

Achromatic interfero coronagraphy

I. Theoretical capabilities for ground-based observations

P. Baudoz¹, Y. Rabbia², and J. Gay¹

¹ Observatoire de la Côte d'Azur, Dpt. Fresnel, UMR-CNRS 6528, BP. 4229, 06304 Nice Cedex 04, France

² Observatoire de la Côte d'Azur, Dpt. Fresnel, UMR-CNRS 6528, Av. Copernic, 06130 Grasse, France

Received March 8; accepted September 29, 1999

Abstract. We present a concept dedicated to the detection of faint companions or other morphological components in the neighbourhood of a star by means of coronagraphy. Basically this method relies on destructive interferences so as to achieve extinction of the on-axis point-like source. In this regard the reported concept is a nulling interferometer designed for work with a single aperture.

Two features are specific to our coronagraph, namely achromaticity and close-sensing. Achromaticity allows flexible choice of a working wavelength and of a large spectral bandwidth. Close-sensing provides the ability to explore around the central source significantly closer than can be achieved by existing coronagraphs.

Though our concept has been initially devised for use on a space-based instrument it is the goal of this paper to show by theoretical analysis that it can be efficient and powerful on ground-based large telescopes equipped with adaptive optics. This especially regards close-sensing capabilities at a level (better than diffraction limit) unreachable by other coronagraphs. The essential limitation regarding detection originates in residual wavefront distortions whose effects are described here using the formalism of adaptive optics.

In this paper the principle, the generic set-up and the limitations are briefly recalled. Algebraic derivations regarding the effect of atmospheric seeing and the use of adaptive optics are given. Theoretical expected detection capabilities for ground-based operation are derived, showing that detection of a companion fainter than the central star by 6 magnitudes with a Signal to Noise Ratio (SNR) of 5 appears to be a reasonable goal from raw data, with a 4-meter class telescope in rather ordinary conditions and with modest integration times. On-going progress in adaptive optics, make it conceivable to reach magnitude differences of 12 under similar conditions.

Send offprint requests to: P. Baudoz,
e-mail: baudoz@obs-nice.fr

Key words: instrumentation: interferometers; atmospheric effects — stars: brown dwarfs; binaries: close

1. Introduction

The study of stellar environment offers a number of exciting topics for which detection and imaging of faint morphological features close to an unresolved source are concerned. This goal relates to both compact features such as faint companions or planets and diluted features such as circumstellar envelopes, protoplanetary disks, jets and other structures. In this paper we consider instead detection of faint companions.

Besides nulling techniques applied to diluted apertures (Bracewell 1978; Woolf & Angel 1997; Mennesson & Marriotti 1997) the use of coronagraphy with compact apertures nowadays concerns the classical Lyot Coronagraphy (Lyot 1930) adapted to stellar studies (Malbet 1996; Beuzit et al. 1997), the Achromatic Interfero Coronagraphy (Gay & Rabbia 1996), and the Phase Mask Coronagraphy (Roddier & Roddier 1997).

The classical Lyot coronagraphy uses an opaque mask located at an intermediate image plane, so as to obscure the center of the field. This technique does not allow a very close-sensing of the star environment since the central area of the field is masked up to a number of Airy rings (say 3 to 10) and provides sensing capabilities ranging between 0.3 and 2 arcsec from the star (Beuzit et al. 1997; Nakajima et al. 1995). Also, removal of the light from the star is not complete because diffraction features in the image plane cannot be totally eliminated even in perfect optical conditions (in spite of using a Lyot stop in the next pupil plane). On the contrary, in similar conditions our coronagraph (simply by its principle) is free from such diffraction features.

In the phase mask approach the opaque mask is replaced by a transparent and dephasing one and should allow the sensing of the source's environment as close as the first Airy dark ring. However, even in perfect optical conditions (as for the Lyot approach) its extinction capability suffers from spurious diffraction features. Moreover extinction is intrinsically chromatic which limits the working spectral bandwidth.

Our coronagraph is likely to be a promising concept since it is able to look very close to the star (as close as a fraction of the first Airy ring), avoids the use of a Lyot stop (which would reduce the effective aperture diameter) and can work at a large spectral bandwidth, which tends to increase detection capabilities and allows spectroscopic study of faint companions.

In Sect. 2 we recall the principle of the concept and the associated generic set-up as well as specific features and inherent limitations in the ideal case. In Sect. 3 departures from the ideal case are considered through a theoretical algebraic derivation. In Sect. 4, expected detection capabilities for ground-based operation are evaluated.

2. Principle and specific features of the concept

The principle has been described elsewhere (Gay & Rabbia 1996; Gay et al. 1997) and is recalled here for the reader's convenience.

Basically our coronagraph is a Michelson interferometer modified by inserting in one arm an achromatic π -dephasing and a pupil π -rotation in its plane. Both effects are obtained by replacing in one arm the flat mirror by a cat's eye. The coronagraph is placed in a collimated beam from the telescope and the beams from the two arms are recombined at a common pupil plane where a focusing element forms an image on a camera. The cat's eye is designed so as to match the pupil planes for each of the two arms. When the source is point-like and set on-axis, our coronagraph delivers a point spread function (PSF in the following) whose maximal intensity is weighted at zero when the Optical Path Difference (OPD in the following) is made zero, so that image plane is then utterly dark. This nulling would affect the same companion, but because of the pupil rotation the light from the companion is unaffected by the destructive interference process. In that case, the off-axis source yields two twin-images (half power each) in the focal plane while the image of the "blinding" central on-axis source has been removed thanks to the nulling process. In Fig. 1, the generic set-up is schematically described and in Fig. 2 the nulling process is described in terms of wavefronts and complex amplitude of the fields. Actually a real set-up would use instead separated beams in each arm with symmetrical double pass of the beamsplitter.

Achromaticity of the dephasing results from the crossing of an additional focus. This property is described

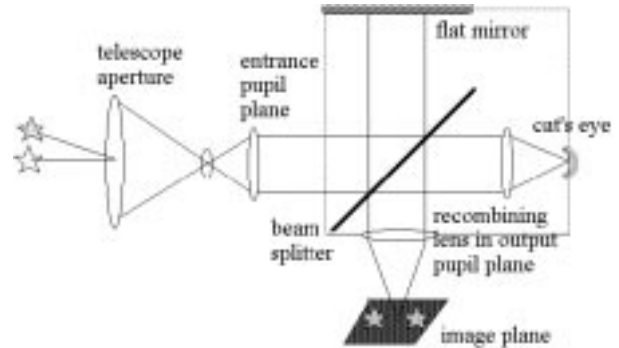


Fig. 1. Schematic of the Generic Set-up of our coronagraph

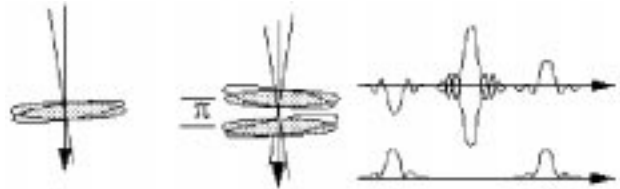


Fig. 2. Left: collected wavefronts, one from the central source the other (tilted) from a companion. Center: wavefronts on the recombining lens. Right: amplitudes and resulting intensity in image plane

in various references (Born & Wolf 1970; Gouy 1890; Boyd 1980; Baudoz et al. 1997).

A point to stress is that the interferometric process does not result in a "dark zone" at the center of image plane, but rather suppresses in the whole image plane all light coming from an on-axis unresolved source. In other words, the system provides the extinction of the central source contribution by weighting it down to zero when OPD is made zero.

Achromaticity is a specific feature of our coronagraph and makes it quite flexible regarding the choice of the working wavelength and of the spectral bandwidth as well. Also it makes it easy to work with two adjacent spectral bands, each providing images on opposite quadrants of a detector. An optimization of the optical coatings (beamsplitter and mirrors) is recommended when going from visible domain to infrared domain so as to keep performance at its best.

The other specific feature of our coronagraph relates to close-sensing capabilities since it allows sensing as close as a fraction (0.5 to 0.3) of the Airy radius, which is better than the diffraction limit set by the aperture and unreached by other coronagraphs. For those interesting features to be fully efficient, two conditions specific to our device must be satisfied:

1. the complex amplitude of the interfering fields must remain fully centro-symmetric up to the recombination stage (to keep effectively "flat" the darkened flat-tint in the recombining pupil plane);

2. the OPD between the two arms of the interferometer must be kept as close to zero as possible (to achieve weighting of PSF at zero).

Few equations are given here since they will be used later in this paper. In image planes, vector coordinates are referred to as α, β or μ (meaning x/F , where x is the linear vector coordinate and F the focal length of the output focusing component), while in pupil planes (aperture and exit pupil) vector coordinates are noted ξ , meaning length/ λ . Since we use circular apertures, polar coordinates will be used: (q, ϕ) in pupil planes and (ρ, θ) in image planes. Notation \hat{X} stands for the Fourier Transform of the distribution X (FT in the following). The complex amplitude of the incoming field is noted $\psi(\xi)$ and the complex transmission of the collecting pupil is noted $P(\xi)$ up to the recombining pupil plane at exit. OPD between the interfering fields is noted “ d ”. The beamsplitter reflexion factor and transmission factor for energy respectively are R and T (typically 0.5 each). The amplitude in the recombining pupil plane is:

$$Q_d(\xi) = \sqrt{R.T.} \cdot \left[\psi(\xi).P(\xi) + \exp(i.\pi).\psi(-\xi).P(-\xi).\exp\left(i.2\pi.\frac{d}{\lambda}\right) \right]. \quad (1)$$

Unless specified, we will work henceforth with OPD at zero.

The complex transmission in pupil planes is given by $P(\xi) = \Pi(\frac{\xi}{D}).\exp(i.\Phi(\xi))$ where $\Pi(\frac{\xi}{D})$ describes the transmission of the perfect pupil of diameter D (Π is unity within the aperture and zero outside). The phase term $\Phi(\xi)$ takes into account all phase corrugations arising up to the recombining pupil plane, which includes corrugations from permanent aberrations met in each arm.

Let us consider a couple of point-like sources (binary star). One component is on-axis (intensity Ω_0), the other (intensity Ω) is off-axis by the angle μ , the intensity distribution is described by

$$\Omega(\alpha, \mu) = \Omega_0.\delta(\alpha) + \Omega.\delta(\alpha - \mu) \quad (2)$$

and the incoming field is expressed as:

$$\psi(\xi, \mu) = \omega_0 + \omega.\exp(-i.2\pi.\xi.\mu) \quad (3)$$

with $\langle |\omega_0|^2 \rangle = \Omega_0$ and $\langle |\omega|^2 \rangle = \Omega$.

The intensity $I(\beta, \mu)$ we have (Re means “real part”):

$$I(\beta, \mu) = R.T.\Omega.[A(\beta - \mu) + A(\beta + \mu) - 2.Re(\hat{P}(\beta - \mu).\hat{P}^*(\beta + \mu))] \quad (4)$$

where $A(\beta)$ is the intensity Point Spread Function normalized to unity, given by the square modulus of \hat{P} and where * denotes complex conjugate, notation $\langle \rangle$ means statistical average.

The centro-symmetry condition $P(\xi) = P(-\xi)$ has totally removed the contribution Ω_0 of the on-axis component. We find two twin-images ($A(\beta - \mu)$ and $A(\beta + \mu)$) of the off-axis source (displayed symmetrically with respect

to the origin of coordinates) and a mixed term containing the product of complex amplitude distributions, one for each interferometer arm.

As soon as the off-axis angle μ is large enough, the amplitude distributions do not overlap and the mixed term cancels. This fact raises the question: how can we decide whether the separation is large enough to cancel the mixed term? In other words: what is the response of the coronagraph versus field coordinate?

This response, the extinction profile versus field coordinate is obtained by integrating over the image plane the energy from an off-axis running point-like source. We introduce $w_0 = 4.R.T.\Omega.S$, which is the energy collected when the coronagraphic effect does not apply. Then, the expression of the integrated intensity normalised to unity at origin is a circular distribution whose radial profile is given by:

$$w(\rho) = \frac{w_0}{2} \cdot \left[1 - \frac{2.J_1(2.\pi.D.\rho)}{2.\pi.D.\rho} \right] \quad (5)$$

where J_1 is the Bessel function of order 1 and D is the diameter of the telescope. Recalling that $\frac{2.J_1(2.\pi.D.\rho)}{2.\pi.D.\rho}$ is the radial profile of the Airy amplitude distribution for a telescope of diameter $2.D$, we see that the extinction profile corresponds to the diffraction limit of a telescope which would have twice the diameter of the one effectively used. In other words, a point-like source located as close as a fraction of the first dark ring of the effective angular Airy pattern can be considered as off-axis and thus can be detected. This is illustrated in Fig. 3, where for comparison the Airy energy profile for a telescope of diameter D also appears.

This extinction profile $w(\rho)$, that we consider as an “extinction lobe as seen from the sky”, is narrower than the Airy lobe of the telescope, from here originates the close-sensing capabilities of our coronagraph, thus enabling it to go beyond the diffraction limit. Let us note that the response remains practically stable once the companion stands outside the “sky hole”.

Even in an ideal situation, limitations apply to the extinction efficiency:

1. The light from a source partially resolved by the telescope is not utterly removed from the image plane, because the extinction profile is zero only at the central point of the field (part of the source covers a non-utterly-darkened field area);
2. When observing from the ground, the radiation travelling along the axis is spread by atmospheric differential dispersion and thus prevents a point-like source to be seen as such (part of the radiation seems to come off-axis). Let us note that this limitation affects any coronagraphic device.

For a partially resolved source, described here by $\Omega(\rho) = \Omega.\Pi(\frac{\rho}{\Theta})$ we find the residual energy by integrating $w(\rho)$ (Eq. 5) up to the angular radius $\Theta/2$ of the source. Setting

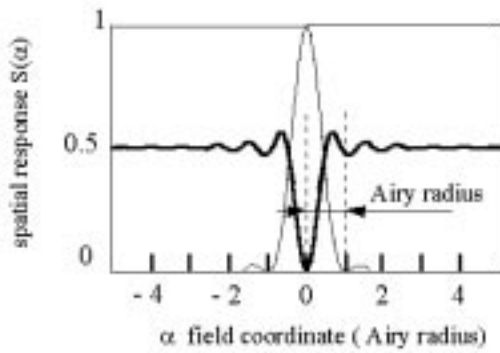


Fig. 3. Radial profile of the response of our coronagraph (response normalized to w_0) with respect to the sky coordinate, counted in Airy radius $1.22\frac{\lambda}{D}$. The normalized Airy profile of the aperture is given for comparison. As close as a fraction of the first zero of the Airy profile of the aperture, the response of the coronagraph has reached its maximum value. At $\lambda = 2.2 \mu\text{m}$ with an aperture diameter of 3.6 m this correspond to close-sensing down to 0.07 arcsec. Closer sensing is possible depending on detection capabilities in terms of the Signal to Noise Ratio

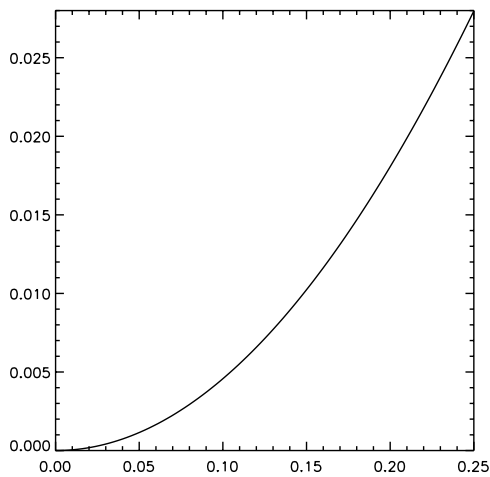


Fig. 4. Integrated residual energy (normalised on the available energy) versus the angular diameter of the partially resolved source (expressed in fraction of $1.22\frac{\lambda}{D}$)

$y = \pi \cdot D \cdot \Theta$ the expression of the residual integrated energy in image plane is:

$$w(y) = \frac{w_0}{2} \cdot \left[1 - \frac{1 - J_0(y)}{\frac{y^2}{4}} \right]. \quad (6)$$

Figure 4 shows the behaviour of w versus the angular diameter Θ counted in Airy angular radius. For example, in an ideal situation, observing with a 2.4 meter telescope at $2.2 \mu\text{m}$ a “large” source with $\Theta = 22 \text{ m arcsec}$ (a tenth of the Airy radius), would lead to a residual integrated energy which amounts to $4.5 \cdot 10^{-3}$ of the total collected energy.

Because of the chromatism of the air optical index, the image of a point-like source is slightly dispersed. Therefore, working with a finite bandwidth, the extinction can be complete only for the central wavelength while the other wavelengths yield point-like sources apparently off-axis. Thus we find a finite residual energy in the image plane, according to the field response of the coronagraph given in Eq. (5). This effect depends on the zenithal distance z of the source, the working central wavelength λ_0 (or the wavenumber $\sigma_0 = \lambda_0^{-1}$) and the spectral bandwidth $\Delta\lambda$ (or $\Delta\sigma$). Let us consider a point-like source at a true zenithal distance z , atmospheric refraction shifts the image by an angle dz whose expression is:

$$dz(\sigma, z) = (n(\sigma) - 1) \cdot \tan(z). \quad (7)$$

Assuming that the source is correctly set on-axis for σ_0 , its image at $h = \sigma - \sigma_0$ will be shifted off-axis by the differential angle $\beta(h, z) = dz(\sigma, z) - dz(\sigma_0, z)$, and some unwanted energy appears in the image plane, according to the spatial response of the coronagraph. We define a residual energy factor $g(h, z)$:

$$g(h, z) = \frac{1}{2} - \frac{1}{2} \cdot \frac{2 \cdot J_1(2 \cdot \pi \cdot D \cdot \sigma_0 \cdot \beta(h, z))}{2 \cdot \pi \cdot D \cdot \sigma_0 \cdot \beta(h, z)}. \quad (8)$$

Again we have assumed a uniform spectral distribution for the source. From Allen (1973) we have $n(\sigma) = 1 + A + B \cdot \sigma^2 + \dots$ where $A = 3 \cdot 10^{-4}$ and $B = 1.6 \cdot 10^{-6} \mu\text{m}^2$, so that we have $\beta(h, z) \approx 2 \cdot B \cdot h \cdot \sigma_0$. The residual energy factor G at a given z is the integral of $g(h, z)$ over the whole bandwidth. Using $y = \frac{h}{\Delta\sigma}$ and $m = 4 \cdot \pi \cdot D \cdot \sigma_0^2 \cdot B \cdot \Delta\sigma \cdot \tan(z)$, we have:

$$G = \frac{1}{2} - \frac{1}{2} \cdot \int_{-\frac{1}{2}}^{+\frac{1}{2}} \frac{2 \cdot J_1(m \cdot y)}{m \cdot y} \cdot dy. \quad (9)$$

Figure 5 shows curves of the residual energy factor G versus z . They refer to observations at $\lambda_0 = 2.2 \mu\text{m}$ with $\Delta\lambda = 0.4 \mu\text{m}$ with telescope apertures of 1.5 m, 3.6 m and 8 m respectively.

This effect becomes more severe when the wavelength decreases, when the bandwidth increases and when the aperture is enlarged. Note that this energy appears spread along the direction of dispersion only. This effect is deterministic and can be removed by observing a suitable reference star. Another way to remove this effect is rely on an Atmospheric Dispersion Compensator (using Risley prisms) in order to balance the atmospheric dispersion at a given z (Breckinridge et al. 1979). It must be noted that this kind of limitation applies to any ground-based coronagraph.

3. Departing from the ideal case

Departure from the ideal case occurs in two scenarios: OPD is not zero and complex transmission up to the recombination pupil plane is not centro-symmetric (which relates to both modulus and phase of the complex transmission function).

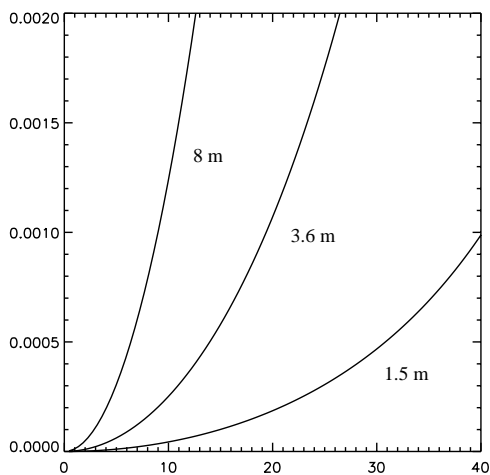


Fig. 5. Integrated residual energy $G(z)$ in the image plane caused by atmospheric differential dispersion at zenithal distance z for apertures of diameter 1.5 m, 3.6 m and 8 m. The full bandwidth of K-band is used ($\lambda_0 = 2.2 \mu\text{m}$, $\Delta\lambda = 0.4 \mu\text{m}$)

3.1. Departing from zero OPD

Since the effect of non-zero OPD is a non-zero weighting of the total energy from the on-axis source, we readily have a specification of the tolerance regarding OPD. Using w for the residual energy and w_0 for the energy available when no extinction occurs, we have, using Eq. (1) and the centro-symmetry condition:

$$w = \frac{w_0}{2} \cdot \left[1 - \cos \left(2\pi \cdot \frac{d}{\lambda} \right) \right]. \quad (10)$$

Selecting a given maximum value M for $\frac{w}{w_0}$ we have to control OPD so as to keep $\frac{d}{\lambda} < \frac{\sqrt{M}}{\pi}$.

A value of $M = 10^{-4}$ maintains $\frac{d}{\lambda} < \frac{1}{400}$. This precision implies that servo-loop of OPD could be performed. A rms precision of $\frac{1}{200}$ is currently achieved in the visible (Reynaud & Delaire 1993) and corresponds to better than $\frac{1}{600}$ at $2.2 \mu\text{m}$. In addition, monitoring OPD variations during the recording of images with suitable exposure-time allows a selection of “best exposures” in terms of lower values of M .

3.2. Departing from centro symmetry

In this situation the condition $P(\xi) = P(-\xi)$ no longer applies. Therefore complex amplitudes at the recombination step no longer mutually destroy (Eq. 1), and some energy from the central source is able to reach the image plane.

Departure related to modulus of the complex transmission relates to essentially obscuration and vignetting in the pupil plane (telescope spider for example), and it is ruled out by using a suitable mask in an intermediate

pupil plane, restoring a centro-symmetric distribution for the modulus.

Departure related to phase of the complex transmission results from phase distortions accumulated up to the recombination pupil plane. Phase distribution conveys both deterministic and stable distortions (aberrations, misadjustment, etc.) and rapidly changing random distortions (atmospheric turbulence). It must be noted that phase distortions described by an even function do not prevent the removal of the on-axis source. This means for example that a defocus and aberrations such as spherical aberration, astigmatism and chromatic aberration still allow the extinction process to work fully. However, they tend to degrade images of the companion, which decreases detection capabilities. On the contrary, such “odd” aberrations as coma and distortion lead to an incomplete destructive interference for the on-axis source. This problem also concerns aberrations which could be different in each arm of the device. Indeed, aberrations and misadjustments can be reduced to an acceptable level by proper optical design and control. Thus we will limit our analysis to the case of turbulence, which for ground-based operation is much more severe than the residual instrumental aberrations. This analysis is based on the description of the phase distortions over the aperture in terms of a weighted sum of Zernike polynomials (Born & Wolf 1970; Noll 1976; Roddier 1990).

Any “odd” distortion generates a spurious halo, a residual volcano-shaped light distribution around the center of image plane. Extension of this halo remains roughly the one of the focal patch. Note that the trouble is not the halo itself but fluctuations of its shape, whether they result from photon noise or from instabilities of the intensity distribution (speckle noise), this latter being largely dominant for ground-based observations. Noise is not uniformly distributed in the focal plane, since it depends on the local illumination level within the halo.

We consider the case of an on-axis point-like source, with intensity Ω_0 leading to the incident field $\psi(\xi) = \omega_0$ with $\langle |\omega_0|^2 \rangle = \Omega_0$. Our starting point is the following Equation describing the complex amplitude in recombining pupil plane with OPD set to zero:

$$Q(\xi) = \sqrt{R.T} \cdot \Pi \left(\frac{\xi}{D} \right) \cdot \omega_0 \cdot [\exp(i \cdot \Phi(\xi)) - \exp(i \cdot \Phi(-\xi))] \quad (11)$$

where $\Phi(\xi)$ describes the phase distortions in the recombination pupil plane. Those turbulence-induced distortions are random and gaussian (central limit theorem). To evaluate the unwanted integrated residual energy we have to work with $E(\xi) = \langle |\exp(i \cdot \Phi(\xi)) - \exp(i \cdot \Phi(-\xi))|^2 \rangle$. Thanks to the Gaussian statistics and using the phase structure function $D_\Phi(\xi) = \langle |\Phi(\mu) - \Phi(\mu + \xi)|^2 \rangle$

(Roddiier 1981), the normalized integrated energy versus the parameter the parameter $p = D/r_0$ is given by:

$$w(p) = \frac{w_0}{2} \cdot \left\{ 1 - 2 \cdot \int_0^1 \exp[-3.44(p \cdot y)^{5/3}] y dy \right\} \quad (12)$$

with $w_0 = 4.R.T.\Omega_0.S$ as stated previously (Eq. 5).

This expression shows that the use of real time compensation of wavefront distortions is mandatory, since even for $p = 1$ we have a residual energy larger than $0.3 w_0$.

3.2.1. Use of real-time correction of phase distortions

In the following analysis we consider the use of real-time wavefront compensation by adaptive optics for both the integrated residual energy and the residual energy distribution. The first provides only a useful numerical figure to summarize the performance relating to extinction and serves to compare various situations. The second is needed to evaluate the Signal to Noise Ratio (SNR) pertaining to the detection of a companion, since this SNR depends on the location of the companion in the image plane via the level of residual light at this location. So, we need to express analytically the distribution in the image plane and not only its integral. A convenient way to express analytically this distribution is to use the small phase distortions hypothesis, leading to $\exp(i \cdot \Phi(\xi)) \approx 1 + i \cdot \Phi(\xi)$. The validity of this assumption will be checked later, but initially, we can refer to the fact that the largest part of the phase distortion effects comes from the tip-tilt of the wavefront (Roddiier 1990). Since this defect is the first which is compensated for in any adaptive optics device, we believe our assumption is relevant.

We first describe the amplitude at the recombination step in the pupil plane, then we evaluate the integrated residual energy and finally we express the residual light distribution in the image plane.

To describe the corrected amplitude at the recombination pupil plane, our starting point is Eq. (11) but since we consider real-time correction of phase distortions we now deal with a residual phase distribution $\Phi_r(\xi) = \Phi(\xi) - \Phi_c(\xi)$ where $\Phi(\xi)$ is the phase function and $\Phi_c(\xi)$ represents the applied corrections. Assuming that Φ_r is small leads to a new expression for the amplitude $Q(\xi)$:

$$Q(\xi) = \omega_0 \cdot \sqrt{R.T} \cdot i \cdot \Pi\left(\frac{\xi}{D}\right) \cdot [\Phi_r(\xi) - \Phi_r(-\xi)] \quad (13)$$

showing that only the odd part of $\Phi_r(\xi)$ is to be taken into account. The even part is automatically corrected “by design”. Introducing Φ_{or} for odd residual phase distortions over the pupil of diameter D (the expression of the amplitude transmission $\Pi(\frac{\xi}{D})$ is now absorbed in Φ_{or}) we write:

$$Q(\xi) = \omega_0 \cdot 2 \cdot \sqrt{R.T} \cdot i \cdot \Phi_{or}(\xi). \quad (14)$$

Thus, describing $Q(\xi)$ we still need to describe $\Phi_{or}(\xi)$. It is usual to describe a phase distribution $\Phi(\xi)$ over a pupil

of radius R_p by a weighted sum of Zernike polynomials (Noll 1976; Roddiier 1990; Conan 1995)

$$\Phi(q, \phi) = \sum_j a_j Z_j\left(\frac{q}{R_p}, \phi\right). \quad (15)$$

Here we use the notation introduced by Noll, where j is a synthetic index based on classical indexes n and k (Born & Wolf 1970) undergoing usual restrictions (n integer, $|k| \leq n$, $n - |k| = \text{even}$):

$$\begin{aligned} Z_{\text{even}j} &= \sqrt{n+1} \cdot R_{nk}(q) \cdot \sqrt{2} \cdot \cos(k \cdot \phi) & k \neq 0 \\ Z_{\text{odd}j} &= \sqrt{n+1} \cdot R_{nk}(q) \cdot \sqrt{2} \cdot \sin(k \cdot \phi) & k \neq 0 \\ Z_j &= \sqrt{n+1} \cdot R_{n0}(q) & k = 0 \end{aligned}$$

the R_{nk} are radial polynomials, defined for $0 \leq q \leq 1$, by the relation:

$$R_{nk}(q) = \sum_{s=0}^{(n-k)/2} \frac{(-1)^s (n-s)!}{s! [(n+k)/2 - s]! [(n-k)/2 - s]!} q^{n-2s}. \quad (16)$$

Using the pupil distribution $W(\frac{q}{R_p})$ giving $\frac{1}{\pi}$ when $q \leq R_p$ and else zero, the following orthogonality relation holds (where we use $R_p = D/2$ and the Kronecker symbol $\delta_{jj'}$)

$$\int_0^{2\pi} \int_0^{R_p} Z_j\left(\frac{q}{R_p}, \phi\right) Z_{j'}\left(\frac{q}{R_p}, \phi\right) q dq d\phi = \delta_{jj'}. \quad (17)$$

For a given phase distribution Φ , the weighting coefficients a_j are retrieved by the relation:

$$a_j = \frac{1}{R_p^2} \cdot \int_0^{2\pi} \int_0^{R_p} \Phi(q, \phi) Z_j\left(\frac{q}{R_p}, \phi\right) q dq d\phi. \quad (18)$$

Applying corrections up to $j = J$, leads to the residual phase distribution:

$$\Phi_r(q, \phi) = \Phi(q, \phi) - \sum_{j=1}^J a_j Z_j\left(\frac{q}{R_p}, \phi\right). \quad (19)$$

This expression is equivalent to a sum of the $a_j \cdot Z_j$ running from $j = J + 1$ up to j infinite. In our concept, only the odd part of the residual phase distribution is to be counted, so that in Φ_{or} (Eq. 15) some j 's are taken out of this infinite summation.

Let us point out that restricting the summation to the odd part of Φ_r does not mean that we restrict the sum to odd values of j , but rather that we restrict it to odd values of n , the radial polynomial order (and therefore to odd values of k). This, for example, automatically eliminates such defects as defocus, 3rd order astigmatism, 3rd order spherical, etc. as already mentioned.

If we want to take into account the case of incomplete correction we write:

$$\Phi_{or}(q, \phi) = \sum_j p_j a_j Z_j\left(\frac{q}{R_p}, \phi\right) \quad (20)$$

where we have introduced the weighting parameter p_j (presence of the defect of index j) to which is assigned the value 0 when correction is fully effective for j , the value 1 when order is not corrected and an intermediate value when correction is not fully effective. For values of j corresponding to even phase defects, p_j automatically has the value 0.

3.2.2. Integrated residual energy under adaptive optics correction

When correction is performed up to index J , the residual energy w_J in image plane can also be found by integrating the intensity in pupil plane:

$$w_J = \int \langle |Q(q, \phi)|^2 \rangle q dq d\phi. \quad (21)$$

Let us disregard for a while the automatic cancellation of even phase distortions, and let us introduce a coefficient ε_j (effectiveness of the correction) defined by $\varepsilon_j^2 = 1 - p_j^2$. Taking into account the weighted Zernike sum and the orthogonality relation we have:

$$w_J = 4.R.T.\Omega_0.\pi.R_p^2. \left[\langle \Phi^2 \rangle - \sum_j \varepsilon_j^2 \langle |a_j|^2 \rangle \right]. \quad (22)$$

For a correction device working up to J , we have a residual energy g_J normalized on the total energy collected on the star (given by $w_0 = 4.R.T.\Omega_0.\pi.R_p^2$) so we have:

$$g_J = \langle \Phi^2 \rangle - \sum_j \varepsilon_j^2 \langle |a_j|^2 \rangle. \quad (23)$$

This expression is quite similar to the one derived by Noll (1976) for the residual variance Δ_J of the phase over the pupil, when correction is achieved up to (and including) $j = J$. The departure from Noll's coefficients comes from the energy removal efficiency ε_j^2 applied to each term of the sum. Correcting up to the radial order N corresponds to a correction up to $J = \frac{(N+1).(N+2)}{2}$. The value $N = 7$ has been effectively reached on a real device (Conan 1995) which corresponds to $J = 36$ that we adopt to evaluate a realistic residual energy. For J greater than 10 ($N > 3$) the residual variance can be approximated by $\Delta_J = 0.2944.J^{-a}.\left(\frac{D}{r_0}\right)^{5/3}$ with $a = \frac{\sqrt{3}}{2}$ (Noll 1976) that we can use to evaluate the residual energy g_J resulting from complete corrections up to the mode J . Since distortions with even radial orders are automatically corrected, the g_J 's are always less than the Noll's Δ_J 's when applied corrections are fully effective as shown in Fig. 6. For a complete correction with $J = 36$ we have $g_{36} \approx 0.006\left(\frac{D}{r_0}\right)^{5/3}$.

So, for a 150 cm telescope diameter and a r_0 of 10 cm (leading to 53 cm at 2.2 μm) we have $g_{36} \approx 0.034$ at 2.2 μm , which corresponds to a rejection by 3.67 magnitudes of the unwanted energy, integrated over the whole image plane.

In the more realistic case of incompletely achieved corrections, we use values of ε_j^2 taken from numerical

Table 1. Correction efficiencies used for radial orders up to $n = 7$ (that is $j = 36$). Even orders remain fully corrected beyond orders displayed in the table

j	1	2, 3	4, 5, 6	7-10	11-15	16-21	22-28	29-36
n	0	1	2	3	4	5	6	7
ε_j^2	1	0.995	1	0.9	1	0.7	1	0.6

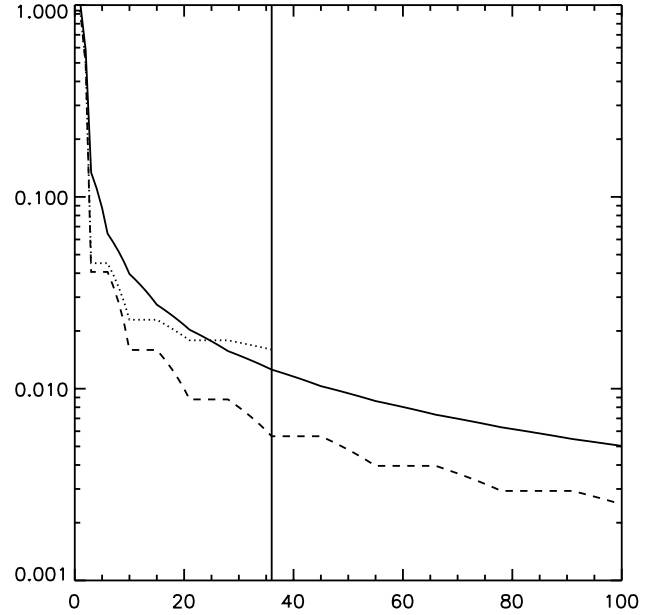


Fig. 6. Noll's residual variances (solid line) and normalised integrated residual energy in the image plane of the coronagraph versus J , the index of the upper corrected mode. Dotted line: incomplete correction. Dashed line: complete correction. This curve constantly runs below the Noll's variances, due to the automatic correction of even modes

simulations based on measured performance of a real servo-correction device (Conan 1995). Values for ε_j^2 are shown in Table 1 (for even orders n , efficiencies ε_j^2 are automatically set to unity, they are not shown beyond $n = 7$ that is $j = 36$)

In this case, we define a figure g_J^u similar to g_J (u for incomplete) and we have $g_{36}^u \approx 0.016\left(\frac{D}{r_0}\right)^{5/3}$. In the same situation as above we have now an integrated rejection of roughly 2.6 magnitudes.

It must be kept in mind that these numerical figures pertain to the total unwanted energy in the image plane, not to the one found at the location of an image of companion. So they must not be interpreted as the limiting magnitude difference that our coronagraph is able to work with. A more relevant evaluation of detection capabilities needs the energy distribution in the image plane to be expressed.

3.2.3. Energy distribution in the image plane

It is more convenient here to come back to the expression $Q(\xi) = \omega_0 \cdot \sqrt{R \cdot T} \cdot 2 \cdot i \cdot \sum_j p_j a_j Z_j(\xi)$ where p_j traces the presence of the defect associated with j and where we have an infinite sum.

The residual energy distribution I_{halo} in the image plane (polar coordinates ρ and θ) is given by the square modulus of the FT of $Q(q, \phi)$ and we have (using \hat{Z} for the FT of Z):

$$I_{\text{halo}}(\rho, \theta) = 4 \cdot RT \cdot \Omega_0 \cdot \pi^2 \cdot R_p^4 \cdot \sum_{j;j'} p_j p_{j'} \langle a_j \cdot a_{j'}^* \rangle \cdot \hat{Z}_j(R_p \cdot \rho, \theta) \hat{Z}_{j'}^*(R_p \cdot \rho, \theta) \quad (24)$$

and the residual energy normalized on the collected energy is:

$$g_{\text{halo}}(\rho, \theta) = \sum_{j;j'} p_j p_{j'} \langle a_j \cdot a_{j'}^* \rangle \cdot \hat{Z}_j(R_p \cdot \rho, \theta) \hat{Z}_{j'}^*(R_p \cdot \rho, \theta) \quad (25)$$

which is more conveniently expressed in using n and k instead of j :

$$g_{\text{halo}}(\rho, \theta) = \sum_{nk;n'k'} p_n p_{n'} \langle a_{nk} \cdot a_{n'k'}^* \rangle \cdot \hat{Z}_{nk}(R_p \cdot \rho, \theta) \hat{Z}_{n'k'}^*(R_p \cdot \rho, \theta) \quad (26)$$

To evaluate g_{halo} we use for \hat{Z}_{nk} the expression (Born & Wolf 1970):

$$\hat{Z}_{nk}(R_p \cdot \rho, \theta) = \pi \cdot R_p^2 \cdot \sqrt{n+1} \cdot (-1)^{(n-k)/2} \cdot i^{-k} \cdot \frac{2 \cdot J_{n+1}(2\pi \cdot R_p \cdot \rho)}{2\pi \cdot R_p \cdot \rho} \cdot \exp(i \cdot k \cdot \theta) \quad (27)$$

and for $\langle a_{nk} \cdot a_{n'k'}^* \rangle$ according to Noll (1976) as above, we use:

$$\langle a_{nk} \cdot a_{n'k'}^* \rangle = \frac{0.046}{\pi \cdot 2^{5/3}} \cdot \left(\frac{D}{r_0} \right)^{5/3} \cdot \sqrt{n+1} \cdot \sqrt{n'+1} \cdot (-1)^{(n+n'-2k)/2} \cdot \delta_{kk'} \cdot \int \rho^{-14/3} J_{n+1}(2\pi \cdot \rho) J_{n'+1}(2\pi \cdot \rho) d\rho \quad (28)$$

where it is apparent that only remain terms with $k = k'$.

Figure 7 shows examples of radial profiles of the residual halo when full correction is applied up to radial orders 3, 5, 7, 9 successively (the higher the order the fainter the halo). Taking into account corrections beyond order 9 would induce fainter and fainter contributions of the halo, raising it further and further from the center of the field.

This behaviour is interesting not only for reducing the level of unwanted energy but also for cleaning progressively from the center, which improves the close-sensing capabilities as early as the lower order is corrected.

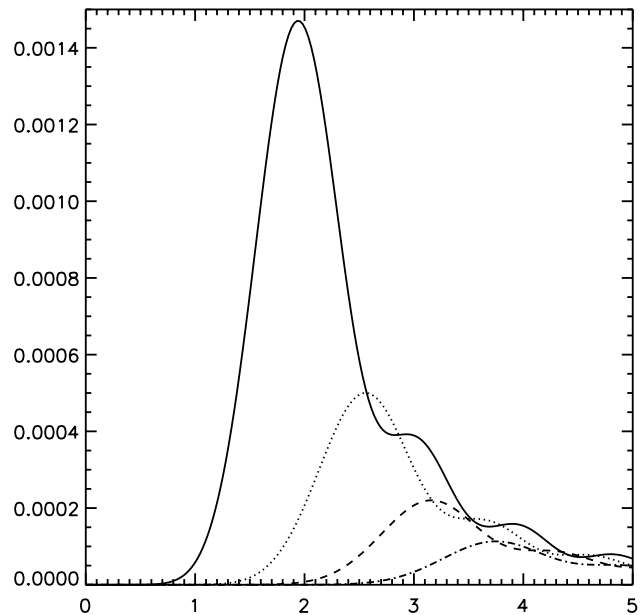


Fig. 7. Radial profiles of residual intensity distribution g_{halo} for $\frac{D}{r_0}(\text{vis}) = 10$ (with imaging in K) normalized on the maximum of the Airy pattern without coronagraph, in the cases of correction of radial orders n 's up to 3, 5, 7, 9. The higher the n , the lower the profile and the wider the central darkening. Horizontal scale: Angular distance from the center of the field with Airy radius as the unit

4. Expected detection capabilities

Detection capabilities primarily relate to the detection of a faint companion and are governed by Signal to Noise Ratio (SNR in the following).

We consider the case of a companion off-axis by β_0 with respect to an on-axis source, both sources being unresolved. We note No and Nc for the number of collected photo-events respectively for on-axis source (magnitude mo) and companion (magnitude $mc = mo + \Delta m$). Using η the quantum efficiency, t_{opt} the optical throughput, $(4 \cdot RT)$ the beamsplitting factor and $F_r(\lambda)$ the reference flux (zero magnitude at wavelength λ), S the collecting area, $\Delta\lambda$ the working spectral bandwidth centered at λ , and τ the exposure-time we have for a single frame:

$$No = [\eta \cdot 4RT \cdot t_{\text{opt}}] \cdot \frac{F_r(\lambda)}{h\nu} \cdot 10^{-0.4mo} \cdot \Delta\lambda \cdot S \cdot \tau \quad (29)$$

$$Nc = No \cdot 10^{-0.4\Delta m} \quad (30)$$

The coronagraph gives two twin-images of the companion. For a given β_0 , the shape of each image is roughly given by the central lobe of the normalized Airy distribution $A(\beta - \beta_0)$ weighted by the spatial response of the coronagraph $w(\beta) = w(\rho)$ with $\rho = |\beta|$ (Eq. 5).

Taking into account the two twin-images, the "signal" for M exposures is given by:

$$\text{signal}(\beta_0) = \frac{Nc}{4} \cdot \sqrt{2} \cdot \sqrt{M} \cdot \int_O w(\beta) A(\beta - \beta_0) d\beta \quad (31)$$

where $\int_{\mathcal{O}}$ means sum over the central lobe (that is a disk of radius $1.22\frac{\lambda}{D}$).

The noise originates in the fluctuations of the unwanted illumination in the image plane comprising residual light from incomplete extinction of the on-axis source (β -dependant) and the contribution of the background. Adding the detector noise contribution we have for a given pixel at β

$$\text{pixnoise}(\beta) = [\text{var}(no(\beta)) + \text{var}_{\text{bg}}(\beta) + \text{ron}^2]^{1/2} \quad (32)$$

where $\text{var}(no(\beta))$ and $\text{var}_{\text{bg}}(\beta)$ are the variances of the number of photo-events counted by the pixel at β , from the residual halo and from the background respectively. Detector contribution is represented by the variance ron^2 of the readout noise.

The background illumination is from a grey body of emissivity μ and temperature T seen within a solid angle ω_p by a pixel. The corresponding variance of background induced photoevents is:

$$\text{var}_{\text{bg}}(\beta) = \frac{\eta \cdot \mu \cdot L_{\lambda}(T) \cdot \Delta\lambda \cdot \omega_p \cdot s \cdot \tau}{h\nu} \quad (33)$$

where $L_{\lambda}(T)$ is the Planck's function and s the area covered by a pixel.

The noise from the residual illumination is a double stochastic process based on Poisson statistics and Rayleigh statistics. From the Mandel's formulae (Goodman 1985) the associated variance includes two terms: one reflects the Poisson noise at a given level of light $\langle no(\beta) \rangle$ the other reflects the fluctuations of this level (speckle noise). So we have:

$$\text{var}(no(\beta)) = \langle no(\beta) \rangle + [\langle no(\beta) \rangle]^2. \quad (34)$$

For ground based observations the speckle noise is the dominant one, except at a sufficiently low level of light.

The distribution $no(\beta)$ is given by: $no(\beta) = g(\beta) \cdot No$, where $g(\beta)$ describes the normalized residual energy (Eq. 26).

To evaluate the total noise we have to consider the sum of variances for the pixels covered by the image of the companion. We can only do that directly, if the fluctuations do not correlate from one pixel to another one, which is questionable considering the speckle noise contribution. Indeed, the exposure-time for a single frame is typically much larger than the coherence-time of the atmosphere, so that a correlation between adjacent pixels is likely to be destroyed. Moreover working at large bandwidth leads to an additional decorrelation.

Therefore the SNR for a companion off-axis by β_0 with respect to an unresolved on-axis source can be written as (β is a vector and b stands for $\text{var}_{\text{bg}}(\beta)$):

$$\text{SNR}(\beta_0) = \frac{Nc}{4} \cdot \frac{\sqrt{2} \cdot \sqrt{M} \cdot \int_{\mathcal{O}} w(\beta) A(\beta - \beta_0) d\beta}{[\int_{\mathcal{O}} [no(\beta) + no^2(\beta) + b + \text{ron}^2] d\beta]^{1/2}}. \quad (35)$$

The dependance on β_0 for the noise stands in the summation area, that is a disk of radius $a_0 = 1.22\frac{\lambda}{D}$ centered at

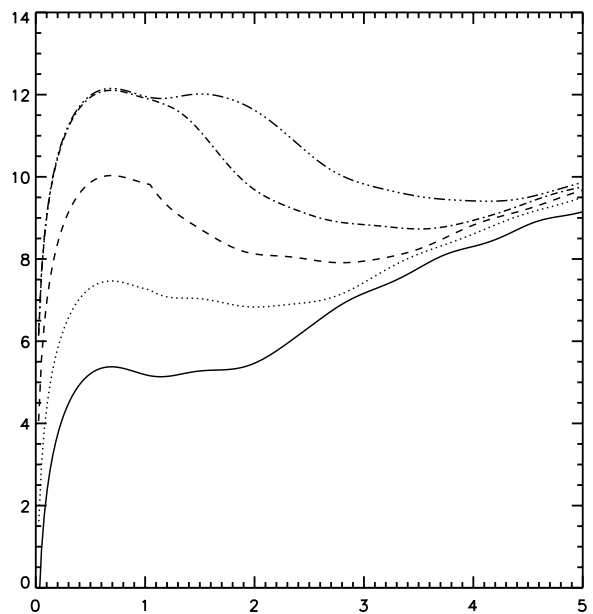


Fig. 8. Expected detectable magnitude difference Δm for the companion versus its angular distance to the on-axis source, in the case of COMPLETE CORRECTIONS (full efficiency) up to radial modes 1, 3, 5, 7, 9 (the higher the mode, the higher the Δm). $D = 12 \cdot r_0(\text{vis})$ (with imaging in K). Other values used for observation parameters are given in the text. Unit for angular distance is the Airy radius

β_0 . This drives us to insert into the integrands the weighting function (pil-box) $\Pi(\frac{\beta - \beta_0}{a_0})$.

Figure 8 and Fig. 9 show examples of expectable detection capabilities (at $RSB(\beta_0) = 5$ for $K_0 = 5$) in terms of detected magnitude differences Δm between on-axis source and companion, when adaptive optics provides real time corrections up to radial modes 1, 3, 5, 7, 9. Figure 8 and Fig. 9 refer to complete (fully efficient) and incomplete correction respectively. The following values of instrumental parameters corresponding to the PUEO adaptive optics device already in operation (Beuzit, private communication, PUEO internet site) and expected observation conditions for our coronagraph.

Working wavelength $\lambda = 2.2 \mu\text{m}$, bandwidth $\Delta\lambda = 0.5 \mu\text{m}$, reference flux $F_r(\lambda) = 3.9 \cdot 10^{-14} \text{ W cm}^{-2} \mu\text{m}^{-1}$, telescope diameter $D = 3.6 \text{ m}$, optical throughput $t_{\text{opt}} = 0.1$, exposure time (single frame) $\tau = 0.1 \text{ s}$, number of frames $M = 5000$, Fried's parameter $r_0 = 30 \text{ cm}$ (that is an $r_0(2.2 \mu\text{m})$ of 140 cm), $R = T = 0.5$, quantum efficiency $\eta = 0.65$, $\text{ron} = 20 \text{ e pix}^{-1} \text{ exposure}^{-1}$, background temperature $T = 280 \text{ K}$, background emissivity $\mu = 0.2$, field of view ω_p corresponding to the central lobe of the Airy pattern of the collecting aperture.

From these profiles, it is apparent that even in the case of incomplete corrections (at a level likely to be nowadays negative) detection capabilities of our Achromatic

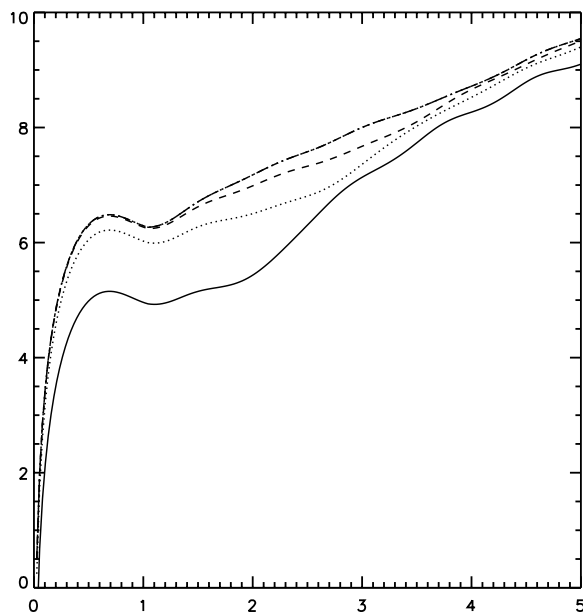


Fig. 9. Similar to Fig. 8 but here in the case of INCOMPLETE CORRECTIONS up to radial modes 1, 3, 5, 7. Values used are the same as for Fig. 8. Note the change in vertical scale (0 to 10)

Interfero Coronagraph are worth considering in the study of the stellar environment and similar topics. For example, clear detection of a companion as close as half the Airy radius, exhibiting a magnitude difference Δm of the order 6 appears as a reasonable goal in rather ordinary observation conditions and with limited integration time (500 s here). Let us point out that the given estimates correspond to raw data and do not include the significant increase of performance usually brought about by appropriate data processing. Moreover, improvements relating to the correction capabilities of the adaptive optics devices (cut-off frequencies, number of actuators, etc.) are still in progress and the relevant profiles are likely to be close to the ones of the full correction case. The major limitation encountered in the correction process comes from the tilt effect, however there are serious hopes for much better capabilities. Thus, owing to the result of the present theoretical analysis it is conceivable that companions with (roughly speaking) Δm up to 12 would be detectable from raw data, as suggested by profiles in Fig. 8. Moreover the reported values of Δm can be increased by observing simultaneously (on the same camera) in two separate and adjacent spectral bands covering the total bandwidth of work, which relaxes the problem of the reference source and the possible random variations of the point spread function (Racine et al. 1999). Of course an additional gain is achievable by using a longer integration time (number of frames) and by using a longer exposure-time for each single frame when

allowed by atmospheric seeing and adaptive optics capabilities.

The large Δm appearing within the Airy radius in the case of complete corrections (Fig. 8) reflects the behaviour of the residual halo, whose central part becomes larger and larger and darker and darker as the quality of the correction becomes better and better, the ultimate limitation to detection being then essentially photon noise. The bump appearing in the Δm profiles about one Airy radius reflects the bump in the field response to extinction (Fig. 3).

5. Conclusion

We have described the concept of the Achromatic Interfero Coronagraph, based on destructive interferences and we have derived analytically its theoretical capabilities relating to the detection of a faint companion close to a star. The two main advantages of this device are its inherent achromaticity (which allows choosing freely both a working wavelength and a spectral bandwidth) and its ability to sense very closely the neighbouring of a star (or other unresolved source), namely at angular distances of a fraction of the first radius of the Airy pattern. This latter feature of our coronagraph is not available in other coronagraphs using Lyot masks (Mouillet et al. 1997) for which sensing is not possible closer than several Airy radii, while achromaticity is not provided by both Lyot Coronagraphs and the Phase Mask Coronagraph (Roddier & Roddier 1997).

Beside these advantages, some specific technical constraints in terms of internal adjustment have to be considered, namely optical quality and accurate nulling of optical path difference. As for every coronagraph, observations from the ground require adaptive optics. The goal of adaptive optics here is to minimize unwanted residual light in the image plane, caused by departure from centro-symmetry of the complex amplitudes of interfering fields.

In such conditions and even when adaptive optics correction is not fully achieved, detection capabilities for ground-based observations makes our coronagraph a powerful tool for the detection of faint sources located in the neighbouring of a star, as close as a fraction of the theoretical angular resolution set by diffraction. Theoretical estimates given in this paper show that at angular distances comparable to the ones explored by other coronagraphs, our coronagraph has comparable detection capabilities. At angular distances, unreachable by other coronagraphs, clear detection of companion 6 magnitudes fainter than a star with K -magnitude of 5 is a truly reasonable goal for a 4-m class telescope equipped with adaptive optics. This magnitude difference can be significantly increased by an appropriate observing strategy (separate and adjacent spectral bands) by longer integration time and by

appropriate data processing. Owing to expected progress on adaptive optics capabilities, detection of a close companion, at an angular distance smaller than the diffraction limit, with Δm up to 12 (and beyond when increasing integration time) seems to be within reach from raw data with our coronagraph.

References

- Allen C.W., 1973, "Astrophysical Quantities". the Athlone Press, London
- Baudoz P., Rabbia Y., Gay J., 1997, in: Proc. ESO Conf. "Science with the VLT Interferometer", Paresce F. (ed.), p. 355
- Beuzit J.L., Mouillet D., Lagrange A.M., Paufique J., 1997, A&AS 125, 175
- Born M., Wolf E., 1970, "Principles of Optics". Pergamon Press, 4th ed.
- Boyd R.W., 1980, JOSA 70, 877
- Bracewell R.N., 1978, Nat 274, 780
- Breckinridge J.B., McAlister H.A., Robinson W.G., 1979, Appl. Opt. 18, 1034
- Conan J.M., 1995, Thesis, University Paris XI
- Gay J., Rabbia, Y., 1996, C.R. Acad. Sci. Paris, 322 Série II b, 265
- Gay J., Rabbia Y., Baudoz P., 1997, C.R. Acad. Sci. Paris, 325 Série II b, 51
- Goodman J.W., 1985, "Statistical Optics". Wiley, New York
- Gouy C., 1890, C.R. Acad. Sci., Paris 110, 1251
- Lyot B., 1930, C.R. Acad. Sci. Paris 191, 834
- Malbet F., 1996, A&AS 115, 161
- Mennesson B., Marriotti J.-M., 1997, Icarus 128, 202
- Mouillet D., Lagrange A.M., Beuzit J.L., Renaud N., 1997, A&A 324, 1083
- Nakajima T., Golimovsky D.A., 1995, AJ 109, 1181
- Noll R.J., 1976, JOSA 66, 207
- PUEO [<http://www.cfht.hawaii.edu/manuals/aob/ao.html>]
- Racine R., Walker G.A.H., Nadeau D., Doyon R., Marois C., 1999, PASP 111, 587
- Reynaud F., Delaire E., 1993, Electronics Lett. 29, 1718
- Roddier F., 1981, in "Progress in Optics", Vol. XIX, p. 281
- Roddier F., Roddier C., 1997, PASP 109, 815
- Roddier N., 1990, Opt. Engin. 29, 1174
- Wolf N.J., Angel J.R.P., 1997, in: Planets Beyond the Solar System and the Next Generation of Space Missions, ASP Conf. Ser. 119, Soderblom D. (ed.), p. 285