Astronomy & Astrophysics

Closed loop performance of a layer-oriented multi-conjugate adaptive optics system

E. Diolaiti[1,2], R. Ragazzoni[3], and M. Tordi[3]

[1] Department of Astronomy, University of Padova, vicolo dell'Osservatorio, 2, 35122 Padova, Italy
[2] Department of Astronomy, University of Bologna, via Ranzani, 1, 40127 Bologna, Italy
[3] Astronomical Observatory of Padova, vicolo dell'Osservatorio, 5, 35122 Padova, Italy

Abstract. The dynamic properties of a layer-oriented multi-conjugate adaptive optics system are examined. It is shown that such a system is stable and closes the loop. A simplified analysis of the achievable correction is presented, in order to define the conditions in which the layer-oriented approach is equivalent to other linear reconstruction schemes. It is found that this approach is equivalent to an optimum one, under certain assumptions, but still retaining the potential advantages of such a novel scheme.

Key words. instrumentation: adaptive optics – telescopes

1. Introduction

Adaptive Optics (AO) techniques are widely used to enhance the angular resolution of large ground based telescopes by correcting in real time the aberrations due to the atmospheric turbulence (Beckers 1993). The basic principle is to measure the induced distortions by a WaveFront Sensor (WFS) coupled to a Guide Star (GS) and to correct the incoming wavefront by means of a Deformable Mirror (DM). The light from the science target, which is generally different from the GS, crosses a different path, especially as far as high altitude layers are concerned, and experiences different aberrations. This effect is commonly referred to as angular anisoplanatism (Roddier et al. 1993; Wilson & Jenkins 1996). In order to achieve an acceptable correction on the science object it is therefore necessary to select a sufficiently close GS. Moreover, due to the demands of current WFSs, the GS must be sufficiently bright. These two requests severly limit the sky coverage, unless Laser Guide Stars (LGS) are used (Foy & Labeyrie 1985; Happer et al. 1994). The other consequences of the anisoplanatism are a small corrected field and a variation of the Point Spread Function (PSF) across the Field of View (FoV) (see Simon et al. 1999; Fusco et al. 2000; Conan et al. 2000).

Multi-Conjugate Adaptive Optics (MCAO) is a technique described for the first time by Beckers (1989) to overcome these limitations and obtain a more uniform correction over a larger field (Ellerbroek & Rigaut 2000). The underlying principle is to use more than one DM to correct different planes in the atmosphere and more than one GS to measure the atmospheric turbulence in different directions, in order to reconstruct its 3-dimensional properties (Tallon & Foy 1990; Ragazzoni et al. 1999; Ragazzoni et al. 2000a). A linear reconstruction algorithm is used to convert the WFS measurements into suitable commands for the DMs.

The closed loop properties of such a system have been extensively studied by Ellerbroek (1994). The linear reconstruction matrix has to be optimized on the basis of the AO system setup, including number of DMs, GSs configuration and statistical properties of the atmospheric turbulence. This matrix is chosen to minimize the residual wavefront error, averaged over the telescope pupil and the FoV. A key point is the possibility to optimize the reconstructor for a specified FoV weighting function $\Omega_\mathrm{F}(\theta)$, in order, for instance, to correct a specific direction or to achieve a uniform correction across a large field.

A different MCAO approach has been described in Ragazzoni (2000) and especially in Ragazzoni et al. (2000b), hereafter denoted by RFM. It is referred to as *layer-oriented* because the WFSs are not coupled to the GSs but are conjugated with specific layers in the atmosphere. The DMs are conjugated with the same layers as the WFSs and each of them applies a correction proportional to what the corresponding WFS measures, optically combining the light from the GSs. This optical combination ensures a very efficient use of the incoming photons and it has been speculated that a more uniform sky coverage is possible even with Natural Guide Stars (NGS).

Each DM coupled to a WFS forms an independent loop: it is then possible to optimize the *spatial* and *temporal* sampling of this loop according to the statistical properties of the corresponding conjugated layer. Furthermore this MCAO approach is simpler to implement, because it does not require a huge global reconstruction matrix, even though an alternative scheme can be implemented to combine the wavefront measurements numerically by a wavefront computer. The properties of layer-oriented MCAO have been investigated by means of numerical simulations (RFM), under realistic hypotheses concerning turbulence, measurement noise and statistical distribution of NGSs. This simulation work strongly suggests that the layer-oriented approach is stable and can indeed correct the atmospheric turbulence. In the present paper, the properties of the layer-oriented approach are investigated by means of analytical methods. The two major issues addressed here are the temporal stability and the achievable correction in a closed loop. A comparison with the *classical* MCAO approach is presented.
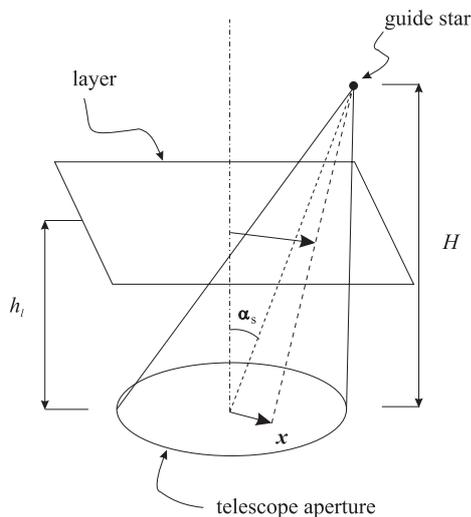

Fig. 1. Geometry used to trace a ray from a GS to a given point on the telescope aperture.

2. Statement of the problem

We assume that the atmospheric turbulence is distributed in L layers, whose altitude from the ground is denoted by h_l. A continuous turbulence profile can be approximated with arbitrary accuracy adopting a suitable number of layers L. The MCAO system has M deformable mirrors, $M \leq L$, conjugated with layers $1, \ldots, M$. According to this convention the layers not conjugated with any DM are denoted by the subscripts $M+1, \ldots, L$; of course there is no relationship between the subscript indicating a given layer and its altitude. The information on the atmospheric turbulence is provided by a set of S reference sources at height H, whose directions are defined by the angles α_s, measured from the optical axis (Fig. 1). We choose to identify the GSs with the vectors $\boldsymbol{\theta}_s$, parallel to the plane of
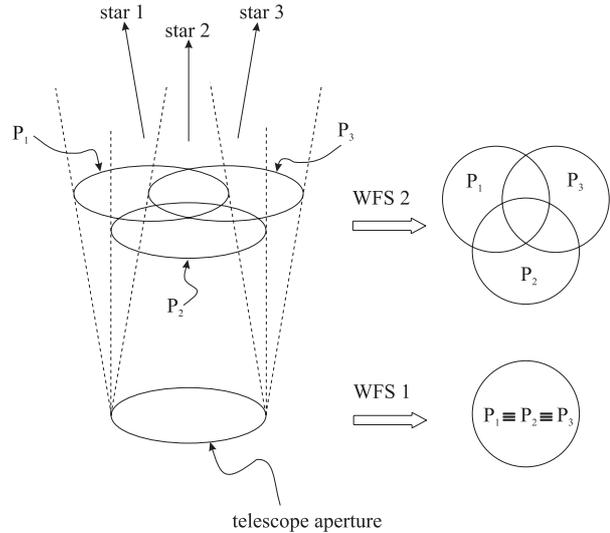

Fig. 2. Simple setup with three GSs and two WFS conjugated with different planes, the lower coincident with the plane of the telescope aperture. The symbols P_1, P_2 and P_3 indicate the footprints of the stars, i.e. the projections of the telescope aperture onto a given plane in the direction of the GSs. On the ground layer the footprints overlap and the conjugated WFS measures the mean of the star wavefronts. On the upper layer the footprints appear displaced and, at any given point, the WFS measures the average wavefront of those stars whose footprint includes that point.

the aperture and with modulus $|\boldsymbol{\theta}_s| = \tan \alpha_s$. A ray from a GS intersects the telescope aperture at $\boldsymbol{x}$ and layer l at $[(1 - h_l/H)\,\boldsymbol{x} + h_l\boldsymbol{\theta}_s]$ (see also Ellerbroek 1994; Johnston & Welsh 1994). In order to simplify the discussion, we assume that the GSs are at an infinite distance from the telescope aperture, so that the point $\boldsymbol{x}$ on the aperture corresponds to $[\boldsymbol{x} + h_l\boldsymbol{\theta}_s]$ on layer l. We will return to this approximation whenever necessary, in order to highlight the differences in the case of LGSs at finite altitude. We notice also that the involved angles α_s are very small in practice, so the approximation $|\boldsymbol{\theta}_s| \approx \alpha_s$ will be used hereafter.

The starlight beams cross different paths in the atmosphere and eventually overlap perfectly on the telescope aperture (Fig. 2). We assume that the entrance pupil of the telescope is located on the aperture plane; this condition, however, does not modify in any form the essential results that we find in this work. The integrated wavefront errors experienced by the stars are different because of the high altitude layers. A WFS conjugated with the telescope aperture *sees* the footprints overlap. The measured wavefront is a linear combination of the single star wavefronts, with weights equal to the normalized intensities of the stars (see RFM), i.e.

$$W\left(\boldsymbol{x}\right) = \frac{\sum_{s=1}^{S} I_s \times W^{(s)}\left(\boldsymbol{x}\right)}{\sum_{s=1}^{S} I_s}. \tag{1}$$

The star wavefronts $W^{(s)}$ are integrated over the corresponding atmosphere columns

$$W^{(s)}\left(\boldsymbol{x}\right) = \sum_{l=1}^{L} w_l\left(\boldsymbol{x} + h_l\boldsymbol{\theta}_{\mathrm{s}}\right) \tag{2}$$

where w_l denotes the phase error introduced by the lth layer. Replacing Eqs. (2) in (1) and assuming for simplicity that the stars have the same intensity, the WFS measurement can be expressed as

$$W\left(\boldsymbol{x}\right) = \frac{1}{S}\sum_{s=1}^{S}\sum_{l=1}^{L} w_l\left(\boldsymbol{x} + h_l\boldsymbol{\theta}_{\mathrm{s}}\right). \tag{3}$$

Notice that the layers are indicated by lower-case letters, whereas the WFS measurement is in capitals. A WFS conjugated with a different height, for instance with the higher layer in Fig. 2, measures a similar combination, the only difference being that the star footprints do not perfectly overlap, as in the case of the ground layer, but are slightly displaced. As a consequence of this partial overlap, there are regions in the high layers that are not illuminated by all the S stars. This edge effect is neglected for simplicity. Under this approximation the measurement of any WFS can be written as

$$W_m\left(\boldsymbol{x}\right) = \frac{1}{S}\sum_{s=1}^{S}\sum_{l=1}^{L} w_l\left[\boldsymbol{x} + (h_l - h_m)\,\boldsymbol{\theta}_{\mathrm{s}}\right] \tag{4}$$

which is just a generalization of Eq. (3). Each term in Eq. (4) can be expressed as the convolution integral of layer $w_l\left(\boldsymbol{x}\right)$ with a suitable Dirac *delta*, which provides the required shift. This allows Eq. (4) to be put in the following compact form

$$W_m\left(\boldsymbol{x}\right) = \sum_{l=1}^{L} \delta_{m,l}\left(\boldsymbol{x}\right) \otimes w_l\left(\boldsymbol{x}\right) \tag{5}$$

where $\otimes$ denotes the convolution integral and

$$\delta_{m,l}\left(\boldsymbol{x}\right) = \frac{1}{S}\sum_{s=1}^{S} \delta\left[\boldsymbol{x} + (h_l - h_m)\,\boldsymbol{\theta}_{\mathrm{s}}\right]. \tag{6}$$

The convolution with the sum of shifted *delta*-functions in Eq. (5) is equivalent to the average of different portions of a given layer: this convolution represents how the mth WFS *sees* the lth non-conjugated layer ($l \neq m$).

In closed loop, the correction $c_m\left(\boldsymbol{x}\right)$ applied by the mth DM affects the measurement of *all* the WFSs, as if it were a wavefront error due to the turbulence. A given point on the DM corresponds to a specific point on the conjugated layer and to an extended region on a nonconjugated plane, of angular size equal to the FoV (Fig. 3). Introducing the DM corrections, the WFS measurements can be re-written as

$$W_m\left(\boldsymbol{x}\right) =$$
$$\sum_{l=1}^{L} \delta_{m,l}\left(\boldsymbol{x}\right) \otimes w_l\left(\boldsymbol{x}\right) - \sum_{l=1}^{M} \delta_{m,l}\left(\boldsymbol{x}\right) \otimes c_l\left(\boldsymbol{x}\right). \tag{7}$$
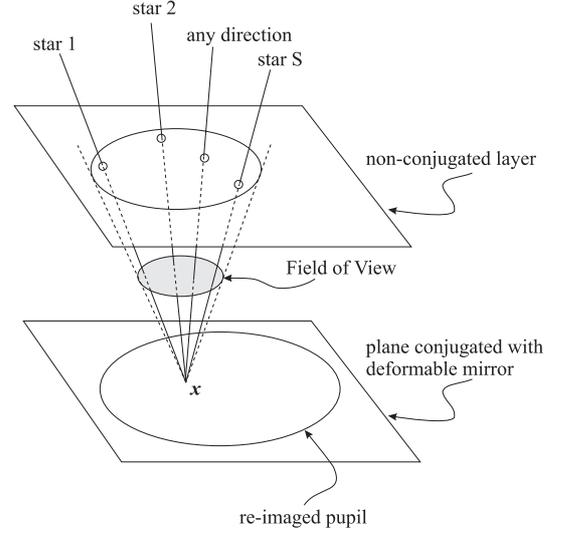

Fig. 3. The bottom plane is optically conjugated with a DM. Each point on this plane corresponds to a region of angular size equal to the FoV on the non-conjugated plane.

In the layer-oriented approach the correction applied by a DM is proportional to the measurement of the corresponding WFS. Actually the instantaneous measurements are filtered in time. Assuming as a linear filter a pure integrator, the dynamic behaviour is described by the system of linear differential equations

$$\frac{1}{\gamma_m}\frac{\mathrm{d}c_m\left(\boldsymbol{x}, t\right)}{\mathrm{d}t} = W_m\left(\boldsymbol{x}, t\right) \tag{8}$$

where γ_m is the gain (assumed to be positive) of the mth independent loop, associated to the mth DM and WFS. Replacing Eqs. (7) in (8) we obtain

$$\frac{1}{\gamma_m}\frac{\mathrm{d}c_m\left(\boldsymbol{x}, t\right)}{\mathrm{d}t} =$$
$$\sum_{l=1}^{L} \delta_{m,l}\left(\boldsymbol{x}\right) \otimes w_l\left(\boldsymbol{x}, t\right) - \sum_{l=1}^{M} \delta_{m,l}\left(\boldsymbol{x}\right) \otimes c_l\left(\boldsymbol{x}, t\right). \tag{9}$$

The general solution of Eq. (9) yields the closed loop dynamic behaviour of the layer-oriented MCAO system. As can be seen from Eq. (9), this solution depends on the temporal evolution of the atmospheric turbulence, whose detailed properties can be known only on a statistical basis. Without introducing any statistical assumption, however, some important issues can be addressed. First, it is possible to study the stability of the system, i.e. the capability to return to the initial conditions when perturbed by an instantaneous input. Furthermore, under the simplifying hypothesis that the response time of the system is much shorter than the evolution time of the turbulence, one may easily study the achievable correction.

3. Steady-state response

The system of Eq. (9) can be handled more easily if all the involved functions are Fourier-transformed with respect to

the spatial variable $\boldsymbol{x}$, so that the convolutions become products in the Fourier domain. The Fourier-transformed system is

$$\frac{1}{\gamma_m}\frac{\mathrm{d}\tilde{c}_m(\boldsymbol{u},t)}{\mathrm{d}t} =$$

$$\sum_{l=1}^{L}\tilde{\delta}_{m,l}(\boldsymbol{u})\,\tilde{w}_l(\boldsymbol{u},t) - \sum_{l=1}^{M}\tilde{\delta}_{m,l}(\boldsymbol{u})\,\tilde{c}_l(\boldsymbol{u},t) \qquad (10)$$

where $\boldsymbol{u}$ is the spatial frequency conjugated to $\boldsymbol{x}$ and the symbol $\tilde{f}(\boldsymbol{u})$ denotes the Fourier transform of $f(\boldsymbol{x})$. The Fourier transform of the superposition of shifted *delta*-functions in Eq. (6) is the average of S complex exponentials that can be interpreted as unit vectors in the complex plane:

$$\tilde{\delta}_{m,l}(\boldsymbol{u}) = \frac{1}{S}\sum_{s=1}^{S}\mathrm{e}^{i2\pi(h_l-h_m)\boldsymbol{\theta}_s\cdot\boldsymbol{u}}. \qquad (11)$$

It should be noticed that the support of the functions $c(\boldsymbol{x},t)$ and $w(\boldsymbol{x},t)$ is not infinite: in the case of the ground layer this support is practically the telescope pupil, whereas for a higher layer it is a superposition of S shifted pupils (see Fig. 2). This fact could be accounted for by convolving the Fourier transforms of the functions in Eq. (10) with the transform of the footprint; this effect, however, is not treated here, for simplicity.

The system of differential Eq. (10) can be put in the matrix form

$$\frac{\mathrm{d}\tilde{\boldsymbol{c}}(t)}{\mathrm{d}t} = \boldsymbol{A}\tilde{\boldsymbol{c}}(t) + \boldsymbol{B}\tilde{\boldsymbol{w}}(t) \qquad (12)$$

where the vectors $\tilde{\boldsymbol{c}}$ and $\tilde{\boldsymbol{w}}$ and the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are defined as

$$\begin{cases} \tilde{\boldsymbol{c}}(t) = \begin{bmatrix} \tilde{c}_1(t) & \tilde{c}_2(t) & \ldots & \tilde{c}_M(t) \end{bmatrix}^{\mathrm{T}} \\[2mm] \tilde{\boldsymbol{w}}(t) = \begin{bmatrix} \tilde{w}_1(t) & \tilde{w}_2(t) & \ldots & \tilde{w}_M(t) & \ldots & \tilde{w}_L(t) \end{bmatrix}^{\mathrm{T}} \\[2mm] \boldsymbol{A} = \begin{bmatrix} -\gamma_1 & -\gamma_1\tilde{\delta}_{1,2} & \ldots & -\gamma_1\tilde{\delta}_{1,M} \\ -\gamma_2\tilde{\delta}_{2,1} & -\gamma_2 & \ldots & -\gamma_2\tilde{\delta}_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ -\gamma_M\tilde{\delta}_{M,1} & -\gamma_M\tilde{\delta}_{M,2} & \ldots & -\gamma_M \end{bmatrix} \\[2mm] \boldsymbol{B} = \begin{bmatrix} -\boldsymbol{A} & \boldsymbol{B}'' \end{bmatrix} \\[2mm] \boldsymbol{B}'' = \begin{bmatrix} \gamma_1\tilde{\delta}_{1,M+1} & \ldots & \gamma_1\tilde{\delta}_{1,L} \\ \gamma_2\tilde{\delta}_{2,M+1} & \ldots & \gamma_2\tilde{\delta}_{2,L} \\ \vdots & \ddots & \vdots \\ \gamma_M\tilde{\delta}_{M,M+1} & \ldots & \gamma_M\tilde{\delta}_{M,L} \end{bmatrix}. \end{cases} \qquad (13)$$

The vector $\tilde{\boldsymbol{w}}$ in the definition above can be partitioned into two blocks: the first of M elements, corresponding to the set of conjugated layers, and the second of size $L-M$, corresponding to the layers not conjugated with any DM or WFS. The matrix $\boldsymbol{A}$ determines the temporal evolution of the system, whereas the matrix $\boldsymbol{B}$ is again composed

of two blocks: one associated to the conjugated layers, the other to the remaining ones. When a very high number L of layers is adopted to model accurately a continuous turbulence profile, the size of the vector $\tilde{\boldsymbol{w}}$ and of the matrix $\boldsymbol{B}''$ increase, but the complexity of the system, in terms of temporal evolution, is always determined by the size of the matrix $\boldsymbol{A}$, i.e. by the number of DMs.

The solution of the system described by Eq. (12) can be found in any textbook on differential equations and is given by

$$\tilde{\boldsymbol{c}}(t) = \mathrm{e}^{\boldsymbol{A}t}\tilde{\boldsymbol{c}}(0) + \int_0^t \mathrm{e}^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\tilde{\boldsymbol{w}}(\tau)\,\mathrm{d}\tau. \qquad (14)$$

Introducing the initial condition $\tilde{\boldsymbol{c}}(0) = 0$, i.e. assuming that the DMs are initially flat, we obtain the steady-state response

$$\tilde{\boldsymbol{c}}(t) = \int_0^t \mathrm{e}^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\tilde{\boldsymbol{w}}(\tau)\,\mathrm{d}\tau. \qquad (15)$$

The *shape* of the DMs at the instant t clearly depends on the history of the atmospheric turbulence.

4. Stability

The first question to be addressed is the stability of layer-oriented MCAO. In general, an AO system is affected by various sources of instability, including measurement noise and loop delay. These factors have already been studied by many authors and the developed theory can be applied as it is to the layer-oriented approach, which is composed of M independent AO loops. In the present work we concentrate on the specific source of instability inherent to layer-oriented. Once the various independent loops are closed, in fact, their correlation, introduced by the particular optical setup, must be taken into account as a possibile cause of instability.

We identify the input of the AO system with the uncorrected atmospheric turbulence $\boldsymbol{w}$. The most natural definition of output is the corrected phase $\boldsymbol{w} - \boldsymbol{c}$. An equivalent formulation is possible, in which the output is identified with the shape $\boldsymbol{c}$ of the DMs (Fig. 4). The concept of stability (see for instance Phillips & Troy Nagle 1984) is illustrated in Fig. 5. An impulse input is *switched on* and immediately *switched off*, perturbing the system and causing it to react. If the output tends asymptotically to zero, the system is said to be stable. The impulse input can modelled in our case as $\boldsymbol{w}(\boldsymbol{x},t) = \boldsymbol{w}_0(\boldsymbol{x})\,\delta(t)$. Replacing the Fourier transform of this expression in Eq. (15), we obtain

$$\tilde{\boldsymbol{c}}(t) = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{B}\tilde{\boldsymbol{w}}_0. \qquad (16)$$

We assume that the matrix $\boldsymbol{A}$ can be diagonalized according to the expression

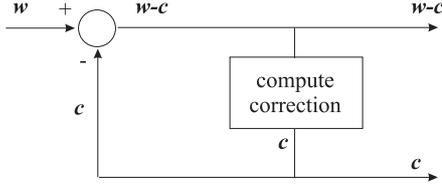$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{-1} \qquad (17)$$

Fig. 4. Block diagram of an AO system. The input is the turbulent phase $\boldsymbol{w}$; the output may be identified with the corrected phase $\boldsymbol{w} - \boldsymbol{c}$ or the DM figure $\boldsymbol{c}$.
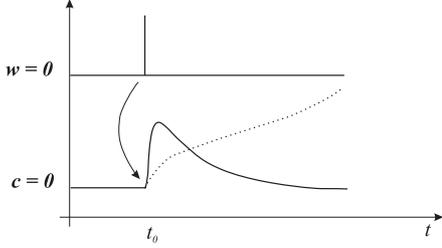

Fig. 5. Top: impulse input. Bottom: impulse response of a stable system (continuous line) and an unstable one (dotted line).

where $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues and $\boldsymbol{U}$ is the matrix formed by the eigenvectors of $\boldsymbol{A}$. Exploiting the Taylor expansion

$$\mathrm{e}^{\boldsymbol{A}t} = \sum_{n=0}^{\infty} \frac{\boldsymbol{A}^n t^n}{n!} \qquad (18)$$

it is straightforward to show that the exponential $\mathrm{e}^{\boldsymbol{A}t}$ can be expressed in the form

$$\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{U} \mathrm{e}^{\boldsymbol{\Lambda}t} \boldsymbol{U}^{-1} \qquad (19)$$

where

$$\mathrm{e}^{\boldsymbol{\Lambda}t} = \begin{bmatrix} \mathrm{e}^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & \mathrm{e}^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathrm{e}^{\lambda_M t} \end{bmatrix}. \qquad (20)$$

Replacing Eqs. (19) in (16), the temporal behaviour becomes

$$\tilde{\boldsymbol{c}}(t) = \boldsymbol{U} \mathrm{e}^{\boldsymbol{\Lambda}t} \boldsymbol{U}^{-1} \boldsymbol{B} \tilde{\boldsymbol{w}}_0. \qquad (21)$$

Each component of the vector $\tilde{\boldsymbol{c}}(t)$ is clearly a linear superposition of exponential terms of the form $\mathrm{e}^{\lambda t}$, therefore the asymptotic properties of the system output are determined by the sign of the eigenvalues of $\boldsymbol{A}$. If the eigenvalues are complex, the solution has an oscillating part and the asymptotic behaviour is determined by the sign of the real part of the eigenvalues. These can be determined solving the characteristic equation

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0 \qquad (22)$$

i.e. finding the solutions of a polynomial of degree equal to the number of DMs (M). The easiest case corresponds to $M = 2$. The cases with $M = 3$ and $M = 4$ can be handled by the well-known Cardano-Tartaglia formulas,

whereas higher order cases can be treated by numerical methods. In this work the situation $M = 2$ is considered, a special though very important case in practice (Fusco et al. 1999). The characteristic equation becomes

$$\lambda^2 + (\gamma_1 + \gamma_2)\lambda + \gamma_1 \gamma_2 \left(1 - |\tilde{\delta}_{1,2}|^2\right) = 0 \qquad (23)$$

and the eigenvalues (real in this case) are

$$\lambda = \left[-(\gamma_1 + \gamma_2) \pm \sqrt{\Delta} \right] / 2 \qquad (24)$$

where

$$\Delta = (\gamma_1 + \gamma_2)^2 - 4\gamma_1 \gamma_2 \left(1 - |\tilde{\delta}_{1,2}|^2\right) \qquad (25)$$

and

$$\tilde{\delta}_{1,2}(\boldsymbol{u}) = \frac{1}{S} \sum_{s=1}^{S} \mathrm{e}^{i2\pi(h_2 - h_1)\boldsymbol{\theta}_\mathrm{s} \cdot \boldsymbol{u}} \qquad (26)$$

is the Fourier transform of

$$\delta_{1,2}(\boldsymbol{x}) = \frac{1}{S} \sum_{s=1}^{S} \delta \left[\boldsymbol{x} + (h_2 - h_1)\boldsymbol{\theta}_\mathrm{s}\right]. \qquad (27)$$

Each term of the summation in Eq. (26) may be regarded as a vector of unit length in the complex plane (see Fig. 6); according to the generalized triangular inequality, the lenght of the sum is not larger than the sum of the individual lengths. Therefore the following condition holds

$$|\tilde{\delta}_{1,2}| \le 1. \qquad (28)$$

The equality sign holds only when the unit vectors are all aligned, or the complex exponentials are in phase. In general, however, $|\tilde{\delta}_{1,2}| < 1$. It is easy to see that this condition implies $\lambda < 0$, therefore all the exponential terms $\mathrm{e}^{\lambda t}$ vanish asymptotically with time. In this case the output of the system (Eq. (21)) fulfills the condition

$$\lim_{t \to \infty} \tilde{\boldsymbol{c}}(t) = 0 \qquad (29)$$

and the system is stable. This result has been derived under the initial approximation of GSs at infinite altitude, but it is valid also with artificial beacons at finite range. In this case it can be shown that the quantities $\tilde{\delta}_{m,l}$, which ultimately determine the temporal evolution of the system, are defined in a slightly different way from Eqs. (11) or (26): the differences $h_l - h_m$ entering the exponents and the terms of the summation are scaled by factors depending on the ranges h_l and h_m and on the altitude H of the GSs. However the considerations reported above are still valid and the system remain stable as long as one of the DMs is not at, or very close, to the beacon (or, in other words $H \approx h_1$ or $H \approx h_2$).

The spatial frequencies for which the equality $|\tilde{\delta}_{1,2}| = 1$ holds are characterized by $\lambda = 0$, i.e. the system output does not vanish but remains limited. The system is said to be marginally stable in this case. These frequencies may
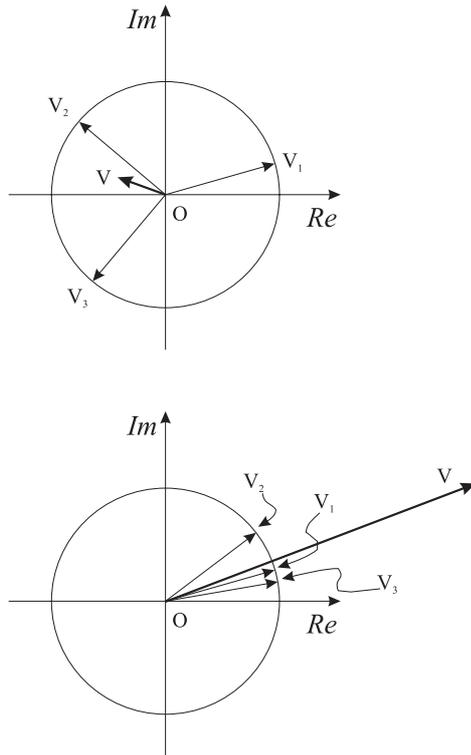
Fig. 6. Unit vectors in the complex plane. The boldface arrow is the sum of the three unit vectors. When the vectors have different phase (top), the length of their sum is small. When they are almost aligned (bottom), the length of the vector sum tends to three times the length of each component.

Fig. 7. Illustration of a waffle mode. The unit vectors which define the directions of the GSs are co-planar. The wavefront error is a sinusoidal pattern, corresponding to the spatial frequency $\boldsymbol{u}$ orthogonal to the plane identified by the GSs. All the starlight beams experience the same wavefront perturbation and no information about its occurrence at different ranges can be retrieved.

be identified with the *waffle modes* of the layer-oriented approach and are defined by the condition

$$2\pi\left(h_2 - h_1\right)\boldsymbol{\theta}_{\mathrm{s}}\cdot\boldsymbol{u} = \mathrm{const} + 2n\pi, \quad n = 0, \pm 1, \pm 2, \ldots \tag{30}$$

which must hold simultaneously for all the GSs. These invisible modes depend on the GS configuration (Fig. 7). In practice, for a realistic configuration, the unique waffle mode is the piston term, corresponding to $\boldsymbol{u} = 0$ in Eq. (30).

Incidentally, we notice that if the GSs had different brightness and one of them were much brighter than the other, the summation in Eq. (26) would practically reduce to a single term and the modulus of $\tilde{\delta}_{1,2}$ would approach unity. This effect can be seen considering again the unit vectors in the complex plane: if very different weights are applied to these vectors, the sum is practically coincident with the more weighted one. In this case the system would be marginally stable. On the other hand this result should be expected: the layer-oriented weights the wavefront measurements by the relative intensities of the GSs and the case of a star much brighter than the others tends to the case with a single reference, where no 3-dimensional mapping of the turbulence is possible. We recall here the possibility, already reported in RFM, of dimming the brightness of such a luminous star.
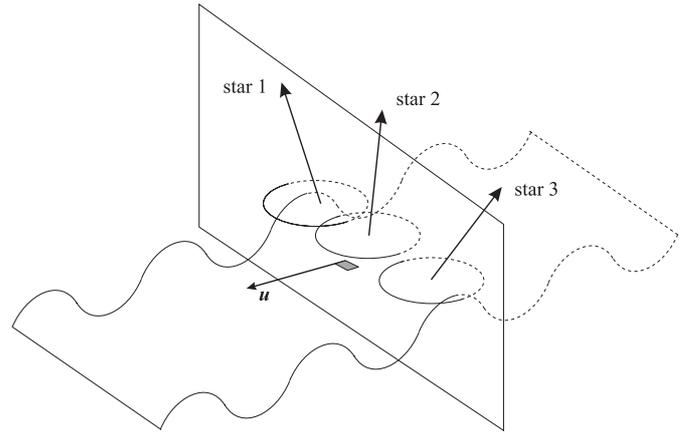
5. Achievable correction

The steady-state response of the MCAO system (Eq. (15)) depends on the atmospheric turbulence history. The system temporal bandwidth in practice is limited by noise, which imposes a finite integration time on the detectors. In order to obtain some essential insights in the layer-oriented approach, we consider the case of a system much faster than the turbulence, a condition which is reasonably met in most practical conditions. Under this approximation we analyze the asymptotic response ($t \rightarrow \infty$). Equation (15) can be recast to the simplified form

$$\tilde{\boldsymbol{c}}\left(t\right) = \left[\int_0^t \mathrm{e}^{\boldsymbol{A}(t-\tau)}d\tau\right]\boldsymbol{B}\tilde{\boldsymbol{w}}. \tag{31}$$

The integral in square brackets can be easily solved and, in the limit $t \rightarrow \infty$, we obtain

$$\lim_{t\to\infty}\int_0^t \mathrm{e}^{\boldsymbol{A}(t-\tau)}\mathrm{d}\tau = -\boldsymbol{A}^{-1}. \tag{32}$$

The asymptotic steady-state response becomes

$$\tilde{\boldsymbol{c}}_\infty = -\boldsymbol{A}^{-1}\boldsymbol{B}\tilde{\boldsymbol{w}}. \tag{33}$$

It should be stressed that the same result can be obtained from Eq. (12) imposing the condition

$$\frac{\mathrm{d}\tilde{\boldsymbol{c}}\left(t\right)}{\mathrm{d}t} = 0. \tag{34}$$

This equation actually defines the so called *static response*, i.e. the correction which gives null WFS measurements in all analyzed planes at a given time; the physical meaning of the two approaches is different, even though the result is the same.

The vector $\tilde{\boldsymbol{w}}$ in Eq. (33) can be partitioned in two blocks (Eq. (13)), one related to the conjugated layers,

the other representing the layers not conjugated with any DM:

$$\tilde{\boldsymbol{w}} = \left[\, \tilde{\boldsymbol{w}}' \;\; \tilde{\boldsymbol{w}}'' \,\right]^{\mathrm{T}}. \tag{35}$$

If we now recall the definition of the matrix $\boldsymbol{B}$ as composed of two blocks $-\boldsymbol{A}$ and $\boldsymbol{B}''$ (Eq. (13)), the asymptotic response can be expressed as

$$\tilde{\boldsymbol{c}}_\infty = \tilde{\boldsymbol{w}}' - \boldsymbol{A}^{-1}\boldsymbol{B}''\tilde{\boldsymbol{w}}''. \tag{36}$$

If the turbulence is distributed only in the layers conjugated with the DMs, i.e. $\tilde{\boldsymbol{w}}'' = 0$, the correction applied by each DM tends to the corresponding conjugated layer and the residual turbulence is zero. In the more general case each DM corrects its own layer exactly, but acts also on *all* the non-conjugated layers. In order to investigate the degree of correction, we consider the special case of $M = 2$ DMs and $L = 3$ layers, two of which are conjugated with the DMs (Fig. 8). The matrices $\boldsymbol{A}^{-1}$ and $\boldsymbol{B}''$ become

$$\begin{cases} \boldsymbol{A}^{-1} = \tilde{\phi} \begin{bmatrix} -1/\gamma_1 & \tilde{\delta}_{1,2}/\gamma_2 \\ \tilde{\delta}_{2,1}/\gamma_1 & -1/\gamma_2 \end{bmatrix} \\[2ex] \boldsymbol{B}'' = \begin{bmatrix} \gamma_1 \tilde{\delta}_{1,3} \\ \gamma_2 \tilde{\delta}_{2,3} \end{bmatrix} \end{cases} \tag{37}$$

where

$$\tilde{\phi} = \frac{1}{1 - |\tilde{\delta}_{1,2}|^2} \tag{38}$$

and the asymptotic correction can be written as

$$\begin{cases} \tilde{c}_{\infty,1} = \tilde{w}_1 + \tilde{\phi}\left(\tilde{\delta}_{1,3} - \tilde{\delta}_{1,2}\tilde{\delta}_{2,3}\right)\tilde{w}_3 \\ \tilde{c}_{\infty,2} = \tilde{w}_2 + \tilde{\phi}\left(\tilde{\delta}_{2,3} - \tilde{\delta}_{2,1}\tilde{\delta}_{1,3}\right)\tilde{w}_3 \end{cases}. \tag{39}$$

Given the properties of the sum of complex exponentials $\tilde{\delta}_{1,2}$ (Eq. (26) and Fig. 6), we may adopt the approximation $\tilde{\phi} \approx 1$. Applying an inverse Fourier transform, the correction becomes

$$\begin{cases} c_{\infty,1} = w_1 + (\delta_{1,3} - \delta_{1,2} \otimes \delta_{2,3}) \otimes w_3 \\ c_{\infty,2} = w_2 + (\delta_{2,3} - \delta_{2,1} \otimes \delta_{1,3}) \otimes w_3 \end{cases}. \tag{40}$$

According to the result (40), each DM corrects exactly its own conjugated layer (w_1 and w_2 respectively), plus the non-conjugated layer denoted by w_3. The correction applied to layer 3 by DM 1, for instance, is the difference of two terms (see Fig. 8). The first is $\delta_{1,3} \otimes w_3$, corresponding to the average of S shifted copies of layer w_3, one copy for each GS: this term can be interpreted as the correction that DM 1 would apply to layer 3 neglecting the effect of DM 2. The second contribution is the double convolution $\delta_{1,2} \otimes (\delta_{2,3} \otimes w_3)$. The term $\delta_{2,3} \otimes w_3$, similar to $\delta_{1,3} \otimes w_3$, is the correction that DM 2 would apply to layer 3 neglecting the effect of DM 1. This correction is *seen* from DM 1 through a further convolution with $\delta_{1,2}$ and it is subtracted from $\delta_{1,3} \otimes w_3$ in order to avoid an over-correction.
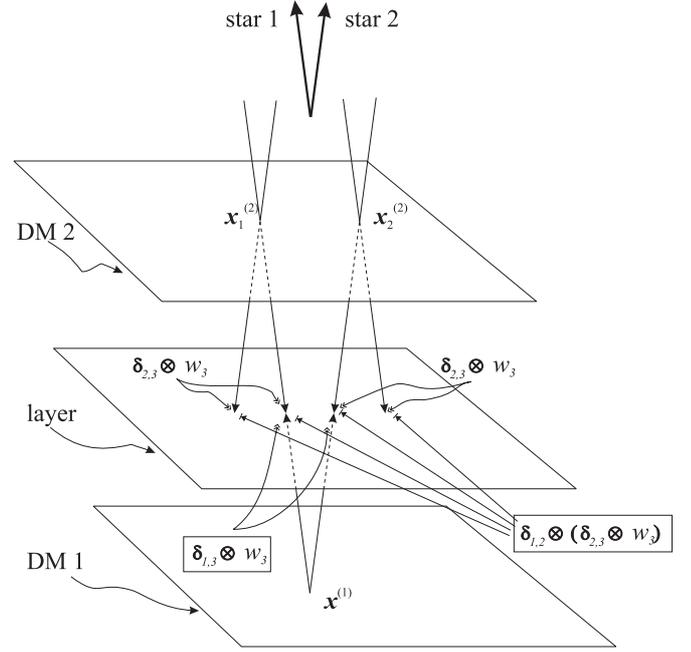

Fig. 8. Simple setup with three layers and two DMs, conjugated with the lower and the upper layers respectively. In a closed loop each DM corrects the corresponding layer and both correct the intermediate layer (w_3). The correction applied on layer 3 by DM 1, for instance, is the difference of the two terms enclosed in the rectangular frames. The term $\delta_{1,3} \otimes w_3$ is the correction that DM 1 would apply to layer 3 in the absence of DM 2. Conversely $\delta_{2,3} \otimes w_3$ is the correction that DM 2 would apply to layer 3 in the absence of DM 1. The latter correction is *seen* by DM 1 through a further convolution by $\delta_{1,2}$.

Apart from the interaction between the DMs, the correction applied to a non-conjugated layer is equal to a spatially smoothed version of the layer itself. This smoothing, which tends to an integration in the limiting case of infinitely high density of GSs, sweeps a region of spatial extent $r \approx \Delta h \Delta \theta$, where Δh is the distance between the conjugated and the non-conjugated plane and $\Delta \theta$ is of the order of the maximum angular displacement between the various GSs. The cutoff frequency is therefore $f_c \approx 1/(\Delta h \Delta \theta)$: a DM conjugated with a plane very distant from the layer corrects only the lowest frequencies of the turbulence power spectrum for such a layer. According to this reasoning, the DM conjugated with the plane nearest to the layer determines the maximum correctable spatial frequency and is supported also by the other DMs in correcting the low order modes: as a by-product of this interaction, the dynamic range requirements on each DM are relaxed.

The simplified analysis presented in this section shows, for instance, that the conjugated layers are corrected perfectly. In practice, a perfect correction is never possible: the system closes the loop and the residual wavefront error approaches a stable value, which is however different from zero.

6. Optimality

Since the conjugated layers are corrected perfectly, in the limit of the adopted approximations, it is interesting to characterize the correction achievable on a plane not conjugated with any DM. If w is the wavefront error introduced by the considered layer and c_m is the correction applied by the mth DM, the MCAO approach analyzed in Ellerbroek (1994) minimizes the merit function

$$
\begin{aligned}
\text{MF} = & \int \mathrm{d}\boldsymbol{x}\,\Omega_\mathrm{A}\left(\boldsymbol{x}\right) \int \mathrm{d}\boldsymbol{\theta}\,\Omega_\mathrm{F}\left(\boldsymbol{\theta}\right) \\
& \times \left[\sum_{m=1}^{M} c_m\left(\boldsymbol{x}+h_m\boldsymbol{\theta}\right)-w\left(\boldsymbol{x}+h\boldsymbol{\theta}\right)\right]^2
\end{aligned} \tag{41}
$$

where $\Omega_\mathrm{A}\left(\boldsymbol{x}\right)$ and $\Omega_\mathrm{F}\left(\boldsymbol{\theta}\right)$ are weighting functions defined over the telescope aperture and the FoV respectively, normalized to unit integral; the layer altitude is denoted simply by h. In practice, MF represents the mean square error between the overall correction applied by the DMs and the turbulent phase w due to the non-conjugated layer; this error is averaged over the telescope aperture and FoV.

When the layer is close to a specific conjugation plane, the effect of the other DMs is negligible. This can be easily understood considering again the special case shown in the previous section and the result expressed by Eq. (40). Assuming that the non-conjugated layer w_3 is very close to the plane conjugated with DM 1, the effect of DM 2, represented to a first approximation by the term $\delta_{2,3}\otimes w_3$, involves spatial smoothing over spatial scales much larger than those concerning DM 1 and is therefore negligible. For the same reason the cross term $\delta_{1,2}\otimes\left(\delta_{2,3}\otimes w_3\right)$ is negligible. This approximation is enforced when the FoV is filled with a suitable number of guide stars, which average distant and uncorrelated portions of the layer. The relevant correction term is therefore $\delta_{1,3}\otimes w_3$. Following the conventions adopted in Fig. 9, this correction can be expressed as

$$
c\left(\boldsymbol{x}\right) = \frac{1}{S}\sum_{s=1}^{S}\delta\left(\boldsymbol{x}+\Delta h\boldsymbol{\theta}_\mathrm{s}\right)\otimes w\left(\boldsymbol{x}\right) \tag{42}
$$

or

$$
c\left(\boldsymbol{x}\right) = \frac{1}{S}\sum_{s=1}^{S}w\left(\boldsymbol{x}+\Delta h\boldsymbol{\theta}_\mathrm{s}\right) \tag{43}
$$

where the non-conjugated layer has been denoted simply by $w\left(\boldsymbol{x}\right)$ and Δh is the distance of the layer from the conjugated plane. We recall here that, in the case of artificial reference sources at finite altitude, this result holds with a slight modification of the cutoff frequency for the spatial smoothing and of the weights applied to the *delta*-functions. In this simplified situation, the merit function Eq. (41) can be recast to the form

$$
\text{MF} = \int \mathrm{d}\boldsymbol{x}\,\Omega_\mathrm{A}\left(\boldsymbol{x}\right) \int \mathrm{d}\boldsymbol{\theta}\,\Omega_\mathrm{F}\left(\boldsymbol{\theta}\right)\left[c\left(\boldsymbol{x}\right)-w\left(\boldsymbol{x}+\Delta h\boldsymbol{\theta}\right)\right]^2 . \tag{44}
$$

No hypothesis is introduced here on the correction, apart from the fact that each point $\boldsymbol{x}$ on the conjugated plane
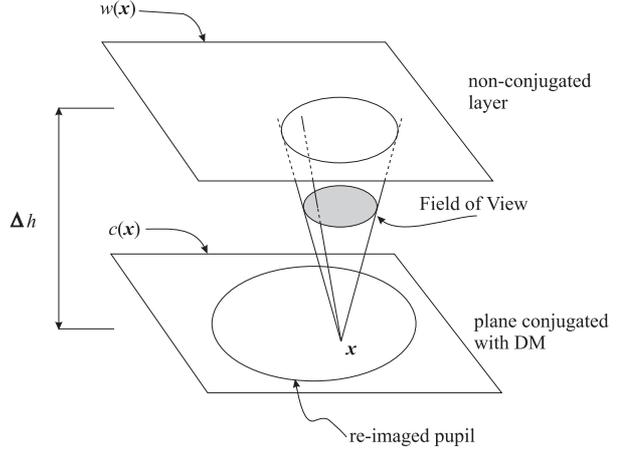

Fig. 9. Geometry used to define the best correction that a DM can apply to a non-conjugated layer.

corresponds to *all* the points on layer w intercepted by a solid angle equal to the FoV (Fig. 9). The optimal correction can be easily found by setting to zero the derivative of MF with respect to c: this condition can be written explicitly as

$$
\int \mathrm{d}\boldsymbol{x}\,\Omega_\mathrm{A}\left(\boldsymbol{x}\right)\left[c\left(\boldsymbol{x}\right)-\int \mathrm{d}\boldsymbol{\theta}\,\Omega_\mathrm{F}\left(\boldsymbol{\theta}\right)w\left(\boldsymbol{x}+\Delta h\boldsymbol{\theta}\right)\right] = 0 \tag{45}
$$

which is fulfilled if

$$
c\left(\boldsymbol{x}\right) = \int \mathrm{d}\boldsymbol{\theta}\,\Omega_\mathrm{F}\left(\boldsymbol{\theta}\right)w\left(\boldsymbol{x}+\Delta h\boldsymbol{\theta}\right) . \tag{46}
$$

The last result coincides with the layer-oriented correction Eq. (43) if the FoV weighting function is a sum of shifted *delta*-functions (possibly weighted by the normalized intensities), one for each reference star:

$$
\Omega_\mathrm{F}\left(\boldsymbol{\theta}\right) = \frac{1}{S}\sum_{s=1}^{S}\delta\left(\boldsymbol{\theta}-\boldsymbol{\theta}_\mathrm{s}\right) . \tag{47}
$$

The optimality of the layer-oriented correction has been shown here in the special case when only one DM is relevant to correct the considered layer. In general it is straightforward to extend this result and show that the merit function in Eq. (41) is minimized by the layer-oriented correction expressed by Eq. (39), which includes the interaction term between the DMs.

This result deserves a discussion. Layer-oriented MCAO is a *linear* reconstruction scheme, implemented optically. It has been shown here that, if the FoV weighting function is a sum of *delta*-functions, one for each GS, then the achieved correction is optimal according to the same criterion adopted in Ellerbroek (1994) to optimize the linear reconstructor embedded in the *classical* MCAO approach. Two linear reconstructors which are optimal according to the same criterion must be *equivalent*. The advantage of the layer-oriented approach is that it is composed of independent loops, one for each WFS coupled to a specific DM. It is therefore possible to optimize the spatial and temporal sampling of each loop according to the properties of the corresponding conjugated layer, which are commonly believed to depend on the altitude.

7. Conclusions

We have presented a simplified analysis of a layer-oriented MCAO system. It has been shown that such a system is stable: if it is perturbed by an impulse input, it returns asymptotically to the initial condition. The system is able to close the loop and the achievable correction has been characterized with a simplified asymptotic analysis. It has been shown that, if the FoV weighting function is a sum of *delta*-functions, one for each GS, the correction is optimal: in this case the layer-oriented approach *coincides* with the *classical* one analyzed in Ellerbroek (1994). Of course if the FoV weighting function includes intermediate directions, not coincident with the GSs, then the two approaches are somewhat different. In this case, however, there might be some additional errors introduced by non-linearity in the WFSs, related to the fact that the loop would not be perfectly closed on the GSs. This topic is beyond the limits of this paper.

The simplified analysis presented here will be soon completed by an extensive simulation work, now in progress. The final most important issue will be to determine the sky coverage achievable with the use of solely NGSs.

References

Beckers, J. M. 1989, in SPIE Proc., 1114, 215

Beckers, J. M. 1993, ARA&A, 31, 13

Conan, J. M., Fusco, T., Mugnier, L. M., et al. 2000, SPIE, 4007, 913

Ellerbroek, B. L. 1994, JOSA A, 11, 783

Ellerbroek, B. L., & Rigaut, F. 2000, SPIE, 4007, 1088

Foy, R., & Labeyrie, A. 1985, A&A, 152, L29

Fusco, T., Conan, J. M., Michau, V., et al. 1999, Opt. Lett., 24, 1472

Fusco, T., Conan, J.-M., Mugnier, L., Michau, V., & Rousset, G. 2000, A&AS, 142, 149

Happer, W., MacDonald, G. J., Max, C. E., & Dyson, F. J. 1994, JOSA–A, 11, 263

Johnston, D. C., & Welsh, B. M. 1994, JOSA A, 11, 395

Phillips, C. L., & Troy Nagle, H. 1984, Digital control system analysis and design (Prentice Hall)

Ragazzoni, R. 2000, ESO Conf. Workshop Proc., 57, 175

Ragazzoni, R., Marchetti, E., & Rigaut, F. 1999, A&A, 342, L53

Ragazzoni, R., Marchetti, E., & Valente, G. 2000a, Nature, 403, 54

Ragazzoni, R., Farinato, J., & Marchetti, E. 2000b, in SPIE Proc., 4007, 1076

Roddier, F., Northcott, M. J., Graves, J. E., McKenna, D. L., & Roddier, D. 1993 JOSA–A, 10, 957

Simon, M., Close, L. M., & Beck, T. L. 1999, AJ, 117, 1375

Tallon, M., & Foy, R. 1990, A&A, 235, 549

Wilson, R. W., & Jenkins, C. R. 1996, MNRAS, 268, 39