative to the SLM is chosen such that the contrast between minimum and maximum throughput is optimized. This contrast was found to be roughly 1:1000. Since the SLM consists of pixels, it is necessary to have a correct sampling of the PSF, which corresponds to at least 2 pixels within the first Airy ring. A magnifying telescopic system enlarges the PSF sampling to 70 μm per λ/D . With this sampling there are 2.2 pixels per λ/D satisfying the Nyquist sampling requirement. The telescopic system consists of two achromatic lenses with a focal length of respectively 1000.0 mm and 100.0 mm. The diameter of both lenses is 50.8 mm. After the second polariser a final lens reimages the pupil on an Allied Vision Technology Pike-F032B camera, with a resolution of 640 by 480 pixels. The pupil is sampled by 135 pixels in diameter.

The prototype only measures one pupil at a time. To measure all four pupils the four amplitude filters have to be applied sequentially to the SLM. An advantage to this set-up is the absence of relative pupil misalignments and flat field effects because the same pixel is used for all four phase measurements. The disadvantage is the limiting frequency at which the SLM can run. The SLM works at 30 Hz, and if four filters have to be applied to the SLM sequentially the wavefront sensing frequency is at most 7.5 Hz. Another effect is the time delay between sending and applying the signal on a pixel. This ultimately limits the operation frequency to roughly 1.5 Hz. The experimental setup is calibrated in the same way that the simulations are calibrated. The DM is operated with Karhunen-Loeve modes with a Kolmogorov power spectrum.

The mode response in Fig. 9 for three different filters corresponds to the quadfoil aberration. The filters that were used to create these responses are the same filters that were used in the simulations in Section 3. From the responses alone it can be seen that the linear filter is the noisiest. The pyramid wavefront sensor already has saturated responses, while the calibrated input is very small. The pyramid could also not be calibrated without using another filter first to flatten the deformable mirror. Because the deformable mirror is not flat when not actuated but in a slight deformed state. These small deviations from flat are already enough to saturate the pyramid. The sigmoid filter had no problems with calibration, because it's dynamic range is extended. This

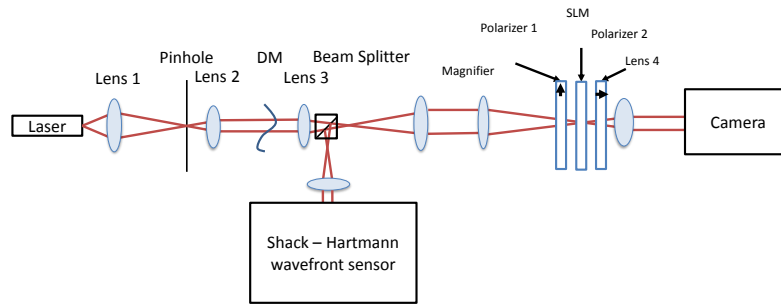


Fig. 8. In this figure the schematic lab set-up is drawn. The focal lengths are not to scale. A laser at 636.3 nm is used as input. The laser is focused at a $10\ \mu\text{m}$ pinhole to simulate a point source. The point source is collimated on a deformable mirror which is used to create and correct pupil aberrations. The new pupil is focused through a beam splitter from which one part goes to the new wavefront sensor and the other to a Shack-Hartmann wavefront sensor. The beam to the new wavefront sensor is magnified such that the spatial light modulator is Nyquist sampling the PSF. Amplitude filtering is done by placing the spatial light modulator between two orthogonal polarisers. The final pupil is then re-imaged onto the camera. The fold mirrors are excluded from this drawing.

already shows one of the advantages of the sigmoid filter, the alignment of the optical system can be relaxed. The sigmoid filter also shows a much sharper response measurement. In the lab measurements the sigmoid filter had the lowest condition number, followed by the linear filter and the pyramid filter had the highest condition number. The calibration amplitudes of the different WFS were; $0.1\ \lambda$ for the pyramid, $1\ \lambda$ for the sigmoid, $10\ \lambda$ for the linear and $8\ \lambda$ for the SHWFS. And the DM was operated with 90 out of the 97 modes.

4.2. Experimental results

The dynamic range of the wavefront sensor is measured by applying an increasing tilt on the deformable mirror and measuring the response. The dynamic range per mode is different, because these WFS's measure slopes and higher order modes have steeper slopes. So the response will saturate at lower input for higher order modes. But the relative dynamic range between the WFS's sensor for a single mode should be the same.

The results for the different filters can be seen in Fig. 10. The linear filter has the largest dynamic range, as was expected from the simulations. The pyramid wavefront sensor has the lowest dynamic range saturating at $1.5\ \lambda$. The dynamic range of our Shack-Hartmann sensor was also measured. The Shack-Hartmann reaches the maximum dynamic range when the spots reach the limit of their sub-aperture. If the size of the aberration increases beyond this dynamic range some of the spots will be lost. The response of the Shack-Hartmann then slowly falls off to zero when all spots are lost. The sigmoid filter is in between the linear filter and the pyramid in terms of dynamic range. The dynamic range is increased by a factor of 10 compared to the pyramid.

Every point in the dynamic range curve is the average of 10 measurements. The previously chosen sigmoid filter had filter parameters $\beta = 1/3$ and $\sigma = 0.05$. To see the influence of the parameters three different step sizes have been used, $\beta = 0.05$, $\beta = 1/3$ and $\beta = 0.8$. All filters had the same width $\sigma = 0.001$. The dynamic range can be seen in Fig. 11. The dynamic range is a strong function of the step size. For $\beta = 0.8$ the filter responds nearly the same as the pyramid with a limiting dynamic range of 1.5λ . The $\beta = 1/3$ sigmoid dynamic range is limited to 25λ and for the $\beta = 0.05$ filter only a lower bound of 125 waves could be measured. This shows that

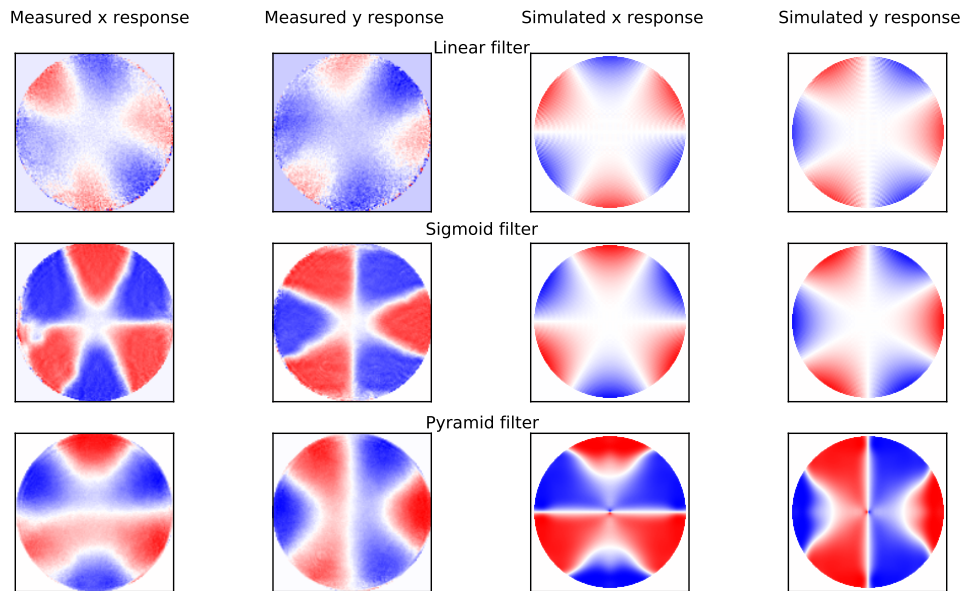


Fig. 9. In these figures the response due to a quadfoil aberration is shown for different filters. The left side corresponds to measured responses from the lab setup. The responses on the right side are simulated responses for each filter. For each measurement the same exposure time is used. The first and third column are the x gradients and the second and fourth columns are the y gradients. Blue corresponds to a saturated negative measurement and red to a saturated positive response. White is a response of zero. The top figure is the measurement from a linear filter. The middle is from the G-ODWFS sigmoid filter. And the bottom figures are the responses from the pyramid.

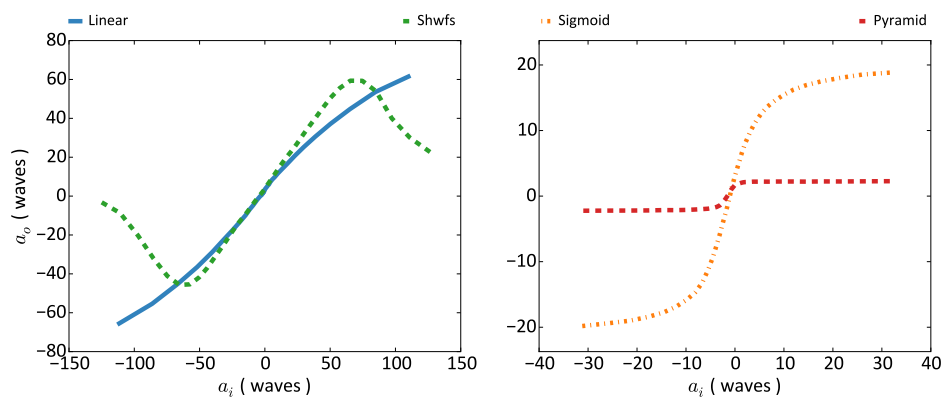


Fig. 10. The responses curves for the tilt mode are shown. On the left the whole range of input tilt is plotted against the response for the SHWFS and the ODWFS. The figure on the right shows the same thing but for the G-ODWFS and the PWFS.

the dynamic range changes non-linear with a change in step size.

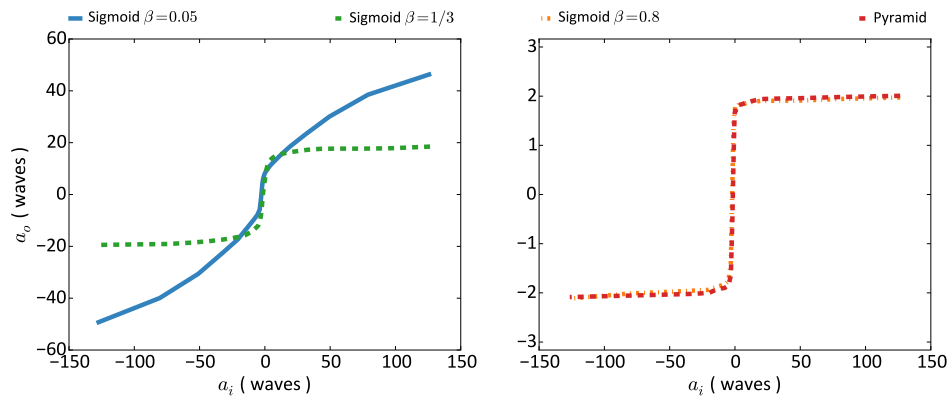


Fig. 11. The response curves for the different step size filters. On the x-axis are the input tilt coefficients and on the y-axis the responses. The figure on the left shows the response for small steps, and the right figure shows the response for large steps.

An estimate for the sensitivity is the Cramer-Rao lower bound [29]. This sets a lower bound on the standard deviation due to photon noise, and so also on the sensitivity. The lower it is the more sensitive a wavefront sensor is. The result for the pyramid, linear and sigmoid filters can be seen in Fig. 12. The pyramid is the most sensitive as was expected. The sigmoid is in between the linear and the pyramid and the linear is the least sensitive. The average mode sensitivity of the sigmoid compared to the pyramid is 3.25. This is very close to the expected value of 3, for a profile with $\beta = 1/3$.

5. Implementation concept

The current test setup can be used in AO systems that operate at a single wavelength and where the operation speed is not important. For many situations the operating frequency is several tens of Hertz to several thousands of Hertz, and the light source is broadband. A concept is proposed here for these cases. The concept is not based on a SLM and will use static optics.

In the low light regime, where high speed AO systems usually are, it is crucial to retain as many photons as possible. Amplitude filters will always block at least 50% of the light and this makes the amplitude filters not a good choice if we need to retain as many photons as possible. The concept proposed here is based on the Yet Another Wavefront sensor (YAWFS) [30], which uses the polarization of light to modulate the amplitude. This makes the WFS more photon efficient because it only changes the polarization of light, there is no blocking any more as is the case for amplitude filters.

If a spatially varying polarization rotator is placed between two polarisers it is possible to make an amplitude modulator by choosing the amount of rotation. By switching the polariser for a polarising beam splitter, both polarizations can be imaged at the same time. For one polarisation the amplitude will be proportional to $\cos(\theta_{\text{out}})$ and for the other it will be $\sin(\theta_{\text{out}})$, here θ_{out} is the polarisation angle after rotation. If the angle of the out coming polarization is limited to $0 \leq \theta_{\text{out}} \leq \frac{\pi}{2}$ then the sine and cosine projections will be mirrored versions of each other.

This creates the two mirrored amplitude filters that are necessary for the G-ODWFS. The hybrid filter that was proposed can not be recreated with this method, because the polarization projections causes non-linear amplitude filter profiles. What can be done is to create a rotation profile where the rotation is linearly varied and has a step in the middle. The difference between

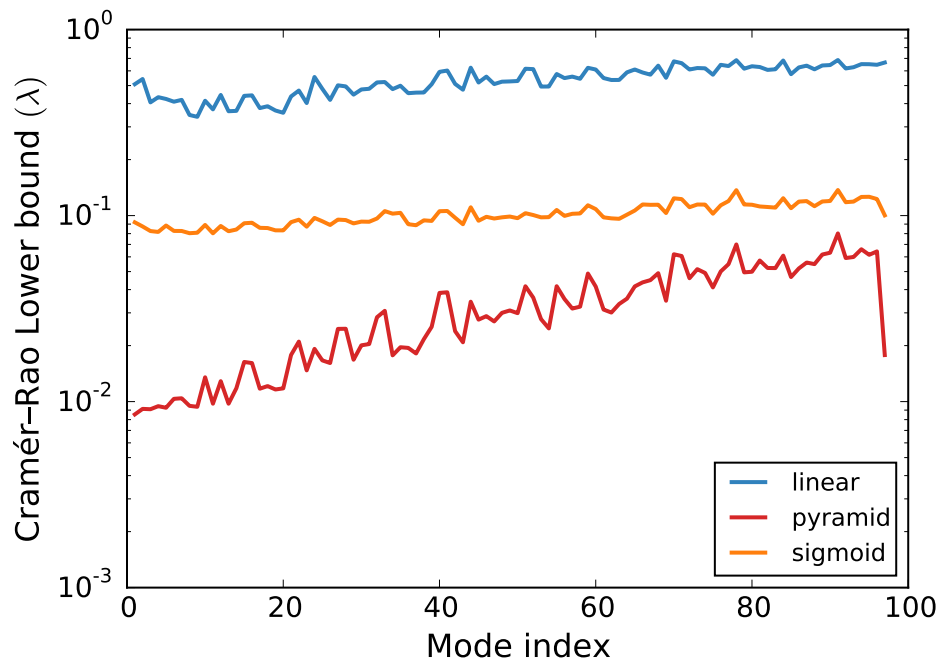


Fig. 12. Here the Cramer-Rao lower bound is shown against the mode index of the deformable mirror. The CR bound is shown in units of λ . The pyramid has the lowest bound and the linear the highest. The sigmoid filter here has a $\beta = 1/3$.

this and the previously proposed hybrid filter is in the continuous part of the filter. It is no longer a linear filter with a step in the middle, but more like a sine filter with a step in the middle. But as shown before in Section 2, it is not necessary to use a linear amplitude filter to retrieve the gradients. The use of a non-linear shape comes at the cost of more higher orders which will increase the diffraction effects.

In the YAWFS the polarization rotator was made with two pieces of quartz crystal with opposite chirality. Left chiral quartz rotates the polarization counter clockwise and right chiral quartz rotates clockwise. Two triangular pieces of quartz with angles of 45 degrees and opposite chirality can create a linearly varying polarization rotator as is shown in the YAWFS. To create arbitrary patterns in crystals is quite difficult. With recent developments in Liquid Crystal(LC) technology it is possible to create arbitrary patterned LC [31, 32]. The direct write method [32] makes it possible to control the shape of the pattern to very high resolution, and can be made broadband by applying several layers of retarders [31]. A spatially varying polarisation rotator can be made with the patterned retarder by making a half-wave plate with spatially varying orientation of the half-wave axis.

The G-ODWFS will need two focal plane filters, one for each direction in which the gradient is required. And the input beam has to be polarised if the patterned retarders are used. Both can be easily achieved by using a Wollaston prism. The Wollaston prism splits the light up into two orthogonal polarised beams, one for each filter. A schematic of the design is shown in Fig. 13. The beam is focused through the first Wollaston which splits the light into two beams. The two beams are filtered by separate filters, one for the x direction and one for the y direction. After the filtering another Wollaston is used to split the two beams in four. These are then collimated on a

camera with a lens. With this setup all four pupils are measured at the same time, and with a high photon efficiency. This makes it possible to use this WFS in a high speed setting.

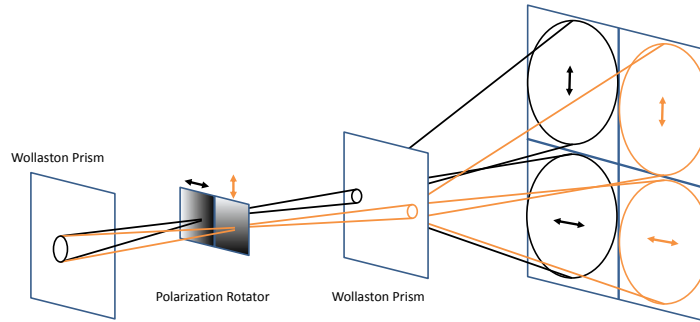


Fig. 13. The optical set-up is fed by an unpolarised focused beam. The first Wollaston prism splits the light in horizontal and vertical polarisation. Then the beams are filtered by the varying half wave plates in the focal plane. Another Wollaston prism will then split the two beams into four. The lens that transforms the focused beams to pupil images is not shown, but should be behind the second Wollaston prism.

6. Conclusion

The theory of optical differentiation is extended to non-linear filters, by expanding the filter shape in terms of Taylor polynomials. This shows that a WFS that filters in the focal plane can only decouple the dynamic range and sensitivity if local changes to the filter shape are applied. This leads to a hybrid filter design for the G-ODWFS, which combines the high sensitivity of a pyramid wavefront sensor and the large dynamic range of an optical differentiation wavefront sensor.

The filter that has been most extensively tested had a step size of $\beta = 1/3$. This should have a sensitivity which would be three times lower than that of a pyramid for small phase aberrations, which was confirmed through measurements of the Cramer-Rao lower bound. The dynamic range for the tilt mode was also measured and estimated to be roughly $25 \lambda/D$. So this filter is three times less sensitive but has a 25 times higher dynamic range. This shows that the trade off between dynamic range and sensitivity is on a different trade off curve than the typical linear trade off curve of conventional slope measuring WFS's. The dynamic range changes non-linearly with a change in step size. For $\beta = 0.05$ only a lower bound of $125 \lambda/D$ could be measured, $\beta = 1/3$ has a dynamic range of $25 \lambda/D$ and the $\beta = 0.8$ has a dynamic range of $1 \lambda/D$ which is comparable to the pyramid.

The dynamical simulations also show that the GDWFS can have a performance comparable or better than the PWFS. To give an accurate picture of the performance of the new WFS compared to a PWFS more simulations and tests have to be done in a wider parameter space and under many different atmospheric conditions. From the current results the generalised optical differentiation wavefront sensor would be an excellent wavefront sensor for high contrast imaging of point sources, because it reaches a high Strehl ratio, has a high temporal bandwidth and keeps the point spread function very stable.

There are also several improvements that could be done. The analysis of the higher order modes can be extended, and the influence of the higher orders should be investigated. If some higher

order modes are not important then the design of the filter can be relaxed. For this work a few filter shapes have been tested, but the filter could be optimised by using numerical optimisation methods. Because the response for large aberrations becomes non-linear, performance can be gained if non-linear phase reconstruction is used. This can be done for the G-ODWFS because the response is monotonically increasing. The non-linear reconstruction can be implemented by using a dynamic gain, which optimizes itself during closed loop operations, or by calibrating the interaction with high precision and use a non-linear function to map the measured coefficients to the input coefficients.

Acknowledgments

The author would like to thank Prof. C.U. Keller (Leiden University) for the helpful discussions, and also the anonymous reviewers for their comments to improve this work.