

# Wave front sensor-less adaptive optics: a model-based approach using sphere packings

**Martin J. Booth**

*Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ,  
United Kingdom*

[martin.booth@eng.ox.ac.uk](mailto:martin.booth@eng.ox.ac.uk)

**Abstract:** Certain adaptive optics systems do not employ a wave front sensor but rather maximise a photodetector signal by appropriate control of an adaptive element. The maximisation procedure must be optimised if the system is to work efficiently. Such optimisation is often implemented empirically, but further insight can be obtained by using an appropriate mathematical model. In many practical systems aberrations can be accurately represented by a small number of modes of an orthogonal basis, such as the Zernike polynomials. By heuristic reasoning we develop a model for the operation of such systems and demonstrate a link with the geometrical problems of sphere packings and coverings. This approach aids the optimisation of control algorithms and is illustrated by application to direct search and hill climbing algorithms. We develop an efficient scheme using a direct maximisation calculation that permits the measurement of  $N$  Zernike modes with only  $N + 1$  intensity measurements.

© 2006 Optical Society of America

**OCIS codes:** (010.1080) Adaptive Optics; (010.7350) Wave-front sensing.

---

## References and links

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** 5788–5792 (2002).
3. A. J. Wright, D. Burns, B. A. Patterson, S. P. Poland, G. J. Valentine and J. M. Girkin, "Exploration of the optimisation algorithms used in the implementation of adaptive optics in confocal and multiphoton microscopy," *Microsc. Res. Technol.* **67** 36–44 (2005).
4. 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** 65–71 (2002).
5. P. N. Marsh, D. Burns and J. M. Girkin, "Practical implementation of adaptive optics in multiphoton microscopy," *Opt. Express* **11** 1123–1130 (2003).
6. 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** 52–54 (2000).
7. W. Lubeigt, G. Valentine, J. Girkin, E. Bente, and D. Burns, "Active transverse mode control and optimization of an all-solid-state laser using an intracavity adaptive-optic mirror," *Opt. Express* **10**, 550-555 (2002).
8. E. Theofanidou, L. Wilson, W. J. Hossack and J. Arlt, "Spherical aberration correction for optical tweezers," *Opt. Commun.* **236** 145–150 (2004).
9. A. C. F. Gonte and R. Dandliker, "Optimization of single-mode fiber coupling efficiency with an adaptive membrane mirror," *Opt. Eng.* **41** 1073–1076 (2002).
10. M. Vorontsov, "Decoupled stochastic parallel gradient descent optimization for adaptive optics: integrated approach for wave-front sensor information fusion," *J. Opt. Soc. Am. A* **19** 356–368 (2002).
11. M. J. Booth, T. Wilson, H.-B. Sun, T. Ota and S. Kawata, "Methods for the characterisation of deformable membrane mirrors," *Appl. Opt.* **44** 5131–5139 (2005).

12. T. Wilson and C. J. R. Sheppard, *Theory and Practice of Scanning Optical Microscopy*, (Academic Press, London, 1984).
  13. M. Born and E. Wolf, *Principles of Optics*, 6th Edition, (Pergamon Press, 1983).
  14. Z. Michalewicz and D. B. Fogel, *How to Solve It: Modern Heuristics*, (Springer, Berlin, 2000).
  15. J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd Edition, (Springer-Verlag, 1998).
  16. T. C. Hales, "An overview of the Kepler conjecture," <http://xxx.lanl.gov/math.MG/9811071> (1999).
  17. R. Kershner, "The number of circles covering a set," *Am. J. Math.* **61** 665–671 (1939).
  18. E. Vitterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory* **45** 1639–1642 (1999).
  19. W. Press, S. Teukolsky, W. Vetterling and B. Flannery, *Numerical Recipes in C*, 2nd Edition, (Cambridge University Press, 1992).
  20. M. A. A. Neil, M. J. Booth and T. Wilson, "New modal wavefront sensor: a theoretical analysis," *J. Opt. Soc. Am. A* **17** 1098–1107 (2000).
  21. R. Noll, "Zernike polynomials and atmospheric turbulence," *J. Opt. Soc. Am.* **66** 207–277 (1976).
- 

## 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 and reflection microscopy [2, 3], two-photon fluorescence microscopy [4, 5], maximisation of second harmonic generation [6], intracavity aberration correction in lasers [7], optical tweezers [8], coupling laser light into an optical fibre [9] and for the characterisation of adaptive optical systems [10, 11]. These applications have employed different schemes for optimisation of the correction element based upon genetic algorithms [3, 4, 5, 6, 9], hill climbing algorithms [3, 5, 7, 8, 9], stochastic gradient descent (Ref. [10] and further references therein), and modal wave front sensing [2, 11].

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 introduce deliberately a particular aberration into the optical system and a photodetector measurement is taken. As the adaptive element is sequentially reconfigured, according to the chosen scheme, a set of corresponding photodetector measurements is compiled. The optimal solution, corresponding to the maximum signal, is derived from this set of measurements using a chosen algorithm. The efficiency of such a strategy depends not only on the choice of algorithm, but also on the parameters used in its implementation. An ill-optimised algorithm might not only be inefficient, and therefore converge too slowly, but also fail to find the correct solution. Most of the previous work in this area has relied upon the application of standard, model-free, stochastic search methods where parameter optimisation was performed either empirically or adaptively, as part of the maximisation algorithm. However, an appropriate mathematical representation of the optical system can provide insight into the optimisation problem, the choice of algorithm and even permit direct calculation of parameters. In particular, if the function being optimised is known, one can take advantage of its topological structure to design algorithms that in general have better convergence than model-free methods.

This paper details an heuristic strategy for such model-based optimisation algorithms using *a priori* knowledge of the optimised function. Rather than provide prescriptions for specific optical applications, we concentrate on a general, mathematically constructive approach that may elucidate algorithm design. The paper is structured as follows. Firstly, we describe the mathematical model of the optical system based upon an aberration expansion in terms of Zernike

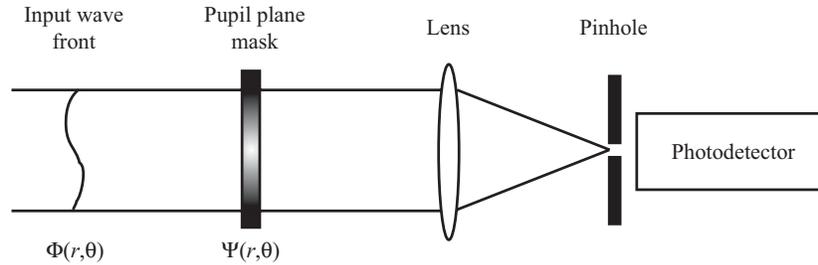


Fig. 1. Schematic diagram of the adaptive system. The input wave front is incident from the left whereupon it passes through the pupil plane of the lens. A phase element in this plane adds a chosen phase aberration to the input wave front. The wave front is then focussed by the lens onto pinhole positioned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

polynomials and we derive a function describing the resulting photodetector signal. This function is shown to have a well defined maximum and spherical symmetry, which can be used to our advantage in algorithm design; in particular, this aids the choice of the aberrations that are sequentially applied by the adaptive element. In Section 4, we use the spherical topology to show how lattice sphere packings assist the choice of candidate solutions to ensure the most efficient coverage of the search space. These optimum lattices are used in Section 5 as the basis for global, exhaustive search algorithms. This is extended in the subsequent section to a more efficient, multi-level exhaustive search. In Section 7, we apply a similar approach to a steepest ascent hill-climbing algorithm, a local search method where a set of candidates is chosen to optimally search the immediate neighbourhood at each stage. By using local properties of the function and following the direction of steepest gradient, this method converges more rapidly than the exhaustive search. In Section 8, we describe a direct maximisation algorithm that has much better convergence properties than the search algorithms. Rather than performing a search through different candidates either globally or locally, this algorithm takes into account the global topology of the function. This enables direct calculation of the position of the maximum using the smallest possible number of function evaluations or, equivalently, the fewest intensity measurements.

The deterministic algorithms presented here differ from many commonly used methods, in which the steps between successive trial solutions adaptively change, usually in a stochastic manner, as the algorithm progresses permitting convergence to a solution with arbitrary precision. We approach the algorithm design by specifying the desired precision and using it in conjunction with the mathematical model to predetermine the necessary sequence of steps. The results show that deterministic, non-adaptive algorithms can be effective in controlling wave front sensorless adaptive optics systems, if they are suitably formulated.

## 2. Mathematical model

In the conceptual measurement system shown in Fig. 1, the input wave front 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 the input wave front is described by the function  $\Phi(r, \theta)$ , where  $r$  and  $\theta$  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 mask, which could in practice be an adaptive element, that subtracts a phase function  $\Psi(r, \theta)$  from the input wave front. Fourier diffraction theory[12] shows that the signal

measured by the photodetector is given by

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

where  $I_0$  is proportional to the incident light power and  $j = \sqrt{-1}$ .

We assume that  $\Phi$  and  $\Psi$  can each be expressed as a series of  $N$  Zernike polynomials[13], each denoted by  $Z_n(r, \theta)$ :

$$\Phi(r, \theta) = \sum_{n=1}^N a_n Z_n(r, \theta) \quad (2)$$

$$\Psi(r, \theta) = \sum_{n=1}^N b_n Z_n(r, \theta) \quad (3)$$

The normalisation and indexing scheme of the Zernike polynomials are explained in Appendix A. The aberration of the input wave front  $\Phi$  can be represented by a  $N$  element vector  $\mathbf{a}$ , whose elements are the Zernike mode coefficients  $a_n$ . Similarly, the correction  $\Psi$  introduced by the adaptive element is represented by the vector  $\mathbf{b}$ , whose elements are the coefficients  $b_n$ . We then define

$$F(\mathbf{c}) = I_0 f(\mathbf{c}), \quad (4)$$

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

$$f(\mathbf{c}) = \left| \frac{1}{\pi} \int_{\theta=0}^{2\pi} \int_{r=0}^1 \exp\{j \sum_{n=1}^N c_n Z_n(r, \theta)\} r dr d\theta \right|^2, \quad (5)$$

where  $c_n$  is the coefficient of the  $n$ th mode. The function  $f(\mathbf{c})$  is independent of the overall intensity and is equivalent to the Strehl ratio [13]. Due to the orthogonality of the Zernike modes, for small  $|\mathbf{c}|$  we find that

$$f(\mathbf{c}) \approx 1 - |\mathbf{c}|^2, \quad (6)$$

where  $|\mathbf{c}|^2$  also represents the variance of the corrected wave front [13]. The system is considered to be well corrected if  $|\mathbf{c}| < \varepsilon$ , where  $\varepsilon^2$  is a small quantity equal to the maximum acceptable wave front variance. Equation (6) provides the equivalent requirement that  $f(\mathbf{c}) > 1 - \varepsilon^2$ . From Eq. (6) we see that  $f(\mathbf{c})$  is isotropic for small arguments - a property that is illustrated further in Fig. 2. It follows that the contours of constant  $f(\mathbf{c})$  are  $N$ -dimensional spheres centred on the origin. This is a different expression of the well known result that the Strehl ratio depends only on the variance of the aberration and not its form[13]. It is also important to note that for large arguments  $f(\mathbf{c})$  becomes small. This is a consequence of the rapid variation of the exponential term in the integral of Eq. (5). The global maximum at  $\mathbf{c} = \mathbf{0}$  is therefore much larger than any of the surrounding local maxima. These properties of  $f(\mathbf{c})$  can be used to our advantage when designing maximisation algorithms.

The discussion in this paper is based upon the measurement of Zernike aberration modes because of their widespread use in optics. However, the results would also be applicable for other, orthonormal, basis functions, for example the normalised eigenmodes of a deformable mirror.

### 3. Strategy for algorithm design

Numerous heuristics have been developed for optimisation problems [14]. These range from simple algorithms, such as exhaustive search or hill climbing methods, to more advanced methods, such as genetic algorithms or simulated annealing. In all cases it is possible to tailor the

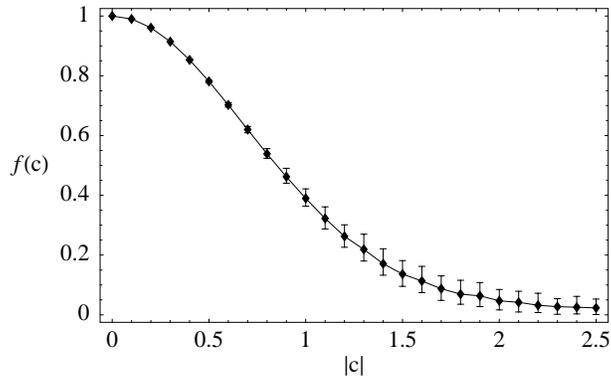


Fig. 2. The photodetector signal calculated as a function of aberration magnitude for random combinations of the six Zernike aberration modes  $i = 4$  to  $i = 9$ . Each data point shows the mean and the 10th and 90th percentiles of a collection of 100 random samples.

algorithm by taking into account *a priori* mathematical knowledge of the problem. The first step in the design of an optimisation algorithm is the specification of the problem. This can be broken down into three components: 1) The objective - the purpose to be fulfilled by the algorithm; 2) The evaluation (or merit) function - a measure indicating the quality of a candidate solution; 3) The representation - a model of the problem that specifies the alternative candidate solutions [14]. We define these components in a general manner that is not specific to a particular algorithm.

The first two components are straightforward to define. Our objective is to find the solution  $\mathbf{b}$  that gives the maximum photodetector signal or  $f(\mathbf{c}) = 1$ . Usually this would be relaxed so that we instead search for a solution  $\mathbf{b}$  that is arbitrarily near to the maximum, such that  $f(\mathbf{c}) > 1 - \varepsilon^2$ . We note that the objective will be fulfilled when  $|\mathbf{a} - \mathbf{b}| < \varepsilon$ . This is equivalent to specifying that  $\mathbf{a}$  lies within an  $N$ -dimensional sphere of radius  $\varepsilon$  centred on the candidate solution  $\mathbf{b}$ . The obvious choice for the evaluation function is the photodetector signal, since this has a well behaved maximum at the correct solution  $\mathbf{b} = \mathbf{a}$ .

The definition of the representation would depend upon the maximisation strategy employed. For a particular maximisation scheme, let us assume that the correct solution  $\mathbf{b} = \mathbf{a}$  lies within a finite region  $\Sigma$  of  $\mathbf{R}^N$ , where  $\mathbf{R}$  denotes the field of real numbers. In a practical system using a global search method, the solution space  $\Sigma$  might be determined by the adaptive element's capabilities - which aberrations it can correct - or the range of input aberrations. When using a local search,  $\Sigma$  might encompass the candidate solutions 'near' to the present estimate. We can define the representation to consist of a set  $B$  of candidate solutions, each represented by a vector  $\mathbf{b}$ , that is a finite subset of the points in  $\Sigma$ .

One could attempt to cover the solution space  $\Sigma$  using a random arrangement of candidate solutions - this would be equivalent to a random search. However, this would not guarantee fulfilling the objective by finding the actual solution; it is possible that such a scheme could miss the maximum entirely. It is also possible that it could over-sample some areas of  $\Sigma$ . One could, of course, increase the numbers of random candidates to give an acceptable probability of success, but this would be at the expense of efficiency. Similar arguments also apply to more advanced stochastic methods.

Another approach would be the use of deterministic algorithms. Since our evaluation function is known and has a well-defined global maximum, we can design such algorithms that

are guaranteed to find the solution. Moreover, through mathematical reasoning we can choose the set  $B$  of candidate solutions to provide the best efficiency. To guarantee finding the correct solution, we need to ensure that every point in  $\Sigma$  lies within a distance  $\varepsilon$  of at least one of the candidate solutions in  $B$ . This would ensure that  $\Sigma$  is ‘covered’ by the overlapping  $N$ -dimensional spheres centred on the candidate solutions. To obtain the best efficiency, we must choose  $B$  to have the fewest candidate solutions whilst still covering  $\Sigma$ . It follows that finding the optimum representation is equivalent to finding the optimum covering of  $\Sigma$  with overlapping  $N$ -dimensional spheres. The search for such sphere packings is a mathematical problem that has been extensively investigated [15]. It is important to note that only in the best case would a random set  $B$  provide the same efficiency as the optimum sphere packing.

#### 4. Sphere coverings and the representation

The problem of sphere coverings involves finding the most efficient way to cover  $\mathbf{R}^N$  with equal, overlapping  $N$ -dimensional spheres (hereafter referred to simply as ‘spheres’). The efficiency of a covering is quantified in terms of its ‘thickness’, which is equivalent to the average number of spheres that cover a point of the space. The covering problem therefore asks for the arrangement of spheres that has the minimal thickness, otherwise termed the thinnest covering. This is non-trivial and has been the subject of lengthy mathematical investigation. Optimal coverings of  $\mathbf{R}^N$ , which include both lattice or non-lattice arrangements of spheres, have only been proven for  $N \leq 3$  and optimal lattice coverings have only been found for  $N \leq 5$  [15, 16]. Conway and Sloane list the best known coverings for  $N \leq 24$ , although they are not necessarily proven to be optimal [15].

We now apply the concepts of sphere coverings to the representation for our optimisation problem. Let us assume that  $\Sigma$  is large so we can apply the best known coverings for infinite regions. We first consider the trivial case  $N = 1$ , when only a single Zernike mode is present. The region bounded by a 1-dimensional sphere of radius  $\varepsilon$  is simply a line segment of length  $2\varepsilon$ . The thinnest covering therefore consists of spheres arranged in a line with their centres spaced by  $2\varepsilon$ ; thus the candidate solutions in  $B$  should consist of integer multiples of  $2\varepsilon$ . Since the spheres do not overlap, this is a perfect covering with a thickness of 1. One can think of this process as dividing up the  $b_1$ -axis into sections of width  $2\varepsilon$  in order to find the section containing the maximum intensity. It is obvious that the maximum error would be  $\varepsilon$ .

Extending this to the case  $N = 2$ , it is tempting to choose each element of a candidate solution  $\mathbf{b}$  to be an integer multiple of  $2\varepsilon$ . The vectors in  $B$  would then point to the vertices of a regular square grid, a scaled version of the integer lattice (usually referred to as  $\mathbf{Z}_N$ ) [15]. This is illustrated in Fig. 3(a). As discussed in Section 3, the correct solution will only be found if it lies within a sphere of radius  $\varepsilon$  centred on a candidate solution  $\mathbf{b}$ . It can be seen that significant portions of the plane, and hence potential solutions, do not lie within one of the spheres. Indeed, only 79% of the plane is covered and therefore the representation is incomplete. We could overcome this by reducing the spacing of the lattice to  $\sqrt{2}\varepsilon$  so that the gaps between the circles disappear. This is illustrated in Fig. 3(b). Although there is now overlap between the spheres, we can be sure that  $\mathbf{a}$  lies within at least one sphere and the representation is complete. However, this scaled integer lattice (with thickness 1.57) does not provide the thinnest possible covering, which is instead given by the hexagonal lattice shown in Fig. 3(c) [17]. The hexagonal covering has thickness 1.21 and is 23% more efficient than the square arrangement.

We can extend this analysis to include more modes. In general, the thickness  $\Theta_1$  of the incomplete integer lattice in  $N$  dimensions is equivalent the ratio between the volume of a sphere and the cube that it inscribes. The volume of a unit radius sphere is given by [15]

$$V_N = \frac{\pi^{N/2}}{\Gamma(N/2 + 1)}, \quad (7)$$

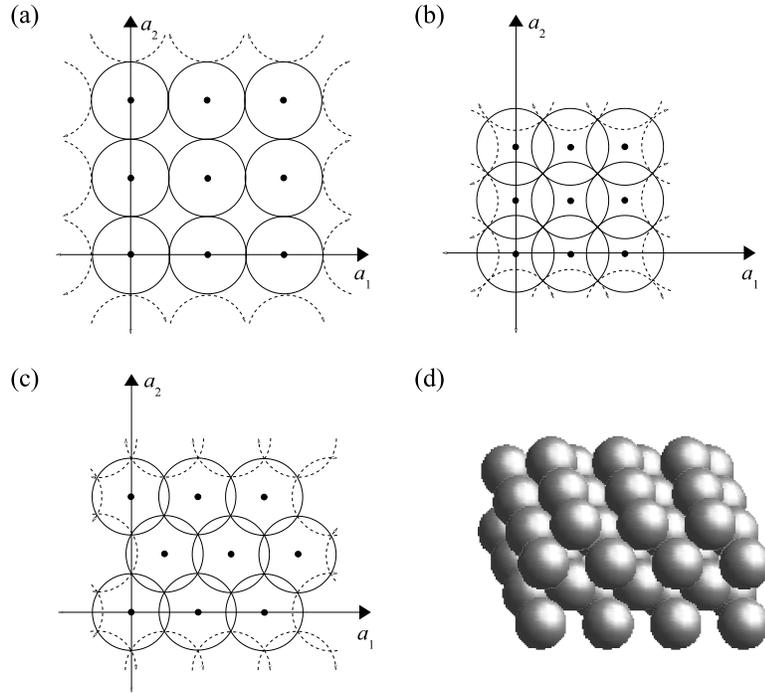


Fig. 3. Sphere packings in two dimensions: (a) incomplete covering based upon an integer lattice with lattice constant  $2\epsilon$ ; (b) complete covering based upon an integer lattice with lattice constant  $\sqrt{2}\epsilon$ ; (c) thinnest possible covering based upon hexagonal lattice. The points at the centre of each circle represent the candidate solutions  $\mathbf{b}$ . Part (d) illustrates the body centred cubic arrangement, the optimal three dimensional covering.

where  $\Gamma$  is the gamma function. Since this would inscribe a cube of volume  $2^N$ , the thickness follows as

$$\Theta_1 = \frac{V_N}{2^N} . \quad (8)$$

To ensure complete coverage, the spacing of the integer lattice must be reduced by a factor  $\sqrt{N}$ . The thickness  $\Theta_2$  is now given by the ratio of the volume of a sphere and its circumscribed cube, which has volume  $(2/\sqrt{N})^N$ . Hence,

$$\Theta_2 = \frac{V_N N^{N/2}}{2^N} . \quad (9)$$

For  $N \leq 23$ , the best known coverings are provided by the lattice known as  $A_N^*$ . In three dimensions,  $A_3^*$  is equivalent to the body centred cubic lattice, illustrated in Fig. 3(d). This covering has thickness [15]

$$\Theta_3 = V_N \sqrt{N+1} \left( \frac{N(N+2)}{12(N+1)} \right)^{N/2} . \quad (10)$$

Figure 4 shows the thickness of these three coverings for different  $N$ . The difference between the thickness of the best integer lattice covering and the optimal known lattice covering in-

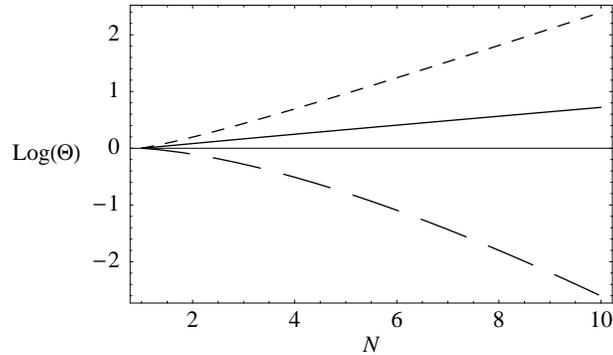


Fig. 4. Base 10 logarithmic plots of the covering thickness for different packings in  $N$  dimensions, corresponding to the efficiency of measurement of  $N$  modes: (i) Incomplete integer lattice covering (long dashed line),  $\Theta_1$ ; (ii) Complete integer lattice covering (short dashed line),  $\Theta_2$ ; (iii) Optimum known lattice covering (solid line),  $\Theta_3$ .

creases markedly with  $N$ . Indeed, for  $N = 10$  their ratio is approximately 50. With the incomplete lattice covering, it is clear that as  $N$  increases, the proportion of the possible solutions that are covered decreases dramatically. This trend continues for larger  $N$ , as can be seen from the asymptotic behaviours of these thicknesses, given by  $\log(\Theta_1) \sim -N \log(N)$ ,  $\log(\Theta_2) \sim N$  and  $\log(\Theta_3) \sim N$ .

The  $A_N^*$  lattice provides the most efficient scheme whilst ensuring that the maximum measurement error does not exceed  $\varepsilon$ . We note that other lattices, known as quantisers, have been found that would alternatively minimise the mean square error of the measurement [15].

## 5. Exhaustive search using sphere coverings

We can illustrate the use of sphere coverings by applying them to an exhaustive search algorithm, in which every candidate solution is checked in turn. This would guarantee finding the globally maximum solution. In most optimisation problems, this would be an impractically slow approach, but it is useful as a baseline for assessment of different algorithms.

In practice,  $\Sigma$  would be a finite region and the optimal sphere coverings might therefore be different to the lattices described above. However, the results for the infinite coverings are likely to be near optimal if  $\Sigma$  is large. We compared representations based upon the integer lattice  $\mathbf{Z}_N$  (with appropriate spacing to ensure coverage) and the lattice  $A_N^*$ . We chose  $\Sigma$  to be a spherical region containing all points within a radius  $\rho = 1.07$  of the origin (this value was arbitrarily chosen to permit direct comparison with results in the subsequent Section). We took  $\varepsilon = \sqrt{0.1}$  so that the final solution would correspond to a Strehl ratio  $f(\mathbf{c}) > 0.9$ . The set  $B$  of candidate solutions consisted of the centres of the spheres required to cover  $\Sigma$ ; this included the lattice points within  $\Sigma$  and an extra 'layer' of spheres necessary to complete the covering near the periphery. We therefore used lattice points within a radius  $\rho + \varepsilon$  of the origin; these were found using an appropriate search algorithm [18]. Since  $\Sigma$  was covered, we could be certain that an exhaustive search would find the correct solution. Hence, we assessed the efficiency of the algorithm by the number of included lattice points: fewer lattice points mean fewer evaluations and greater efficiency. The total numbers of candidate solutions,  $K$ , for  $N \leq 6$  are shown in Fig. 5 (for the calculations throughout this paper we take  $N = 1$  to be Zernike mode  $i = 4$ ,  $N = 2$  to include modes  $i = 4$  and  $i = 5$ ,  $N = 3$  to include modes  $i = 4$  to  $i = 6$ ,

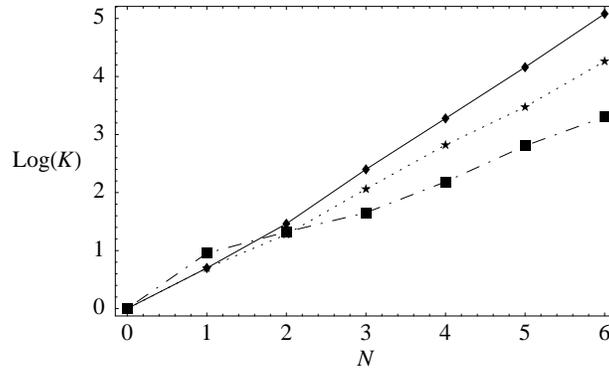


Fig. 5. Base 10 logarithmic plots of the number of evaluations,  $K$ , required for the exhaustive search in  $N$  dimensions, based upon (i) Complete integer lattice (diamonds); (ii) the  $A_N^*$  lattice (stars); (iii) three-level branch and bound exhaustive search using the  $A_N^*$  lattice (squares).

etc.). As expected, the lattice  $A_N^*$  provided a more efficient algorithm than  $\mathbf{Z}_N$  and in both cases  $K$  increased exponentially with  $N$ .

## 6. Exhaustive search with branch and bound

A more efficient algorithm was obtained by combining the exhaustive search with a branch and bound technique [14]. Rather than searching directly through a space  $\Sigma$  partitioned into spheres of radius  $\varepsilon$ , we performed a coarse search using spheres of larger radius before performing progressively more local searches. The total number of evaluations was therefore reduced whilst still obtaining the same final accuracy. It can be seen from Fig. 2 that the variation in  $f(\mathbf{c})$  is small for  $|\mathbf{c}| \approx 1$  so a search algorithm based upon spheres of this radius or smaller would still be appropriate. Again we chose  $\Sigma$  to be a spherical region of radius  $\rho = 1.07$  centred on the origin and we took  $\varepsilon = \sqrt{0.1}$  (note that  $\rho \approx 1.5^3 \varepsilon$  and was chosen to permit a convenient three level search). For the first, coarsest search we chose to cover  $\Sigma$  with spheres of radius  $1.5^2 \varepsilon = 0.71$  and we proceeded to find the candidate with the maximum intensity. For the second search, the sphere centred on this candidate was covered with smaller spheres of radius  $1.5 \varepsilon = 0.47$ . The final search used spheres of radius  $\varepsilon = 0.32$ , yielding a result with the same accuracy as the simple exhaustive search algorithm. Since the ratio of the search space radius and the covering sphere radius was the same at each step, the same number of evaluations was performed at each of the three levels. Such a three level search, based upon  $A_N^*$ , was performed for different combinations of modes for  $N \leq 6$ . For each  $N$ , one thousand random solutions were tested by evaluating  $f(\mathbf{c})$  using Eq. 5. In each case, the solution was found within the required tolerance of  $\varepsilon$ . The total numbers of evaluations,  $K$ , for each search are shown in Fig. 5. For  $N > 2$ , the branch and bound method was significantly more efficient than a simple exhaustive search.

## 7. Steepest ascent hill climbing

Whilst guaranteed to find the global solution, the exhaustive search algorithm is usually slow since it requires evaluation over the whole search space. The steepest ascent hill climbing (SAHC) algorithm is attractive since it exploits local information and concentrates on the more promising regions of the search space. Their drawback is that it converges on locally optimal, rather than globally optimal, solutions. The SAHC algorithm starts from the present candidate

solution and evaluates all candidate solutions in its neighbourhood. It then selects the candidate with the highest evaluation function, which becomes the new present solution, and the process repeats. The algorithm terminates when the evaluation of the present solution is higher than its neighbours. As with the exhaustive search, we can take advantage of sphere coverings to optimise the parameters of the SAHC algorithm.

The problem requires definition of the neighbourhood - the set of candidate solutions near to the present solution that are tested at each step. If the candidates are too close to the present solution or there are too many, the algorithm will have slow convergence; if they are too widely separated, then potential solutions will be missed. The neighbourhood of the present solution should therefore be adequately, but thinly covered. There are various ways to define such a neighbourhood. Most simply, we could take it to be the surface of the sphere of radius  $\varepsilon$  centred on the present solution. This surface needs to be covered by the spheres centred on the adjacent candidate solutions. Just as the intersection of two spheres forms a circle, the intersection of two  $N$ -spheres forms an  $(N-1)$ -sphere. The problem of covering this neighbourhood is therefore equivalent to the problem of covering the surface of an  $N$ -sphere with an arrangement of  $(N-1)$ -spheres. The minimum number of neighbouring candidates that would be required whilst still spanning  $N$  dimensions is  $N+1$  and, if regularly spaced, they would be positioned at the vertices of an  $N$ -dimensional regular simplex (i.e. an equilateral triangle for  $N=2$ , a regular tetrahedron for  $N=3$ , etc.). The vectors representing the simplex vertices are derived in Appendix B. This arrangement covers the neighbourhood only if the distance,  $s$ , from the present solution to the candidates satisfies  $0 < s \leq 2\varepsilon/N$  (see Appendix C). The efficiency of such a SAHC algorithm is illustrated in Fig. 6 for  $s = 2\varepsilon/N$  and  $N \leq 6$ . In each trial a randomly oriented vector of magnitude 1.5 was used as the initial candidate. For each data point,  $K$  was taken as the mean number of evaluated candidates from one hundred trials. Since the step size  $s \propto 1/N$  and the number of evaluations per step varies as  $N$ , we expect  $K$  to vary as  $N^2$ ; this quadratic dependence is confirmed by Fig. 6. When  $s$  was increased beyond  $2\varepsilon/N$ , it was observed that the algorithm failed to find the solution using some initial candidates.

This approach guarantees convergence of the SAHC algorithm, but the efficiency can be improved further. For example, when using the simplex arrangement of candidates there is strong directional dependence, that is convergence is much faster for initial candidates in the direction of the simplex vertices. This dependence can be reduced by reversing the orientation of the simplex at each step (Fig. 6). Further improvement could also be achieved by randomising the orientation of the simplex at each step and by combining the algorithm with a branch and bound approach to perform optimisation on progressively finer scales.

It is useful to note the difference between the method presented here and the commonly used 'simplex method' for function maximisation [19]. This latter method uses variable step sizes and orientations of the simplex to adapt to the local shape of the function as the algorithm progresses. Because of the known topology of  $f(\mathbf{c})$ , we were able to use fixed step sizes to achieve convergence to the desired precision.

## 8. Direct maximisation

The search algorithms described in the previous sections are based upon direct comparisons between candidate solutions and at each stage the best solution estimate is taken as the candidate with the best evaluation. The accuracy of this estimate is therefore limited by the coarseness of the distribution of the candidates. If further refinement is necessary then another search must be performed at a finer scale. This approach is useful when one has little information about the function being optimised. However, when the form of the function is known, it is often possible to obtain an estimate with an accuracy greater than the coarseness of the candidate distribution. As a clear example, the extremum of a quadratic function can be found exactly

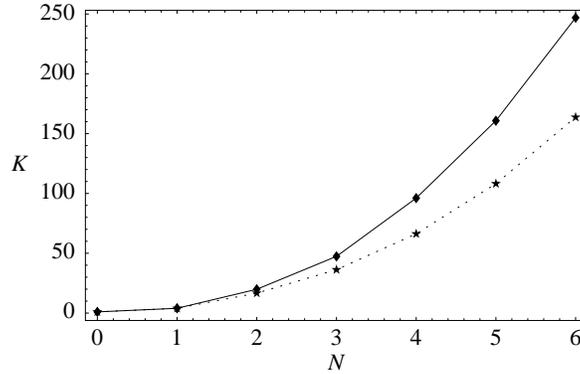


Fig. 6. Linear plots of  $K$ , the mean number of evaluations required for the hill climbing algorithm in  $N$  dimensions, using (i) a simplex with fixed orientation (diamonds); (ii) a simplex with reversals at each step (stars).

with three, arbitrary evaluations (as long as the evaluations are not coincident). In the present problem, aside from a translation, the form of the function  $F(\mathbf{c})$  is known entirely. We can take advantage of this *a priori* knowledge in constructing a scheme to find the maximum directly, rather than using a search strategy. This leads to significantly faster optimisation than the search algorithms.

It can be seen in Fig. 2 that the function  $F(\mathbf{c})$  has a well defined maximum and is isotropic in the surrounding region. Since its value also becomes small away from the maximum, a good estimate of the location of the maximum can be found from a first moment (centre of mass) calculation. This estimate, denoted by the vector  $\mathbf{W}$ , could be formulated as

$$\mathbf{W} = \frac{\int \cdots \int_{\Sigma} \mathbf{b} F(\mathbf{a} - \mathbf{b}) dV}{\int \cdots \int_{\Sigma} F(\mathbf{a} - \mathbf{b}) dV} \approx \mathbf{a} , \quad (11)$$

where  $\int \cdots \int$  represents a  $N$ -dimensional integral over a suitably large region  $\Sigma$  and  $dV$  is the volume element at  $\mathbf{b}$ . In practice, Eq. (11) must be approximated using a discrete numerical integration scheme where each integrand evaluation corresponds to a photodetector measurement. We therefore calculate  $\mathbf{W}$  as

$$\mathbf{W} = \frac{\sum_{m=1}^M \gamma_m \mathbf{b}_m F(\mathbf{a} - \mathbf{b}_m)}{\sum_{m=1}^M \gamma_m F(\mathbf{a} - \mathbf{b}_m)} , \quad (12)$$

where  $\gamma_m$  are integration weights that depend upon the particular numerical integration scheme [19]. The vectors  $\mathbf{b}_m$  are the  $M$  integration abscissæ, the locations where the integrand is evaluated, each of which represents the aberration introduced by the adaptive element for a particular intensity measurement. If a large number of appropriately distributed abscissæ are used then  $\mathbf{W} \approx \mathbf{a}$ . A simple integration scheme could, for example, involve the regular distribution of the  $\mathbf{b}_m$  throughout  $\Sigma$ , based upon a lattice such as  $A_N^*$  or  $Z_N$ , using weights  $\gamma_m = 1$ . This would be a multidimensional analogue of block integration with one variable. More advanced schemes, such as Gaussian quadrature, might alternatively be employed [19].

The simplest way to distribute the  $\mathbf{b}_m$  for measurement of  $N$  modes would be at the  $N + 1$  vertices of a regular simplex. In this case, since a small number of abscissæ are used, the approximation errors in Eq. (12) become significant and the approximate equality  $\mathbf{W} \approx \mathbf{a}$  no longer holds. However, with suitable choice of the vector length  $|\mathbf{b}_m|$  there remains a linear

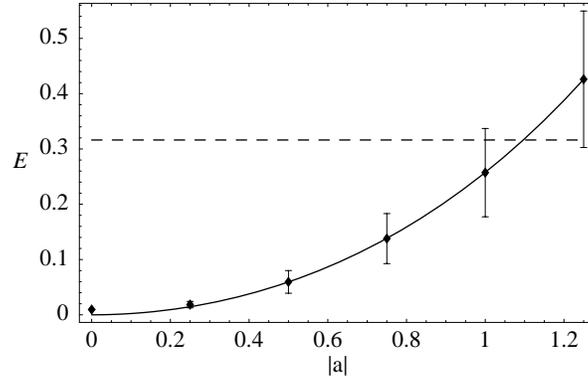


Fig. 7. Aberration measurement accuracy for direct estimation using a simplex arrangement of abscissæ for  $N = 6$  and  $|\mathbf{b}_m| = 0.5$ . The data points show the mean and standard deviation from a sample of 1000 input aberrations. The solid line is an even polynomial fit. The dashed line shows the measurement tolerance of  $E = \sqrt{0.1}$ .

relationship between  $\mathbf{W}$  and  $\mathbf{a}$  so we can calculate  $\mathbf{a}$  as

$$\mathbf{a} \approx \mathbf{S}^{-1} \mathbf{W} , \quad (13)$$

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

$$S_{ik} = \left. \frac{\partial W_i}{\partial a_k} \right|_{\mathbf{a}=\mathbf{0}} . \quad (14)$$

The variables  $W_i$  and  $a_k$  are the corresponding elements of  $\mathbf{W}$  and  $\mathbf{a}$ , respectively, and  $W_i$  is calculated using Eq. (12). A similar matrix was introduced for the closely related method of modal wave front sensing, where Neil *et al.* showed that the matrix was sparse and diagonally dominated with approximately equal diagonal elements [20]. These same properties were noted in the calculation of Eq. (14).

We can investigate the measurement accuracy by defining the measurement error  $E$  as

$$E = |\mathbf{S}^{-1} \mathbf{W} - \mathbf{a}| . \quad (15)$$

As an example, Fig. 7 shows the variation of  $E$  with  $|\mathbf{a}|$  for  $|\mathbf{b}_m| = 0.5$  and  $N = 6$ . Each data point shows the mean and standard deviation from a sample of 1000 randomly oriented input aberrations of a given magnitude. The mean value of  $E$  was within the measurement tolerance of  $\sqrt{0.1}$  over the range  $|\mathbf{a}| \leq 1.09$ . The same analysis was performed for other combinations of modes for  $N \leq 6$ . In all cases  $E$  showed a similar dependence on  $|\mathbf{a}|$  and the ranges over which the mean value of  $E$  was within the measurement tolerance all lay between  $|\mathbf{a}| \leq 1.05$  and  $|\mathbf{a}| \leq 1.26$ .

The direct maximisation scheme permits the measurement of  $N$  Zernike modes using only  $N + 1$  intensity measurements. This is the minimum number of measurements possible to retrieve the  $N + 1$  unknown parameters in the system (i.e., the  $N$  modal coefficients and the total intensity,  $I_0$ ).

## 9. Conclusions

Through construction of an appropriate mathematical model and by heuristic reasoning, we have designed and analysed deterministic aberration measurement algorithms for a wave front

sensorless adaptive optics system. By using knowledge of the maximised function's topology, we were able to optimise algorithms to provide higher efficiency. The results indicated that deterministic, non-adaptive algorithms could be effective in controlling these systems, if suitably formulated.

We demonstrated the relative effectiveness of the different algorithms as the number of aberration modes,  $N$ , was increased. The number of measurements required in the exhaustive search methods increased exponentially with  $N$ , whereas the SAHC method required a number that increases as  $N^2$ . The direct maximisation method was significantly more efficient, requiring only  $N + 1$  measurements for  $N$  modes, over the range of input aberrations used here. This improvement over the other methods was possible since the calculation effectively took into account *a priori* knowledge about the form of the function being maximised. This is applicable to any system where the aberration can be accurately represented by the  $N$  orthonormal modes.

The discussion in this paper was based around the simple optical configuration of Fig. 1. However, similar mathematical models, demonstrating the same spherical symmetry, can be obtained for many other optical systems including those described in the Introduction. With minor adjustments, the results presented here are also applicable to these other systems.

### Acknowledgments

The author is a Royal Academy of Engineering/EPSRC Research Fellow. Thanks are due to M. Schwertner for helpful discussions concerning this paper.

### Appendix A: Zernike polynomials

The Zernike polynomials used in this paper are listed in Table 1. They are defined such that the variance of each polynomial is normalised to unity [21]. Since it represents phase, a coefficient of value 1 corresponds to a wave front variance of 1 rad<sup>2</sup>. The mode indexing schemes, using the single index  $i$  or the dual indices  $(n, m)$ , are explained by Neil *et al.* [20]

Table 1. Zernike mode definitions

Index			Zernike mode	Name
$i$	$n$	$m$		
4	2	0	$\sqrt{3}(2r^2 - 1)$	Defocus
5	2	2	$\sqrt{6}r^2 \cos(2\theta)$	Astigmatism
6	2	-2	$\sqrt{6}r^2 \sin(2\theta)$	Astigmatism
7	3	1	$2\sqrt{2}(3r^3 - 2r) \cos(\theta)$	Coma
8	3	-1	$2\sqrt{2}(3r^3 - 2r) \sin(\theta)$	Coma
9	3	3	$2\sqrt{2}r^3 \cos(3\theta)$	Trefoil

### Appendix B: Simplex construction

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

$$\sum_{n=1}^{N+1} \mathbf{b}_n = \mathbf{0} . \quad (16)$$

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

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

where  $(\mathbf{b}_n)_m$  represents the  $m$ th element of  $\mathbf{b}_n$ . The symmetry of the simplex means that the coefficients  $(\mathbf{b}_n)_1$  for  $n > 1$  are equal, so we can determine the first coordinate of the remaining vectors as

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

We choose 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

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

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

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

### Appendix C: Calculation of step size for SAHC method

An  $N$ -dimensional sphere of radius  $\varepsilon$  is centred at the origin. Its surface, denoted by  $T$ , must be covered by the  $N + 1$  overlapping spheres, also of radius  $\varepsilon$ , centred on the vertices of a regular simplex whose centroid is positioned at the origin. The distance from the origin to each simplex vertex is  $s$  and the vertices are described by the vectors  $s\mathbf{b}_n$ , where we use the vectors derived in Appendix B. There are  $N + 1$  similar points on  $T$ , given by the vectors  $-\varepsilon\mathbf{b}_n$ , that are equidistant from the adjacent vertices (the directions of these vectors correspond, in  $N = 2$ , to the midpoints of the triangle's edges, and in  $N = 3$ , to the centroids of the tetrahedron's faces). If we ensure that one of these points is covered, we can conclude that the whole of  $T$  will be covered. To satisfy this, the distance between the point  $-\varepsilon\mathbf{b}_1$  and the vertex  $s\mathbf{b}_2$  must be less than  $\varepsilon$ . Hence,

$$|s\mathbf{b}_2 - \varepsilon\mathbf{b}_1|^2 = s^2 - \frac{2\varepsilon s}{N} + \varepsilon^2 \leq \varepsilon^2 \quad . \quad (21)$$

For the inequality to be satisfied,

$$s \left( s - \frac{2\varepsilon}{N} \right) \leq 0 \quad , \quad (22)$$

from which it follows that

$$0 \leq s \leq \frac{2\varepsilon}{N} \quad . \quad (23)$$