Direct measurement of Zernike aberration modes with a modal wavefront sensor

Martin J. Booth

Department of Engineering Science, University of Oxford, U.K.

ABSTRACT

It is often convenient to represent a wavefront aberration by the superposition of several aberration modes, for example, using the set of Zernike polynomials. In many practical situations the total aberration can be accurately represented by a small number of such modes. It is therefore desirable to be able to measure directly the modal content of the wavefront. The modal wavefront sensor allows us to do just this. This sensor can be applied to closed-loop aberration correction in adaptive systems and to direct absolute measurement of modal aberration coefficients. One implementation offers the possibility of an adaptive optics system where we have disposed altogether of a separate wavefront sensor. We present here extension of the modal wavefront sensing theory and explore the options for optimisation of the design. We investigate the linear measurement range of the sensor and the performance in closed-loop systems.

Keywords: Wave front sensing, Zernike polynomials, adaptive optics

1. INTRODUCTION

The ability to measure the phase properties of wave fronts has many applications in optical science and engineering and various methods of wave front sensing have been developed for this purpose. Among these methods, the most prominent are interferometry and Hartmann-Shack wave front sensors.[1] Both methods can provide detailed wave front information but are sometimes over-complex or inappropriate for simple applications. Interferometers, for example, whilst providing highly detailed wave front information, can be bulky and require careful use. They are also limited by the requirement of a coherent light source. Hartmann-Shack sensors are versatile but are not necessarily optimised for particular applications. In many cases, what one is interested in is the modal content of a wave front, that is the amplitude of the aberration modes from which the wave front is composed. With traditional sensors, the usual approach is to calculate the modal content from indirect measurements of properties of the wave front. For example, in the Hartmann-Shack sensor, local measurements of the wave front phase gradient are used to reconstruct the wave front.[2]

In several practical applications, it is known that only a small number of aberration modes are significant. It is therefore desirable to be able to measure these modes directly and it was for this purpose that the modal wave front sensor (MWFS) was introduced.[3] The modal content of a wave front is usually expressed in terms of the Zernike polynomials. These functions form an orthogonal set over the unit circle and are convenient since the lower order functions are equivalent to conventional aberrations such as astigmatism and coma. Although the concept of the MWFS is also applicable to other sets of functions, it is particularly useful for the Zernike polynomials.

The MWFS uses the technique known as wave front biasing. This is the deliberate addition of chosen amounts of aberration modes to the input wave front. When the wave front is focussed by a positive lens the intensity of the resulting focal spot is altered by the bias. By using appropriate combinations of biases and focal spot measurements it is possible to generate outputs signals dependent on the modal content of the input wave front. Conceptually, the MWFS consists of several sensors that each respond to an aberration mode of interest. It has been shown theoretically that a sensor for a particular aberration mode generates an output signal that is proportional to the modal amplitude in the input wave front.[3] It was also shown that some small

Further author information:
E-mail: martin.booth@eng.ox.ac.uk
Full address: Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, United Kingdom

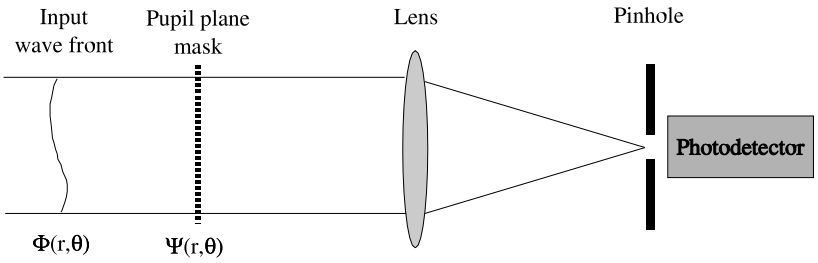

Figure 1. Schematic diagram of the wave front sensor measurement system. The input wave front is incident from the left whereupon it passes through the pupil plane of the lens. A phase element in this plane adds a chosen phase aberration to the input wavefront. The wave front is then focussed by the lens onto pinhole positioned at the nominal focus where the intensity signal is measured by the photodetector.

cross sensitivities between certain modes exist, for example between different orders of spherical aberration. This small degree of cross sensitivity is however predictable and can be easily accommodated by simple data processing.

The MWFS has been demonstrated experimentally in closed loop adaptive systems.[4,5] Different MWFS implementations have been investigated. One implementation allows the wave front biases and focal intensity measurements to be performed sequentially. The biases can be introduced using an adaptive aberration correction element such as a deformable mirror of spatial light modulator. This method is particularly useful since it permits the implementation of wave front sensing effectively without a separate wave front sensor. The sequential biasing method has been implemented in an adaptive optics system for correction of specimen induced aberrations in confocal and two-photon fluorescence microscopy.[3,5] It has also been shown that the MWFS is particularly useful in microscopy since it only responds to light emitted from the focal plane unlike other forms of wave front sensor.[6]

Another MWFS implementation allows the simultaneous measurement of a number of aberration modes. This implementation uses a binary phase, computer generated hologram to produce an array of biased focal spots. The focal intensities are then measured by an array of photodetectors or a CCD camera. This method has been demonstrated in a closed-loop correction system.[4]

In this paper, we consider the theory operation of the MWFS and investigate methods of optimising its performance. We discuss methods for choosing the bias aberrations and introduce new ways of calculating the output signal from the sensor. We show that these developments result in improved measurement accuracy of the sensor. This has beneficial implications for the use of the MWFS in both open-loop wave front measurement and closed-loop adaptive optics systems.

2. THEORY OF MODAL WAVE FRONT SENSING

2.1. Modal wave front sensor configuration

Our analysis is based upon the conceptual measurement system shown in Figure 1. The input wave front we wish to measure is incident from the left. A positive lens focuses the wave front onto an infinitely small pinhole photodetector situated at the nominal focal point of the lens. The phase aberration of this input wave front is described by the function $\Phi(r,\theta)$, where r and θ are polar coordinates in the pupil plane of the lens. The coordinates are normalised such that the pupil has a radius of 1. In the pupil plane of the lens is a phase element, which could in practice be an adaptive or fixed element, that adds a phase function $\Psi(r,\theta)$ to the input wave front. Using Fourier-based diffraction theory,[7] we can show that the signal measured by the photodetector is given by

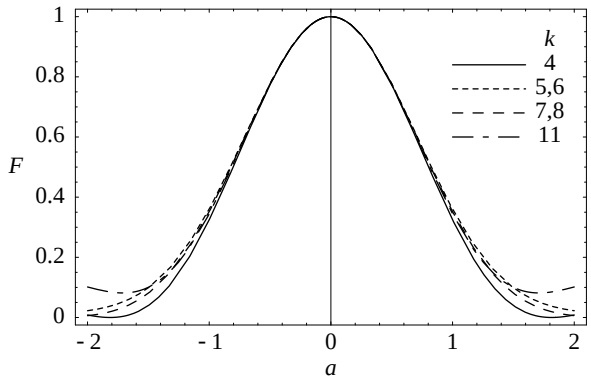

Figure 2. The photodetector signal calculated as a function of input mode amplitude for different Zernike aberration modes: $k = 4$, defocus; $k = 5, 6$ astigmatism; $k = 7, 8$ coma; $k = 11$, spherical. The two astigmatism modes, and similarly the two coma modes, are grouped together since the phase functions differ only by a rotation about the optic axis and hence the photodetector responses are identical. For small amplitudes, each response is identical irrespective of the Zernike mode. The value of I_0 was taken to be 1.

$$F = I_0 \left| \frac{1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} \exp\{j\Phi(r,\theta) + j\Psi(r,\theta)\} r \mathrm{d}r\mathrm{d}\theta \right|^2 , \tag{1}$$

where I_0 is a multiplying factor that is proportional to the incident light power and $j = \sqrt{-1}$. We assume that $\Phi(r,\theta)$ and $\Psi(r,\theta)$ have respective mean values of zero. This assumption is reasonable since only the phase variations across the wavefront, rather than its absolute phase, would have an effect on the measured intensities. Let us consider the case when the input wavefront contains an amount a of single Zernike aberration mode, Z_k:

$$\Phi(r,\theta) = aZ_k(r,\theta) \ . \tag{2}$$

The Zernike modes and their algebraic forms are shown in Table 1. If $\Psi(r,\theta) = 0$, then the photodetector signal is given, as a function of a, by:

$$F(a) = I_0 \left| \frac{1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} \exp\{jaZ_k(r,\theta)\} r \mathrm{d}r\mathrm{d}\theta \right|^2 . \tag{3}$$

The function $F(a)$ is even with respect to the mode amplitude a. It is also useful to note that for small a the function $F(a)$ has the same form irrespective of the choice of Zernike mode[8]:

$$F(a) \approx I_0(1 - a^2) \ . \tag{4}$$

This Independence of the Zernike aberration mode is an important consequence of using Noll's normalisation.[9] Figure 2 shows the function $F(a)$ for some common aberration modes. It is apparent that the curves are all similar for small values of a. As a consequence of this similarity, we find that generally the properties of each aberration mode sensor are also similar.

For convenience, we shall rewrite Equation 3 as

$$F(a) = I_0 f(a) \ , \tag{5}$$

where

$$f(a) = \left| \frac{1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} \exp\{jaZ_k(r,\theta)\} r \mathrm{d}r\mathrm{d}\theta \right|^2 \tag{6}$$

is not dependent upon the overall intensity.

2.2. Modal wave front sensor optimisation

A number of approaches can be taken to optimise the design of the MWFS. We extend here the previous theoretical modelling of the MWFS in order to elucidate these approaches and to introduce new algorithms for calculation of the sensor output signals. Because of the similarity of the responses to the Zernike modes, the results are generally applicable. As illustrative examples, we calculate the optimum parameters for modes 2 to 22.

2.2.1. Wave front biasing

The modal wave front sensor uses the process of wave front biasing. This is the deliberate addition to the input wave front of a fixed amount of the aberration mode to be measured. This could be implemented using the pupil plane phase mask shown in Figure 1. In the simplest implementation, an appropriate aberration bias is introduced and the photodetector signal is measured. An equal but opposite bias is then introduced and a second measurement is made. The difference between these measurements is then proportional, within a certain range, to the amount of the measured aberration mode contained in the input wavefront. Mathematically, we can examine this process as follows. Consider that the input wave front contains one Zernike mode such that $\Phi(r,\theta) = aZ_k(r,\theta)$. Firstly, we choose an appropriate bias amplitude b and we apply the first aberration bias by setting $\Psi(r,\theta) = bZ_k(r,\theta)$. We represent the resulting photodetector signal as

$$g_+(a) = F(a+b) \ . \tag{7}$$

When the equal and opposite aberration bias is applied, such that $\Psi(r,\theta) = -bZ_k(r,\theta)$, the photodetector signal is given by

$$g_-(a) = F(a-b) \ . \tag{8}$$

For later convenience, we also introduce the function

$$g_0(a) = F(a) \ , \tag{9}$$

where no bias is applied. Taking account of the even symmetry of the function $f(a)$, the right hand sides of equations 7, 8 and 9 can be expanded as Taylor series to give

$$g_+(a) = I_0\left[f(b) + af'(b) + \frac{a^2}{2}f''(b) + \frac{a^3}{6}f'''(b) + O(a^4)\right] \tag{10}$$

$$g_-(a) = I_0\left[f(b) - af'(b) + \frac{a^2}{2}f''(b) - \frac{a^3}{6}f'''(b) + O(a^4)\right] \tag{11}$$

$$g_0(a) = I_0\left[f(0) + \frac{a^2}{2}f''(0) + O(a^4)\right] \ , \tag{12}$$

where $O(a^n)$ represents series terms of order n and higher. Although the exact form of the function $f(.)$ is different for each Zernike mode, for small arguments the response is similar, as can be seen in Figure 2.

2.2.2. Type A sensor output signal

In the original WFS implementation, as introduced by Neil et al.,[3] the output signal was derived as

$$\begin{aligned} W_A(a) &= g_-(a) - g_+(a) \\ &= -I_0\left[2af'(b) + \frac{1}{3}a^3f'''(b) + O(a^5)\right] \ . \end{aligned} \tag{13}$$

We refer to this as the type A output signal. As was shown by Neil et al.,[3] for small values of the input mode amplitude a, this output signal is proportional to a. The expressions in Equation 13 allow us to examine how the bias amplitude b can be chosen in order to optimise the operation of the WFS. One approach is to maximise the value of $f'(b)$ to increase the sensitivity of the sensor in the linear region. The optimum value of b depends upon the Zernike mode in question (see Table 1); for modes 2 to 22 all values lie in the range 0.61 to 0.75.

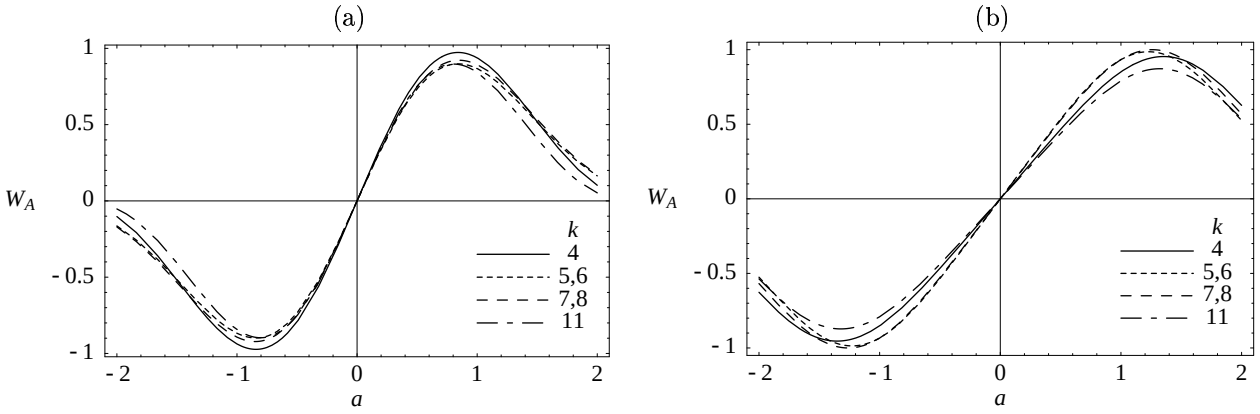

Figure 3. Output signals calculated as a function of input mode amplitude for different Zernike aberration modes: $k = 4$, defocus; $k = 5, 6$ astigmatism; $k = 7, 8$ coma; $k = 11$, spherical. The value of I_0 was taken to be 1. In (a) the bias amplitude was chosen to maximise sensitivity, whereas in (b) it was chosen to maximise linearity of the sensor.

Alternatively, we could choose b in order to maximise the linearity of the WFS by setting the third order term in Equation 13 to zero. In this case the optimum bias values lie in the range 1.08 to 1.34 for modes 2 to 22 (see Table 1). Examples of the resulting response curves are shown in Figure 3. Sensors designed in this way have been demonstrated in practical systems.[4, 5, 10, 11]

It is important to note that the output signal of Equation 13 depends upon I_0 and thus direct measurement of the input amplitude a is only possible if we have explicit knowledge of the total intensity. This could be measured, for example, with a large area detector. However, this could add to the complexity of a practical system. Furthermore, the output signal $W_A(a)$ is also dependent upon the presence of other aberration modes in the input wavefront. As an illustrative example, let us consider that two Zernike modes are present: the mode we wish to measure, Z_k of amplitude a, and another mode, Z_l with amplitude a_l. The photodetector output $F(a, a_l)$ is then given by

$$F\left(a, a_l\right) = I_0 f\left(a, a_l\right) \tag{14}$$

where

$$f\left(a, a_l\right) = \left| \frac{1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} \exp\{jaZ_k\left(r, \theta\right) + ja_l Z_l\left(r, \theta\right)\} r \mathrm{d}r \mathrm{d}\theta \right|^2 \tag{15}$$

and the output signal becomes

$$W_A\left(a, a_l\right) = F\left(a - b, a_l\right) - F(a + b, a_l)\ . \tag{16}$$

As an illustrative example, Figure 4 shows the response of a mode 4 (defocus) sensor optimised for sensitivity and linearity when different amounts of mode 5 (astigmatism) are present. It can be seen the the presence of the other mode reduces the sensitivity and would lead to an erroneous measurement of a. This has obvious implications for the useful range of open-loop aberration measurement and the bandwidth of closed-loop aberration correction systems.

2.2.3. Type B sensor output signal

In the previous section some limitations of the type A output signal were demonstrated. However, through simple changes in the definition of the output signal, it is possible to circumvent both the dependence on the overall intensity and the effect of other aberration modes. Consider the following type B output signal derived

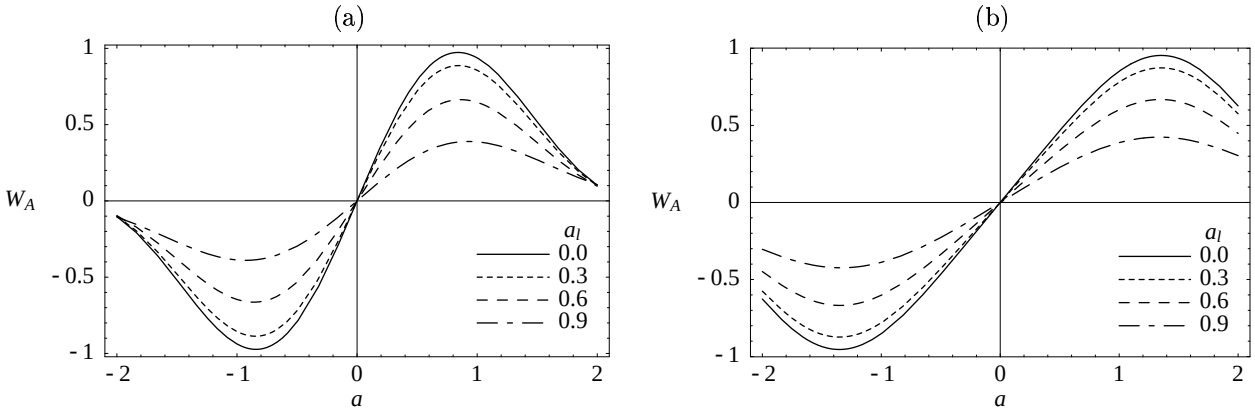

Figure 4. Output signals of a mode 4 (defocus) sensor optimised for (a) sensitivity and (b) linearity when an amount a_l of mode 5 (astigmatism) is also present in the input wave front. The sensitivity of the sensor is affected by the presence of another mode. The value of I_0 was taken to be 1.

from the same two photodetector measurements $g_-(a)$ and $g_+(a)$:

$$\begin{aligned} W_B(a) &= \frac{g_-(a) - g_+(a)}{g_-(a) + g_+(a)} \\ &= -\frac{2af'(b) + \frac{1}{3}a^3 f'''(b) + O(a^5)}{2f(b) + a^2 f''(b) + O(a^4)} \,. \end{aligned} \tag{17}$$

We could eliminate the second order term on the denominator by choosing b such that $f''(b) = 0$. This is, of course, equivalent to our previous requirement that $f'(b)$ be maximum. The optimum bias values are therefore the same as those shown in Table 1. Selecting b according to this criterion simplifies Equation 17 to

$$W_B(a) \approx -\frac{af'(b)}{f(b)} \,. \tag{18}$$

Example responses for this sensor output signal are shown in Figure 5. This output signal is linear in a and independent of I_0. Moreover, it is resilient to the presence of other aberration modes. We investigate this property in the next Section. Since W_B requires the same two measurements used in the calculation of W_A but has these superior properties, in most applications W_B would be a more appropriate output signal.

2.2.4. Type C sensor output signal

It is possible to use also the unbiased intensity measurement $g_0(a)$ in the derivation of the output signal of the wave front sensor. This type C output signal would take the form

$$\begin{aligned} W_C(a) &= \frac{g_-(a) - g_+(a)}{g_-(a) + \gamma g_0(a) + g_+(a)} \\ &= -\frac{2af'(b) + \frac{1}{3}a^3 f'''(b) + O(a^5)}{2f(b) + \gamma f(0) + a^2\left[f''(b) + \frac{\gamma}{2}f''(0)\right] + O(a^4)} \,, \end{aligned} \tag{19}$$

where γ is a constant. Here we find optimal values for both b and γ by firstly choosing b according to the optimum linearity criterion, $f'''(b) = 0$. We then choose the value of γ according to $\gamma = -2f''(b)/f''(0)$ so that the second order term on the denominator is eliminated. These calculated values of γ are shown in Table 1. The output signal then becomes

$$W_C(a) \approx -\frac{2af'(b)}{2f(b) + \gamma} \,, \tag{20}$$

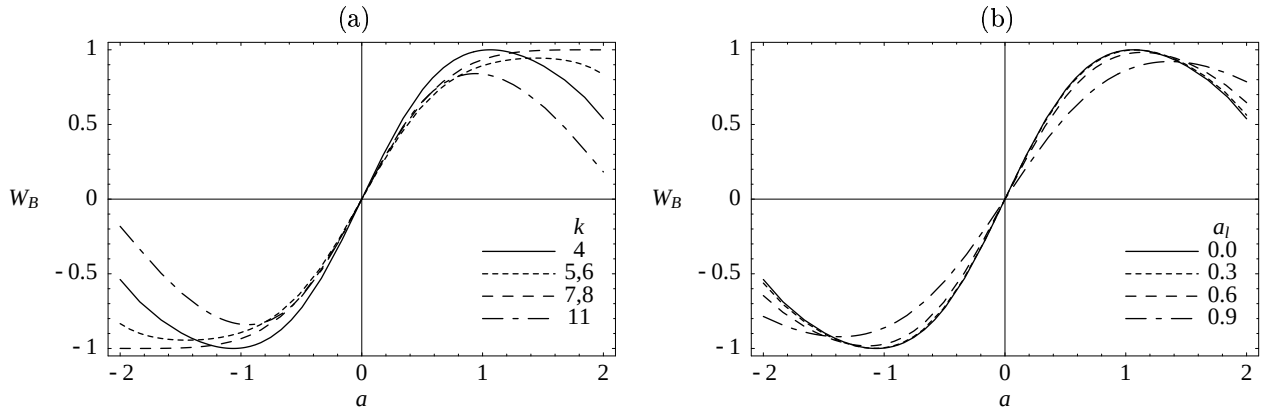

Figure 5. (a) Type B output signals calculated as a function of input mode amplitude for different Zernike aberration modes: $k = 4$, defocus; $k = 5, 6$ astigmatism; $k = 7, 8$ coma; $k = 11$, spherical. The response is independent of I_0. (b) The response of a mode 4 (defocus) sensor when different amounts a_l of mode 5 (astigmatism) are present. Comparison with the corresponding type A curves in Figure 4(a) shows that the type B output signal possesses greater independence of other modes.

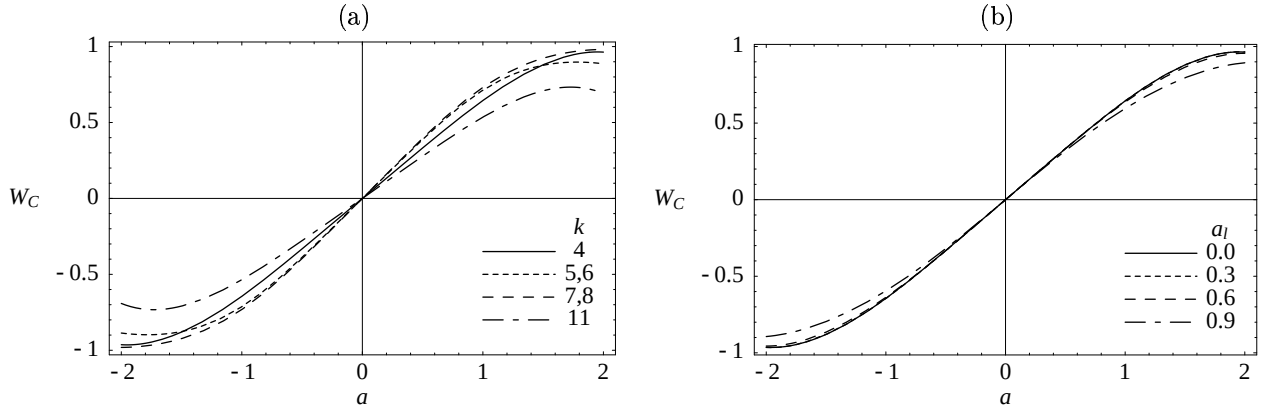

Figure 6. (a) Type C output signals calculated as a function of input mode amplitude for different Zernike aberration modes: $k = 4$, defocus; $k = 5, 6$ astigmatism; $k = 7, 8$ coma; $k = 11$, spherical. The response is independent of I_0. (b) The response of a mode 4 (defocus) sensor when different amounts a_l of mode 5 (astigmatism) are present. It can be seen that the type C output signal possesses independence of other modes.

where we have used the fact that $f(0) = 1$. As with W_B, this output signal is linear in a, independent of I_0 and resilient to the presence of other aberration modes. Example responses are shown in Figure 6.

2.3. Rejection of other aberration modes

In Figure 4 it is clear that the output signal W_A is adversely affected by the presence in the input wave front of aberration modes other than that being measured. One major benefit of using the output signals defined in Equations 17 and 19 is that they are independent of other modes. In this Section we investigate this property further.

If we perform the Taylor expansions of the functions $g_+(a)$ and $g_-(a)$ about the points $a = c$ and $a = -c$ respectively, we obtain the expressions

$$g_+(a) = I_0 \left[f(c) + (a + b - c) f'(c) + ... \right] \tag{21}$$

$$g_-(a) = I_0 \left[f(c) - (a - b + c) f'(c) + ... \right] \ . \tag{22}$$

Table 1. Zernike mode definitions and optimum parameter values derived according to the criteria described in the text. For each Zernike mode there are two values listed for the optimum bias: (1) derived for maximum sensitivity, (2) derived for maximum linearity. The mode indexing scheme is explained by Neil et al.[3]

Index			Zernike mode		Optimum parameters		
k	n	m		Name	b (1)	b (2)	γ
2	1	1	$2r\cos(\theta)$	Tip	0.741	1.308	1.128
3	1	-1	$2r\sin(\theta)$	Tilt	0.741	1.308	1.128
4	2	0	$\sqrt{3}(2r^2-1)$	Defocus	0.750	1.337	1.211
5	2	2	$2\sqrt{3}r^2\cos(2\theta)$	Astigmatism	0.710	1.228	0.960
6	2	-2	$2\sqrt{3}r^2\sin(2\theta)$	Astigmatism	0.710	1.228	0.960
7	3	1	$2\sqrt{2}(3r^3-2r)\cos(\theta)$	Coma	0.721	1.256	0.987
8	3	-1	$2\sqrt{2}(3r^3-2r)\sin(\theta)$	Coma	0.721	1.256	0.987
9	3	3	$2\sqrt{2}r^3\cos(3\theta)$	Trefoil	0.678	1.171	0.891
10	3	-3	$2\sqrt{2}r^3\sin(3\theta)$	Trefoil	0.678	1.171	0.891
11	4	0	$\sqrt{5}(6r^4-6r^2+1)$	Spherical	0.721	1.291	1.247
12	4	2	$\sqrt{10}(4r^4-3r^2)\cos(2\theta)$	Astigmatism	0.714	1.242	0.939
13	4	-2	$\sqrt{10}(4r^4-3r^2)\sin(2\theta)$	Astigmatism	0.714	1.242	0.939
14	4	4	$\sqrt{10}r^4\cos(4\theta)$		0.647	1.124	0.866
15	4	-4	$\sqrt{10}r^4\sin(4\theta)$		0.647	1.124	0.866
16	5	1	$2\sqrt{3}(10r^5-12r^3+3r)\cos(\theta)$	Coma	0.711	1.237	0.918
17	5	-1	$2\sqrt{3}(10r^5-12r^3+3r)\sin(\theta)$	Coma	0.711	1.237	0.918
18	5	3	$2\sqrt{3}(5r^5-4r^3)\cos(3\theta)$	Trefoil	0.703	1.220	0.887
19	5	-3	$2\sqrt{3}(5r^5-4r^3)\sin(3\theta)$	Trefoil	0.703	1.220	0.887
20	5	5	$2\sqrt{3}r^5\cos(5\theta)$		0.619	1.081	0.860
21	5	-5	$2\sqrt{3}r^5\sin(5\theta)$		0.619	1.081	0.860
22	6	0	$\sqrt{7}(20r^6-30r^4+12r^2-1)$	Spherical	0.732	1.301	1.022

The type B output signal of Equation 17 can then be expressed as*

$$\begin{aligned} W_B(a) &= \frac{g_-(a)-g_+(a)}{g_-(a)+g_+(a)} \\ &\approx -\frac{af'(c)}{f(c)}\ , \end{aligned} \tag{23}$$

with the requirement that a is small and $b \approx c$. We now consider the form of function $f(.)$ when two aberrations are present in the input wave front: the mode to be measured, Z_k, and another mode, Z_l with amplitude a_l. Using the definition of Equation 15 it can be shown that for small values of c and a_l

$$f(c,a_l) \approx (1-c^2)(1-a_l^2)\ , \tag{24}$$

from which it follows that

$$f'(c,a_l) \approx -2c(1-a_l^2)\ , \tag{25}$$

*Although this expression is similar to that in Equation 18, we find that the constraints placed upon allowable values of b are less stringent in this case.

where the prime denotes the derivative with respect to the first argument. It is then straightforward to show that

$$W_B(a) \approx \frac{2ac}{1-c^2} , \tag{26}$$

where we see that the output signal is linear in a and not dependent upon a_l.

3. MEASUREMENT RANGE OF THE M.W.F.S.

The performance of a wave front sensor can be measured in several ways including linearity, measurement range or temporal bandwidth. Some influencing factors are dependent upon the implementation, for example, the choice of photodetector will influence the noise performance. Other aspects are inherent in the design of the sensor, for example, the measurement range. In this Section we investigate the measurement capabilities of the MWFS based upon the design specifications discussed in the previous Section. Two methods of implementation, sequential biasing and simultaneous biasing, are considered.

3.1. Sequential biasing method

This implementation of the MWFS permits the design of adaptive optics systems without a dedicated wave front sensor. If we consider the set-up of Figure 1 with an adaptive phase element, e.g. a deformable mirror, placed in the pupil plane, we have an adaptive system that can sequentially apply the necessary wave front biases to the input wave front. This allows us to take the photodetector intensity measurements sequentially. The measurement and correction for a type B sensor proceeds as follows. Firstly, to measure the mode Z_k we apply the positive bias and measure the intensity g_+. Then we apply the negative bias and measure the intensity g_- (If we were using output signal W_C we should also measure g_0 with no bias applied). From these values, we then calculate the output signal as

$$a_{out} = \frac{W_B}{S_k} = \frac{1}{S_k}\left(\frac{g_- - g_+}{g_- + g_+}\right) , \tag{27}$$

where the sensitivity of the mode k sensor, S_k, is calculated as

$$S_k = \left.\frac{\partial W_B}{\partial a}\right|_{a=0} . \tag{28}$$

Within the correct measurement range, a_{out} is equal to the amplitude a_{in} of the input mode. This mode is then corrected by adding the phase $-a_{out}Z_k(r,\theta)$ to the input wave front using the adaptive element. Correction of other modes proceeds sequentially in a similar manner. When using a type C output signal, the corresponding calculations would be given by

$$a_{out} = \frac{W_C}{S_k} = \frac{1}{S_k}\left(\frac{g_- - g_+}{g_- + g_0 + g_+}\right) , \tag{29}$$

where the sensitivity is

$$S_k = \left.\frac{\partial W_C}{\partial a}\right|_{a=0} . \tag{30}$$

It is important to ascertain the range of aberrations that can be corrected using the sequential method. We can represent the modal content of the input wave front by a vector of Zernike mode coefficients, $\mathbf{a}$. Given a particular combination of modes in the input wave front, it is straightforward to simulate the correction process to calculate how many measurements are required. The results are dependent upon the exact combination of modes. However, we found that the average number of required measurements depends primarily on the magnitude of the coefficient vector, i.e. $|\mathbf{a}|$, particularly for $|\mathbf{a}| < 2$. As an example, results are shown in Figure 7 for correction sequences involving the eight Zernike modes 4 to 11. Each point on the graph was calculated as the average of 100 measurements for randomly generated coefficient vectors of the given amplitude. The aberration modes were measured sequentially in numerical order from 4 to 11. Hence, when $N = 8$ each mode had been measured and corrected once; this is one correction cycle. As the target for the simulation, we specified that the corrected Strehl ratio should be greater than 0.9. This is equivalent to specifying a wave front standard deviation of less than $\lambda/20$, where λ is the wavelength. We see in both Figures that $N = 0$ for $|\mathbf{a}| < 0.3$; this is a consequence of

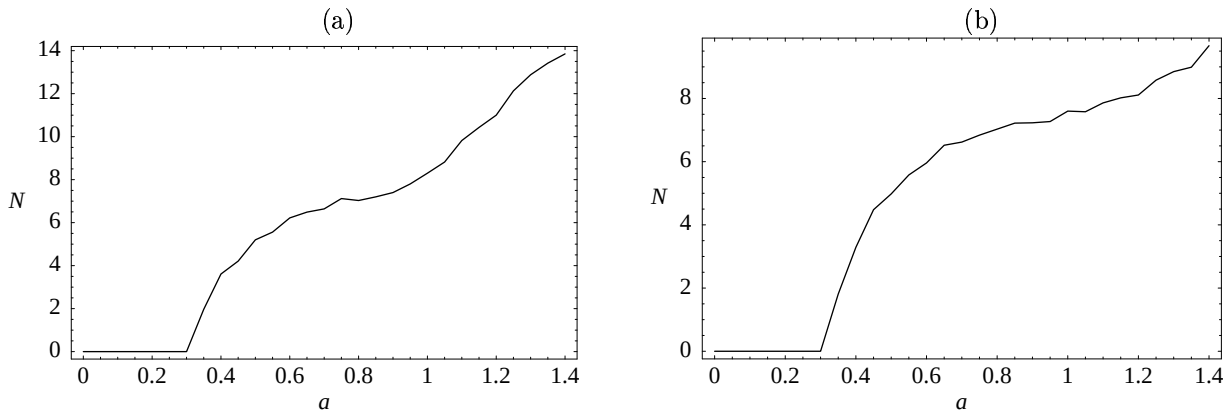

Figure 7. Convergence characteristics of the sequential MWFS measuring the eight modes from $k = 4$ to $k = 11$ inclusive: (a) using type B output signals, (b) using type C output signals. N represents the average number of aberration mode measurement/corrections required starting from a modal coefficient vector of magnitude $|\mathbf{a}|$. The modes were measured in order, from $k = 4$ to $k = 11$. The coefficient vector magnitudes corresponding to $N = 8$ (one correction cycle) are (a) $|\mathbf{a}| = 0.96$ and (b) $|\mathbf{a}| = 1.12$.

the Strehl ratio target - when the magnitude is below this value, the wave front is already corrected. The values of $|\mathbf{a}|$ for which, on average, one correction cycle is sufficient are $|\mathbf{a}| = 0.96$, using a type B output signal, and $|\mathbf{a}| = 1.12$, using a type C output signal. One should note that the each modal measurement in the type B case requires two intensity measurements, whereas three are required for type C.

We also see from the results that the measurement and correction scheme still converges even when $|\mathbf{a}|$ is greater than the values discussed. As such, the sequential MWFS can still be used with larger input aberrations if more correction cycles can be accommodated.

The implementation of the MWFS can be optimised for a particular application so that the sensor only measures the aberration modes that are typically present in the input wave front. Furthermore, the sequence of correction can be optimised. If, for example, the input wave front were dominated by a particular mode, it would be sensible to correct this mode first. In this way, the MWFS can be tailored to fit each application.

3.2. Simultaneous biasing method

In this implementation, the operation of the MWFS is multiplexed allowing simultaneous measurement of each mode of interest. This method has been demonstrated experimentally using a fixed computer generated hologram to produce multiple biased spots from a single input wave front.[3,4] If we consider the set-up of Figure 1, we can place such an element in the pupil plane and replace the single pinhole photodetector with an array of detectors, one for each of the biased spots. The output signals are derived from these photodetector measurements.

In the simultaneous method, we take account of the small cross sensitivity between certain modes by calculating a sensitivity matrix $\mathbf{S}$, the elements of which are given by

$$S_{ik} = \left.\frac{\partial W_i}{\partial a_k}\right|_{a_k=0} , \tag{31}$$

where W_i is the relevant output signal from a mode i sensor and a_k is the amplitude of the input mode Z_k. As was shown by Neil et al.,[3] this matrix is sparse and diagonally dominated. The mode amplitude vector $\mathbf{a}_{out}$ is calculated as

$$\mathbf{a}_{out} = \mathbf{S^{-1}w} , \tag{32}$$

where the elements of the vector $\mathbf{w}$ are the modal sensor output signals.

To investigate the measurement range of this sensor, we considered an input wave front containing the eight Zernike modes 4 to 11 and a MWFS designed to measure these same eight modes. We modelled the process as

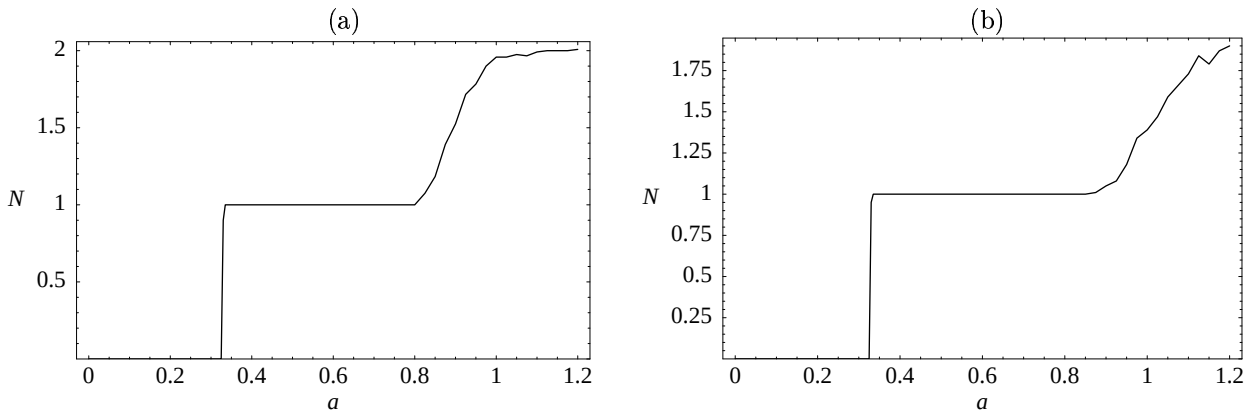

Figure 8. Convergence characteristics of the simultaneous MWFS measuring the eight modes from $k = 4$ to $k = 11$ inclusive: (a) using type B output signals, (b) using type C output signals. N represents the average number of aberration mode measurement/corrections required starting from a modal coefficient vector of magnitude $|\mathbf{a}|$. In (a) the largest magnitude for which $N = 1$ is $|\mathbf{a}| = 0.79$; in (b) it is $|\mathbf{a}| = 0.94$.

a correction system with a target Strehl ratio of 0.9. For a given combination of modes in the input wave front, we first calculated the vector of output signals, $\mathbf{w}$. Using Equation 32 we found the amplitude vector $\mathbf{a}_{out}$ and removed the corresponding modes from the input wavefront. If the target Strehl ratio was not met, the process was repeated. The results are shown in Figure 8 for sensors using both type B and type C output signals. Each point on the graph was calculated as the average of 100 measurements for randomly generated coefficient vectors of the given amplitude. We notice again that $N = 0$ for $|\mathbf{a}| < 0.3$ in both cases as a consequence of the target criterion. We also see that, for the type B sensor, only one measurement was required for $|\mathbf{a}| < 0.79$. For the type C sensor the corresponding relationship was $|\mathbf{a}| < 0.94$.

The results show that this implementation of the MWFS is capable of direct open-loop measurement of several Zernike modes when $|\mathbf{a}|$ is less than the values discussed. If this sensor were used in a closed-loop correction system, it would also provide convergence for larger values of $|\mathbf{a}|$ if more correction cycles could be accommodated. Again this sensor could be designed such that it only measured the aberration modes that are typically present in the input wavefront.

4. CONCLUSION

Using a simple model for the biasing and detection process in the MWFS we were able to examine theoretically the operation of the sensor. By expanding the mathematical representation of the detected signals as power series we showed that different approaches could be taken towards optimising the performance of the sensor. In one case, we maximised the sensitivity, whereas another involved the optimisation of the sensor linearity. By using different definitions of the sensor output signal it was possible to remove the dependence on the overall intensity and the presence of other aberration modes. The measurement range of the MWFS was investigated by modelling an adaptive aberration correction system for both sequential biasing and simultaneous biasing implementations. It was found that performance of the sensor depends primarily on the magnitude of the vector of modal coefficients. The sensor was shown to be linear for coefficient vector magnitudes below a particular value. In closed-loop systems the sensor would allow convergence for larger input aberrations.

REFERENCES

1. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes*, Oxford University Press, 1998.
2. B. Platt and R. Shack, "History and priciples of shack-hartmann wavefront sensing," *J. Refract. Surg.* **17**, pp. 573–577, 2001.

3. M. A. A. Neil, M. J. Booth, and T. Wilson, "New modal wavefront sensor: a theoretical analysis," *J. Opt. Soc. Am. A* **17**(6), pp. 1098–1107, 2000.
4. M. A. A. Neil, M. J. Booth, and T. Wilson, "Closed-loop aberration correction by use of a modal zernike wave-front sensor," *Opt. Lett.* **25**(15), pp. 1083–1085, 2000.
5. M. J. Booth, M. A. A. Neil, R. Juškaitis, and T. Wilson, "Adaptive aberration correction in a confocal microscope," *Proc. Nat. Acad. Sci.* **99**(9), pp. 5788–5792, 2002.
6. M. J. Booth, M. A. A. Neil, R. Juškaitis, and T. Wilson, "New modal wave-front sensor: application to adaptive confocal fluorescence microscopy and two-photon excitation fluorescence microscopy," *J. Opt. Soc. Am. A* **19**(10), pp. 2112–2120, 2002.
7. T. Wilson and C. Sheppard, *Theory and Practice of Scanning Optical Microscopy*, Academic Press, London, 1984.
8. M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, 6th ed., 1983.
9. R. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66**, pp. 207–277, 1976.
10. M. A. A. Neil, R. Juškaitis, M. J. Booth, T. Wilson, T. Tanaka, and S. Kawata, "Adaptive aberration correction in a two-photon microscope," *J. Microsc.* **200**(2), pp. 105–108, 2000.
11. M. A. A. Neil, R. Juškaitis, M. J. Booth, T. Wilson, T. Tanaka, and S. Kawata, "Active aberration correction for the writing of three-dimensional optical memory devices," *Appl. Opt.* **41**(7), pp. 1374–1379, 2002.