Wave front sensorless adaptive optics, modal wave front sensing, and sphere packings

Martin J. Booth

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

ABSTRACT

We investigate the properties of a class of adaptive optics systems that do not employ a wave front sensor but rather optimise a photodetector signal by appropriate control of an adaptive element. Such wave front control methods have already been implemented in various applications. It is often convenient to represent a wave front 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 shown that the design of wave front sensor-less adaptive optics systems based upon Zernike modes is related to the mathematical problem of sphere packing. This involves the arrangement of spheres in multiple dimensions, where the coordinate for each dimension corresponds to a Zernike mode amplitude. This observation permits optimisation of the systems providing considerable increases in efficiency over schemes that take no account of the geometries involved. We combine this approach with modal wave front sensing to provide efficient, direct measurement of Zernike aberration modes.

Keywords: Wave front sensing, Zernike polynomials, adaptive optics, sphere packings

1. INTRODUCTION

A conventional adaptive optics system employs a wave front sensor to measure aberrations, an adaptive correction element to remove the aberrations, and a control system that processes the sensor signals in order to drive the correction element.[1] Another approach permits the design of adaptive optics systems without a distinct, separate wave front sensor but rather a single photodetector. In general, the operation of these systems is based upon the maximisation of the photodetector signal, or a related measure, by adaption of the correction element. Such wave front sensorless adaptive systems have been implemented in confocal fluorescence microscopy,[2] two-photon fluorescence microscopy,[3,4] maximisation of second harmonic generation,[5] optical tweezers[6] and coupling laser light into an optical fibre.[7] These applications have employed different schemes for optimisation of the correction element based upon genetic algorithms,[3–5,7] hill climbing algorithms[4,6,7] and modal wave front sensing.[2] Although the detail of each of these approaches is different, the basis of operation is the same. In each case, the adaptive element is configured to add a particular aberration to the input wave front and a photodetector measurement is taken. According to some chosen scheme, the adaptive element is sequentially reconfigured and the corresponding photodetector measurements are taken. The final correction is ultimately derived, using the relevant algorithm, from this set of photodetector measurements. The two major design choices are therefore (i) what algorithm to use, and (ii) what sequence of aberrations to apply with the adaptive element. Of course, these two choices are usually somewhat linked. In this paper, we discuss what requirements an optimal correction scheme might have. We use a simple model of a wave front sensorless adaptive optics system and a direct search algorithm to show that problem of finding an optimal scheme is closely related to the mathematical problem of sphere packings in multiple dimensions. Using this connection, we argue that considerable gains in efficiency can be obtained by taking account of this spherical geometry in the design of more advanced schemes. This is illustrated by applying the concepts to the modal wave front sensor.[8–10]

2. WAVE FRONT SENSORLESS ADAPTIVE OPTICS

Our analysis is based upon the conceptual measurement system shown in Figure 1. The input wave front is incident from the left. A positive lens focuses the wave front onto an infinitely small pinhole photodetector

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

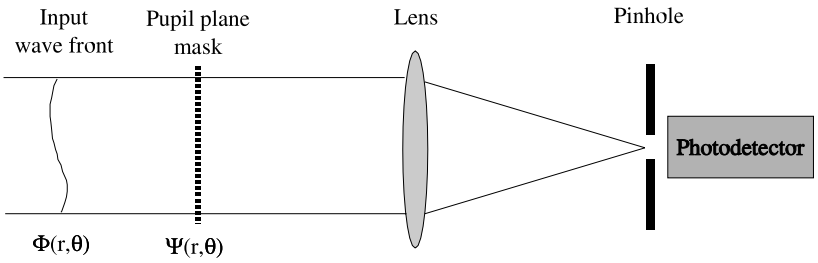

Figure 1. Schematic diagram of the wave front sensor measurement system. The input wave front is incident from the left whereupon it passes through the pupil plane of the lens. A phase element in this plane adds a chosen phase aberration to the input wavefront. The wave front is then focussed by the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

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 element, that subtracts a phase function $\Psi(r,\theta)$ from the input wave front. Using Fourier diffraction theory,[11] we can show that the signal measured by the photodetector is given by

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

where I_0 is a multiplying factor that is proportional to the incident light power and $j = \sqrt{-1}$. We assume that Φ and Ψ 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:

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

and the correction element introduces an amount b of the same mode:

$$\Psi(r,\theta) = bZ_k(r,\theta) . \tag{3}$$

The Zernike modes, their algebraic forms and the indexing scheme used in this paper are explained by Neil et al..[8] The photodetector signal is given, as a function of $c = a - b$, by:

$$F(c) = I_0 \left| \frac{1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^{1} \exp\{jcZ_k(r,\theta)\} r \mathrm{d}r\mathrm{d}\theta \right|^2 . \tag{4}$$

It is also useful to note that for small c the function $F(c)$ has the same form irrespective of the choice of Zernike mode[12]:

$$F(c) \approx I_0(1 - c^2) . \tag{5}$$

This independence of the Zernike aberration mode is an important consequence of using Noll's normalisation.[13] Figure 2 shows the function $F(c)$ for some common aberration modes. It is apparent that the curves are all similar for small values of c. We also note that, using this normalisation, the quantity c^2 is equivalent to the variance of the aberrated wave front, $\Phi - \Psi$.

For convenience, we shall rewrite Equation 4 as

$$F(c) = I_0 f(c) , \tag{6}$$

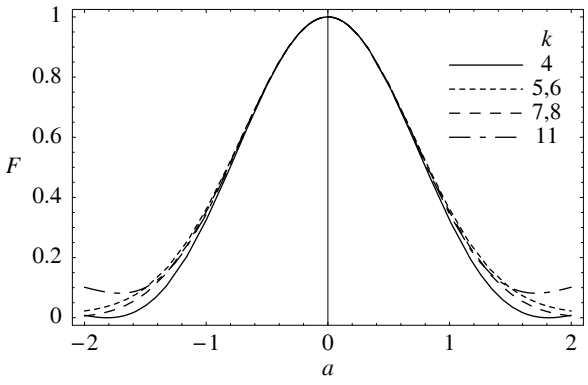

Figure 2. The photodetector signal calculated as a function of input mode amplitude for different Zernike aberration modes: $k = 4$, defocus; $k = 5, 6$ astigmatism; $k = 7, 8$ coma; $k = 11$, spherical. The two astigmatism modes, and similarly the two coma modes, are grouped together since the phase 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.

where

$$f\left(c\right) = \left|\frac{1}{\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{1}\exp\{jcZ_k\left(r,\theta\right)\}r\mathrm{d}r\mathrm{d}\theta\right|^2 \tag{7}$$

is not dependent upon the overall intensity and is equivalent to the Strehl ratio (the ratio of the on-axis intensity of an aberrated focus to that of the unaberrated case).

It is straightforward to extend the preceeding expressions to represent situations when multiple aberration modes are present in the wave front. Firstly, if K modes are present, we can represent the modal content of the input wave front by a K element vector of Zernike mode coefficients, $\mathbf{a}$. Similarly, the correction introduced by the adaptive element is represented by the vector $\mathbf{b}$. We then define

$$F\left(\mathbf{c}\right) = I_0 f\left(\mathbf{c}\right) \ , \tag{8}$$

where $\mathbf{c} = \mathbf{a} - \mathbf{b}$ and

$$f\left(\mathbf{c}\right) = \left|\frac{1}{\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{1}\exp\{j\sum_{k=1}^{K}c_k Z_k\left(r,\theta\right)\}r\mathrm{d}r\mathrm{d}\theta\right|^2 \tag{9}$$

and c_k is the coefficient of the kth mode. Due to the orthogonality of the modes, for small $|\mathbf{c}|$ we find that

$$f(\mathbf{c}) \approx \mathbf{1} - |\mathbf{c}|^{\mathbf{2}} \ . \tag{10}$$

In this case, $|\mathbf{c}|^2$ is equivalent to the variance of the corrected wave front. By measuring $f(\mathbf{c})$, we can calculate the variance ϵ as

$$\epsilon \approx 1 - f(\mathbf{c}) \ . \tag{11}$$

An individual measurement using the system of Figure 1 therefore provides us with one piece of information: the degree to which the input aberration Φ is corrected by the phase Ψ introduced by the adaptive element. Furthermore, since $\sqrt{\epsilon} = |\mathbf{a} - \mathbf{b}|$ we can also consider this measurement to provide the Euclidean distance between $\mathbf{a}$ and $\mathbf{b}$. All correction schemes are constructed around sequences of such measurements.

The relationship of Equation 10 was derived using the optical configuration of Figure 1. However, similar relationships can be obtained for many similar optical systems including those described in the Introduction. Our following discussion is based upon this particular optical system but the results are equally applicable to other systems.

3. RELATIONSHIP TO SPHERE PACKINGS

From Equation 10 we see that the form of $f(\mathbf{c})$ is dependent only on the variance of the aberration, irrespective of the particular mode combination. Alternatively, we could say that $f(\mathbf{c})$ is independent of the direction of the vector $\mathbf{c}$. It is clear therefore that, for small $|\mathbf{c}|$, all points on the contours of constant $f(\mathbf{c})$ are equidistant from the origin and thus take the form of K-dimensional spheres in the space spanned by $\mathbf{c}$. K-dimensional spheres are simply the multidimensional analogues of circles in two dimensions and spheres in three dimensions. A circle (or sphere) of radius r can be defined as the locus of all points in a two (or three) dimensional space that are situated a distance r from its centre. By extension of this definition, K-dimensional sphere is simply the locus of all points in K-dimensional space that are located a distance r from its centre. Since the wave front variance $\epsilon = |\mathbf{c}|^2$, it follows that the contours of constant ϵ also form K-dimensional spheres.

We can illustrate the consequences of this spherical geometry using a simple scheme for adaption of the correction element. We apply a sequence of different aberrations (or biases) using the adaptive element and take the corresponding photodetector measurements. Our solution is taken as the setting for which the maximum photodetector signal was obtained (i.e. the aberration that provides best correction of the input aberration). It would be impractical to sequence through the continuum of possible combinations of aberrations so we have to choose a sensible subset as the biases. This choice could be helped by specifying a tolerance within which the measurement should be made. We specify this tolerance as a wave front variance of $\epsilon_0 = 0.1$ rad^2; this is equivalent to a standard deviation of 1/20 wavelength. A measurement within this tolerance is considered an acceptable approximation to the solution. As we have shown, this variance is easily calculated from the photodetector measurement. We then take a sequence of measurements using a set of biases that are chosen according to the following requirements:

1. All possible input aberrations should be covered, thus ensuring that the globally optimum correction is found.

2. The minimum necessary number of individual measurements should be taken, whilst still satisfying the above condition.

Let us first consider the case when only a single Zernike mode is present in the wave front (Equation 4). The wave front would be corrected within tolerance when the signal $f(c) \geq 1 - \epsilon_0$. We should therefore choose a sequence of values for the bias b according to the requirements stated above. The most efficient way to do this is to space the biases uniformly so that only one will be sufficiently close to a to give $f(c) \geq 1 - \epsilon_0$. This could be done by choosing the values of b to be at integer multiples of $2\sqrt{\epsilon_0} = 0.6$. One can think of this as dividing up the b-axis into sections of width 0.6 and finding in which section the maximum intensity lies.

We now extend this process to include more modes, as modelled in Equation 9. We take a sequence of biases $\mathbf{b}$ in order to find a solution that lies close enough to the random vector $\mathbf{a}$ to give an error variance $\epsilon \leq \epsilon_0$ rad^2. Again, we need to choose the biases in order to satisfy our two requirements. Let us consider the case $n = 2$. It is tempting to extend the previous scheme to more dimensions and space the biases such that each component of $\mathbf{b}$ is an integer multiple of $2\sqrt{\epsilon_0}$. The bias vectors would then point to the vertices of a regular square grid, a scaled version of the integer lattice.[14] This approach would however be misguided; as we shall show, there is no guarantee that the correct solution would be found. From the above results we saw that a bias $\mathbf{b}$ is an acceptable solution only if $\mathbf{a}$ lies within a circle of radius $\sqrt{\epsilon_0}$ centred on $\mathbf{b}$. Fig. 3(a) illustrates this principle when the bias vectors form the two dimensional integer lattice. It can be seen that significant portions of the plane do not lie within one of the circles. Indeed it is straighforward to show that only 79% of the plane is covered. Hence, this arrangement of biases provides an incomplete covering that would not satisfy the first requirement and hence would not necessarily find an acceptable approximation to $\mathbf{a}$. We could overcome this by reducing the spacing of the lattice to $\sqrt{(2\epsilon_0)}$ so that the 'holes' between the circles disappear. This is illustrated in Fig. 3(b). Although there is overlap between the circles, we can be sure that $\mathbf{a}$ lies within at least one circle and that our first requirement is satisfied. However, to satisfy requirement 2 we must find the *most* efficient arrangement of biases that involves the mimimum overlap between the circles. This is given by the hexagonal arrangement shown in Fig. 3(c).[15] Using the hexagonal arrangement, the average number of circles that contain

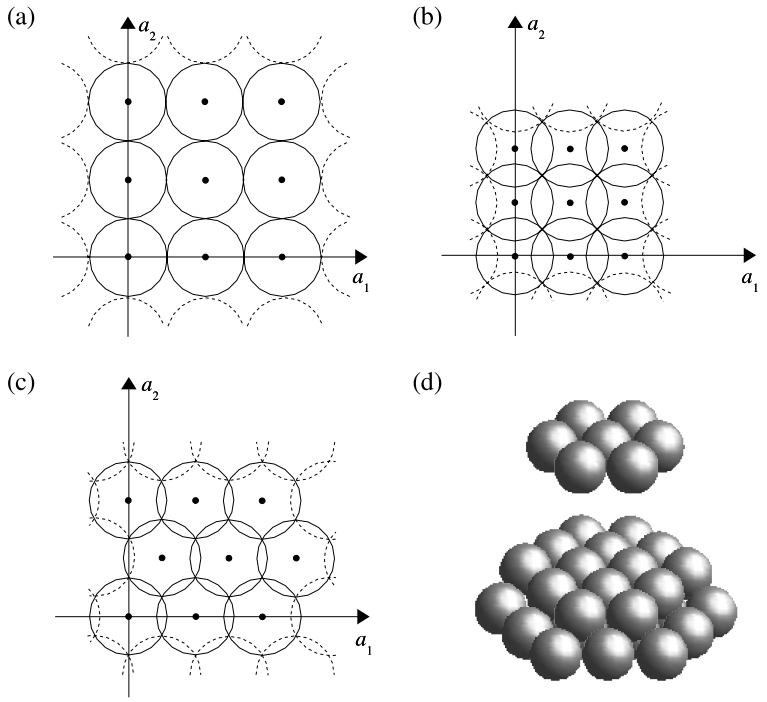

Figure 3. Circle packings for the sensing of two aberration modes: (a) incomplete covering based upon an integer lattice with lattice constant $2\sqrt{\epsilon_0}$; (b) complete covering based upon an integer lattice with lattice constant $\sqrt{2\epsilon_0}$; (c) thinnest possible covering based upon hexagonal lattice. The points at the centre of each circle represent the bias vectors **b**. Part (d) illustrates the face centred cubic arrangement, the three dimensional extension of the hexagonal packing.

a point in the plane - the 'thickness' of the covering - is 1.21.[14] For the square grid arrangement of Fig. 3(b), the thickness is 1.57. The hexagonal arrangement is therefore 23% more efficient than the square arrangement.

This principle can be extended to deal with more modes. When three modes are present, the problem involves the packing of spheres in three-dimensions. In this case, which is closely linked to the famous Kepler conjecture, the optimum packing for this problem is the face centred cubic structure (Fig. 3(c)). When more modes are present, the problem involves the packing of K-dimensional spheres in K-dimensional space. Although visualisation becomes difficult, the same principles still apply: To satisfy requirement 1, all points in space should be covered by at least one K-dimensional sphere. To satisfy requirement 2, the thickness of the covering should be as small as possible. This problem of sphere coverings is non-trivial and has been the subject of lengthy mathematical investigation. Indeed, optimal coverings, which include both lattice or non-lattice arrangements, have only been proven for $K \leq 3$ and optimal lattice coverings for $K \leq 5$.[14, 16] Fortunately, even when one takes advantage of the best *known* coverings, the influence on the efficiency of the wave front sensing scheme is striking. This is illustrated in Fig. 4, which shows the thickness of different coverings in K dimensions. The difference between the thickness of the best integer lattice covering and the optimal known lattice covering increases with K. Indeed, for $K = 10$ the ratio is approximately 50. For comparison, we have also shown the thickness of the incomplete lattice covering, based upon an integer lattice with the lattice constant $2\sqrt{\epsilon_0}$ derived from the one dimensional case. It is clear that as K increases, the proportion of the possible input aberrations that are covered decreases dramatically.

The results shown in Fig. 4 clearly demonstrate the importance of the arrangement of the aberration biases in our simple wave front sensing scheme. Incorrect arrangements may cause certain input aberrations to be missed by the measurement/correction scheme. Moreover, even if all possible input aberrations are covered, the

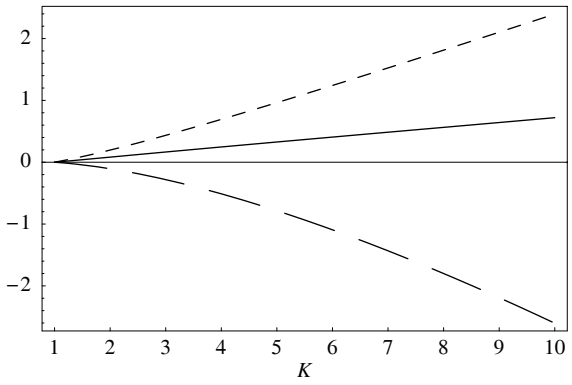

Figure 4. Base 10 logarithmic plots of the covering thickness for different packings in K dimensions, corresponding to the efficiency of measurement of K modes: (i) Optimum known lattice covering (solid line); (ii) Incomplete integer lattice covering with lattice constant $2\sqrt{\epsilon_0}$ (long dashed line); (iii) Complete integer lattice covering with lattice constant $2\sqrt{(\epsilon_0/K)}$ (short dashed line).

efficiency of the scheme can be far from optimum. It should be expected that more advanced wave front sensing schemes could take advantage of the approach described here.

4. MODAL WAVE FRONT SENSING AND SPHERE PACKINGS

As an example of the application of the sphere packing results to more advanced wave front sensing schemes, we combine the previous results with a modal wave front sensor (MWFS).[8, 10, 17] The MWFS measures directly the amount of each Zernike modes in the input wave front. 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 proportional to the modal coefficients of the input wave front. The MWFS has been demonstrated experimentally in closed loop adaptive systems.[2, 9] 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 or spatial light modulator, as in the set-up of Fig. 1. The simplest way to pack spheres locally is in the form of a simplex. In two dimensions, the simplex is an equilateral triangle; in three dimensions, it is a regular tetrahedron. K-dimensional simplices are simply the extensions of these forms. The generation of the required bias vectors is explained in the Appendix. We note that the measurement of K Zernike modes requires a minimum of $K+1$ photodetector measurements, since there are $K+1$ degrees of freedom in the system: the K modal coefficients and the overall intensity, I_0. We therefore arrange the $K+1$ biases to be the vertices of the K-dimensional simplex. This implementation is the most efficient way to measure a given set of aberration modes using the optical system of Figure 1.

One way of calculating the output signals from the MWFS is

$$\mathbf{W}_{out} = \frac{\sum_{n=1}^{N} \mathbf{b}_n F(\mathbf{a} - \mathbf{b}_n)}{\sum_{n=1}^{N} F(\mathbf{a} - \mathbf{b}_n)} \quad , \tag{12}$$

where the elements of the vector $\mathbf{W}_{out}$ represent the measurement of each modal coefficient and $F(.)$ is defined in Equation 8. As before, the vectors $\mathbf{b}_n$ represent the N bias aberrations applied in sequence by the adaptive element. Since each vertex of the simplex is equidistant from the origin, the length of each $\mathbf{b}_n$ is taken to be b_{max}. The measured aberration vector $\mathbf{a}$ is calculated as

$$\mathbf{a} \approx \mathbf{S}^{-1}\mathbf{W}_{out} \quad , \tag{13}$$

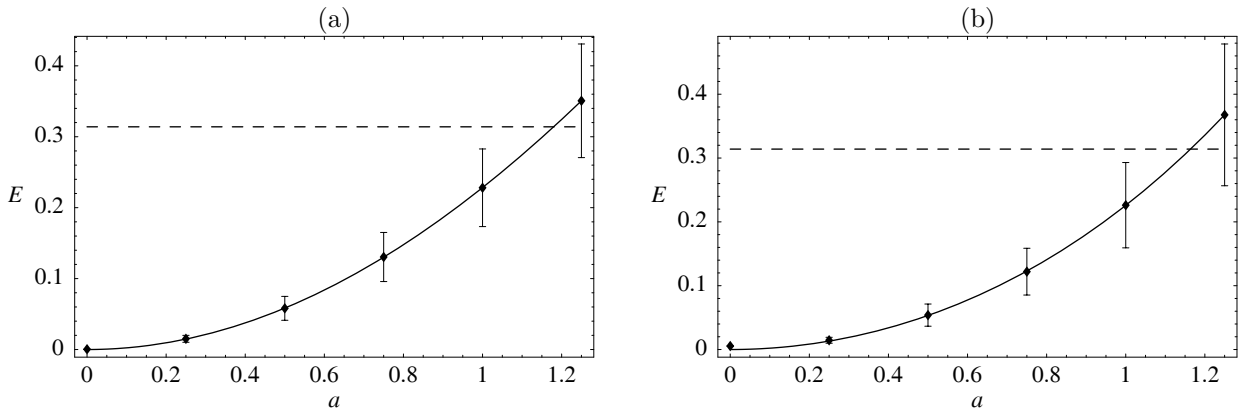

Figure 5. Measurement accuracy of multiple mode, simplex biased MWFS configurations as a function of input aberration vector magnitude. (a) for modes 5, 6, and 11 with $b_{max} = 0.5$. (b) for modes 4 to 8 with $b_{max} = 0.5$. The data points show the mean and standard deviation using a sample of 1000 input wave fronts. The solid line is a fourth order even polynomial fit. The dashed line shows the measurement tolerance of $\pi/10$.

where we use the sensitivity matrix $\mathbf{S}$, whose element S_{ik} is defined and calculated as

$$S_{ik} = \left. \frac{\mathrm{d}W_{out,i}}{\mathrm{d}a_k} \right|_{a_k=0} \quad . \tag{14}$$

This matrix is, in general, sparse and diagonally dominated.[8] We can investigate the measurement accuracy by defining the measurement error E as

$$E = \left| \mathbf{S}^{-1}\mathbf{W}_{out} - \mathbf{a} \right| \quad . \tag{15}$$

Since it is difficult to visualise the measurement range in multiple dimensions, we choose an alternative illustration. If the bias vectors are evenly distributed about the origin, we expect that the measurement range of the sensor be evenly distributed. If the input aberration $\mathbf{a}$ is represented by the product of its magnitude and unit direction vector, i.e. $\mathbf{a} = |\mathbf{a}|\hat{\mathbf{a}}$, then we would expect the MWFS measurement accuracy to depend primarily on $|\mathbf{a}|$. We therefore take a statistical sample using input vectors of fixed magnitude but with randomly oriented directions. We take the acceptable measurement range to be that for which $E \leq \pi/10$ radians. This is equivalent to a wave front standard deviation of 1/20 wavelength. As examples, in Figure 5(a) we show the measurement accuracy of such sensors for the simultaneous measurement in a system using the Zernike modes 5 and 6 (astigmatism) and 11 (spherical). Figure 5(b) shows the equivalent results for a sensor measuring modes 4 to 8. Each data point shows the mean and standard deviations calculated from 1000 samples. By ensuring that the mean error is less than $\pi/10 \approx 0.3$ radians, we find that these sensors can measure aberrations in the range $|\mathbf{a}| < 1.1$.

The simplex biasing scheme uses only $K + 1$ biases and photodetector measurements to measure K modes. The previously employed schemes used a minimum of $2K$ biases to measure K modes.[8, 10, 17] The new scheme therefore uses approximately half the number of biases and photodetector measurements that were previously required. The simplex scheme will therefore lead to more efficient MWFS implementations for the measurement of multiple Zernike modes.

5. CONCLUSIONS

By establishing a connection between wave front sensorless adaptive optics systems and sphere packings, we have provided a tool for the analysis optimisation of such adaptive systems. This connection illustrates issues that are neither obvious nor trivial, particularly when the adaptive system must correct for multiple aberration modes. The results show that this approach has significant benefit when applied to a simple correction scheme

and we could expect that benefits will also be apparent in more advanced schemes. As an illustration of this, we showed that the sphere packing concept could be applied to a modal wave front sensor. This implementation of the sensor was possible with half the number of biases previously used.

APPENDIX

The construction of a set of $K+1$ unity magnitude vectors $\mathbf{b}_n$ that represent the vertices of a regular K-dimensional simplex proceeds as follows. Firstly, note that when the centre of mass of the simplex is at the origin

$$\sum_{n=1}^{K+1} \mathbf{b}_n = \mathbf{0} \quad . \tag{16}$$

We choose the initial vector to have a non-zero coordinate only in the first dimension, such that $\mathbf{b}_1 = (1, 0, 0 \ldots)$. Using equation 16 we find that

$$\sum_{n=2}^{K+1} (\mathbf{b}_n)_1 = -1 \quad , \tag{17}$$

where $(\mathbf{b}_n)_m$ represents the mth element of $\mathbf{b}_n$. Since the coefficients $(\mathbf{b}_n)_1$ for $n > 1$ are equal, we can determine the first coordinate of the remaining vectors as

$$(\mathbf{b}_n)_1 = -\frac{1}{K} \qquad n > 1 \quad . \tag{18}$$

We consider the second vector $\mathbf{b}_2$ to have only two non-zero coordinates in the first two dimensions and hence lies in the plane defined by the first two coordinates. The second coordinate is simply obtained by

$$(\mathbf{b}_2)_2 = \sqrt{1 - (\mathbf{b}_2)_1^2} = \sqrt{1 - \left(\frac{1}{K}\right)^2} \quad . \tag{19}$$

Continuation of this approach, including an extra dimension in each consecutive vector, leads to a general result for calculation of the vector coordinates as

$$\begin{aligned} (\mathbf{b}_n)_m &= \frac{-(\mathbf{b}_m)_m}{K - m + 1} && m < n \\ &= \sqrt{1 - \sum_{p=1}^{m} (\mathbf{b}_m)_p^2} && m = n \\ &= 0 && m > n \quad . \end{aligned} \tag{20}$$

REFERENCES

1. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes*, Oxford University Press, 1998.
2. 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.
3. L. Sherman, J. Y. YE, O. Albert, and T. B. Norris, "Adaptive correction of depth-induced aberrations in multiphoton scanning microscopy using a deformable mirror," *J. Microsc.* **206**(1), pp. 65–71, 2002.
4. P. N. Marsh, D. Burns, and J. M. Girkin, "Practical implementation of adaptive optics in multiphoton microscopy," *Opt. Expr.* **11**, pp. 1123–1130, 2003.
5. O. Albert, L. Sherman, G. Mourou, T. B. Norris, and G. Vdovin, "Smart microscope: an adaptive optics learning system for aberration correction in multiphoton confocal microscopy," *Opt. Lett.* **25**(1), pp. 52–54, 2000.
6. E. Theofanidou, L. Wilson, W. J. Hossack, and J. Arlt, "Spherical aberration correction for optical tweezers," *Opt. Commun.* **236**, pp. 145–150, 2004.

7. A. C. F. Gonte and R. Dandliker, "Optimization of single-mode fiber coupling efficiency with an adaptive membrane mirror," *Opt. Engin.* **41**, pp. 1073–1076, 2002.
8. 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.
9. 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.
10. M. J. Booth, "Direct measurement of zernike aberration modes with a modal wave front sensor," *Proc. S.P.I.E., 'Advanced Wavefront Control: Methods, Devices, and Applications'* **5162**, pp. 79–90, 2003.
11. T. Wilson and C. Sheppard, *Theory and Practice of Scanning Optical Microscopy*, Academic Press, London, 1984.
12. M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, 6th ed., 1983.
13. R. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66**, pp. 207–277, 1976.
14. J. H. Conway and N. J. A. Sloan, *Sphere Packings, Lattices and Groups*, Springer-Verlag, 2nd ed., 1993.
15. R. Kershner, "The number of circles covering a set," *American J. Math.* **61**, pp. 665–671, 1939.
16. T. C. Hales, "An overview of the Kepler conjecture," *http://xxx.lanl.gov/* , p. math.MG/9811071, 1999.
17. M. J. Booth, M. A. A. Neil, 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.