Available online at www.sciencedirect.com
ScienceDirect
C. R. Physique 8 (2007) 323–332
COMPTES RENDUS PHYSIQUE
http://france.elsevier.com/direct/COMREN/

Optical techniques for direct imaging of exoplanets/Techniques optiques pour l'imagerie directe des exoplanètes

A theoretical look at coronagraph design and performance for direct imaging of exoplanets

Olivier Guyon

Subaru Telescope, National Astronomical Observatory of Japan, 650 North A'ohoku Place, Hilo, HI 96720, USA

Available online 30 May 2007

Abstract

New coronagraph concepts are being proposed at an increasing rate. In this article, I look at the intrinsic performance each concept can theoretically offer for direct detection of exoplanets. Fundamental physics sets quantifiable limits on the performance of coronagraphs. This theoretical work provides valuable insight into how to design high performance coronagraphs, explains many of the characteristics of particular designs and reveals the importance of stellar angular diameter on coronagraph performance. This study shows that the PIAA coronagraph appears to be extremely attractive for direct imaging of exoplanets. Recent results obtained in the laboratory with this coronagraph are briefly presented. ***To cite this article: O. Guyon, C. R. Physique 8 (2007).***

Résumé

Analyse théorique de l'architecture et des performance des coronographes pour l'imagerie directe d'exoplanètes. De nouveaux coronographes sont proposés a un rythme de plus en plus soutenu. Dans cet article, j'étudie la performance théorique qu'ils offrent pour la detection directe d'exoplanètes. Les lois fondamentales de la physique imposent des limites sur la performance des coronagraphes et ces limites sont comparées aux coronographes existants. Ce travail théorique permet de mieux comprendre comment concevoir un coronographe haute performance et explique de nombreuses caracteristiques de plusieurs types de coronographes. Cette nouvelle perspective illustre tout particulierement l'importance de la taille angulaire de l'étoile. Cette étude montre que le coronagraph PIAA est tout particulierement attrayant pour l'imagerie directe d'exoplanètes. De recents resultats obtenus en laboratoire sont présentés. ***Pour citer cet article : O. Guyon, C. R. Physique 8 (2007).***

Keywords: Coronagraph; Exoplanets; Telescope

Mots-clés : Coronographe ; Exoplanètes ; Téléscope

1. Introduction

The pace at which new coronagraph designs appear seems to be ever accelerating, as new ideas are frequently developed and previous coronagraph designs are branching out in broader families. In Table 1, I have attempted to provide a snapshot of current coronagraph designs. Fortunately, they can be grouped in 3 broad (but somewhat overlapping) categories:

E-mail address: guyon@naoj.org.

1631-0705/$ – see front matter
doi:10.1016/j.crhy.2007.02.005

Table 1
List of coronagraphs able to achieve 10^{10} PSF contrast within 5 λ/d

Coronagraph	Abrev.	Reference	Design(s) adopted
'Interferometric' coronagraphs			
Achromatic Interferometric Coronagraph	AIC	[1]	
Common-Path Achromatic Interferometer-Coronagraph	CPAIC	[2]	(= AIC)
Visible Nulling Coronagraph, X–Y shear (4th order null)	VNC	[3]	Shear distance = ±0.3 pupil radius
Pupil Swapping Coronagraph	PSC	[4]	Shear distance = 0.4 pupil diameter
Pupil apodization			
Conventional Pupil Apodization and Shaped-Pupil	CPA	[5]	Prolate ($r = 4.2\lambda/d$, 8% throughput)
Achromatic Pupil Phase Apodization	PPA	[6]	$\phi = \phi_2(x) + \phi_2(y)$; $a = 2$; $\epsilon = 0.01$
Phase Induced Amplitude Apodization Coronagraph	PIAAC	[7]	Prolate apodization
Phase Induced Zonal Zernike Apodization	PIZZA	[8]	Not simulated
Improvement on the Lyot concept with amplitude masks			
Apodized Pupil Lyot Coronagraph	APLC	[9]	$r = 1.8\lambda/d$
Apodized Pupil Lyot Coronagraph, N steps	APLC_N	[10]	$(N, r) = (2, 1.4); (3, 1.2); (4, 1.0)$
Band limited, 4th order	BL4	[11]	$\sin^4$ intensity mask, $\epsilon = 0.21$
Band limited, 8th order	BL8	[12]	$m = 1$, $l = 3$, $\epsilon = 0.6$
Improvement on the Lyot concept with phase masks			
Phase Mask	PM	[13]	with mild prolate pupil apod.
4 quadrant	4QPM	[14]	
Achromatic Phase Knife Coronagraph	APKC	[15]	(=4QPM)
Optical Vortex Coronagraph, topological charge m	OVC_m	[16]	$m = 2, 4, 6, 8$
Angular Groove Phase Mask Coronagraph	AGPMC	[17]	(=OVC)
Optical Differentiation	ODC	[18]	mask: $x \times \mathrm{e}^{-(x/10)^2}$

– *Interferometric coronagraphs (AIC, CPAIC, VNC, PSC)*: The coronagraph design is inspired from interferometry; a single telescope pupil is used to extract several beams which are coherently mixed to produce a null;
– *Pupil apodization (CPA, PPA, PIAAC, PIZZA)*: These designs create an apodized pupil without sharp edges, which is then used to form a high contrast image;
– *Improvement on the Lyot design (APLC, BL4/8, PM, 4QPM, APKC, OVC, AGPMC, ODC)*: This very rich family of coronagraphs uses a combination of focal and pupil plane masks to selectively reject starlight.

The 'external occulter' concept, where a physical mask is placed far in front the telescope to mask the star, is not considered in this study, and is not listed in Table 1.

The goal of this article is to understand how these coronagraphs compare, and to also quantify the fundamental limits of coronagraphy imposed by fundamental physics. The analysis given in this paper is detailed in [19]. Recent results obtained in the laboratory with the highly efficient PIAA coronagraphs are also shown.

2. Fundamental limits of coronagraph performance & comparison to existing coronagraph designs

2.1. Theoretical limit of coronagraph performance for a monochromatic point source

Before attempting to find fundamental limits and compare coronagraph performances, a performance metric needs to be defined. I propose here to use the useful throughput, which is the maximum amount of planet light which can be collected (most often by summing up pixels in a focal plane) without collecting at the same time a greater quantity of starlight. This definition is somewhat arbitrary (maybe we can still use planet light which is superimposed on much brighter starlight?) and overly simple (it does not describe well what happens in the presence of exodozi for example). It is, however, very helpful because it can be applied to any coronagraph design, and, although not perfect, is still a good representation of coronagraph performance.

What is the theoretical limit of coronagraph performance? Clearly, there must be such a limit: it must be impossible to have an optical system remove most of the starlight while transmitting most of the planet light if the planet-star

separation is much less than 1 λ/d. This limit can be accurately computed by recognizing optical systems are linear in complex amplitude. The complex amplitude in any point of the system (for example a pixel in the focal plane) is a linear function of the complex amplitude in the entrance pupil of the telescope. Fourier transforms (commonly used in coronagraph modeling) and diffraction propagation are linear in complex amplitude. The linear relationship between input (pupil) and output (final focal plane) is entirely fixed by the coronagraph design. We can now think of a coronagraph as a linear system which has to filter out a flat non-tilted wavefront in the pupil plane (= on axis point source). The wavefront of an off-axis source can then be decomposed as the sum of an on-axis wavefront (which the coronagraph will remove) and a residual, which gets smaller as the point source is moved closer to the optical axis. This residual contains all of the energy that can be transmitted by the coronagraph. This simple framework allows us to now quantify a fundamental limit on coronagraph performance. Using the same framework, we can even prove that one could build such a coronagraph with a large number of beamsplitters, so this limit can be reached. It is unfortunately unknown if a simple practical way of building this coronagraph exists.

2.2. The stellar angular size problem

A quick glance a Fig. 1 paints a very rosy picture of the coronagraph world: the theoretical limit shows that we should be able to get 50% throughput at 0.5 λ/d separation. We even happen to have several coronagraphs at hand which are getting fairly close to it. Incidentally, one of these (the AIC) is achromatic by design and optically simple to build.

Checking (see Fig. 2) what happens when taking into account the fact that stars are not points but small disks reveals a true problem in the coronagraph world: none of the coronagraphs which appeared to be 'high performance' can cope with the small stellar angular diameter. The star diameter (which is 1% of the planet-star separation for the Sun–Earth system) simply prevents high contrast: the coronagraphic leak for a point at the edge of the star far exceeds $1e{-}10$ for these 'high performance' coronagraphs. Worse: even the theoretical limit obtained with our linearity framework is hugely degraded by making the star a fraction of λ/d. It would seem odd that making the star 0.01 λ/d in radius moves the theoretical inner working angle limit from 0.5 λ/d to more than 1 λ/d. This can, however, be explained by our framework: when the star gets a little bigger, the coronagraph needs to reject more than just the flat on-axis wavefront. The number of 'wavefront modes' which need to be rejected grows, and the wavefront from an off-axis source at 1 λ/d starts to 'look like' these additional modes. Now that stellar diameter is taken into account, the most promising coronagraphs appear to be the PIAA and OVC. Even though they can function cope with stellar angular diameter, their performance is still somewhat degraded as the star gets bigger (from 0.01 λ/d to 0.1 λ/d).

The BL8 and CPA coronagraphs, even though they offer significantly lower performance than the PIAA or OVC, appear to be largely unaffected by stellar angular size.

2.3. Zodiacal and exozodiacal light & coronagraph design

So far, we have developed an accurate model of coronagraph performance for looking at a single planet around a partially resolved star. The useful throughput tells us how many photons we can hope to detect from the planet in a given exposure time with a given telescope diameter. We have, however, not included zodiacal or exozodiacal backgrounds, which can easily overwhelm the planet signal. Unlike starlight (which can be removed by the coronagraph), not much can be done about background light. For distant targets, the planet is intrinsically fainter, but both the zodi and exozodi have the same surface brightness: the effect of zodi and exozodi is especially strong for distant targets. Although the level of background light is independent of coronagraph design, the amount of background light mixed with the planet image is not. Ideally, this amount is the equal the background level integrated over a 1 λ/d disk (the size of the plane image in a conventional telescope). Unfortunately, many coronagraphs do not preserve the full angular resolution of the telescope, or produce non-ideal off-axis PSFs (for example, PSFs in which a large amount of flux is scattered in large 'wings'). This effect can be quantified by the diffraction efficiency factor (DEF), which tells us how much more background light is mixed with the planet image compared to a normal diffraction limited imaging system. DEFs for the coronagraphs studied in this paper range from a perfect 1.0 (PM, 4QPM, OVC, PIAA) to values equal (AIC, PSC) or less than (VNC, BL4, CPA, PPA, BL8) 0.5.

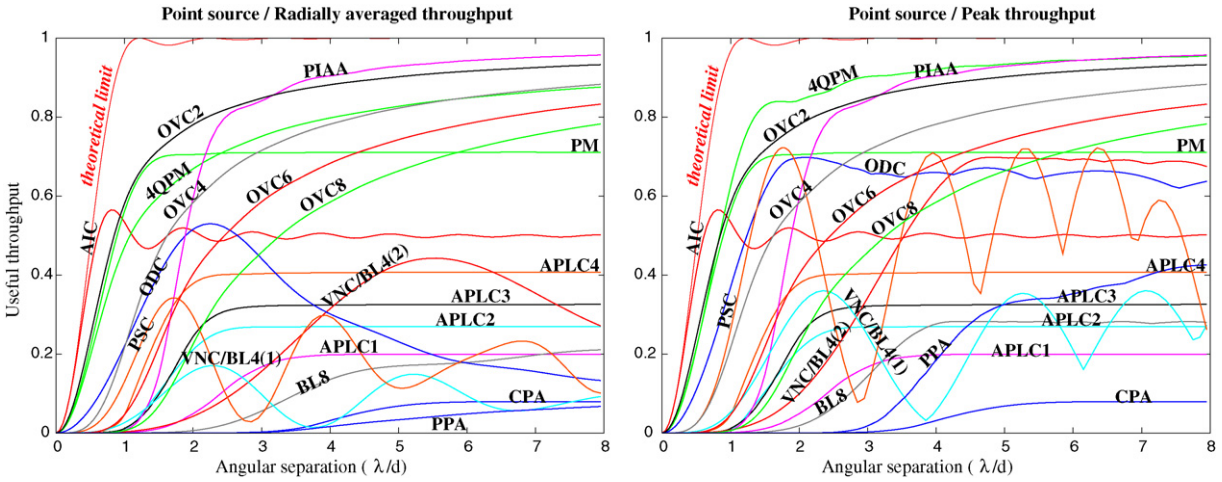

Fig. 1. Throughput, at the 10^{10} contrast level, of the coronagraphs listed in Table 1 as a function of angular separation. On the left, the useful throughput has been radially averaged for coronagraphs with 'preferential' directions (BL4, BL8, 4QPM, ODC, VNC, PSC). On the right, the peak throughput is shown, assuming that the telescope orientation is optimal. The theoretical limit discussed in Section 2.1 is shown in red. The central source is assumed here to be monochromatic and infinitively small.

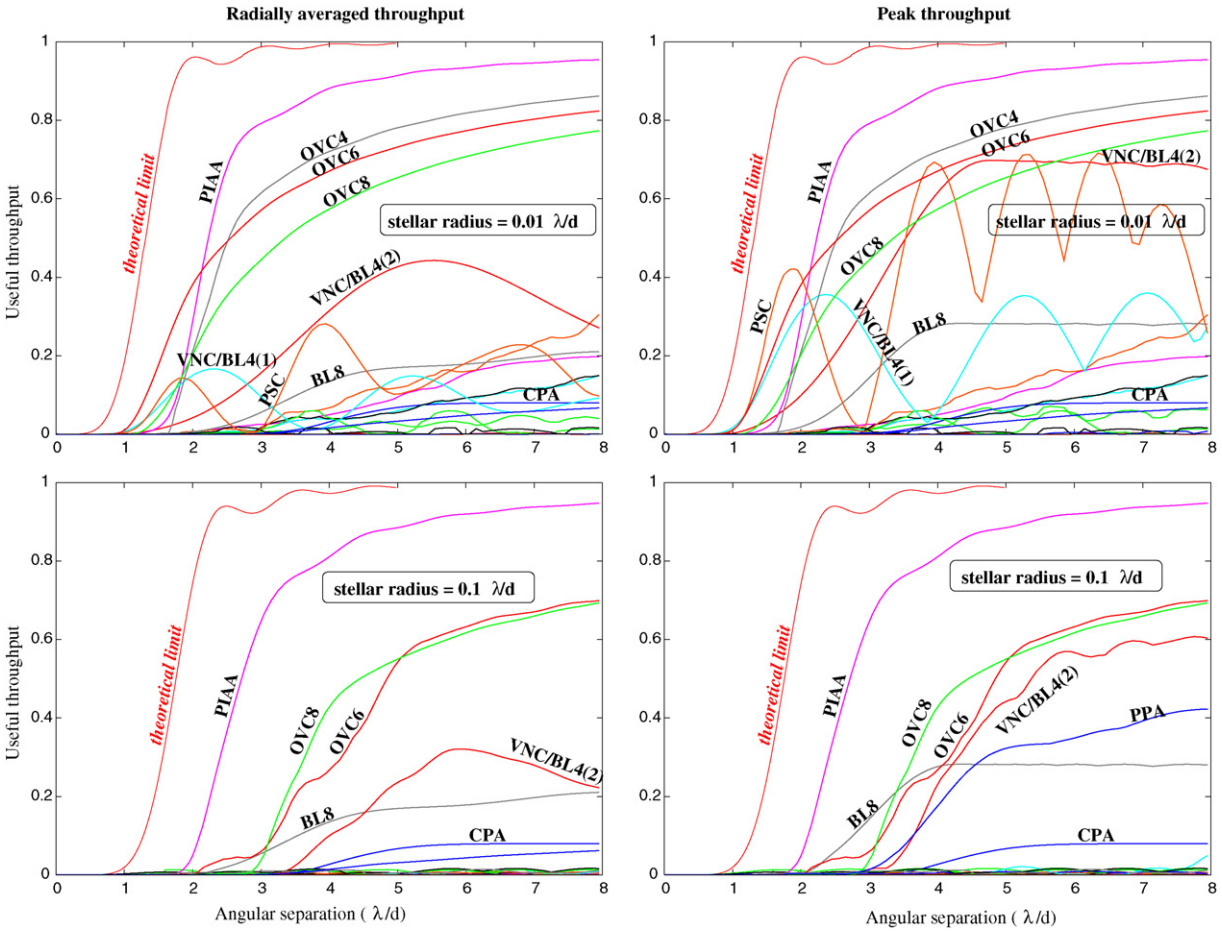

Fig. 2. Same as Fig. 1, but with an extended central source (top: 0.01 λ/d radius disk, bottom: 0.1 λ/d radius disk).

2.4. Wavefront control

From the coronagraph point of view, wavefront errors appear to be a fatality: if the telescope upstream of the coronagraph has a 'wiggly' primary mirror, the coronagraph contrast is going to be degraded. It would also seem that 'high performance' coronagraphs (PIAA, OVC) with smaller IWA, since they are more sensitive to low order aberrations, have a more serious problem with low order wavefront errors. Mid-spatial frequencies which create 'speckles' at the planet's position affect all coronagraphs equally, since the speckle contrast is simply a function of the magnitude of the wavefront aberration. In fact, the coronagraph has a central role in the wavefront control scheme for the telescope, as it provides the precious photons for wavefront sensing:

- Low IWAs enables active closed-loop control of low order aberrations. Low order aberrations are therefore much smaller in a 'high performance' coronagraph than in a large-IWA coronagraph;
- High throughput maximizes the number of photons available for wavefront sensing, increasing the temporal bandwidth of the wavefront control loop. With a higher throughput coronagraph, the requirements on the stability of the telescope optics can be relaxed.

3. Coronagraph performance simulations for exoplanet imaging

The theoretical model and performance metrics used so far should provide a fairly complete and accurate picture of coronagraph performances. To verify the conclusion obtained, and to derive an accurate estimate of the coronagraph performance for exoplanet imaging, a Monte Carlo imaging simulation of coronagraphic observations was developed. A target list was first established, assuming each target star has an Earth-like planet. Coronagraphic images were computed for each target and possible planet position along its orbit. Sample images can be seen for star HIP56997 in Fig. 3.

Fig. 3 nicely illustrates the coronagraphic characteristics quantified in the previous sections:

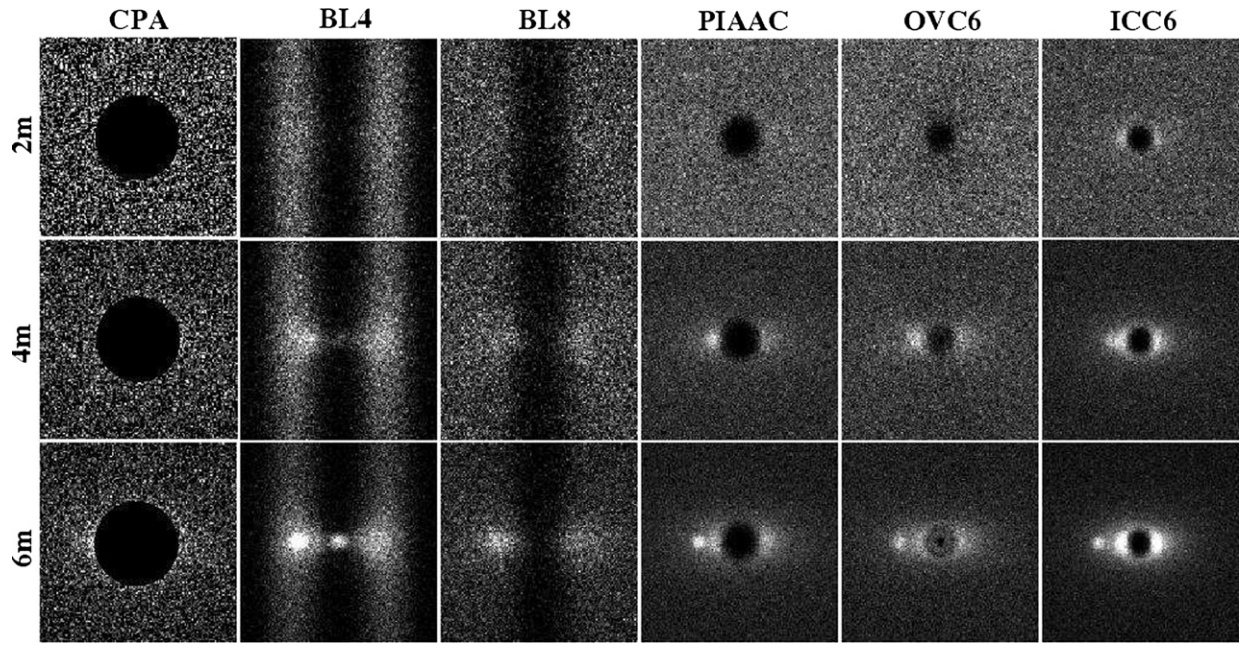

Fig. 3. Simulated 4 hours exposures of HIP 56997 and an hypothetical Earth-type planet at maximum elongation for telescope sizes ranging from 2 to 6 m. HIP 56997 is a G8 type main sequence star at 9.54 pc. Each image assumes a perfect detector, minimum zodiacal background ($m_V = 22.95$ zodiacal background for this 29 deg ecliptic latitude source), a 1 zodi exozodi cloud, a 25% telescope + camera throughput, and a 0.1 μm bandpass centered at 0.55 μm. The system inclination for this particular simulation was arbitrarily set at $i \approx 59$ deg. Each image is $20 \times 20\ \lambda/d$, and the planet-star separation is 80 mas.

– *Throughput*: The CPA, BL8, and to a lesser extent VNC/BL4(2) suffer from low coronagraphic throughput. As a result, the planet's image, even if well outside the coronagraph mask's influence, appears noisy (few photons detected). The PIAAC, OVC6 and ICC6, on the other hand, enjoy nearly 100% throughput: the planet image is brighter and less noisy.
– *Angular resolution*: The CPA, BL8, and to a lesser extent VNC/BL4(2) have poorer angular resolution: the planet image is larger and more zodi/exozodi light is mixed with it.
– *Ability to work at small angular separation*: None of the coronagraph tested can detect the planet on a 2 m telescope. Detection appears feasible on a 4 m telescope with the VNC/BL4(2), PIAAC, OVC6 and ICC6, but requires a 6 m telescope with the BL8. Finally, an 8 m telescope would be needed for detection with a CPA.
– *Sensitivity to stellar angular size*: For all 6 coronagraphs, starlight leaks are very small.

Images similar to the ones showed in Fig. 3 have been computed for all possible positions of the planet along its orbit. With photon noise the only source of noise, the exposure time required to detect the planet (SNR = 7 detection threshold) is computed for each planet position. For a given exposure time, the planet detection probability can then be measured as the fraction of the planet orbit along which it can be detected. This analysis quantifies the large performance differences between coronagraph designs. To reach a 50% detection probability for HIP56997's planet, the CPA requires 4.3 hours of exposure time, while the PIAAC requires 5.8 min. Comparison of such exposure times suggests that a PIAAC or OVC6 is equivalent to a CPA or BL8 on a telescope 2 to 3 times larger.

This exposure time vs. detection probability analysis was repeated for each star of the sample, for which an 'exposure time required to reach 50% planet detection probability' was assigned. For a given telescope/coronagraph combination, the sample stars can then be ranked from the easiest (shorter exposure time) to the hardest (longer exposure time). A given total "open shutter" observation time can then be optimally assigned to maximize the number of targets sampled at the 50% detection probability level. Results of this analysis are shown in Fig. 4.

Fig. 4 confirms the results obtained so far:

(i) The PIAAC and OVC appear to be the most promising coronagraphic designs;
(ii) The theoretically ideal coronagraph ICC6, thanks to its smaller IWA, would enable a larger sample size than allowed by either PIAAC or OVC. For a small sample size, however, most the targets do not require this small IWA, and the PIAAC and OVC performance is not largely different from the theoretical optimal coronagraph;

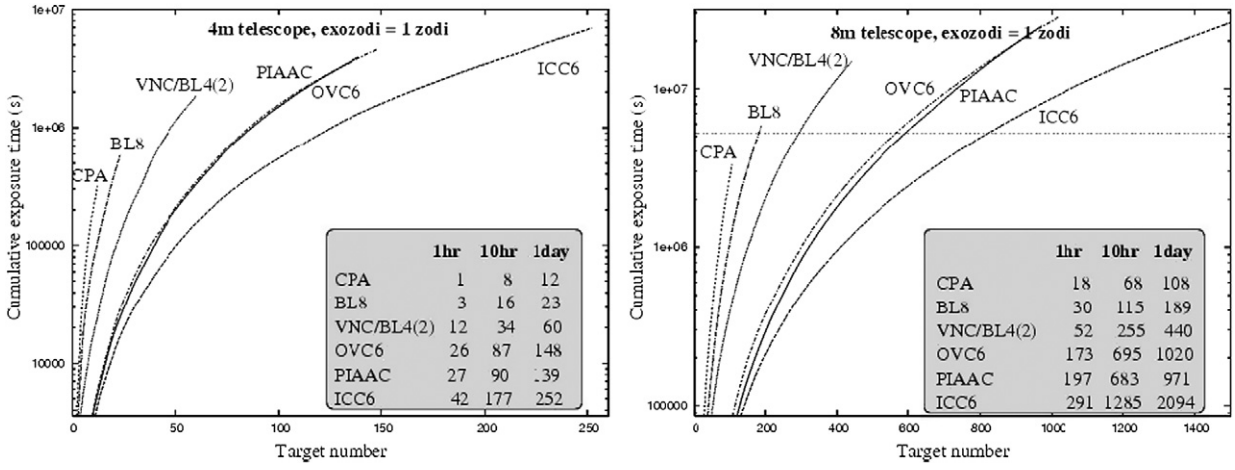

Fig. 4. Total cumulative exposure times required to reach a 50% planet detection (SNR = 7) probability for a single observation as a function of number of targets. For all simulations, a 25% throughput in the 0.5–0.6 μm band is adopted, and targets are ordered with increasing exposure time. Each curve terminates when the required exposure time per target reaches 1 day. Results are shown for a 4 m telescope (left) and an 8 m telescope (right). The number of accessible targets for which the required exposure times are less than 1 hour, 10 hour and 1 day are listed for each case in the Grey boxes. In the 8 m telescope plots, the horizontal line corresponds to a 2 months 'shutter open' cumulative exposure time, and may be considered as a practical limit on the number of targets that can be visited for a 3 to 5 year mission.

(iii) The CPA and BL8 coronagraphs would require a telescope about twice as large to reach the same exoplanet detection productivity than the PIAAC or OVC.

4. PIAA coronagraph

The PIAA coronagraph has been identified in the previous sections as one of the most attractive techniques for direct imaging of exoplanets. It is described in more detail in this section, and laboratory results are presented.

4.1. PIAA coronagraph principle

Conventional apodization coronagraphs use masks to apodize the telescope pupil. These masks unfortunately remove most of the planet light and greatly reduce the telescope angular resolution. An alternative solution is to produce the apodized pupil by geometrical redistribution (remapping) of the flux in the pupil plane rather than selective absorption. The PIAAC performs this lossless amplitude apodization with 2 aspheric optics; the resulting pupil then yields a high contrast PSF in which starlight can be removed by a small focal plane occulting disk. The geometric remapping introduced by the aspheric PIAA mirrors limits the 'clean' field of view in the focal plane: PSFs for sources at more than 10 λ/d from the optical axis are heavily distorted, which has the undesirable effect of mixing more exozodiacal + zodiacal light with the planet image. A set of correcting optics may be added after the focal plane occulter to restore a clean PSF across a reasonable field of view (up to 100 λ/d in radius). A schematic representation of the PIAAC is shown in Fig. 5, which also shows off-axis PSFs in both the 'intermediate' focal plane (where the focal plane occulter is located) and the final focal plane (where field of view is restored).

Thanks to the lossless apodization, the PIAA coronagraph offers simultaneously:

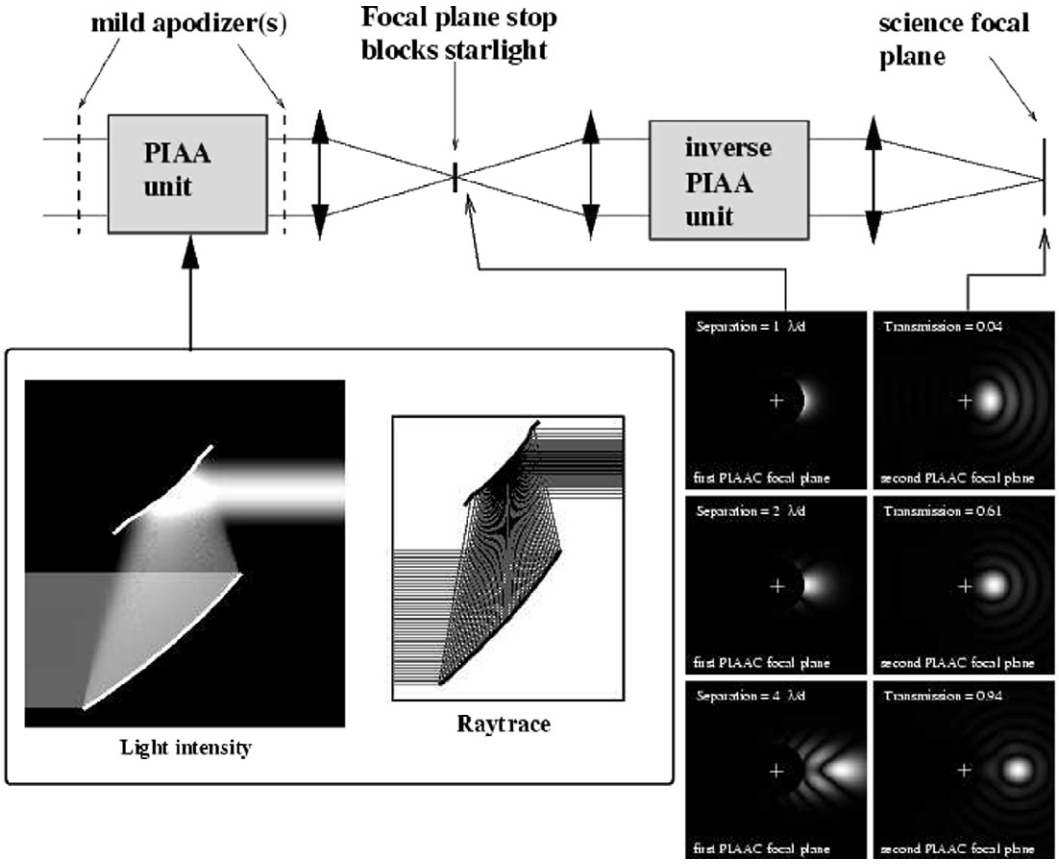

Fig. 5. PIAA coronagraph principle.

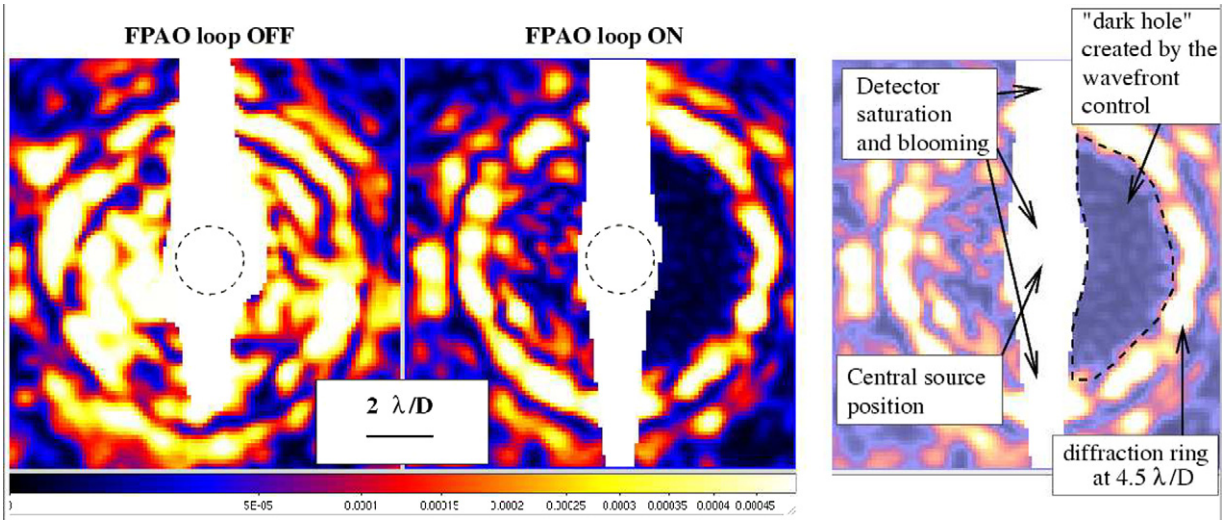

Fig. 6. Preliminary results from the PIAA coronagraph laboratory demonstration. A heavily stretched version of our laboratory PSF image (left) shows that almost all starlight is concentrated within 1.5 λ/D radius (the large vertical structure is due to charge bleeding on the detector). The light beyond 1.5 λ/D is due to phase aberrations in the system and is greatly reduced thanks to wavefront correction using focal-plane Adaptive Optics (FPAO) as shown in the middle image. At the position of the first Airy ring, the contrast is 2.5e−6. The large bright (1e−3 contrast) ring visible in the image is at 4 to 5 λ/d from the central star.

– nearly 100% throughput;
– 1 λ/d angular resolution;
– <2 λ/d inner working angle;
– full 360 deg search area;
– good achromaticity.

4.2. Laboratory demonstration

Our laboratory prototype includes a monochromatic light source (single mode fiber at HeNe), immediately followed by the 2 PIAA aspheric mirrors (which are designed for a $f/15$ diverging input beam and deliver a $f/15$ output converging beam). All optics (PIAA and reimaging optics) are within an enclosure which provides some thermal stability. A 1024 actuators (32×32) MEMs-type actuators, driven by custom-built 16-bit high voltage drivers, is used to correct for residual wavefront aberrations. A phase diversity algorithm is used for wavefront control (the diversity is introduced by the DM itself). In the initial configuration of our experiment, no focal plane mask was blocking starlight, and the apodizer was not located in a pupil plane. Despite these limitations, a 2.5e−6 contrast was reached at 1.5 λ/d from the optical axis (see Fig. 6). In the new optical layout (see Fig. 7), a pupil plane is made available for the conventional apodizer, and a focal plane is accessible for the focal plane occulter.

5. Conclusion

The analysis presented in this article ignores important challenges which must be overcome to reach the performance levels described in this paper: wavefront control and chromaticity. More fundamentally, the assumption that detection is photon noise limited is optimistic: ETPs are embedded in an exozodiacal light unlikely to be perfectly smooth and symmetrical. Since good angular resolution may help to distinguish between a planet and exozodiacal light features (such as arcs or rings), it is important to avoid using coronagraphs that clip or strongly attenuate the edge of the telescope pupil if at all possible. The results obtained in this work should therefore be cautiously considered as upper limits on what may be achieved with a particular coronagraph/space telescope design for the specific goal of imaging ETPs around nearby stars.

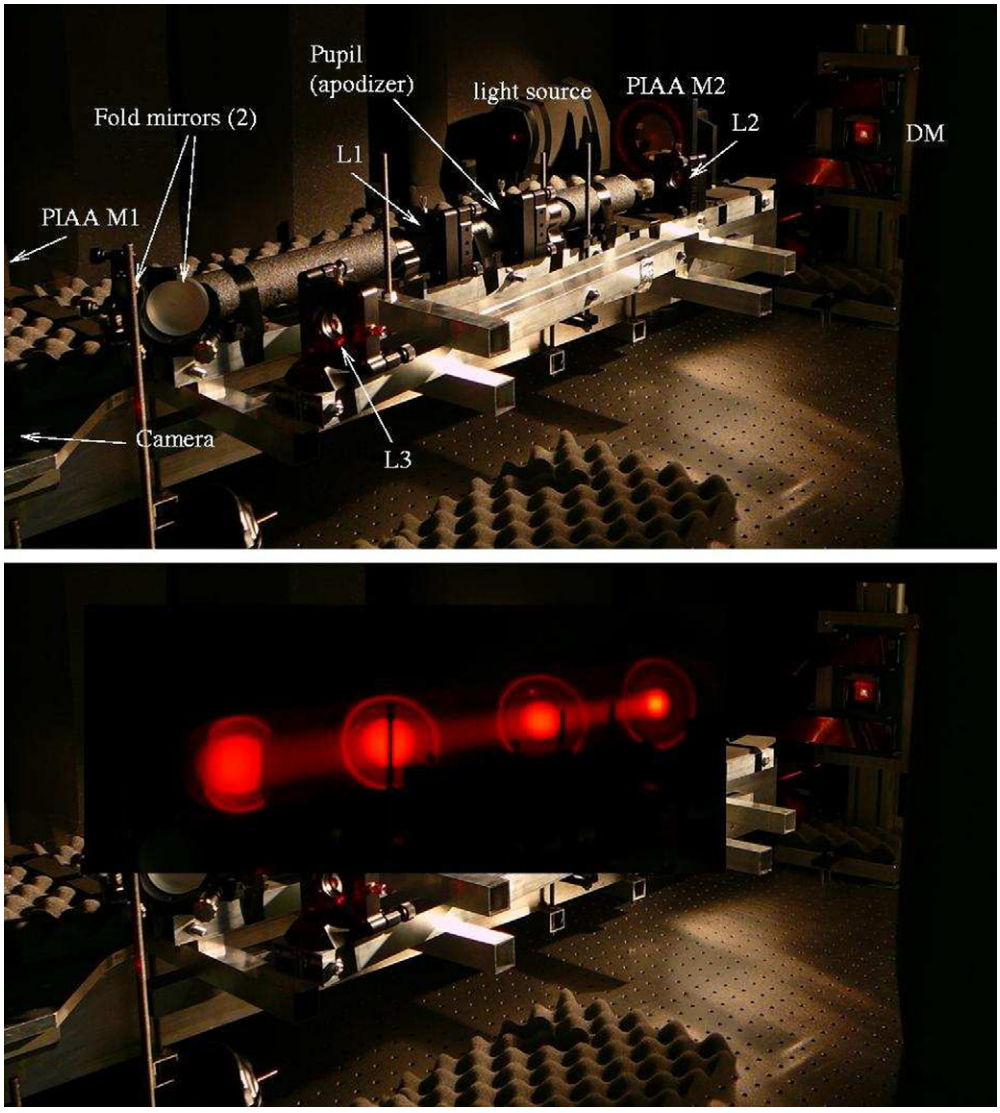

Fig. 7. Layout of the PIAA laboratory demonstration. In the lower image, the beam can be seen to be apodized between the 2 PIAA mirrors.

Acknowledgements

This work was carried out with the support of the National Astronomical Observatory of Japan and JPL contract numbers 1254445 and 1257767 for Development of Technologies for the Terrestrial Planet Finder Mission.

References

[1] P. Baudoz, Y. Rabbia, J. Gay, Astron. Astrophys. Suppl. 141 (2000) 319.
[2] A.V. Tavrov, Y. Kobayashi, Y. Tanaka, T. Shioda, Y. Otani, T. Kurokawa, Opt. Lett. 30 (2005) 2224.
[3] B.P. Mennesson, M. Shao, B.M. Levine, J.K. Wallace, D.T. Liu, E. Serabyn, S.C. Unwin, C.A. Beichman, Proc. SPIE 4860 (2003) 32.
[4] O. Guyon, M. Shao, Publ. Astron. Soc. Pacific 118 (2006) 860.
[5] N.J. Kasdin, R.J. Vanderbei, D.N. Spergel, M.G. Littman, Astrophys. J. 582 (2003) 1147.
[6] W. Yang, A.B. Kostinski, Astrophys. J. 605 (2004) 892.
[7] O. Guyon, Astron. Astrophys. 404 (2003) 379.
[8] F. Martinache, J. Opt. A 6 (2004) 809.
[9] R. Soummer, C. Aime, P.E. Falloon, Astron. Astrophys. 397 (2003) 1161.

[10] C. Aime, R. Soummer, Proc. SPIE 5490 (2004) 456.
[11] M.J. Kuchner, W.A. Traub, Astrophys. J. 570 (2002) 900.
[12] M.J. Kuchner, J. Crepp, J. Ge, Astrophys. J. 628 (2005) 466.
[13] F. Roddier, C. Roddier, Publ. Astron. Soc. Pacific 109 (1997) 815.
[14] D. Rouan, P. Riaud, A. Boccaletti, Y. Clénet, A. Labeyrie, Publ. Astron. Soc. Pacific 112 (2000) 1479.
[15] L. Abe, F. Vakili, A. Boccaletti, Astron. Astrophys. 374 (2001) 1161.
[16] D.M. Palacios, Proc. SPIE 5905 (2005) 196.
[17] D. Mawet, P. Riaud, O. Absil, J. Surdej, Astrophys. J. 633 (2005) 1191.
[18] J.E. Oti, V.F. Canales, M.P. Cagigal, Astrophys. J. 630 (2005) 631.
[19] O. Guyon, E.A. Pluzhnik, M.J. Kuchner, B. Collins, S.T. Ridgway, Astrophys. J. Suppl. Ser. 167 (2006) 81.