THEORETICAL LIMITS ON EXTRASOLAR TERRESTRIAL PLANET DETECTION WITH CORONAGRAPHS

O. Guyon, E. A. Pluzhnik, M. J. Kuchner, B. Collins, and S. T. Ridgway

Received 2006 May 25; accepted 2006 July 7

ABSTRACT

Many high-contrast coronagraph designs have recently been proposed. In this paper, their suitability for direct imaging of extrasolar terrestrial planets is reviewed. We also develop a linear algebra based model of coronagraphy that can both explain the behavior of existing coronagraphs and quantify the coronagraphic performance limit imposed by fundamental physics. We find that the maximum theoretical throughput of a coronagraph is equal to 1 minus the nonaberrated noncoronagraph PSF of the telescope. We describe how a coronagraph reaching this fundamental limit may be designed, and how much improvement over the best existing coronagraph design is still possible. Both the analytical model and numerical simulations of existing designs also show that this theoretical limit rapidly degrades as the source size is increased: the “highest performance” coronagraphs, those with the highest throughput and smallest inner working angle (IWA), are the most sensitive to stellar angular diameter. This unfortunately rules out the possibility of using a small IWA (<\lambda/d) coronagraph for a terrestrial planet imaging mission. Finally, a detailed numerical simulation that accurately accounts for stellar angular size, zodiacal and exozodiacal light is used to quantify the efficiency of coronagraph designs for direct imaging of extrasolar terrestrial planets in a possible real observing program. We find that in the photon noise–limited regime, a 4 m telescope with a theoretically optimal coronograph is able to detect Earth-like planets around 50 stars with 1 hr exposure time per target (assuming 25% throughput and exozodi levels similar to our solar system). We also show that at least two existing coronagraph design can approach this level of performance in the ideal monochromatic case considered in this study.

Subject headings: planetary systems — techniques: high angular resolution — telescopes

1. INTRODUCTION

Direct imaging and characterization (through low-resolution spectroscopy) of extrasolar terrestrial planets (ETPs) is an exciting but challenging scientific application of coronagraphy. While the angular separation is well within the reach of a midsized telescope in the visible, the high star-planet contrast (about 10^{10} in the visible) requires a wave front stability that can only be obtained with a space telescope. The low apparent luminosity of an Earth-like planet (m_v \approx 29 for an Earth at 10 pc) requires the precious planetary photons to be well isolated from the stellar light; otherwise, photon noise from the stellar light drives the exposure time unreasonably high and speckles from the stellar image, if not calibrated, prevent detection. A system capable of delivering 10^{10} contrast at less than 0\arcsec is therefore required for efficient detection and characterization of ETPs around a reasonably large (\approx 30) sample of nearby stars. The purpose of this work is to identify and quantify how fundamental physics imposes hard limits on what coronagraphs can achieve and to compare these limits with what existing coronagraph designs can achieve.

Coronagraph designs simulated in this work are introduced and § 2, where they are tentatively grouped in a four families. The configuration adopted for each design in this study is also described. In § 3, a metric to quantify coronagraph performance is proposed, allowing existing coronagraph designs to be evaluated in the “ideal” case (monochromatic light, unresolved central star). A theoretical limit of coronagraph performance in this ideal case is also derived. In the next two sections, we closely look at the most fundamental (unavoidable) deviations from this ideal case and their effects on coronagraphic performance: stellar angular size (§ 4) and presence of a zodiacal/exozodiacal background (§ 5). Finally, a detailed simulation of coronagraphic observations of a sample of nearby stars is performed in § 6 to derive the efficiency of coronographs for detection of ETPs.

2. BRIEF DESCRIPTION OF EXISTING CORONAGRAPH DESIGNS

The coronagraphs studied in this work (see Table 1) are all the coronographs known to us to theoretically achieve a 10^{10} PSF contrast within 5\lambda/d of the central source, as illustrated in Figure 1. It is impossible to compare all the different kinds of coronagraph designs that have been suggested in all their details. But without these concrete examples, we could not communicate our new perspective and illustrate what are the fundamental physical effects limiting coronagraphic performance. We offer this brief summary of previous work as historical background prior to using these designs. More detailed descriptions of these coronographs can be found in the references given in this section.

We only consider in this paper unobstructed circular pupils. We note that several techniques have been proposed to adapt “unfriendly” pupil shapes (central obstruction, spiders) to specific coronagraphs (Soummer 2005; Murakami & Baba 2005).

2.1. “Interferometric” Coronagraphs (AIC, VNC, PSC)

These coronagraphs look much like nulling interferometers: they rely on interferometric combination of discrete beams derived from the entrance pupil.

AIC.—The Achromatic Interferometric Coronagraph (Gay & Rabbia 1996; Baudouz et al. 2000) uses a beam splitter to


### Coronagraphs Able to Achieve $10^{10}$ PSF Contrast within $5 \lambda/d$

| Coronagraph                                | Abbreviation | Reference                                      | Design(s) Adopted                      |
|--------------------------------------------|--------------|------------------------------------------------|----------------------------------------|
| Achromatic Interferometric Coronagraph     | AIC          | Baudoz et al. (2000)                           |                                       |
| Common-Path Achromatic Interferometer-Coronagraph | CPAIC        | Tavrov et al. (2005)                          | (=AIC)                                |
| Visible Nulling Coronograph, X-Y shear (fourth-order null) | VNC         | Mennesson et al. (2003) | Shear distance =±0.3 pupil radius         |
| Pupil Swapping Coronograph                 | PSC          | Guyon & Shao (2006)                           | Shear distance =0.4 pupil diameter       |

#### Pupil Apodization

| Conventional Pupil Apodization and Shaped-Pupil\(^b\) | CPA          | Kasdin et al. (2003) | Prolate \((r = 4.2 \lambda/d, 8\% throughput)\) |
|-------------------------------------------------------|--------------|----------------------|-----------------------------------------------|
| Achromatic Pupil Phase Apodization                    | PPA          | Yang & Kostinski (2004) | \(\phi = \phi_x(x) + \phi_y(y); \ a = 2; \ c = 0.01\) |
| Phase Induced Amplitude Apodization Coronograph       | PIACCESS     | Kuchner & Traub (2002) |sin\(^2\) intensity mask, \(c = 0.21\)        |
| Phase Induced Zonal Zernike Apodization              | PIzza        | Martinache (2004)     | Not simulated                                  |

#### Improvement on the Lyot Concept with Amplitude Masks

| Phase Mask                                      | PM           | Roddier & Roddier (1997) | With mild prolute apod.                      |
|-------------------------------------------------|--------------|--------------------------|----------------------------------------------|
| 4 Quadrant Phase Mask                            | 4QPM         | Rouan et al. (2000)      |                                              |
| Achromatic Phase Knife Coronagraph               | APKC         | Abe et al. (2001)        | (=4QPM)                                       |
| Optical Vortex Coronograph, topological charge \(m\) | OVC\(m\)  | Palacios (2005)          | \(m = 2, 4, 6, 8\)                           |
| Angular Groove Phase Mask Coronograph            | AGPM\(C\)   | Mawet et al. (2005)      | (=AGPM)                                      |
| Optical Differentiation                           | ODC          | Ott et al. (2005)        | Mask: \(x \times \exp(-\alpha/\lambda d^4)\) |

\(a\) The Visible Nulling Coronograph (VNC) and Band-limited fourth-order (BL4) coronagraphs belong to the same class of pupil-shearing fourth-order coronagraphs, and are simply two ways of achieving the same result. They can be designed to have exactly the same performance. In this table, the VNC is chosen with a small IWA and two orthogonal shear directions, while the BL4 is designed with a larger IWA and two shears in the same direction. To reflect this similarity, they are referred to as VNC/BL4(1) for the small IWA option (listed as VNC in this table) and VNC/BL4(2) for the large IWA option (listed as BL4 in this table).

\(b\) The CPA design adopted here is a continuous apodization (rather than binary apodization/shaped pupil) that maximizes the radially averaged performance at \(\approx 4 \lambda/d\). More optimal designs exist in other conditions: CPA with high contrast at specific position angles for observations at \(\approx 3 \lambda/d\) or high throughput CPA for observations at \(\approx 4 \lambda/d\).

\(c\) CPA, APLC, APLC\(N\): \(r\) is the radius, in \(\lambda/d\), of the mask within which the circular prolate function is invariant to a Hankel transform. This parameter is half of the mask diameter \(a\) defined in Soummer et al. (2003a).

\(d\) ODC: \(x\) is in \(\lambda/d\). Maximum mask transmission at \(7 \lambda/d\). Lyot pupil mask radius =0.85 times pupil radius.

---

**PSC.**—In the Pupil Swapping Coronograph (Guyon & Shao 2006), parts of the pupils are geometrically swapped prior to destructive interferometric combination, thus avoiding the throughput loss due to the shear in the VNC. The PSC design considered in this work is the fourth-order PSC described in Guyon & Shao (2006).

#### 2.2. Pupil Apodization Coronagraphs (CPA, PPA, PIACCESS, PIZZA)

The pupil complex amplitude can be modified to yield a PSF suitable for high-contrast imaging, a property used by many coronagraph concepts.

**2.2.1. Conventional Pupil Apodization (CPA) with Amplitude Masks**

Apodization can be performed by a pupil plane amplitude mask (Conventional Pupil Apodization, or CPA), which can be continuous (Jacquinot & Roisin-Dossier 1964; Nisenson & Papaliolios 2001; Gonsalves & Nisenson 2003; Aime 2005b) or binary (Kasdin et al. 2003; Vanderbei et al. 2003a, 2003b, 2004; Aime 2005b). Apodization by Mach-Zehnder type pupil plane interferometry was also suggested by Aime et al. (2001) to produce a continuous apodization, but it is not explored in this work.

Within the CPA “family,” we adopt for our comparative study a prolate spheroidal apodized pupil design (Slepian 1965; Aime...
et al. 2002; Kasdin et al. 2003; Soummer et al. 2003a; Vanderbei et al. 2003b), optimally tuned, according to the performance criteria defined in §3.1, for a 4λ/d angular separation. In this design, the apodizer transmits 8% of the light, and at 4λ/d separation, exactly half of this light is detectable according the criteria defined in §3.1. This apodization is nearly identical to the general prolate apodization function shown in Figure 8 of Vanderbei et al. (2003a).

2.2.1.1. Design Trade-offs

In CPAs, trade-offs exist between IWA, throughput, and “discovery space” (fraction of the field of view usable for planet detection). A description of these trade-offs can be found in Kasdin et al. (2003), and they have been extensively explored for the design of binary masks/shaped pupils (Vanderbei et al. 2003a, 2003b, 2004). For example, in the checkerboard-mask design proposed in Figure 1 of Vanderbei et al. (2004), the PSF surface brightness reaches 10^{-10} at 2.8λ/d, but at this separation, the companion needs to be on the diagonal to be observed, and only a quarter of its light then falls in a “black” quadrant. In this particular small-IWA design, at 2.8λ/d, the peak throughput is then 3.8% for a square pupil, and the radially averaged throughput is ≈2%. At large (>4λ/d) separation, the average and peak throughputs are 15.1%. Our prolate spheroidal apodization is not the optimal design if peak throughput is to be maximized (in this case, the “single prolate” shaped pupil described in Kasdin et al. 2003 seems much superior). Our choice is therefore not optimal if the companion’s position is previously known, in which case it could be intentionally placed at a given position angle.

2.2.1.2. Binary versus Continuous Apodization

Binary pupil masks have the advantage of being achromatic and significantly easier to manufacture than continuous transmission masks. This technological advantage, however, comes at a non-negligible cost in planet detection performance:

1. In the PSF delivered by a nonobstructed, nonapodized circular aperture (Airy pattern), 84% of the energy transmitted by the pupil is contained in the central diffraction spot; if the pupil is apodized with a continuous mask, this fraction is close to 100% (but the spot is larger, and less light is transmitted by the mask). In the PSF delivered by binary apodization masks, this fraction is typically about 50%; there is a factor of ≈2 difference between

![Simulated monochromatic images of a 10^{10} contrast system, 1λ/d to 4λ/d separation, for some of the coronagraphs listed in Table 1. The central source was assumed to be a single point (angular diameter =0). The white number in each image is the coronagraph throughput for the off-axis source. The pixel scale is the same for all images, but the brightness scale is not. The companion is moving on a diagonal (rather than horizontal) line for the VNC, PSC, PPA, and 4QPM coronagraphs.](image)
the mask throughput and the “Airy throughput” (Kasdin et al. 2003) useful for detection. Within the central diffraction spot of a planet’s image, the planet flux is proportional to the Airy throughput, while the background light contribution is proportional to the mask throughput. Binary masks therefore “amplify” the amount of background light mixed with the planet image. This effect, discussed in § 5.1 is also found in sparse aperture interferometers, where additional background light is collected from the secondary lobes.

2. The difference between mask throughput and Airy throughput is diffracted in regions of the PSF where planet detection is impossible. In a properly designed mask, this unwanted light avoids a large central region around the PSF’s center.

As the number of openings or slits in a binary mask increases, these two effects become less of a concern and the mask asymptotically becomes identical to a continuous mask.

2.2.2. Pupil Phase Apodization (PPA)

If only one-half of the focal plane is considered, an amplitude pupil apodization can be replaced by a phase-only apodization (pupil phase apodization, or PPA), just as phase corrections in the pupil can only cancel focal plane speckles in one-half of the field of view. High-contrast imaging with phase apodization was proposed by Yang & Kostinski (2004), who found solutions for broadband imaging and obtained contrast/throughput performances similar to amplitude apodization designs (although only over a quarter of the field of view). Codona & Angel (2004) independently computed a PPA solution for the Hubble Space Telescope pupil to suppress diffraction in half of the field of view. The PPA design adopted in this work uses a square pupil mask and the sum of two identical phase apodizations (eq. [6] in Yang & Kostinski 2004), one in the $x$-direction, the other in the $y$-direction.

Although the PPA mask has a 100% throughput within the square aperture, it exhibits large phase slopes at its edges, which are necessary to suppress diffraction in a quarter of the field but significantly increase the intensity of diffracted light in other parts of the field, reduce the amount of light in the central core of the PSF and broaden the PSF core. Just as for binary aperture apodization masks, the PPA’s large ratio between the pupil mask throughput (here equal to 0.6 area of the largest square that fits within the pupil) and the light concentrated in the PSF core ($\approx 0.3$) is problematic in the presence of zodiacal and exozodiacal backgrounds.

2.2.3. Phase-Induced Amplitude Apodization Coronagraph (PIAAC)

This coronagraph uses a lossless amplitude apodization of the pupil performed by geometric redistribution of the light rather than selective absorption (Guyon 2003; Traub & Vanderbei 2003; Guyon et al. 2005; Vanderbei & Traub 2005; Martinache et al. 2006; Vanderbei 2006; Pluzhnik et al. 2006). The same beam remapping technique is known as beam shaping in laser science (Shealy 2002) and radio astronomy.

The geometrical remapping of the telescope pupil amplifies phase slopes, therefore enabling observations at angular separations $\approx 3$ times smaller than with a conventional amplitude apodization. Two designs are considered in this work:

PIAA.—The detection occurs in the first focal plane after the beam remapping. Starlight therefore falls on the detector but is concentrated in a single diffraction spot. This design is adopted for practical reasons in §§ 3 and 4 to avoid having to tune a focal plane mask diameter.

PIAAC.—A focal plane mask occults starlight in the first focal plane. An inverse beam remapping unit then restores the original pupil to recover field of view. While a focal plane mask size needs to be chosen for this design, it is adopted in § 6 to maintain a sharp planet image over the zodi and exozodi backgrounds.

2.2.4. PIZZA

An interesting alternative to CPA techniques is Phase Induced Zonal Zernike Apodization (PIZZA; Martinache 2004), which achieves the pupil amplitude apodization without compromising the telescope’s throughput or angular resolution. PIZZA, which is not simulated in this work, may be too chromatic to be practically suitable for imaging ETPs. The IWAs delivered by PIZZA, CPA, and PPA are similar.

2.3. Improvements of the Lyot Coronagraph Design (APLCs, BLs, PM, 4QPM, OVCs, ODC)

In a Lyot coronagraph, the telescope pupil and the focal plane mask are not perfectly fitted one for another, resulting in strong stellar leaks (much higher than the $10^{-10}$ goal) within the coronagraph’s Lyot stop. For a Lyot coronagraph to deliver high-contrast performance, the complex amplitude in the “intermediate” focal plane (equal to the Fourier transform of the entrance pupil multiplied by the focal plane occulter) needs to be “band-limited”: its Fourier transform (=complex amplitude in the Lyot pupil plane) should have nearly zero power in a frequency range that is then selectively transmitted by the Lyot pupil stop. Improved performance over the original Lyot design can therefore be obtained by either of the following methods:

1. Adapting the pupil to the hard edge focal plane mask. This is achieved in the Apodized Pupil Lyot Coronagraph, APLC, described in Soummer et al. (2003a), where the entrance pupil of a Lyot coronagraph with a hard edge focal plane occulter is optimally apodized. A slightly different approach, explored by Vanderbei et al. (2004, see their Fig. 3), is to apodize the pupil after the hard edge focal plane occulter. As suggested by Aime & Soummer (2004), the output of an APLC can be used as the input of a second stage APLC: these are the multistep APLCs (noted APLCN, where $N$ is the number of steps).

2. Adapting the focal plane mask to the telescope pupil. This is achieved by using “band-limited masks” in the focal plane: the fourth-order band-limited coronagraph, BL4 (Kuchner & Traub 2002) and eighth-order band-limited coronagraph, BL8 (Kuchner et al. 2005). A visible nuller interferometer (VNC) equivalent to the BL4 coronagraph design adopted in this work could be designed, and we therefore refer to this BL4 design as VNC/BL4(2).

Phase mask coronagraphs (PM, 4QPM, OVC, ODC) introduce phase shifts in the focal plane. This allows smaller IWAs than offered by APLCs and BLs.

PM.—The Phase Mask coronagraph (Roddier & Roddier 1997) uses a circular $\pi$-shifting focal plane mask, and, to allow $10^{-10}$ contrast, a mild pupil amplitude apodization (Guyon & Roddier 2000; Soummer et al. 2003a). The achromatic phase knife coronagraph (Abe et al. 2001) is another form of 4QPM.

4QPM.—The 4-Quadrant Phase Mask (Rouan et al. 2000) uses a focal plane mask that shifts two out of four quadrants of the image by $\pi$. The achromatic phase knife coronagraph (Abe et al. 2001) is another form of 4QPM.

OVC.—In the Optical Vortex Coronagraph (Palacios 2005; Foo et al. 2005; Swartzlander 2006) and the Angular Groove Phase Mask Coronagraph (Mawet et al. 2005), a focal plane vortex phase mask replaces the four-quadrant phase mask of the 4QPM, thus avoiding the “dead zones” of the 4QPM. In ($r, \theta$) polar coordinates, the mask phase is equal to $m\theta$, where $m$ is the
topological charge. OVCs with high topological charges \((m > 2)\) exhibit low sensitivity to low-order aberrations.

**ODC.**—The Optical Differentiation Coronagraph (Oti et al. 2005) is using a phase and amplitude focal plane mask.

2.4. Other Coronagraph Designs and Techniques Not Studied in This Work

Some of the designs listed in Table 1 may be combined in series. The multistep APLC concept (Aime & Soummer 2004) illustrates how such a “cascade” can simultaneously improve IWA and throughput by designing each step for a lower contrast level. Nishikawa et al. (2005) showed how a nulling interferometer can benefit from a CPA, and a similar combination may also be advantageous on a circular pupil (VNC-CPA for example). The entrance pupil apodization required for the APLCs may be performed by a PIAA unit to improve IWA and throughput. Although combinations of coronagraphic techniques are not studied in this paper, the results obtained on individual designs can provide a good insight into which combinations may be most beneficial.

Pupil replication (Greenaway et al. 2005) has been proposed to transform a phase slope (equivalent to angular separation on the sky) into phase discontinuities in the replicated pupil. For the off-axis planet, this creates secondary diffraction peaks far from the central source image. The authors pointed out that it could enhance the performance of a coronagraph with which it would be combined. Riaud et al. (2005), in a sensitivity study of pupil replication, suggested to use it after a 4QPM. The tip-tilt sensitivity of both techniques (pupil replication introduces phase steps in the wave front of an off-axis source; see § 4.1 for the 4QPM) may be a concern for direct imaging of terrestrial planets.

A coronagraphic effect can be obtained by placing an occulting screen, at least as big as the telescope diameter, between the star and the telescope (Cash et al. 2005). This external occulter needs to be designed to not diffract starlight within the telescope aperture, needs to be placed far from the telescope (on a separate spacecraft), and the telescope-occulter-star alignment needs to be maintained during observations. While maintaining the alignment is especially challenging, a “conventional” imaging telescope can be used without the need for exquisite control of wave front aberrations.

3. CORONAGRAPH PERFORMANCE FOR A MONOCHROMATIC POINT SOURCE

3.1. A Coronagraph Performance Metric: the “Useful Throughput”

Since we aim at quantitatively comparing the performance of coronagraph designs, we need a “fair” performance metric: one that can be applied uniformly to all coronagraph designs. Coronagraph performance is usually quantified with IWA (usually defined as the minimal angular distance at which the throughput for the planet is half of the maximal throughput), throughput (measured at large angular separation), and search area (fraction of the focal plane image within which a planet can be detected). The IWA is somewhat arbitrary, as there is no well-defined stellar-to-companion light ratio threshold beyond which companion light suddenly becomes unmeasurable. IWA is coupled with throughput: for example, a coronagraph with a 10% maximum throughput achieving 5% throughput at \(1 \lambda/d\) has a smaller IWA than a coronagraph with a 100% maximum throughput achieving 30% throughput at \(1 \lambda/d\). In this example, the larger IWA coronagraph outperforms the smaller IWA coronagraph. Search area is also somewhat arbitrary and often cannot be quantified as a single number. For example, diffraction from four spider vanes in two perpendicular directions (we assume here that the width of the diffraction spikes created by the spiders is equal to the IWA of the coronagraph) covers \(\approx 100\%\) of the field of view at the IWA and \(\approx 0\%\) at large angular separations. Metrics commonly used to quantify coronagraph performances (IWA, throughput, and search area) are therefore difficult to use to directly compare coronagraph designs.

A simpler, more direct measure of how well a coronagraph can separate planet light from star light, the “useful throughput,” is proposed here. It is defined as the amount of planet light that can be used toward detection, and is expressed as a function of the planet position relative to the star and the planet/star contrast. When averaged over all position angles, it conveniently quantifies the coronagraph performance with a single number (making it possible to directly compare coronagraph performances), function of the angular separation. This function accurately includes the effects of IWA, throughput, and search area, and is a good estimation of the coronagraph performance.

To compute this quantity, we must first define when planet light in the coronagraph detector plane becomes “unusable” because of excessive starlight mixed with it. Theoretically, direct detection of companions embedded in much brighter starlight is possible if the starlight component on the image is well characterized. Practically, due to calibration problems and photon noise (exoplanets are faint), companion light mixed with much brighter starlight is usable. There is, however, no well-defined stellar-to-companion light ratio threshold beyond which companion light becomes unmeasurable: the criteria we will use to define the “useful throughput” (hereafter throughput) is therefore somewhat arbitrary.

As illustrated in Figure 2, the useful throughput is the maximum fraction of the planet’s light gathered by the telescope that can be separated from starlight. The “separation” criteria used is that, integrated over the area of the focal plane considered, there is as much planet light as there is residual stellar light. Physically, this area may be a single pixel or a group of pixels on a focal plane array detector.

To compute the useful throughput of a coronagraph at a particular location in the image plane, we first model the images \(I_s(x, y)\) and \(I_p(x, y)\) produced respectively by a star and a much fainter
planet in the coronagraph. Pixels are then sorted in decreasing order of $I_p(x, y)/I_s(x, y)$; for the first pixels in this sorted list, a large fraction of the light collected originates from the planet, while the last pixels of this list are dominated by starlight. The number $n$ is chosen such that, over the first $n$ pixels of this list, the total starlight collected is equal to the total planet light collected. This total amount of planet light, divided by the flux in planet light that is collected by the entrance aperture, is the *useful throughput*. We note that if, on every pixel of the image, starlight exceeds planet light, then the useful throughput will be zero.

### 3.2. Useful Throughput of Existing Coronagraph Designs

Figure 3 shows the useful throughput for the coronagraphs listed in Table 1. For small angular separations ($<1 \lambda/d$), the AIC is the most efficient coronagraph, with a 50% useful throughput at $0.5 \lambda/d$. Based on our assumptions, the phase mask coronagraphs (especially OVC2, PM, and 4QPM) and PIAA are the best choices from $0.5 \lambda/d$ outward. Many coronagraphs in this list never achieve a very high throughput because of strong entrance pupil apodization (CPA, APLC, and to a lesser degree APLCN), beam splitting (PSC, VNC), or pupil diaphragming (VNC/BL4, BL8). The range of performance is very large: the AIC offers an IWA about 8 times smaller than the CPA, and the PIAA's throughput is about 10 times the CPA's throughput beyond $5 \lambda/d$.

Many of the coronagraph designs studied in this paper were developed in the last few years, and there is therefore hope that a large number of other coronagraph designs will be discovered in the near future. It is particularly interesting to wonder how much of the upper left area of Figure 3 (high throughput at small angular distance) can/will be accessed. In the following section, we derive a fundamental limit of coronagraph throughput (shown in Fig. 3) using a “universal” model of coronagraphs.

### 3.3. Theoretical Upper Limit for the Coronagraph Throughput

#### 3.3.1. Coronagraph Model

We consider an optical system, shown in Figure 4, in which light enters through a pupil (which may or may not be circular)
that we represent by a finite (but large) number $N$ of regularly spaced elements. Similarly, the output of this optical system, defined by the points at which light is detected, blocked, or exits the system, is represented by $N$ points. A coherent light source (such as a point source at infinity) corresponds to a distribution of complex amplitude at the pupil of this system, written here as a vector $A$ of $N$ complex numbers. Although $A$ is written as a vector, it represents a two-dimensional (2D) distribution of complex amplitude across the pupil: each of its coefficients is the complex amplitude of the incident wave across one of the $N$ elements in the entrance pupil. We note that this description is only valid to the extent that complex amplitude is constant across each pupil element—this limitation is equivalent to a field of view limit in the focal plane, and our model is therefore not suitable for point sources at more than $\approx \sqrt{N}/(2d)$ from the optical axis. For convenience, we choose to normalize this vector such that the total light intensity in the pupil is unity ($|A|^2 = 1$).

Optical systems are (usually) linear in complex amplitude: the complex amplitude anywhere in the system can be written as a linear combination of complex amplitudes in the entrance pupil. For example, Fourier transforms and Fresnel propagation simulations are linear operations commonly performed to compute coronagraphic images. The output complex amplitude vector $B$ can therefore be written as

$$B(\alpha) = UA(\alpha),$$

(1)

where $U$ is a $N \times N$ complex matrix that entirely describes the optical system and $\alpha$ is the point source position in the sky. The output vector $B$ includes light that is blocked by masks (turned into heat), exits the system (for example, reflected back to the sky) or falls on the detector (these are the only coronagraph outputs that can be measured and therefore used for planet detection). Conservation of energy in the system imposes $|B|^2 = |A|^2$: $U$ is therefore a complex unitary matrix, function of the coronagraph optics, but independent of the entrance wave front (and therefore independent of the source position $\alpha$ in the sky). Wave front control elements, if present, are also part of $U$.

As illustrated in Figure 4, in a coronagraph, the goal is to create system output(s) for which an on-axis source will produce very little light, but a close-by off-axis source’s flux will be transmitted. Mathematically speaking, a matrix $U$ needs to be chosen such that some coefficients of $B(\alpha)$ (we denote $n < N$ the number of such coefficients) are very small for $\alpha = 0$ but relatively high for other values of $\alpha$. This set of coefficients corresponds to a subspace $\Omega$ of dimension $n$ (shown as the bottom part of vector $B$ on Fig. 4).

The coronagraph optimization problem can be further restricted by fixing a 2D vector $A(\alpha_p)$, where $\alpha_p$ is the planet off-axis position at which the coronagraph’s throughput is to be maximized. In this case, a change of coordinate system (rotation) within the subspace $\Omega$ can concentrate all the planet flux within $\Omega$ on a single coefficient. This change of coordinate system is a function of $\alpha_p$ and aligns, within $\Omega$, the 2D vector $A(\alpha_p)$ with one of the directions of the new coordinate system. Physically, this is equivalent to placing multiple beam splitters and phase shifters to concentrate all of the planet light within $\Omega$ in a single output beam: this would be possible thanks to the fact that planet light is fully coherent.

This norm-preserving coordinate change is itself a unitary matrix that can be integrated within $U$ (by direct multiplication with the “old” $U$), in which case only a single column of this new $U$ matrix is now relevant to the problem: we denote $C$ this column ($C$ is a vector). Since $U$ is unitary, $|C| = 1$. In this new basis, $|C \cdot A(\alpha)|^2$ is the square absolute value of one of the coefficients of the new vector $B$ and is equal to the coronagraph throughput at the sky position $\alpha$: it is therefore the quantity to be minimized for $\alpha = 0$ and maximized for $\alpha = \alpha_p$.

3.3.2. Coronagraphic Throughput Upper Limit in the Ideal Case

We use in this section the base described at the end of the previous section, where the subspace $\Omega$ corresponds to a single coefficient of $B$. In this “optimized coronagraphic projection,” the energy in this single coefficient is the useful throughput of the coronagraph, as long as, within this single “pixel,” the stellar flux is smaller than the planet flux. Excluding all other coefficients of $B$, we can now equivalently refer to the coronagraph throughput or useful throughput.

We denote $\epsilon$ the on-axis coronagraph throughput (no more than $10^{-10}$ for a system designed to image ETs in the visible) and $A(0)$ the input complex amplitude for an on-axis source. We note that $\epsilon$ is a throughput, not a contrast, and should therefore ideally be no more than $10^{-10}$ times the coronagraph throughput for the planet. It is now possible to design an ideal theoretical coronagraph for detecting a source at the position $\alpha_p$ by choosing a complex vector $C$ of norm 1 such that:

1. $|C \cdot A(0)|/|C \cdot A(\alpha_p)| < \sqrt{\epsilon}$: this is the coronagraphic “contrast” at position $\alpha_p$.
2. $|C \cdot A(\alpha_p)|$ is as large as possible: this is the square root of the coronagraph intensity throughput at position $\alpha_p$.

This way of looking at the problem shows that optimizing a coronagraph is equivalent to choosing a projection axis $C$ that is “mostly” perpendicular to $A(0)$ (constraint [1]) and as colinear as possible with $A(\alpha_p)$ (constraint [2]). A glance at Figure 5 reveals that if $A(\alpha_p)$ and $A(0)$ are close (i.e., the planet position $\alpha_p$ is close to the optical axis), the coronagraphic throughput cannot be high. As shown in Figure 5, the input vector $A(\alpha_p)$ can be decomposed into an “along optical axis” component and a “perpendicular to optical axis” component:

$$A(\alpha_p) = A_\| (\alpha_p) + A_\perp (\alpha_p),$$

(2)

where

$$A_\perp (\alpha_p) = |A(\alpha_p) \cdot A(0)| A(0)$$

(3)

is the projection of $A(\alpha_p)$ onto $A(0)$ (“along optical axis”).

Using this decomposition, the linearity of the coronagraph in complex amplitude, and the $|C \cdot A(0)| < \sqrt{\epsilon}$ constraint, an upper
bound can be established for the square root of the coronagraphic throughput:

$$|A(\alpha_p) \cdot C| < |A_\parallel(\alpha_p) \cdot A(0)|/\sqrt{\epsilon} + |A_\perp(\alpha_p) \cdot C|.$$  

(4)

Since $\epsilon \ll 1$ and $||C|| = 1$,

$$|A(\alpha_p) \cdot C| < ||A_\perp(\alpha_p)||.$$  

(5)

The upper limit given by equation (5) is shown in Figure 6 for a “normal” circular pupil (unobstructed, unapodized) as well as for Gaussian-apodized pupils and a pupil with a large 80% central obstruction. Not surprisingly, amplitude apodization decreases the throughput at all distances and also pushes further the point at which the maximum throughput is reached (decrease in angular resolution). With a large central obstruction, the throughput reaches its maximum value quicker (slight increase in angular resolution).

This result shows that a coronagraph, regardless of its on-axis throughput (as long as it is much smaller than 1 of course), cannot have a throughput exceeding 50% at 0.5$\lambda/d$ for a circular pupil. Curiously, the upper limit derived by equation (5) oscillates with distance, especially in the large central obstruction case. This is due to a slight similarity in the wave front at $0 \lambda/d$ and $1.6 \lambda/d$ for the circular pupil case. In the extreme case of a two-pupil interferometer of baseline $b$, the same effect would lead to a $\lambda/b$ periodicity of the theoretical maximum throughput.

Since $||A_\perp||^2 = ||A(\alpha_p)||^2 - ||A_\parallel||^2$ (right angle in Fig. 5), $||A(\alpha_p)|| = 1$ by definition, and $||A_\perp(\alpha_p)||$ is the intensity of the nonaberrated noncoronagraphic PSF at position $\alpha_p$ (eq. [3]): the maximum throughput of a coronagraph is equal to 1 minus the nonaberrated noncoronagraphic PSF of the telescope. In the case of a nonobstructed circular pupil, this throughput is 1 minus the Airy pattern, as can be seen in Figure 6.

This last property is a direct consequence of the linearity in complex amplitude of the coronagraph. Since the coronagraph needs to remove the central star, it also removes the “flat nonlittled component” of any incident wave front, which, expressed as a function of the source position on the sky, is the Airy function for an unobstructed pupil.

3.3.3. Can This Upper Limit Be Reached in an Optical System?

With the above analysis, the vector C that would be required to reach the upper limit at a given position $\alpha_p$ can be computed.

Each of the $N$ elements of the pupil can be treated as a small independent beam, and we could construct the optical system with a series of beam splitters (not necessarily 50%/50%) and phase shifts. Each beamsplitter combines two beams and produces two output beams. Since $|C \cdot C| = 1$, the problem is equivalent to concentrating all the flux into a single beam if the entrance vector ($N$ beams) is equal to $C$. It is relatively easy to do incrementally: at each step two beams are combined with a phase shift + beam splitter such that all the light is sent into one output beam (this is possible since the beams are fully coherent between them and we know exactly their amplitudes and phases). $N - 1$ beam splitters are required in this scheme. If the coronagraph is only intended to work close to the optical axis, $N$ could be quite small ($N$ scales as the square of the outer working angle), resulting in an optical system that could possibly be built. Unfortunately, this coronagraph is only guaranteed to reach the theoretical limit at a single planet position $\alpha_p$.

A better way to build the “ideal” coronagraph is to have the light of an on-axis point source entirely concentrated in a single output. This would ensure that, for large values of $N$, the coronagraphic throughput (now measured on $N - 1$ beams) is equal to the theoretical limit simultaneously at all source positions. Special attention would be given to the arrangement of the beam splitters that only affect the $N - 1$ “useful” outputs in order to allow the source position to be easily retrieved from comparisons between their intensities.

4. Sensitivity to Tip-Tilt: The Stellar Angular Diameter Problem

4.1. Effect of Stellar Angular Diameter in Existing Coronagraph Designs

The sensitivity of some coronagraphs to tip-tilt errors has been investigated by several authors (see, e.g., Lloyd & Sivaramakrishnan 2005; Shaklan & Green 2005; Sivaramakrishnan et al. 2006).

A Sun-like star at 10 pc is 1 mas across: with an 8 m visible telescope its apparent radius is about $0.1 \lambda/d$. Many coronagraphs cannot maintain a 10$^7$ contrast on such an extended central source. Even if the telescope diameter were reduced (or the wavelength increased) such that the planet, at maximum elongation, would be at $2 \lambda/d$, the stellar radius would still be $0.01 \lambda/d$.

Figure 7 shows how coronagraphic performance (useful throughput) decreases as the stellar radius increases and can be used to somewhat arbitrarily divide coronagraphs into three groups:

1. High sensitivity to stellar radius (AIC, 4QPM, PM, OVC2, ODC). These coronagraphs cannot maintain high contrast even with a $0.01 \lambda/d$ stellar radius. They would be very difficult to use for ETI imaging.

2. Moderate sensitivity to stellar radius (VNC and BL4 with small IWA/large shear: VNC/BL4(1), PSC, APLCs, OVC/ with large value of $N$). These coronagraphs may be used efficiently on “small” diameter telescopes (for which the stellar radius is about $0.01 \lambda/d$ or less).

3. Low sensitivity to stellar radius (PIAAC, VNC and BL4 with large IWA/small shear, BL8, CPA). These coronagraphs are suitable for direct imaging and characterization of ETIs, as they tolerate stellar radii of $0.1 \lambda/d$.

The robustness of VNC/BL4(1) compared to the VNC/BL4(2) is entirely due to the larger IWA chosen in this study for the VNC/BL4(2). For both the BL4 and VNC, there is a IWA versus sensitivity to stellar radius trade-off.

For most coronagraphs, increasing the stellar diameter first compromises contrast at small angular separations. Even the
VNC/BL4(2) and PIAAC, which are relatively insensitive to angular size, lose throughput at small angular separations in the 0.1λ/d radius case.

The CPA and BL8 stand out as being extremely robust to stellar angular size: their throughput curves are almost identical from 0 to 0.1λ/d stellar radius. The CPA owes its relative immunity to tip-tilt errors to the field-invariance of its PSF. The sensitivity of the BL8 to tip-tilt errors was studied in detail by Shaklan & Green (2005), who also find that the BL8 can tolerate relatively high levels of focus, coma, and astigmatism.

Direct comparison between Figures 3 and 7 shows a very disappointing result: the best coronagraphs in the point source case (AIC, PM, 4QPM, OVC2) all perform very poorly when the stellar angular size is considered. Increasing the stellar size from 0.1λ/d to 0.1λ/d has moved the smallest angular separation at which 50% throughput is reached from 0.5λ/d (for the AIC) to slightly more than 2λ/d (for the PIAA). Is this behavior imposed by fundamental physics and therefore unavoidable? Or is there hope to find a low-IWA coronagraph that is not so dramatically sensitive to stellar angular size?

4.2. Throughput Limit Imposed by Stellar Angular Size

In this section, we quantify how the angular extent of the source affects the coronagraph throughput. We assume that the coronagraph is observing a disk of uniform surface brightness and angular radius rs. Using the same framework as in § 3.3, the star is now modeled by a series of vectors $A_k = A(\alpha_k)$, where the $\alpha_k$ values are uniformly distributed on the stellar disk ($|\alpha_k| < r_s$) and $k = 0 \ldots N_s - 1$. An orthonormal base can be built from the vectors $A_k$ with the following process:

1. The vector $M_i$ ($i = 0$ initially) of this base is chosen to be colinear with the vector $A_k$ with the highest norm.
2. All vectors $A_k$ are replaced by their projection on a hyper-plane perpendicular to $M_i$.
3. $i$ is incremented, and we return to step 1 with the new vectors $A_k$.

This algorithm ensures that, in the orthonormal base obtained, the coefficients of the vectors $A_k$ decrease very rapidly. We denote $m_i, i = 0 \ldots N - 1$ the maximum absolute value of the coefficient $i$ in this new base across all $A_k$. The 2D representation of the first five vectors $M_i$ are shown in Figure 8. The same modes can be obtained by linear expansion of the pupil complex amplitude $e^{i(\omega x + \gamma y)}$ for small values of $u$ and $v$ [here pupil plane coordinates are $(x, y)$ and point source angular coordinates on the sky are $(\alpha, \delta)$].

As previously shown, a coronagraph that needs to cancel the light across the stellar disk can be represented by a vector $C$ of norm 1. We denote $c_i, i = 0 \ldots N - 1$ the coefficients of $C$ in the new base of modes $M_i$. To ensure that the coronagraph cancels the light of the star by a factor $\epsilon$, the following constraint on each $c_i$ can be imposed:

$$|c_i| < \sqrt{\epsilon/m_i}. \quad (6)$$

This constraint is not rigorously equivalent to stating that the integrated starlight transmitted by the coronagraph is smaller than $\epsilon$, but we have verified numerically that it yields a stellar throughput that is in most cases within 20% of $\epsilon$. This can be explained by the very rapid decrease of $m_i$ with $i$ (especially for...
smaller stars) and the rapid quadratic increase of stellar leaks if $|c_i|$ goes above $\sqrt{\epsilon/m_i}$. For modest stellar sizes (below $\lambda/d$), equation (6) constrains only the first few values of $c_i$ (the others are constrained by $|c_i| < 1$). We note that adopting a zero stellar size will result in $m_0 = 1$ and $m_i = 0$ for $i > 0$. In this case, equation (6) only imposes a constraint for $i = 0$. To obtain the maximum throughput as a function of angular distance, the vector $A(c_i)$ is decomposed in the new base, and its first coefficients are multiplied by the values of $|c_i|$ derived in equation (6).

The results obtained, shown in Figure 9 and also overplotted in Figure 7, confirm the large effect of a small increase in stellar size on the coronagraphic IWA. In the point source case, the 50% throughput can be reached at $0.5\lambda/d$; For a $0.1\lambda/d$-radius source, the 50% throughput limit moves to almost $2\lambda/d$. This behavior is contrary to a “geometrical optics” intuition, which would suggest that increasing the stellar size by $0.1\lambda/d$ moves IWA by $0.1\lambda/d$. In the linear algebra model of coronagraphy presented in this work, this behavior is, however, understandable: as the stellar diameter increases, a higher number of modes needs to be removed from the wave front; each mode removed increases the coronagraph’s “blind spot” by a $(\lambda/d)^2$ area in the focal plane.

A coronagraph that follows exactly the limit shown in Figure 9 can be assembled from discrete beam splitters and phase shifts as described in § 3.3.3. An example of such a design is shown in Figure 10. A $c_0 = 0$ coronagraph is often referred to as a second-order null coronagraph, while $c_0 = c_1 = c_2 = 0$ ensures a fourth-order null (and so on).

4.3. Coronagraphic Imaging at Small Angular Separation ($<\lambda/D$)

In this section, our linear algebra-based model is used to predict and explain how different coronagraphs behave at small angular separation. These results allow us to understand how coronagraphic performance degrades as stellar angular diameter increases, and allow us in § 4.4 to determine if wave front control techniques can mitigate this problem.

4.3.1. Second-Order Coronagraphs (IAC, PM, 4QPM)

The smallest-IWA coronagraphs considered in this study (IAC, PM, 4QPM) all remove mode M0 (see Fig. 8) from the incoming wave front but “transmit” (at least partially) modes M1 and M2. At very small angular separations, the “transmitted” wave front is therefore dominated by a combination of modes M1 and M2. The coefficients of this combination are proportional to the off-axis distance, resulting in a coronagraphic residual light intensity proportional to the square of the source’s angular separation (second-order coronagraphs). It is therefore expected that for these coronographs, the complex amplitude in the final image is also
proportional to the source coordinates: the complex amplitude for a source at position \((u, v)\) is the opposite of the complex amplitude for a source at position \((-u, -v)\). At small angular separations, our theoretical analysis therefore predicts that the images (square modulus of the complex amplitudes) of two point sources on opposite sides of the optical axis are identical and that, as the source moves away from the optical axis, its image is simply multiplied by a constant but does not change structure. Both these predictions are confirmed by numerical simulations, as shown in Figure 11 (top three rows).

4.3.2. Other Small-IWA Coronagraphs (ICC6, OVC6, PIAAC)

Just as in the case of second-order coronagraphs, the linear algebra-based theory outlined in the previous sections can predict/explain the behavior of other coronagraphs at small angular separation.

The Interferometric Combination Coronagraph (ICC6), which will be described in § 5.3, perfectly removes the first six modes shown in Figure 8, but leaves other modes untouched. The “transmitted” wave front of this theoretical coronagraph at small angular separation is therefore dominated by the \(-i(xu + yv)^3/6\) term in the polynomial expansion of \(e^{i(xu + yv)}\), corresponding to modes \(x^3, y^3, xy^2,\) and \(x^2y\) (not shown in Fig. 8). We can therefore predict that along any direction, its PSF is a fixed image multiplied by the sixth power of source angular separation and that PSFs for two point sources on opposite sides of the optical axis will be identical. A similar behavior is seen in Figure 11 for the OVC6, which also removes the first six modes shown in Figure 8. Since this coronagraph also attenuates further modes, the PSF structures and throughput are, however, different from the ICC6, which is designed to perfectly cancel the first six modes and fully transmit the others.

Finally, the PIAAC is a more complex (but also more representative of “real world” coronagraphs) case: it strongly attenuates, but does not fully remove, modes 1 to 6 of Figure 8. At very small angular separation (up to \(\approx 10^{-3} \lambda/d\)), mode \(M0\) dominates the coronagraphic residual: the PSF is independent of source position. In the \(10^{-3}\) to \(10^{-2} \lambda/D\) range, a small residual in modes \(M1\) and \(M2\) dominates and the coronagraph behaves as a second-order coronagraph (see § 4.3.1 above). At a 0.1\(\lambda/d\) separation and beyond, no single mode is dominating, and several partially transmitted modes interfere and/or are together to form the image: the PSF clearly shows on which side of the optical axis the source is, a behavior that cannot be achieved if the image is dominated by a single mode since individual modes are either symmetric or anti-symmetric in complex amplitude (and therefore always symmetric in intensity).

4.4. Speckle-nulling Techniques, Stellar Size, and Pointing Errors

Active elements (deformable mirrors) can be used to create a coherent “anti speckle halo” that destructively interferes with the unwanted residual focal plane speckles. These speckle-nulling techniques, first proposed by Malbet et al. (1995), were more recently studied by several authors (Codona & Angel 2004; Labeyrie 2004; Guyon 2005; Borde & Traub 2006) and successfully tested in laboratory (Trauger et al. 2003).

4.4.1. Can Speckle-nulling Techniques Mitigate Coronagraphic Leaks Produced by Stellar Angular Size?

The star can be modeled as a set of point sources uniformly distributed on the stellar disk (these sources are not mutually coherent). We first assume that a second-order coronagraph is used,
and we consider only one of these point sources (not at the center of the disk) that contributes to the coronagraphic leakage by adding an image on the focal plane. This particular image is fully coherent and can therefore be canceled by a speckle-nulling technique. The shape/state of the active optical element(s) (DM, for example) in the speckle control device will be such that the phase and amplitude of the “antispeckle halo” is the opposite (same amplitude, π phase offset) of the leakage: destructive interference between the two will leave no light in the focal plane. We now decide to “freeze” the state of the speckle-nulling device and move the point source on the other side of the optical axis. This multiplies the complex amplitude of the coronagraphic leakage by −1 (see § 4.3.1) but does not change the “antispeckle halo” (which is independent of source position within the small angular displacement considered). The leakage and “antispeckle halo” will therefore add rather than subtract because of a phase mismatch, and the leakage intensity will be 4 times higher than without speckle-nulling. Every point on the stellar disk corresponds to an equally bright point on the other side of the optical axis: speckle-nulling cannot reduce coronagraphic leaks in a second-order coronagraph.

The same “phase mismatch” argument applies to sixth-order coronagraphs such as ICC6. Independently of this “phase mismatch” effect (which would not apply to a fourth-order coronagraph dominated by modes M3 to M5), coronagraphic leaks increase as the point on the stellar disk moves away from the optical axis, and there is therefore a similar “amplitude mismatch” effect when stellar points at different radii are considered. The stellar angular size problem described in § 4.1 truly is a fundamental limit.

4.4.2. Pointing Errors and Low-Order Aberrations

Pointing errors (or any higher order optics aberrations), on the other hand, produce a coherent leak that can be suppressed by wave front control techniques. Low-order aberrations can be measured from the coronagraphic leaks they produce. A more optimal strategy is to “catch” aberrations before they start to increase coronagraphic leaks by utilizing “for free” the large amount light that is blocked by the coronagraph. This can usually be done by reimaging of the light reflected by the focal plane occulter (or, for phase mask coronagraphs, the light outside the geometric pupil). We note that speckle sensing and nulling usually require a small spectral bandwidth, as the speckle coherence is low in white light: the observing spectral band may have to be split with dichroics in small “bins” prior to applying speckle-nulling.

5. ZODIACAL AND EXOZODIACAL LIGHT

5.1. Mixing of Incoherent Background Light with the Planet’s Light: The Diffractive Efficiency Factor

The “useful throughput” introduced in § 3.1 accurately quantifies the coronagraphic performance in a system composed solely of a star and a single planet. A planetary system is, however, observed against incoherent zodiacal and exozodiacal light backgrounds. For simplicity, we assume here that both background components are constant across the field (this is an approximation for the exozodiacal light) and that they can be numerically subtracted to the photon noise level.

We consider a planet + background system (no central star) in which β is the ratio between total planet light and total background light within the planet image (which we define here, for convenience, as the smallest focal plane area containing half of the planet’s PSF light). We denote β0 this value for a non-coronagraphic imaging telescope (Airy pattern). Coronagraphs unfortunately tend to decrease β (β < β0) by a factor that we denote as the diffractive efficiency factor (DEF):

1. Pupil amplitude apodization and/or clipping in the Lyot plane increases the angular size of the planet image, therefore mixing more background light with the planet light.
2. The AIC delivers double images of the planet, therefore dividing β by 2.
3. Coronagraphs delivering PSFs in which a significant fraction of the light is outside the main diffraction core decrease β.

Three factors determine the DEF: the width w of the PSF core (more exactly, the square root of the “useful” PSF area), the fraction f of the PSF energy in the diffraction core, and the total throughput T. The collectible planet energy is proportional to fT, while the background flux collected in the same focal plane area is proportional to w2T; the DEF is therefore proportional to f/w2.

The diffractive efficiency factors of most of the coronagraphs in Table 1 are listed in Table 2, along with other key characteristics. The DEF is especially low on the CPA and BL8 due to their poor angular resolution. As discussed in § 2.2.1, a CPA with a binary mask (shaped pupil) could “amplify” background levels by another factor of 2: the combined effect is then equivalent to dividing the planet/background ratio by ≈7.

5.2. Background Level

The zodiacal light background level is about mV = 22.5 arcsec−2 (can be less in some directions), and an Earth-like planet (6400 km radius, 0.33 visible albedo, orbiting a Sun-like star at 1 AU), at maximum elongation, is mV = 24.1 + 5 log (d), where d is the system distance in parsecs. If we assume that the zodiacal, exozodiacal, and planet spectra are identical, the
In a classical noncoronographic imaging telescope, the system matrix $U$ shown in Figure 4 is a Fourier transform, offering full angular resolution imaging and DEF = 1 but poor coronagraphic performance: we denote this “imaging” matrix as $U_t$. On the other hand, a “coronagraphic” matrix $U_c$ optimized for coronagraphic performance, such as the one shown in Figure 10, efficiently isolates starlight but suffers from poor imaging capability and high diffractive efficiency factor.

The same iterative method used to produce the “vectors” $C$, $X$, and $Y$ in Figure 10 can be used to produce any predefined unitary matrix. Unitary matrices can also be cascaded: the input of a matrix, instead of being the telescope pupil, can be the output of another matrix. For example, it is possible to cascade a “pupil to focal plane” matrix $U_f$ with a “focal plane to pupil” $U_f^{-1}$ matrix, with no net result (the product of the two matrices is the unity matrix).

One such cascade is shown in Figure 12: it is designed to produce a “coronagraphic intermediate step” where all starlight is easily masked in the “coronagraphic intermediate step,” with minimal impact on the planet light (unless the planet is very close to the star) and background light (which is always spread across all outputs). The sequence of optical steps to perform is therefore:

1. Select the transformation $U_c$ that best isolates stellar light from planetary light.

\[d_z(\text{pc}) = \frac{2.3 \times \text{DEF} \times D(\text{m})}{\sqrt{(1+2z) \lambda(\mu\text{m})}}, \tag{7}\]

where $z$ is the exozodiacal cloud dust content, normalized to our own system, and DEF is the diffractive efficiency factor described in §5.1. For example, with $z = 1, \text{DEF} = 0.5, \lambda = 0.55 \mu\text{m}$ and $D = 4 \text{ m}, d_z = 4.5 \text{ pc}$: for all but a few targets, photon noise from the background level will be the main source of noise and will drive the exposure time required for successful detection of ETPs. At longer wavelength, the PSF size increases, therefore gathering more zodi + exozodi background and reducing $d_z$.

### 5.3. Can an “Ideal” Coronagraph Be Built with DEF = 1?

In §4.2, we have demonstrated that a coronagraph that delivers theoretically ideal throughput on a partially resolved star can be designed using beam splitters and phase shifts. The example shown in Figure 10 achieves this goal by confining starlight in a few interferometric outputs (as the required contrast and/or stellar diameter increases, the number of such outputs grows). However, no care was given to the distribution of planet light across the remaining outputs, and it is likely that background and planet light are highly mixed, resulting in high diffractive efficiency factor and poor imaging capabilities. We show here that a coronagraph can be designed to simultaneously deliver optimal throughput for a partially resolved star (see §4.2), full angular resolution imaging and DEF = 1.
In the first step, multiplication of the complex amplitude in the pupil plane by the coronagraphic matrix $U_c$ concentrates starlight in a few outputs. This operation can be inverted by multiplication by $U_c^{-1}$, therefore restoring the telescope pupil. Multiplication by $U_i$ produces a clean “image” of the planet. At each step, distribution of stellar (vector on the left), planet (center vector) and background (right vector) light is shown in gray scale: from white for no light to black for maximum light. The “ideal” coronagraph is obtained by simply masking the few outputs containing starlight in the “coronagraphic outputs” intermediate step.

2. Block starlight in the few outputs where it has been concentrated.

3. Undo transformation 1 to recover the original telescope pupil ($U_c^{-1}$).

4. Select the transformation that concentrates the planet light in a single bright diffraction peak (imaging $U_i$).

In an optimized implementation of this scheme, where the number of optical elements is to be minimized, there is no need to separate steps 2 and 3, as the intermediate product (the original telescope pupil) is not used.

An identical scheme is used in the PIAAAC design, although this coronagraph falls short of the theoretically ideal performance derived in this work. The PIAAAC first concentrates stellar light in a very tight diffraction core (step 1), which is then blocked by a mask (step 2). The pupil remapping used to perform step 1 unfortunately “scrambles” the planet image beyond $\approx 10 \lambda/d$. A second remapping unit is therefore then used to undo the first remapping (step 3). Finally, an image is formed (step 4).

Since the only currently known solution to build the ideal coronagraph requires many beam splitters, this type of coronagraph is referred to as the Interferometric Combination Coronagraph (ICC$N$, where $N$ is the number of modes removed in the coronagraphic intermediate step) in this work.

We have used the above method to simulate an ICC$6$, designed to remove the six modes shown in Figure 8. The resulting coronagraphic throughput is of the eighth-order (just as for the BL8 coronagraph), with an IWA of 1.32 $\lambda/d$. This particular design is well suited for direct imaging of ET$P$s and accurately represents the absolute limit of what coronagraphy can achieve for this purpose. The useful throughput versus angular separation of the ICC$6$ is virtually identical to the “theoretical limit” shown in the upper panel (stellar radius = 0.01 $\lambda/d$) of Figure 7.

The performance of this “ideal” coronagraph (as well as the performance of a few other coronagraphs) is modeled in more detail in the following section. Table 2 concludes the previous sections by listing the main characteristics of the coronagraphs studied in this paper.

6. MONTE CARLO SIMULATIONS

In the previous sections, coronagraph performance has been quantified by the useful throughput and the diffractive efficiency factor. Together, these two quantities could be used to evaluate, for a given target (star + planet + background), the signal-to-noise (SNR) as a function of exposure time. This simple model can unfortunately be quite inaccurate:

1. Exozodiacal light is not a smooth background (it is significantly brighter closer to the star and in a high inclination system, the exozodi disk can be a narrow linear feature): its effect cannot be accurately modeled by the diffractive efficiency factor. Planets are brighter when in nearly full phase, but are then closer to the star and therefore over brighter portions of the exozodi cloud.

2. In the presence of background light (zodiacal+exozodiacal light), faint parts of the PSFs that contribute to the useful throughput can have a negligible contribution to the detection SNR: the spatial distribution of the planet light affects the SNR.

While the useful throughput and the DEF provided us with valuable insight into coronagraph performances, accurate estimation of the detection sensitivity requires explicit computation of images delivered by the coronagraph.

In this section, the performance of six promising coronagraphs identified in the previous sections [CPA, PIAAC, BL8, OVC6, ICC6, and VNC/BL4(2)] is quantified for the direct detection of Earth-type planets. Details of this simulation are given in § 6.1. As the goal of this study is to show what can be potentially achieved with various coronagraphs, very optimistic assumptions have consistently been made. They are listed in Table 3 and further detailed in the next section.

6.1. Description of the Simulation

6.1.1. Initial Sample of Targets

A list of potential targets was first created by selecting stars within 100 pc in the Hipparcos catalog. The effective temperature of each star was computed from the $B-V$ color index, using the color index versus temperature law of main-sequence stars. Stars with effective temperature between 4060 and 7200 K (types F, G, and K) were then selected, and those listed as subarcsecond separation doubles removed. Giants were excluded from our sample using the stellar bolometric luminosity/effective temperature criteria outlined in Figure 13.

This sample now contains 16,921 stars. Each star is assumed to have an extrasolar planet (ET$P$). In this section, ET$P$ is used to mean a standard Earth analog: Lambert phase function, radius =6400 km, albedo =0.33, circular orbit at 1 AU times the square root of the star’s bolometric luminosity (computed assuming the star is a perfect blackbody radiator). The inclination of the orbit, assumed to be coplanar with the exozodi disk, is set to $\pi/3$ for all systems: this is the median value expected for a random system orientation.

A final selection criteria is to reject systems for which the planet, at maximum elongation, is within $\lambda/d$ of the star: none of the coronagraphs studied in this work can efficiently detect planets within this angular separation. For an 8 m telescope observing in the 0.5–0.6 $\mu$m band, this final sample contains 3429 stars at an average distance of 51 pc.
6.1.3. Exozodiacal Light

The exozodiacal cloud is assumed to be a thin disk with dust optical depth, as seen from above the 2D disk, varying as \( r^{-0.34} \) (surface brightness \( \propto r^{-2.34} \)). This power law index is derived from modeling of our own zodiacal cloud in the Earth neighborhood (Kelsall et al. 1998). For a 1 zodi exozodiacal disk, the dust optical depth at the location of the planet is chosen to be identical to the vertical optical depth of our zodiacal cloud at 1 AU. In this model, the 0.55 \( \mu \)m surface brightness of a face-on exozodi cloud at the location of the planet is thus a function of the ratio between the star bolometric luminosity (which defines where the planet is) and the star luminosity at 0.55 \( \mu \)m (which defines how much the star illuminates the dust). The inner edge of the exozodi cloud is set at 0.02 AU times the square root of the star’s bolometric luminosity (dust sublimation), and its outer edge is 5 times the planet-star distance. The image of the exozodiacal cloud delivered by the telescope/coronagraph is computed with the same 0.2\( \lambda /d \) sampling as used for the zodiacal background image.

6.1.4. Detection Metrics: Signal-to-Noise Ratio (SNR) and Detection Probability

It is assumed that the stellar diameter, exozodi cloud, zodiacal background, and telescope/coronagraph characteristics are perfectly known. The planet detection SNR is therefore driven by photon noise and is proportional to the square root of the exposure time. The photon noise (from the planet, star, zodiacal, and exozodi images) and the signal (the planet image) are evaluated for each pixel of the image. The planet detection SNR is obtained by the quadratic sum of the pixel SNRs: this optimal weighting scheme favors pixels with large planet flux and low zodi + exozodi + star flux.

We assume here that at the time of the observation, the position of the planet along its orbit is unknown. For a given exposure time, the planet detection SNR is computed at each point along the planet orbit. A limit on the SNR is adopted (SNR > 7) for successful detection, and the number of points along the planet orbit for which this limit is reached yields the planet detection probability. This probability, also referred to as completeness (Brown 2004, 2005), is therefore a function of exposure time.

A related metric used in this study, derived from the planet detection probability, is the single-exposure integration time required to reach a given planet detection probability.

6.2. Example: HIP 56997

Figure 14 shows some example frames obtained by the simulation code for the six coronagraphs selected. These images nicely illustrate coronagraphic characteristics quantified in the previous section:
Useful throughput at large separation.—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.

**Fig. 14.**—Simulated 4 hr exposures of HIP 56997 and an hypothetical Earth-type planet at maximum elongation for telescope sizes ranging from 2 to 12 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° 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 ≈ 59°. Each image is (20 × 20);/d, and the planet-star separation is 80 mas.

**TABLE 4**
THEORETICAL EXPOSURE TIMES REQUIRED TO REACH 20% AND 50% DETECTION (SNR = 7) PROBABILITY ON HIP 56997 IN THE 0.5–0.6 μM BAND

| Telescope Diameter (m) | CPA    | VNC/BL4(2) | BL8   | PIAAC | OVC6  | ICC6  |
|------------------------|--------|------------|-------|-------|-------|-------|
|                        | 20% Detection Probability |        |       |       |       |       |
|                        |        |            |       |       |       |       |
| 2                      | a      | 6.9 hr     | 14 minutes | 44.8 minutes | 5.7 minutes | 8.6 minutes | 3.78 minutes |
| 4                      |        | 12.4 hr    | 2.1 hr | 11.4 hr | 34 minutes | 48 minutes | 14.3 minutes |
| 6                      |        | 63 minutes | 4.8 minutes | 17 minutes | 2.1 minutes | 3.1 minutes | 110 s |
| 8                      |        | 12 minutes | 2.7 minutes | 4.3 minutes | 45 s | 56 s | 42 s |
| 12                     |        |            |       |       |       |       |       |

|                        | 50% Detection Probability |        |       |       |       |       |
|                        |        |            |       |       |       |       |
| 2                      | 6.9 hr | 14 minutes | 44.8 minutes | 5.7 minutes | 8.6 minutes | 3.78 minutes |
| 4                      |        | 12.7 hr    | 2.3 hr  | 2.2 hr  | 44.1 minutes |       |       |
| 6                      |        | 63 minutes | 4.8 minutes | 17 minutes | 2.1 minutes | 3.1 minutes | 110 s |
| 8                      |        | 12 minutes | 2.7 minutes | 4.3 minutes | 45 s | 56 s | 42 s |
| 12                     |        |            |       |       |       |       |       |

* Exposure times above 1 day are not shown in this table, as they are practically unrealistic.
Ability to work at small angular separation.—None of the coronagraph tested can detect the planet on a 2 m telescope with the exposure time, wavelength, and throughput used in Figure 14. 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 is needed for detection with a CPA.

Sensitivity to stellar angular size.—The CPA and BL8 are extremely robust to stellar angular size: starlight leaks are virtually nonexistent even on the 12 m telescope. With the VNC/BL4(2), PIAAC, OVC6, and ICC6, starlight is visible on the 12 m telescope (equivalent to a 6 m telescope observing the same system at 5 pc), although it is still fainter that the exozodiacal contribution.

Images similar to the ones shown in Figure 14 can be generated for each possible position of the planet along its orbit. These images have been used to compute for HIP 56997 the exposure time required to reach 20% and 50% detection (SNR = 7) probability. Results, in Table 4, show very large discrepancies between

Fig. 15.—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) with exozodi levels ranging from 1 zodi (top) to 10 zodi (bottom). The number of accessible targets for which the required exposure times are less than 1 hr, 10 hr and 1 day are listed for each case in the gray 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.
coronograph designs. For example, on a 6 m telescope, the PIAAC requires 44 times less exposure time to reach a 50% detection probability than the CPA. For this target, the performance of the PIAAC, OVC6, and ICC6 coronographs are similar, while the other three coronographs are significantly inferior (especially the CPA). For the VNC/BL4(2), it was assumed that the preferential direction of the coronagraph was optimally aligned with the system orientation.

6.3. 4 and 8 m Diameter Telescopes Observing in the 0.5–0.6 μm Band

In the previous section, a single target was adopted while the telescope size was variable. Here we adopt fixed telescope sizes and evaluate the telescope/coronagraph combination performance on a sample of nearby stars, defined in 6.1.1. Results are shown in Figure 15 for the 4 and 8 m telescopes.

In a purely background-limited detection, exposure time is expected to scale linearly with the background level: approximately 7 times longer for the 10 zodi exozodi case than for the 1 zodi exozodi case (assuming the local zodi background corresponds to half a zodi of exozodiacal light). The fact that this is generally true in Figure 15 confirms the importance of the DEF in coronagraph designs. For coronographs with low DEFS, such as PIAAC, OVC6, and ICC6, increasing the exozodi, however, results in moderate increase of exposure time for nearby targets (non-background-limited detections). This effect is most clearly visible in the 8 m telescope case for the easiest targets.

Figure 15 compares how many targets fall below the 1 hr, 10 hr, and 1 day required exposure time levels with different exozodi contents. With the 4 m telescope, the CPA and BL8 cannot access the “habitable zone” of targets beyond ∼5 pc, resulting in a very small number of potential targets (8 and 16 targets respectively for CPA and BL8 if “potential target” is defined as one that requires less than 10 hr exposure time to reach the 50% detection probability). The VNC/BL4(2), thanks to its higher throughput, can access 34 targets. With this telescope diameter (4 m), the stellar angular size is sufficiently small to allow improved performance over CPA and BL8, which are more robust to stellar angular size. In increasing order of performance, the OVC6, PIAAC, and ICC6 offer from 87 to 177 targets in the 1-zodi case. It is interesting to note that the performance of the OVC6 and PIAAC is very similar, while the “ideal” ICC6 offers almost twice as many targets as the PIAAC. This is mostly due to the benefit of smaller IWA in the ICC6 over the PIAAC: 1.32λ/d versus 1.88λ/d, which increases the IWA-accessible volume by 2.9. The planet intrinsic brightness and the background level prevent access to distant systems, and there is therefore little gain in reducing the IWA below 1.88λ/d.

With any of the three “top” coronagraph designs simulated in this work (PIAAC, OVC6, and ICC6), a 4 m telescope could, with a 10 hr per observation limit and a 1-zodi assumption on the exozodi, perform a ∼100 target survey for ETPs. The total exposure time to complete this survey, assuming six observations of 10 hr for each target, would be well under a year, and therefore sounds realistically achievable. Under the same assumptions, strong exozodi (10 zodi) would cut in half the sample size.

Figure 15 shows that doubling the telescope diameter (from 4 to 8 m) multiplies the number of accessible targets by about 8 for all exozodi content, coronagraph type, and exposure time limit combinations tested. The relative efficiencies of the six coronographs tested stay the same, with, in order of decreasing performance, the ICC6, PIAAC, and OVC6 able to access several hundred targets in the 1-zodi case (with a 1 hr exposure time limit criteria), the VNC/BL4(2) about 50 targets and the BL8 and CPA a few tens of targets.

In this study, we have used an Earth-analog as a prototype for Earth-like planets. The broad concept of candidate Earth-like (potentially habitable) planets likely extends to include planet diameters to several times smaller and larger than Earth. Orienting a survey to optimize detection for a different planet radius would of course change the observation times for each system and the number of systems that could be surveyed. Similarly, an observing program can be optimized for any target list and planet description to maximize the productivity with respect to total detections or other criterion.

7. CONCLUSION

We have quantified the performance limit imposed by fundamental physics on the direct imaging of ETPs. A solution to build an “optimal” coronagraph reaching this limit was proposed, but may be technologically difficult to implement. Fortunately, we have also demonstrated that two existing coronagraph designs (PIAAC and OVC6) yield performances reasonably close (within a factor of 2 in number of accessible targets) to this fundamental limit. These recently proposed designs could image ETPs around up to about 100 stars with a 4 m telescope, or several hundred stars with an 8 m telescope.

For the coronagraphs to perform as detailed in this study, two serious technological challenges must, however, be solved:

Wave front control at the sub-Å level.—Thanks to the large number of stellar photons available, recent detector developments, and the expectation that a space telescope can be designed to offer a relatively stable wave front (with little rapidly varying aberrations, as opposed to atmospheric turbulence for ground-based telescopes), it may be reasonable to hope for this level of correction. Several authors have recently explored promising wave front control schemes free of aliasing and noncommon path aberrations (Codona & Angel 2004; Labeyrie 2004; Guyon 2005; Bordé & Traub 2006), and contrast levels obtained in laboratory (Trauger et al. 2003) are approaching the desired performance.

Chromatic effects.—Chromatic effects can seriously limit the coronagraph’s usable spectral band. A large number of solutions have been proposed to “achromatize” coronagraphs, many of them specific to a single coronagraph design (Abe et al. 2001; Soummer et al. 2003b; Aime 2005a; Mawet et al. 2005; Pluzhnik et al. 2006; Swartzlander 2006). In addition to coronagraph-induced chromatic effects, the beam delivered by the telescope is expected to be chromatic in amplitude and optical path length differences (nonideal coatings, chromatic diffraction propagation between optics). In general, chromatic effects can be mitigated by reducing chromaticity in individual components (beam splitters, phase-shifters, apodization masks) or/and by splitting the spectral band into several narrower bands with dichroics.

More fundamentally, the assumption that detection is photon-noise–limited seems highly 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. We also note that several low-throughput/large IWA coronographs studied in this paper can be more favorably designed for ground-based applications, where contrast requirements are relaxed by about 4 orders of magnitude.

The linear algebra approach to coronagraphy presented in this work successfully explains key characteristics and fundamental
limits of coronagraphy. While we have focused our study on the fundamental limitations of coronagraphy (due to stellar angular size and background contamination), this model can be extended to evaluate the sensitivity of coronagraphs to wave front aberrations other than tip-tilt. The same model can also be used to derive fundamental limits and guide the design of coronagraphs for more “exotic” pupil shapes, from segmented pupils telescopes to short or long baseline interferometric arrays.

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. We thank Lyu Abe and Remi Soummer for providing comments and helping the authors to understand trade-offs for the Apodized Pupil Lyot Coronagraph. We thank Stuart Shaklan for his useful comments and review of this work.

REFERENCES

Abe, L., Vakili, F., & Boccaletti, A. 2001, A&A, 374, 1161
Aime, C. 2005a, PASP, 117, 1012
———. 2005b, A&A, 434, 785
Aime, C., & Soummer, R. 2004, Proc. SPIE, 5490, 456
Aime, C., Soummer, R., & Ferrari, A. 2001, A&A, 379, 697
———. 2002, A&A, 389, 334
Baudouz, P., Boccaletti, A., Rabbia, Y., & Gay, J. 2005, PASP, 117, 1004
Baudouz, P., Rabbia, Y., & Gay, J. 2000, A&AS, 141, 319
Bordé, P. J., & Traub, W. A. 2006, ApJ, 638, 488
Brown, R. A. 2004, ApJ, 607, 1003
———. 2005, ApJ, 624, 1010
Cash, W. C., Kasdin, N. J., Seager, S., & Arenberg, J. 2005, Proc. SPIE, 5899, 274
Codona, J. L., & Angel, R. 2004, ApJ, 604, L117
Foo, G., Palacios, D. M., & Swartzlander, G. A. 2005, Opt. Lett., 30, 3308
Gay, J., & Rabbia, Y. 1996, C. R. Acad. Sci. Paris, 322, Ser. II b, 265
Gonsalves, R., & Nisenson, P. 2003, PASP, 115, 706
Greenaway, A. H., Spaan, F. H. P., & Mourai, V. 2005, ApJ, 618, L165
Guyon, O. 2003, A&A, 404, 379
———. 2005, ApJ, 629, 592
Guyon, O., Pluzhnik, E. A., Galicher, R., Martinache, F., Ridgway, S. T., & Woodruff, R. A. 2005, ApJ, 622, 744
Guyon, O., & Roddier, F. J. 2000, Proc. SPIE, 4006, 377
Guyon, O., & Shao, M. 2006, PASP, 118, 860
Jacquinot, P., & Roisin-Dossier, B. 1964, Prog. Opt., 3, 29
Kasdin, N. J., Vanderbei, R. J., Spergel, D. N., & Littman, M. G. 2003, ApJ, 582, 1147
Kelsall, T. et al. 1998, ApJ, 508, 44
Kuchner, M. J., Crepp, J., & Ge, J. 2005, ApJ, 628, 466
Kuchner, M. J., & Traub, W. A. 2002, ApJ, 570, 900
Labeyrie, A. 2004, in Astronomy with High Contrast Imaging, ed. C. Aime & R. Soummer (Paris: EAS), 3
Levasseur-Regourd, A. C., & Dumont, R. 1980, A&A, 84, 277
Lloyd, J. P., & Sivaramakrishnan, A. 2005, ApJ, 621, 1153
Malbet, F., Yu, J. W., & Shao, M. 1995, PASP, 107, 386
Martinache, F. 2004, J. Opt. Am., 6, 809
Martinache, F., Guyon, O., Pluzhnik, E. A., Galicher, R., & Ridgway, S. T. 2006, ApJ, 639, 1129
Mawet, D., Riaud, P., Absil, O., & Surdej, J. 2005, ApJ, 633, 1191
Mennesson, B. P., Shao, M., Levine, B. M., Wallace, J. K., Liu, D. T., Serabyn, E., Unwin, S. C., & Beichman, C. A. 2003, Proc. SPIE, 4860, 32
Murakami, N., & Baba, N. 2005, PASP, 117, 295
Nisenson, P., & Papaliolios, C. 2001, ApJ, 548, L201
Nishikawa, J., Kotani, T., Murakami, N., Baba, N., Itoh, Y., & Tamura, M. 2005, A&A, 435, 379
Oti, J. E., Canales, V. F., & Cavigal, M. P. 2005, ApJ, 630, 631
Palacios, D. M. 2005, Proc. SPIE, 5905, 196
Pluzhnik, E. A., Guyon, O., Ridgway, S. T., Martinache, F., Woodruff, R. A., Blain, C., & Galicher, R. 2006, ApJ, 648, 1246
Riaud, P., Mawet, D., & Absil, O. 2005, ApJ, 628, L81
Roddier, F., & Roddier, C. 1997, PASP, 109, 815
Rouan, D., Riaud, P., Boccaletti, A., Clenot, Y., & Labeyrie, A. 2000, PASP, 112, 1479
Shaklan, S. B., & Green, J. J. 2005, ApJ, 628, 474
Shealy, D. L. 2002, Proc. SPIE, 4770, 28
Sivaramakrishnan, A., Soummer, R., Sivaramakrishnan, A. V., Lloyd, J. P., Oppenheimer, B. R., & Makidon, R. B. 2006, ApJ, in press
Slepian, D. 1965, J. Opt. Soc. Am., 55, 1110
Soummer, R. 2005, ApJ, 618, L161
Soummer, R., Aime, C., & Falloon, P. E. 2003a, A&A, 397, 1161
Soummer, R., Dohlen, K., & Aime, C. 2003b, A&A, 403, 369
Swartzlander, G. A., Jr. 2006, Opt. Lett., 31, 2042
Tavrov, A. V., Kobayashi, Y., Tanaka, Y., Shioda, T., Otani, Y., Kurokawa, T., & Takeda, M. 2005, Opt. Lett., 30, 2224
Traub, W. A., & Vanderbei, R. J. 2003, ApJ, 599, 695
Trauger, J. T., Moody, D., Gordon, B., Gursel, Y., Ealey, M. A., & Bagwell, R. B. 2003, Proc. SPIE, 4854, 1
Vanderbei, R. J. 2006, ApJ, 636, 528
Vanderbei, R. J., Kasdin, N. J., & Spergel, D. N. 2004, ApJ, 615, 555
Vanderbei, R. J., Spergel, D. N., & Kasdin, N. J. 2005a, ApJ, 590, 593
———. 2005b, ApJ, 599, 686
Vanderbei, R. J., & Traub, W. A. 2005, ApJ, 626, 1079
Yang, W., & Kostinski, A. B. 2004, ApJ, 605, 892