National Aeronautics and Space Administration

Coronagraph Workshop 2006

Coronagraph Workshop Papers
28–29 September, 2006
Pasadena, CA

Edited by W.A. Traub

This publication was prepared by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Preface

Today's grand quest of astrophysics is to detect and characterize exoplanets. This new branch of astrophysics has stimulated a series of astonishingly clever optical solutions to the problem of observing a faint planet close to a bright star.

The first potentially viable technique for suppressing starlight, but preserving planet light, was the nulling interferometer, invented by Bracewell in 1978. This idea inspired people to think of numerous types of arrays of spacecraft that could collect and phase-shift light in order to null out the star efficiently. These designs culminated in the Terrestrial Planet Finder Interferometer (Beichman et al., 1999) in the US, and the Darwin project (Fridlund, 2000) in Europe. For several years it was widely believed that this general architecture, with multiple spacecraft, was the only practical way to suppress starlight sufficiently to see a companion planet. Although these are still exciting and viable designs, we now have several alternatives, as we will explore in this volume.

The idea of using a telescope incorporating a blocking disc (to suppress the light of a bright source but still be able to see faint structures nearby) was invented by Bernard Lyot in 1933 to observe the solar corona. Since that time all descendents of Lyot's telescope have been called coronagraphs, whether the blocker is external or internal or accomplished virtually by means of a pi phase shift. Lyot's design worked well enough for the sun and corona, where the angular separation (~0.5 degree) was relatively large and the contrast ($\sim 10^{-6}$) relatively modest, but today's needs are greater.

Exoplanet observations will demand much more of a coronagraph. For an Earth-Sun twin at 10 pc, for example, the separation is smaller (~0.1 arcsec) and the contrast more extreme ($\sim 10^{-10}$), both by factors of tens of thousands.

The main technical issues of an exoplanet coronagraph are: (1) the Airy diffraction rings must be suppressed; (2) the scattered light speckles must be suppressed; (3) the design must be able to tolerate a finite-sized star and pointing error; (4) the design must work with a finite wavelength band; (5) the design must be able to tolerate realistic background zodiacal light levels; (6) the design must support spectroscopy; (7) the design will preferably allow general astrophysical observations in addition to exoplanets.

Further practical issues include fabrication, testing, validation, risk, cost, flight operations, observing efficiency, and scientific yield.

To meet these challenges, roughly a score of new optical designs for coronagraphs have been proposed, none of which yet appear in any textbooks, but most of which are discussed in this volume.

By summer of 2006 it was clear that many exoplanets had been discovered, and that it was likely (but not yet proven) that there were many terrestrial planets waiting to be discovered in the habitable zones around nearby stars. It was also clear that there were many new coronagraph designs that could potentially detect and characterize these terrestrial planets. What was lacking, however, was an evaluation of how these designs measured up to the 7 criteria listed above.

To begin to answer these questions, many of the world's leading coronagraph experts gathered in Pasadena in the fall of 2006 to explain, compare, and evaluate these designs. Soon after the workshop these experts also provided written versions of their presentations, which are collected here.

The good news from this workshop is that there are several types of coronagraph that are theoretically attractive, and that there is at least one laboratory result showing that the direct imaging of Earth-like exoplanets is nearly within reach.

The current volume is a fall-2006 snapshot of the state of coronagraphic art. We open with a group photo of participants, with a list of names and email addresses, to facilitate community connections. The papers then follow in three major sections.

The first section contains four overview papers: a theoretical analysis of coronagraph types; the optical requirements of coronagraphs; a review of current laboratory testbeds (adapted by invitation from a paper delivered at the TPF-Darwin meeting in November 2006); and an invited commentary on the state-of-the-art of coronagraphy.

The second section contains 21 papers on lessons learned with coronagraphs in the lab, at ground-based telescopes, and expected in space.

The third section contains 6 papers that focus on the theoretically best types of coronagraphs: the shaped pupil design, the nulling coronagraph, band-limited mask coronagraphs, the optical vortex coronagraph, the phase-induced amplitude apodization (or pupil mapping) coronagraph, and the external occulter coronagraph.

This volume is JPL Publication 07-02, available by request from the editor. An on-line version of this volume, as well as the presentation charts from which the written papers were derived, are available at http://planetquest.jpl.nasa.gov/TPF/tpf-c_workshopDocs.cfm.

Wesley A. Traub, editor
15 June 2007

REFERENCES

Beichman, C.A., N.J. Woolf, and C.A. Lindensmith, eds., 1999, Terrestrial Planet Finder, JPL Publication 99-3, http://planetquest.jpl.nasa.gov/TPF/tpf_book/index.cfm.

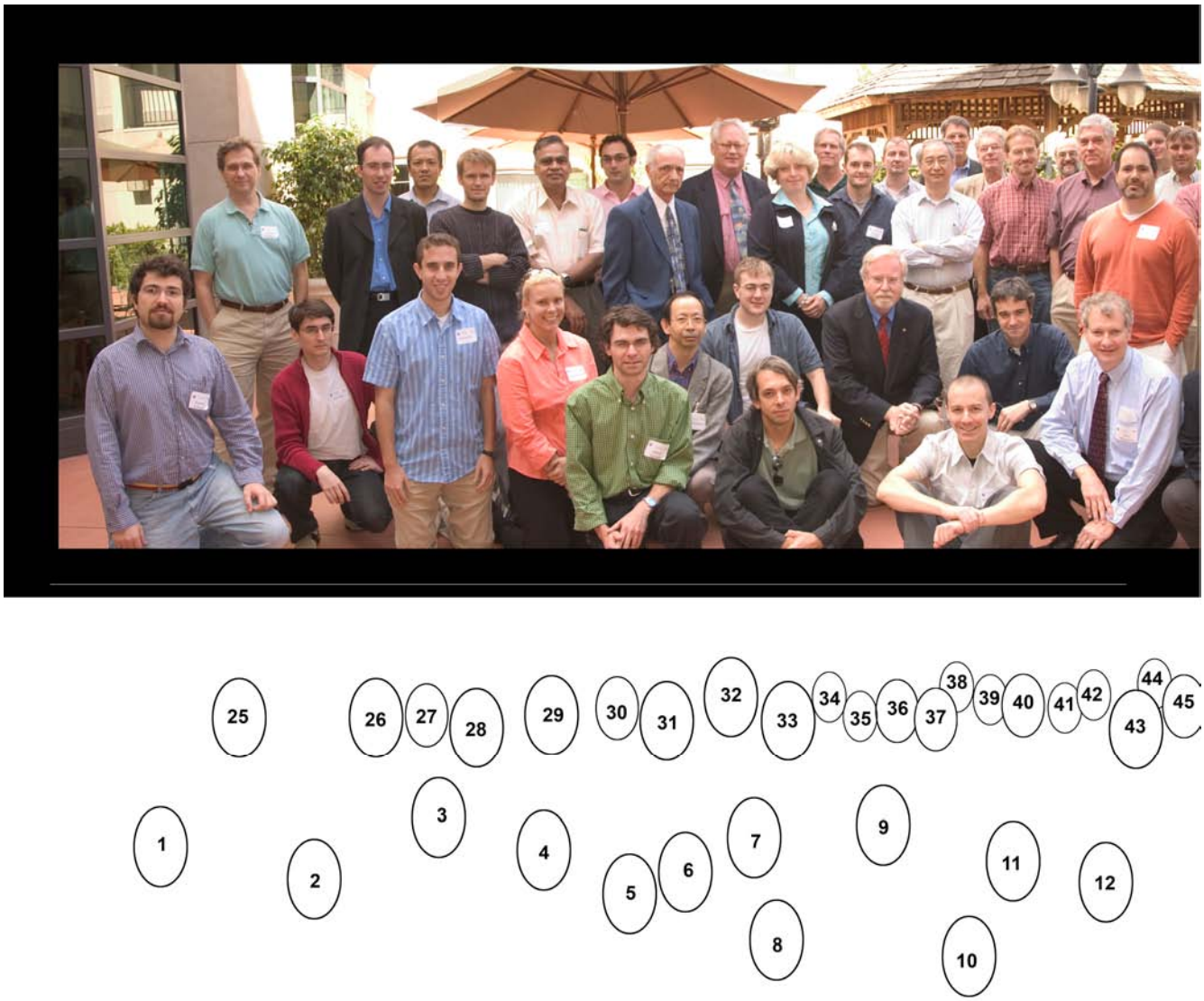
Bracewell, R., 1978, Detecting nonsolar planets by spinning infrared interferometer, Nature **274**, pp.780–781.

Fridlund, M., 2000, in DARWIN: The Infrared Space Interferometer, ESA SCI-12, Noordwijk, ESA, 47.

Lyot, B., 1933, The study of the solar corona without an eclipse, J. Royal Astronomical Society of Canada, **27**, pp. 265–280.

Today 2½ years after the Coronagraph Workshop met, the papers in this volume are still highly relevant. This edition includes important corrections to the author lists of two papers, plus a handful of fixed typos. WAT, 18 June 2009

Coronograph Workshop Attendees Photo

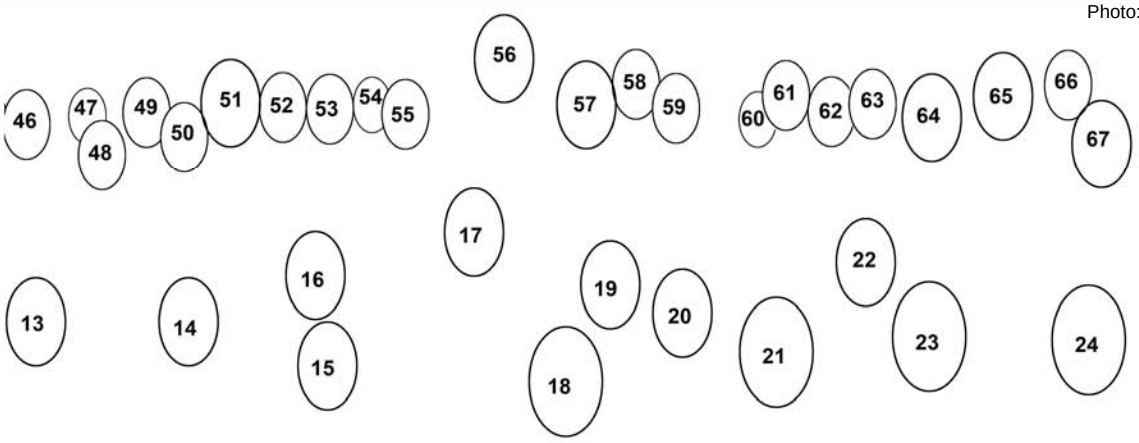


Attendees at the Exoplanet Detection with Coronagraphs 2006 Workshop. See the following list of attendees for the key to names (photo).

First Name	Last Name	Photo #
Egemen	Kolemen	1
Lyu	Abe	2
Jason	Kay	3
Sarah	Hunyadi	4
Robert	Egerman	5
Jun	Nishikawa	6
Eric	Cady	7
Olivier	Guyon	8
Wesley	Traub	9
Joseph	Carson	10
Mark	Swain	11
Gary	Matthews	12
Jon	Arenberg	13
Stuart	Shaklan	14
Jian	Ge	15
Motohide	Tamura	16

First Name	Last Name	Photo #
Ian	Crossfield	17
Justin	Crepp	18
Ruslan	Belikov	20
Richard	Lyon	21
Dwight	Moody	22
Grover	Swartzlander	23
David	Palacios	24
Robert	Vanderbei	25
Pascal	Borde	26
Shinichiro	Tanaka	27
Pierre	Baudoz	28
Bala	Balasubramanian	29
Remi	Soummer	30
Dominick	Tenerelli	31
Robert	Brown	32
Marie	Levine	33

Photo: G. Swartzlander



First Name	Last Name	Photo #
Tom	Roellig	34
Rocco	Samuele	35
Kent	Wallace	36
Michael	Shao	37
Michael	Devirian	38
Jean	Schneider	39
Brian	Kern	40
Marty	Levine	41
John	Trauger	42
Marc	Kuchner	43
Karl	Stapelfeldt	45
Roger	Angel	46
John	Krist	47
Beth	Biller	48
Zlatan	Tsvetanov	49
Allan	Eisenman	50
Steven	Beckwith	51

First Name	Last Name	Photo #
Joe	Pitman	52
Laurent	Pueyo	53
Chuck	Bowers	54
Amir	Give'on	55
Frantz	Martinache	56
Dimitri	Mawet	57
Charley	Noecker	58
Anthony	Boccaletti	59
Eugene	Pluzhnik	60
Jeremy	Kasdin	61
Gene	Serabyn	62
Bob	Woodruff	63
Martin	Lo	64
Chuck	Lillie	65
Steven	Ridgway	66
Peggy	Park	67

Last Name	First Name	Organization	E-mail
Abe	Lyu	NAOJ	abe@optik.mtk.nao.ac.jp
Angel	Roger	Univ of Arizona	rangel@as.arizona.edu
Arenberg	Jon	NGST	jon.arenberg@ngc.com
Balasubramanian	Bala	JPL	kbala@jpl.nasa.gov
Baudoz	Pierre	Observatoire de Meudon	baudoz@mesiog.obspm.fr
Beckwith	Steven	STSci	svwb@stsci.edu
Beichman	Chas	JPL / MSC / Caltech	chas@mail.jpl.nasa.gov
Belikov	Ruslan	Princeton	rbelikov@Princeton.EDU
Biller	Beth	Univ of Arizona	bbiller@coatlicue.as.arizona.edu
Boccaletti	Anthony	Observatoire de Meudon	anthony.boccaletti@obspm.fr
Borde	Pascal	Harvard CFA	pborde@cfa.harvard.edu
Bowers	Chuck	GSFC	Charles.W.Bowers@nasa.gov
Breckinridge	James	JPL	jbreckin@mail.jpl.nasa.gov
Brown	Robert	STSci	rbrown@stsci.edu
Cady	Eric	Princeton	ecady@princeton.edu
Carson	Joseph	JPL	Joseph.C.Carson@jpl.nasa.gov
Cash	Webster	Univ of Colorado	cash@origins.colorado.edu
Clampin	Mark	GSFC	mark.clampin@nasa.gov
Codona	John	Univ of Arizona	jcodona@as.arizona.edu
Cohen	Eri	JPL	Eri.J.Cohen@jpl.nasa.gov
Coulter	Dan	JPL	Daniel.R.Coulter@jpl.nasa.gov
Crepp	Justin	Univ of Florida	jcrepp@astro.ufl.edu
Crossfield	Ian	JPL	Ian.J.Crossfield@jpl.nasa.gov
Dekany	Richard	Caltech	rgd@astro.caltech.edu
Devirian	Michael	JPL	devirian@jpl.nasa.gov
Egerman	Robert	ITT	Robert.Egerman@itt.com
Eisenman	Allan	JPL	Allan.R.Eisenman@jpl.nasa.gov
Enya	Keigo	JAXA	enya@ir.isas.jaxa.jp
Ftaclas	Christ	Univ of Hawaii	ftaclas@ifa.hawaii.edu
Ge	Jian	Univ of Florida	jge@astro.ufl.edu
Give'on	Amir	Caltech	agiveon@caltech.edu
Green	Joe	JPL	Joseph.J.Green@jpl.nasa.gov
Guyon	Olivier	Subaru Telescope, NOAJ	guyon@subaru.naoj.org
Hong	John	JPL	John.H.Hong@jpl.nasa.gov
Hoppe	Dan	JPL	Daniel.J.Hoppe@jpl.nasa.gov
Hunyadi	Sarah	JPL	Sarah.Hunyadi@jpl.nasa.gov
Johnson	Bill	JPL	william.r.johnson@jpl.nasa.gov
Kasdin	Jeremy	Princeton	jkasdin@Princeton.EDU
Kay	Jason	Princeton	jkay@princeton.edu
Kern	Brian	JPL	Brian.D.Kern@jpl.nasa.gov
Kolemen	Egemen	Princeton	ekolemen@princeton.edu
Krist	John	JPL	krist@jpl.nasa.gov
Kuchner	Marc	GSFC	mkuchner@milkyway.gsfc.nasa.gov
Kuhnert	Andy	JPL	Andreas.C.Kuhnert@jpl.nasa.gov
Labeyrie	Antoine	OHP	labeyrie@obs-hp.fr
Lawson	Peter	JPL	Peter.R.Lawson@jpl.nasa.gov
Levine	Marie	JPL	marie.levine@jpl.nasa.gov
Levine	Marty	JPL	bmlevine@s383.jpl.nasa.gov
Lillie	Chuck	NGST	chuck.lillie@ngc.com

Last Name	First Name	Organization	E-mail
Lo	Martin	JPL	martin.lo@jpl.nasa.gov
Lyon	Richard	GSFC	Richard.G.Lyon@nasa.gov
Macintosh	Bruce	LLNL	bmac@igpp.ucllnl.org
Martinache	Frantz	Cornell	frantz@astro.cornell.edu
Matthews	Gary	ITT	Gary.Matthews@itt.com
Mawet	Dimitri	Liege	mawet@astro.ulg.ac.be
Milman	Mark	JPL	mark.milman@jpl.nasa.gov
Moody	Dwight	JPL	dwight.moody@jpl.nasa.gov
Nielsen	Eric	Steward Observatory, Univ of Arizona	enielsen@omoikane.as.arizona.edu
Nishikawa	Jun	NAOJ	jun.nishikawa@nao.ac.jp
Noecker	Charley	Ball Aerospace	mcnoecke@ball.com
Palacios	David	JPL	david.palacios@jpl.nasa.gov
Park	Peggy	JPL	Peggy.H.Park@jpl.nasa.gov
Pitman	Joe	Lockheed Martin	joe.pitman@lmco.com
Pluzhnik	Eugene	NAOJ	epluzhnyk@keck.hawaii.edu
Pueyo	Laurent	Princeton	lpueyo@princeton.edu
Ridgway	Steven	NOAO	ridgway@noao.edu
Roellig	Tom	NASA Ames	troellig@mail.arc.nasa.gov
Samuele	Rocco	NGST	rsamuele@huey.jpl.nasa.gov
Schneider	Jean	Observatoire de Meudon	Jean.Schneider@obspm.fr
Serabyn	Gene	JPL	gene.serabyn@jpl.nasa.gov
Shaklan	Stuart	JPL	shaklan@s383.jpl.nasa.gov
Shao	Michael	JPL	mshao@s383.jpl.nasa.gov
Sidick	Erkin	JPL	Erkin.Sidick@jpl.nasa.gov
Soummer	Remi	American Museum of Natural History	rsoummer@amnh.org
Stapelfeldt	Karl	JPL	krs@exoplanet.jpl.nasa.gov
Swain	Mark	JPL	swain@mpia-hd.mpg.de
Swartzlander	Grover	Univ of Arizona	grovers@optics.arizona.edu
Tamura	Motohide	NOAJ	hide@optik.mtk.nao.ac.jp
Tanaka	Shinichiro	JAXA	stanaka@ir.isas.jaxa.jp
Tenerelli	Dominick	Lockheed Martin	dominick.tenerelli@lmco.com
Traub	Wesley	JPL	Wesley.A.Traub@jpl.nasa.gov
Trauger	John	JPL	John.T.Trauger@jpl.nasa.gov
Tsvetanov	Zlatan	NASA Headquarters	zlatan.tsvetanov@nasa.gov
Unwin	Steven	JPL	Stephen.C.Unwin@jpl.nasa.gov
van Zyl	Jakob	JPL	Jakob.J.vanZyl@jpl.nasa.gov
Vanderbei	Robert	Princeton	rvdb@Princeton.EDU
Wallace	Kent	JPL	James.K.Wallace@jpl.nasa.gov
Werner	Michael	JPL	mwerner@sirtfweb.jpl.nasa.gov
Wilson	Daniel	JPL	daniel.w.wilson@jpl.nasa.gov
Woodruff	Bob	Lockheed Martin	robert.a.woodruff@lmco.com

Table of Contents

1. Overview Papers

2. Lessons Learned with Coronagraphs

3. Reviews of Leading Coronagraph Types

Theoretical Analysis of Coronagraphs

Olivier Guyon
Subaru Telescope

ABSTRACT

Many high contrast coronagraph concepts have recently been proposed, and many more will likely be developed in the near future. In this paper, I look at the intrinsic performance each concept can theoretically offer (with perfect wavefront, perfect optics). I also show that simple physics sets fundamental limits on the performance of coronagraphs—these limits are quantified and compared to existing designs. This theoretical work provides valuable insight into how to design high performance coronagraphs and explains many of the characteristics of particular designs. This new perspective reveals the importance of stellar angular diameter on coronagraph performance.

The simulated observation a sample of target stars reveals very large performance differences between coronagraph designs considered for exoplanet detection. These performance differences are equivalent to changing the telescope diameter by a factor 2 to 3. This analysis shows that by adopting a high performance coronagraph design, the TPF-C science goals which were previously thought to require an ~8m telescope, could be achieved with a 4m telescope.

TABLE 1
CORONAGRAPHS ABLE TO ACHIEVE 10^{10} PSF CONTRAST WITHIN 5 λ/d

Coronagraph	abrev.	reference	Design(s) adopted
"Interferometric" Coronagraphs			
Achromatic Interferometric Coronagraph	AIC	Baudoz et al. (2000)	
Common-Path Achromatic Interferometer-Coronagraph	CPAIC	Tavrov et al. (2005)	(=AIC)
Visible Nulling Coronagraph, X-Y shear (4^{th} order null)[a]	VNC	Mennesson et al. (2003)	Shear distance = ± 0.3 pupil radius
Pupil Swapping Coronagraph	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[c]($r = 4.2\lambda/d$, 8% throughput)
Achromatic Pupil Phase Apodization	PPA	Yang & Kostinski (2004)	$\phi = \phi_2(x) + \phi_2(y)$: $a = 2$: $\epsilon = 0.01$
Phase Induced Amplitude Apodization Coronagraph	PIAAC	Guyon (2003)	Prolate apodization
Phase Induced Zonal Zernike Apodization	PIZZA	Martinache (2004)	Not simulated
Improvement on the Lyot concept with amplitude masks			
Apodized Pupil Lyot Coronagraph	APLC	Soummer et al. (2003)	$r = 1.8\lambda/d$
Apodized Pupil Lyot Coronagraph, N steps	$APLC_N$	Aime & Soummer (2004)	$(N, r) = (2, 1.4)$: $(3, 1.2)$: $(4, 1.0)$
Band limited, 4^{th} order[a]	BL4	Kuchner & Traub (2002)	$\sin^4$ intensity mask, $\epsilon = 0.21$
Band limited, 8^{th} order	BL8	Kuchner et al. (2005)	$m = 1, l = 3, \epsilon = 0.6$
Improvement on the Lyot concept with phase masks			
Phase Mask	PM	Roddier & Roddier (1997)	with mild prolate pupil apod.
4 quadrant	4QPM	Rouan et al. (2000)	
Achromatic Phase Knife Coronagraph	APKC	Abe et al. (2001)	(=4QPM)
Optical Vortex Coronagraph, topological charge m	OVC_m	Palacios (2005)	$m = 2, 4, 6, 8$
Angular Groove Phase Mask Coronagraph	AGPMC	Mawet et al. (2005)	(=OVC)
Optical Differentiation	ODC	Oti et al. (2005)	mask: $x \times \exp -(x/10)^2$ d

[a] The Visible Nulling Coronagraph (VNC) and Band limited 4^{th} order (BL4) coronagraphs belong to the same class of pupil-shearing 4^{th} order coronagraphs, and are simply 2 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 2 orthogonal shear directions, while the BL4 is designed with a larger IWA and 2 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) which 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 $> 4\lambda/d$.

[c] CPA, APLC, $APLC_N$: r is the radius, in λ/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. (2003).

[d] ODC: x is in λ/d. Maximum mask transmission at $7\lambda/d$. Lyot pupil mask radius = 0.85 times pupil radius.

An inventory of current coronagraph designs

The pace at which new coronagraph designs appear seem to be ever accelerating, but I have attempted to provide a snapshot of current coronagraph design in Table 1. Fortunately, they can be grouped in 3 broad categories:

- Inteferometric coronagraphs: The coronagraph design is inspired from interferometry; the single telescope pupil is used to extract several beams which are coherently mixed to produce a null
- Pupil Apodization: These design create an apodized pupil without sharp edges, which is then used to form a high contrast image
- Improvement on the Lyot design: 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.

Fundamental limits of coronagraph performance & comparison to existing coronagraph designs

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 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.

The stellar angular size problem

A quick glance a Figure 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. It looks as if TPF-C must be possible with a ~2m telescope! Checking (see Figure 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 moved 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.

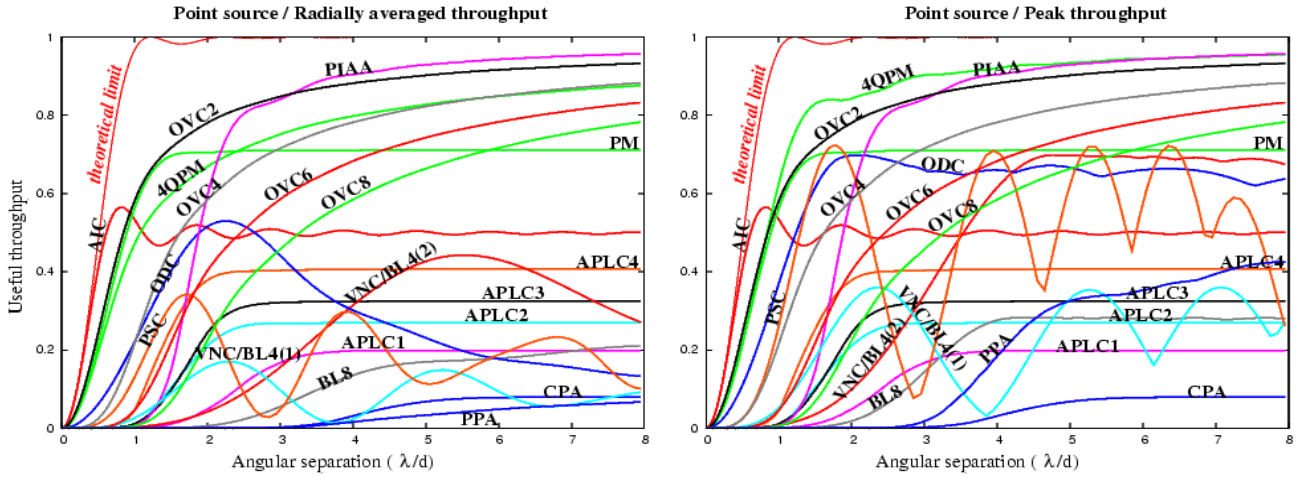


Figure 1: *Useful throughput, at the 1e-10 contrast level, of several coronagraph designs. The theoretical limit of useful throughput is also shown.*

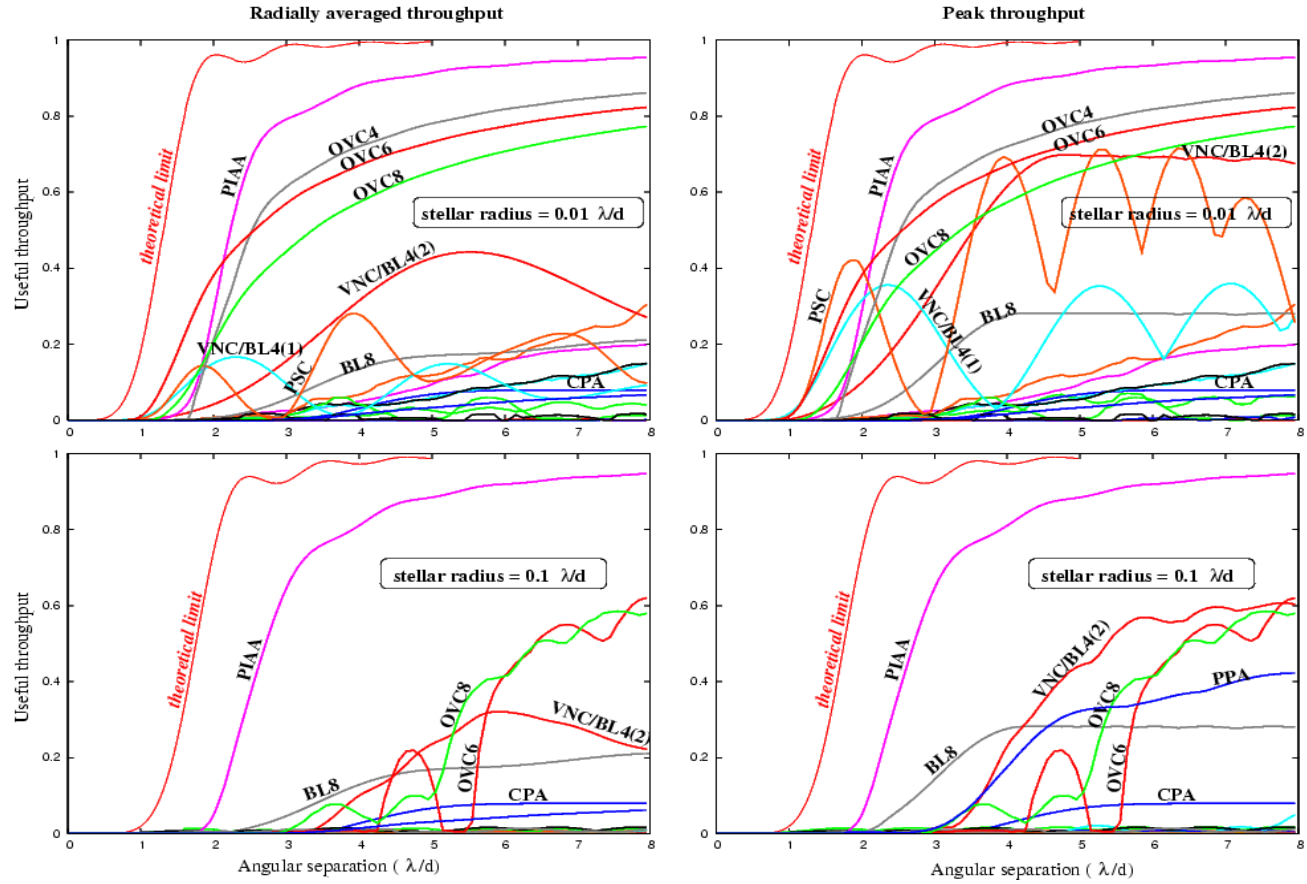


Figure 2: *Same as figure 1, but with non-zero angular stellar diameter. Note that the PPA coronagraph was omitted in the upper plots, but included in the lower plots (it is not affected by stellar angular diameter).*

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 degrades 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 unaffected by stellar angular size.

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 instrisically 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 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 are listed in Table 2.

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:

(1) 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
(2) 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.

Coronagraph designs characteristics: Summary

Table 2 lists, for the coronagraphs of Table 1, the values obtained for the performance metrics used in this study: IWAs and throughput derived from the useful throughput vs. angular separation curves (and how stellar angular diameter affects the IWA), angular resolution and DEFs. The last line of table shows the performance of the "ideal" coronagraph constructed with the theoretical framework used in this paper.

This table shows that the most promising coronagraph designs are the PIAA and OVC. The BL8, CPA, and VNC/BL4 also appear to be suitable for exoplanet detection, but have a lower efficiency (lower throughput, higher DEF, larger IWA and poor angular resolution).

Coronagraph performance 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 sample of target was first established, assuming each sample 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 Figure 3.

TABLE 2
CORONAGRAPH CHARACTERISTICS FOR 10^{10} CONTRAST

Coronagraph name	Throughput average	Throughput peak	Angular resolution[a]	diffractive eff. fact.[b]	IWA[c]	IWA(5%) with stellar radius[d] $0\ \lambda/d$	$0.001\lambda/d$	$0.01\lambda/d$	$0.1\lambda/d$
AIC	0.50	0.57	1.0	0.5	0.38	0.15	6.14	> 8.0	> 8.0
VNC/BL4(1)	0.09	0.36	1.65[e]	0.37[e]	1.49	1.27	1.27	1.43	> 8.0
PSC	0.18	0.73	1.0[f]	0.5[f]	1.13	0.79	0.79	3.24	> 8.0
CPA	0.08	0.08	1.81	0.31[g]	4.20	4.42	4.42	4.42	4.47
PPA	0.08[h]	0.33[h]	1.35	0.18	3.91	3.29	3.29	3.30	4.02
APLC	0.20	0.20	1.22	0.67	2.41	2.01	2.17	4.02	> 8.0
$APLC_2$	0.27	0.27	1.11	0.81	1.64	1.26	1.78	4.71	> 8.0
VNC/BL4(2)	0.35	0.70	1.19	0.71	3.02	2.17	2.20	2.21	3.77
BL8	0.26	0.28	1.86	0.29	2.96	3.04	3.04	3.04	3.05
PM	0.71	0.71	1.0	1.0	0.69	0.23	2.94	3.59	> 8.0
4QPM	1.0	1.0	1.0	1.0	0.84	0.28	2.39	> 8.0	> 8.0
OVC_2	1.0	1.0	1.0	1.0	0.84	0.21	2.21	> 8.0	> 8.0
OVC_4	1.0	1.0	1.0	1.0	1.62	0.62	0.66	1.74	> 8.0
OVC_6	1.0	1.0	1.0	1.0	2.45	1.02	1.02	1.19	4.32
OVC_8	1.0	1.0	1.0	1.0	3.26	1.41	1.42	1.42	3.34
ODC	-[i]	0.70	1.18[j]	0.72[j]	0.97	0.50	4.59	> 8.0	> 8.0
PIAA	1.0	1.0	1.0	1.0	1.88	1.28	1.38	1.66	2.0
ICC_6	1.0	1.0	1.0	1.0	**1.32**	**0.76**	**0.76**	**0.76**	**1.76**

[a] Square root of the smallest focal plane area containing half of the planet's light. Normalized to 1.0 for the Airy pattern of an non-coronagraphic telescope. In the case of the AIC, the angular resolution is measured on one of the 2 images.

[b] Diffractive efficiency factor measured at $20\lambda/d$ unless otherwise specified.

[c] Angular separation at which the planet's useful throughput first reaches half of the peak throughput. This is the standard definition of inner working angle (IWA): it assumes that the planet is favorably placed (no radial average) and that the star is point-like.

[d] Angular separation at which the planet's radially averaged useful throughput first reaches 5% for the stellar radius shown.

[e] Measured at $2.36\lambda/d$.

[f] Measured at $1.77\lambda/d$.

[g] Higher if the apodization is performed by a binary mask

[h] Only the main PSF diffraction peak is considered in the throughput

[i] The outer part of the field is not transmitted in the ODC

[j] Measured at $2\lambda/d$.

Figure 3 nicely illustrates the coronagraphic characteristics quantified in the previous sections:

- **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 2m telescope. Detection appears feasible on a 4m telescope with the VNC/BL4(2), PIAAC, OVC6 and ICC6, but requires a 6m telescope with the BL8. Finally, an 8m telescope is needed for detection with a CPA.

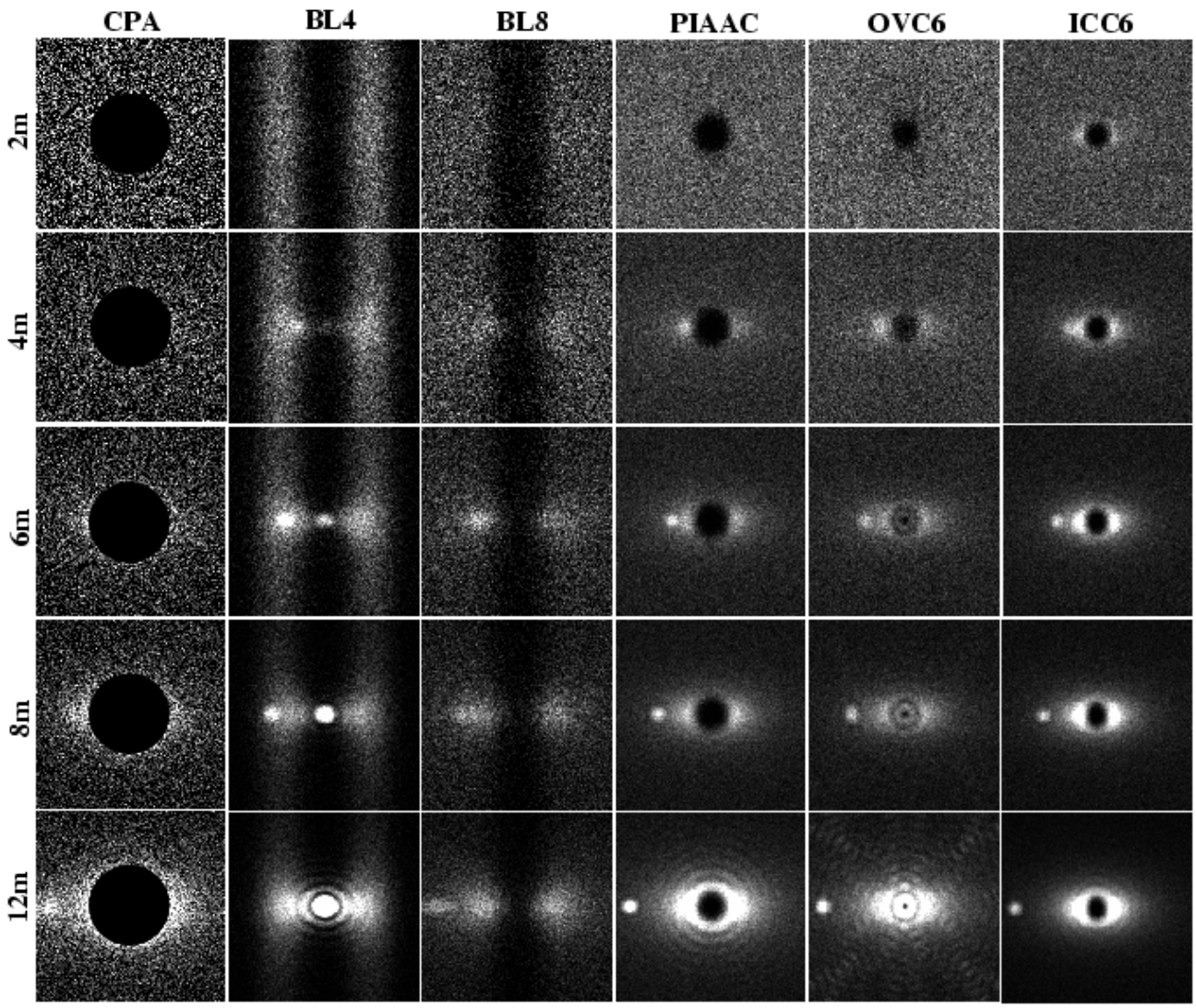


Figure 3: *Simulated 4 hours exposures of HIP56997 and a hypothetical Earth-type planet at maximum elongation for telescopes sizes ranging from 2m to 12m. HIP56997 is a G8 tye main sequence star at 9.54pc. Each image assumes a perfect detector, mV=22.95/sqarcsec zodiacal background, a 1 zodi exozodi cloud, a 25% telescope+camera throughpu, and a 0.1 micron bandpass centered at 0.55 micron. The system inclination for this particular simulation was arbitrarily set at i=59 deg. Each image is 20 × 20 λ/d, and the planet-star separation is 80 mas.*

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

Images similar to the ones showed in Figure 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. These results are shown for HIP56997 in the table below for a selection of coronagraph designs, including the theoretically ideal ICC6 coronagraph.

TABLE 3

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

Tel. Diam.	CPA	VNC/BL4(2)	BL8	PIAAC	OVC6	ICC6
20% detection probability						
2m	... [a]	...	...	...	...	12.4 hour
4m	...	2.1 hour	11.4 hour	34 min	48 min	14.3 min
6m	6.9 hour	14 min	44.8 min	5.7 min	8.6 min	3.78 min
8m	63 min	4.8 min	17 min	2.1 min	3.1 min	110 sec
12m	12 min	2.7 min	4.3 min	45 sec	56 sec	42 sec
50% detection probability						
2m	...	...	...	...	...	...
4m	...	12.7 hour	...	2.3 hour	2.2 hour	44.1 min
6m	...	59 min	5.0 hour	14 min	19.8 min	12.23 min
8m	4.3 hour	18 min	1.8 hour	5.8 min	7.6 min	5.4 min
12m	31 min	55 min	19 min	2.0 min	2.4 min	113 sec

[a]Exposure times above 1 day are not shown in this Table, as they are practically unrealistic.

This table shows 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 exposure times in this table 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 Table 4.

Table 4 confirms the results obtained so far:
- The PIAAC and OVC appear to be the most promising coronagraphic designs.
- 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
- 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.

The analysis presented in this paper ignores important challenges which must be overcome to reach the performance levels described in this paper:
- wavefront control
- 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.

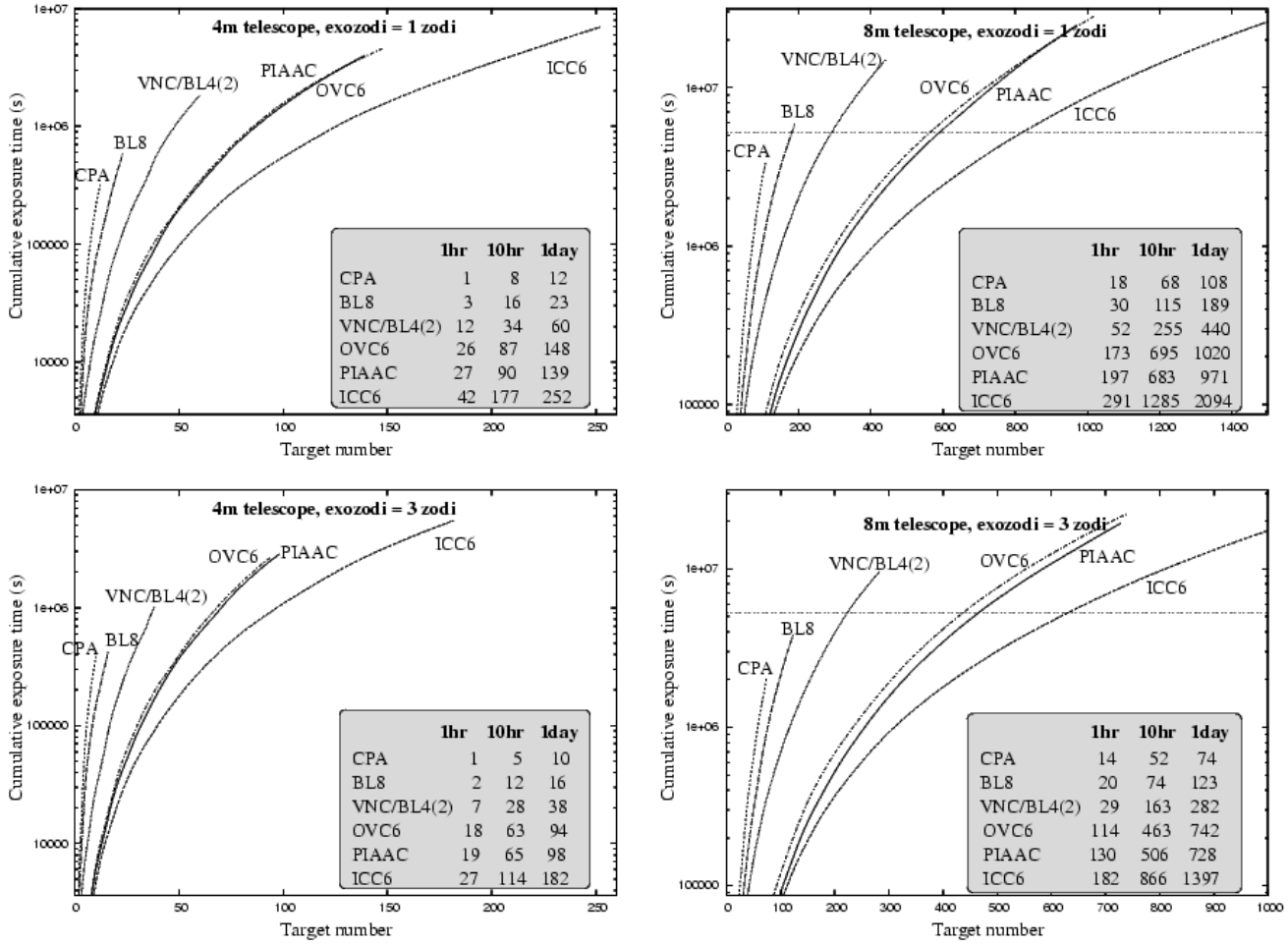


Figure 4: *Total cumulative exposure times required to reach a 50% planet detection (SNR=7) probability for a single observation as a function of number of targets. For all simulations, a 25% throughput in the 0.5–0.6 micron band is adopted, and targets are ordered with increasing exposure time. Each curve terminates when the required exposure time per target reaches 1 day. Results are shown for a 4m telescope (left) and an 8m telescope (right) with exozodi levels of 1 zodi (top) and 3 zodi (bottom). 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 8m 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.*

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.

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.

A more detailed report of this study, including the full list of references, can be found in:
"Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs" Guyon, O., Pluzhnik, E.A., Kuchner, M.J., Collins, B., Ridgway, S.T. ApJS, 167, 81 (2006) astro-ph/0608506

Terrestrial Planet Finder Coronagraph Optical Requirements

Stuart B. Shaklan
Jet Propulsion Laboratory, California Institute of Technology

ABSTRACT

We discuss the optical requirements for the TPF-C mission. The requirements flowdown begins with science and relies on extensive mission studies to determine the necessary spatial resolution and minimum star-to-planet contrast ratio. This in turn leads to engineering requirements imposed on the system, e.g., mirror shape and stability, pointing control, etc. Our work in this area has led to four important lessons: 1) sequential wave front controllers are preferred because they relax both amplitude uniformity and surface flatness requirements; 2) Uncontrollable high-spatial frequencies look manageable using existing technology; 3) Transmissive masks placed in the image plane have challenging surface power spectral density requirements; and 4) working at 2 or 3 λ/D is much harder than 4 λ/D resulting in several important system architecture challenges.

I. Requirements Flowdown

The reader is referred to the TPF-C Science and Technology Definition Team Report[1] to view the full text of the science requirements. The science requirements cover not only terrestrial planet detection and characterization, but also Jovian and transition planets, protoplanetary disk science, disk/planet interactions, and a wide range of general astrophysics.

The main requirement that has so far driven the TPF-C baseline design is: TPF-C shall have an excellent chance (95%) of detecting at least one planet that is potentially habitable, assuming that ten percent or more of all target stars have such a planet ($\eta_\oplus$ = 0.10). Equivalently, TPF-C shall find ~30 potentially habitable planets if all target stars have one such planet.

We have performed extensive mission modeling to determine the spatial resolution (specifically, the inner working angle, IWA) and limiting star-to-planet contrast ratio (the minimum useful delta-magnitude, Δmag, between star and planet) that, combined with collecting area, overhead, and other parameters, ensures that the mission will meet this detection requirement. The modeling is described in detail by Brown et al. in several papers[2-7]. The mission design is called Flight Baseline 1 (FB-1)[8]; it is a room temperature telescope having an 8×3.5 m primary mirror, operating in an earth-trailing or L2 orbit. The telescope is surrounded by an efficient v-groove thermal shroud to isolate it from solar heating, and the primary mirror and instrumentation are actively heated to ~ 300 Kelvin. FB-1 employs either a band-limited 8^{th}-order mask,[9] or a shaped pupil mask[10].

Notably, several key additional requirements will likely drive the IWA and Δmag to more challenging values. These include disambiguation from background sources, spectral characterization, and orbit determination. For now, we assume that the FB-1 design is adequate to meet all the requirements and carry margin that may be used to meet the more challenging ones.

Our mission studies show that to sample 30 habitable zones in a 3 yr mission, assuming that 1/3 of the time is used for planet detection, we require Δmag = 25.5, IWA =4 λ/D for lambda = 600 nm, and a bandwidth of 100 nm. These values can all be traded against each other, and they are a function of the mask and system throughput, assumed exozodiacal level, and other parameters. We choose these values as our starting point and in the rest of this document study the engineering requirements flowed down to the telescope and optical surfaces. It will become evident by the concluding section that we have attempted to balance the requirements to create an FB-1 design that can be built without major new leaps in technology.

II. Wave Front Control System

One proposed TPF-C wave front control configuration employs two DMs in a symmetric Michelson configuration first proposed by Littman[11]. This architecture (Fig. 1) allows independent control of phase and

amplitude variations. Phase is controlled when the two DMs have identical piston, while amplitude is controlled (without affecting phase) when they have opposite pistons (one 'pushes' while the other 'pulls'). The system then controls both the phase and amplitude components of the wave front and clears out a dark hole that is symmetric about the optical axis. It can be shown[11] that the amplitude generated by the pair of DMs varies as $1/\lambda^2$.

In the sequential configuration[12], one DM is placed in the pupil (DM_p), while a second one, DM_{np}, is located a distance z_{DM} downstream. With one or more DMs located away from the pupil, we have a propagation lever that provides a wavelength-independent amplitude control term.[12,13] DM_{np} is commanded to create a wavelength-independent amplitude that is a broad-band null of both wavelength-independent amplitude non-uniformites and propagated phase-to-amplitude terms. Alternatively, the two DMs could be located on either side of the pupil; as with the Michelson, the configuration provides independent phase and amplitude control. See references 12 and 13 for the detailed mathematics showing that the controller generates a wavelength-independent amplitude via propagation from DM_{np} to the pupil.

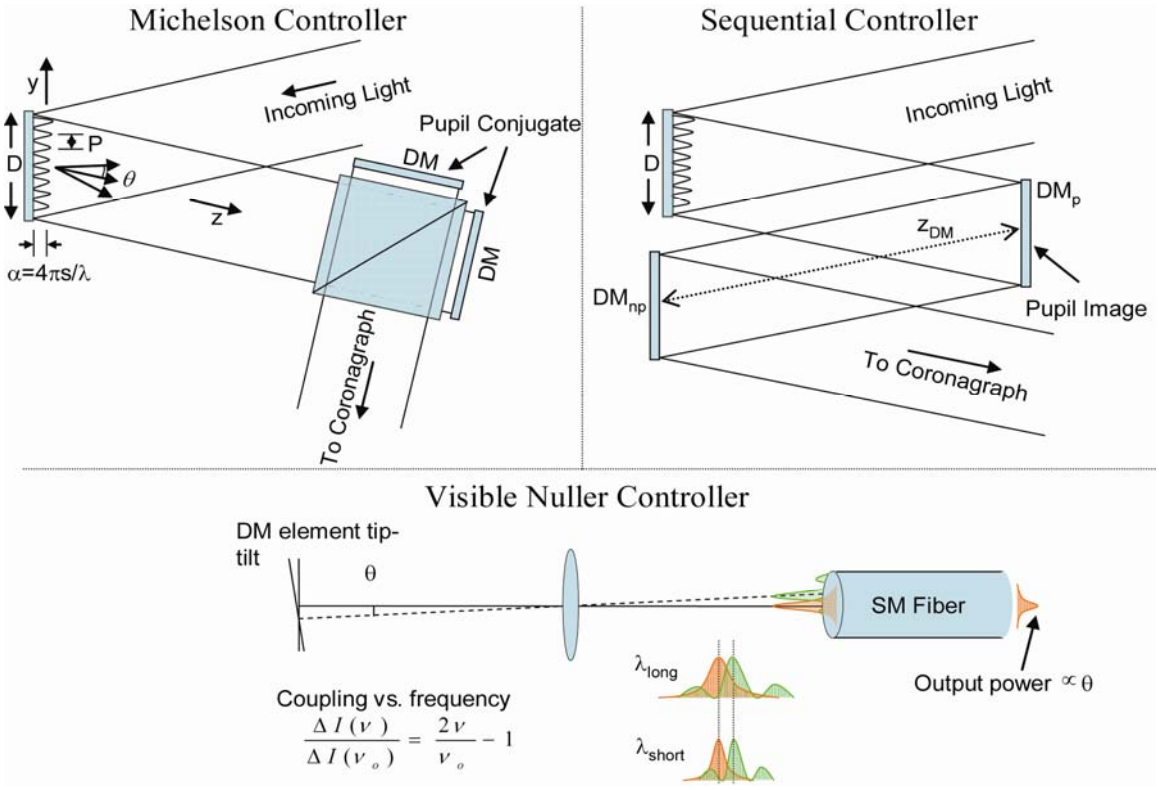


Figure 1. Wave front control systems. Upper left: In the Michelson controller, DMs move in opposite direction to control amplitude with 1/wavelength squared dependence.. Upper right: In the sequential configuration, a DM located some distance away from the pupil is actuated, resulting in wavelength-independent control in the pupil. In the visible nuller, tip-tilt actuators move a spot across the core of a single-mode fiber. Amplitude control varies as 1/2*wavelength.

A third approach is shown in the lower panel of Figure 1. The visible nuller[14] (VN) is a coronagraph alternative proposed for TPF-C. It is a shearing interferometer followed by an array of lenslets mated to an array of single-mode (SM) fiber optics. The fiber optics serve as local spatial filters of the wave front. The DMs in the VN have segmented face sheets with one segment per lenslet. The segments have 3 degrees of freedom: piston for wave front phase control, and tip/tilt for amplitude control. The latter is achieved by tilting the beam so that the spot formed by the lenslet is offset relative to the core of the SM fiber. As shown in Figs. 1, the controlled amplitude varies with wavelength. The diameter of the Airy spot grows linearly with wavelength, as does the dimension of the mode of the fiber Thus starlight-to-fiber coupling efficiency versus tilt is a function of wavelength, with longer wavelengths exhibiting less change in the output power for a given tilt than shorter wavelengths.

In summary, we have identified three wave front control configurations. All three control phase through surface deformation. This carries a 1λ dependence. For amplitude control, the Michelson has $1/\lambda^2$ dependence, the Visible Nuller has 1/(2λ dependence, and the sequential configuration is wavelength-independent.

The wave front controller is placed between the telescope and the coronagraph, as shown in Figure 2. Optics following the telescope secondary mirror fold and collimate the beam to a 10 cm diameter. The primary mirror (PM) is imaged onto the DM (or one of the DMs for the sequential controller). As seen from the pupil image, the other optics in the system appear from a few cm (secondary mirror) to many meters away. As shown in refs. 12 and 13, corrugated wavefronts propagating from the distant optics arrive at the pupil and exhibit wavelength dependence that is not necessarily matched to that of the wave front controller. Phase corrugations arrive with 1/λ dependence—these are fully correctable by any of the controllers. Some fraction of the phase corrugations become amplitude-corrugations and these are wavelength-independent at the pupil. Only the sequential controller, which itself uses propagated phase to generate amplitude, can fully compensate this light. Reflectivity variations (which we assume to be wavelength independent, e.g., due to gray blemishes on the optics) arrive and are wavelength-independent, again controllable only by the sequential controller. Some of the reflectivity variations propagate into phase terms proportional to λ (not 1/λ): these are not correctable by any two-mirror system and serve as the contrast-limiting terms for the sequential controller. Figure 3 shows the required optical surface characteristics to obtain sufficiently high contrast for the three different controllers, while Figure 4 shows the amplitude uniformity requirement.

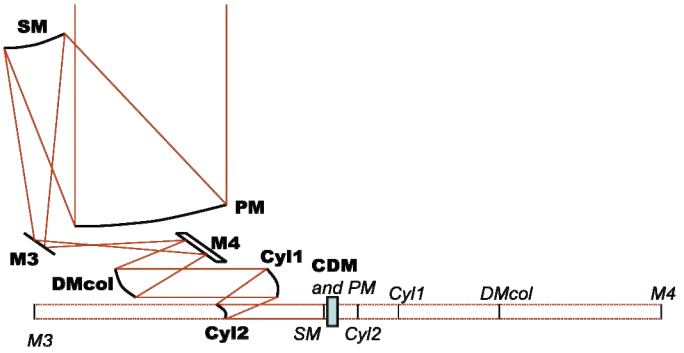


Figure 2. TPF optical layout up to the coarse DM (not to scale). Bold font: optics position. Italics: Image of the optic as seen from the coarse DM. PM: Primary Mirror. SM: Secondary Mirror. M3 and M4: flat fold mirrors. DMcol: Collimating mirror. Cyl1 and Cyl2: cylindrical telescope to circularize the beam. CDM: coarse DM. The PM is imaged onto the CDM. The beam is collimated to a 10 cm diameter at the CDM. Hence the images of all the optics are also 10 cm in diameter.

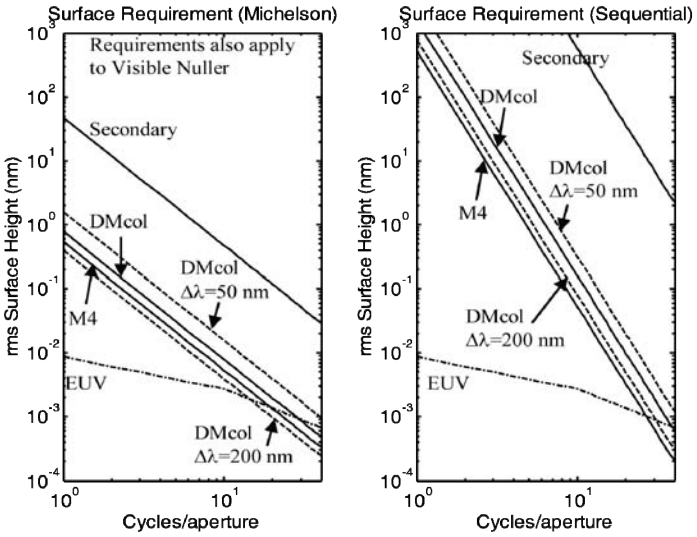


Figure 3. Surface figure requirements for Contrast = 10^{-12} for the Michelson, Visible Nuller, and sequential configurations. The curves all assume Δλ = 100 nm except otherwise indicated. The EUV curve is derived from a measured PSD and represents the state-of-the-art.

Figures 3 and 4 reveal our **first lesson:** the sequential wave front controller relaxes both the optical surface height and reflectivity uniformity requirements. In the case of the former, the size of the dark hole and/or optical bandwidth is increased. For the latter, the sequential control system is crucial because the required uniformity (a few × 10^{-5} r.m.s. at any spatial frequency for the Michelson or VN configurations, for 10^{-12} contrast) is significantly beyond the state-of-the-art for low-spatial frequencies on large optics. A higher scattered light level would have to be tolerated, at the expense of tightened requirements elsewhere in the system, and longer integration times.

In addition, the sequential controller, when implemented with 3 deformable mirrors (one at the pupil, one on each side) provides a full broad-band control redundancy against the failure of any one DM.[12]

III. High Spatial Frequencies Not Controllable by the DMs

We have to this point discussed surface spatial frequencies falling within the control bandwidth of the DM, that is spatial frequencies below *N/2* cycles/aperture, where the square DM has *N* elements per side.

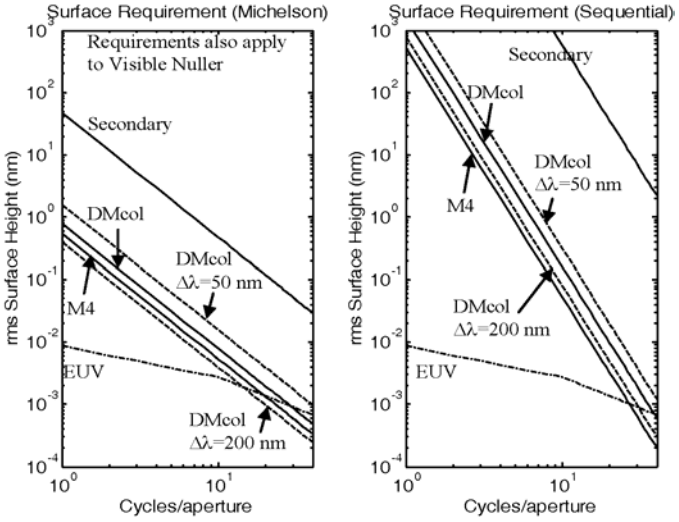


Figure 4. Surface figure requirements for C = 10^{-12} for the Michelson, VN, and sequential configurations. The Michelson and VN requirement is the horizontal line. The curves assume a 1 cycle/aperture bandwidth (= 1 speckle) and Δλ = 100 nm. Dashed line: amount reflectivity can change for a 30 nm piston of DM_{np} at a distance of 3 m from the pupil.

For the VN, that is the end of the story because the SM fibers filter all spatial frequencies above *N*/2. For the other configurations light with higher spatial frequencies (e.g., *N*/2+5 and *N*/2+12 cycles/aperture) mixes together and folds into the dark hole (at 12-5=7 cycles/aperture in this example). Give'on et al. (2006) refer to this as 'frequency folding' and discuss optimal monochromatic control using band-limited DMs[15]. In broadband light, we must deal with the λ^{-2} wavelength dependence of the pure amplitude term arising from mixing of phase corrugations. Assuming the folded light can be sensed independently of other scattered light, it can be perfectly corrected by the Michelson with its λ^{-2} amplitude dependence, but not by the sequential configuration (wavelength-independent amplitude). Here we show how the effect arises and compute the residual field for the sequential configuration.

In ref. 13, we showed that the residual contrast using the sequential controller is given by the

$$\begin{aligned}\left\langle \bar{C}(\xi)\right\rangle &= \frac{1}{6R^2}\left(\frac{4\pi}{\lambda_o}\right)^4 \sum_{n=N/2}^{\infty}\sum_{m=N/2}^{\infty} \sigma_n^2\sigma_m^2\,\delta\delta\left(\xi-(m-n)\right) \\ &= \frac{1}{6R^2}\left(\frac{4\pi}{dk\lambda_o}\right)^4 \sum_{m=N/2}^{\infty} PSD_m PSD_{m+\xi}\end{aligned} \tag{1}$$

where the surface variance in a spatial bandwidth dk cycles/m is related to the surface Power Spectral Density (PSD) by

$$\sigma_m^2 = \frac{PSD(k)}{dk\cdot dk} \tag{2}$$

and other quantities are: spectral resolution $R=\lambda_o/\delta$, and ξ, *m*, and *n* are spatial frequencies (cycles/aperture). For convenience in notation we have written the convolution as a one-dimensional quantity. However, the sum is performed over the 2-dimensional overlap of the circularly symmetric PSD functions. Note that the summation starts at *n*=*N*/2; only the spatial frequencies not controlled by the DM are folded into the dark hole.

The results of high-frequency folding for two large optics—a Very Large Telescope (VLT) primary mirror and the Hubble Space Telescope (HST) primary mirror—are shown in Figure 5. Data for the PSD of these optics is taken from Borde & Traub (2006)[16]. Surprisingly the VLT mirror at high-spatial frequencies is a very good surface, with just a 4 nm r.m.s. surface above 40 cycles/aperture. If this telescope were used uncompensated in

TPF-C, the frequency folding contrast would be 10^{-11}. After compensation at the central frequency of a 100 nm bandpass, the contrast in the sequential DM configuration is ~ 10^{-13}. This is acceptable and shows that current technology produces 8-m optics with PSDs that meet the TPF-C high-spatial frequency requirements.

We have now arrived at our **second lesson:** the frequency folding terms caused by high spatial frequencies (we evaluated frequencies > 48 cycles/aperture, requiring a 96 × 96 DM) on an existing 8 m telescope lead to acceptably small scattered light when controlled by a sequential wavefront controller. (They are fully controllable by a Michelson and filtered by the VN.) We have not evaluated print-through effects in a highly lightweighted mirror that would likely appear in an 8-m version of TPF-C. An HST-quality primary mirror is also acceptable. Note, however, that these results apply to a 100nm bandwidth: attempting to correct much larger bandwidths will raise the scattered light level above the exozodiacal background.

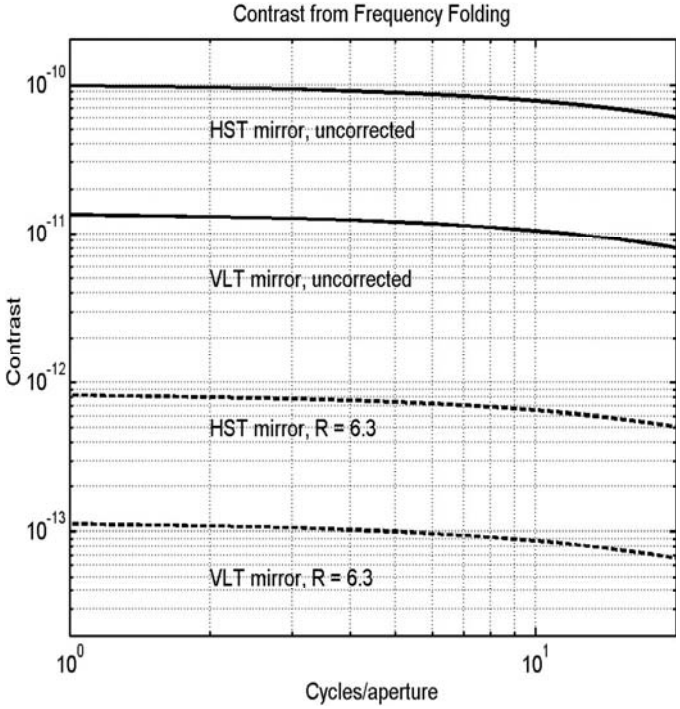


Figure 5. Contrast from frequency folding for spatial frequencies above 48 cycles per aperture, for an 8-m VLT primary and the 2.4 m HST primary. The uncompensated effect is above the required level of 10^{-12} for both mirrors. The sequential DM configuration provides about ~100× reduction of the contrast when it compensates the center of a 100 nm bandpass centered at 633 nm. Both mirrors are acceptable after compensation. The frequency folding effect can be perfectly compensated by the Michelson configuration and is not present in the Visible Nuller.

IV. Image Plane Mask Requirements

In the above discussion, wavelength dependent errors are compensated by wavelength-dependent controllers. The control happens at or near the pupil plane, so there is no spatial scaling with wavelength. When the wavefront corrugations reach the image plane, they appear at radial position given by $Nf\lambda/D$, where N is the spatial frequency (cycles/aperture), f is the focal length, D is the aperture diameter, and λ is the wavelength. Consider now an error in an image plane mask, e.g., a pinhole or a surface height bump. Its position is fixed in the image plane—in this respect a mask error looks like a planet. At a given wavelength, the incident field can be modified by the DM to place more or less light or modify the phase of the light at the location of the mask defect, recreating an ideal field at the output of the mask. But the compensation is radially scaled by λ, and decreases in amplitude as λ^{-1}. Thus image plane mask compensation using a pupil-plane or any non-image-plane DM is intrinsically bandwidth-limited. Mask requirements in broadband light have been defined by Lay et al[17] and are summarized below. We focus on the phase transmission—we ignore material dispersion, and assume that the λ^{-1} phase-amplitude generated at the DM compensates the λ^{-1} transmission through the mask substrate. In this analysis, only the radial scaling contributes to mask leakage.

Mask leakage requirements were addressed by evaluating the intensity of light leaking through a set of uncorrelated phase transmission errors represented by a grid of Gaussian bumps having random amplitudes. The width and spacing of the Gaussian bumps was scaled to study the requirements on different spatial frequencies. Table 1 shows the surface rms that produces a contrast of 10^{-11}, for a range of offsets $|\mathbf{x}_0|$ and bump widths F, and assuming a substrate refractive index $n = 1.5$ (appropriate for Silica). Two competing effects underlie the dependence on offset angle: (1) the point spread function from the input aperture is increasing rapidly as one approaches the position of the star; (2) the relative shift between the point spread functions at two different wavelengths increases linearly with angular offset from the star. In the scenario evaluated here, the former effect is dominant, and the contrast deteriorates close to the star. Table 1 also shows that the scale size of the mask

errors is important. First consider the effect of a very narrow Gaussian (small *F*). The product of the incident field and the mask transmission error is dominated by high spatial frequencies that are blocked by the Lyot stop (i.e., 'smeared' out in the convolution process). A broad Gaussian (high *F*) spans a similar region of sidelobes for the point spread function at two different wavelengths and the DM correction becomes more effective. It is the intermediate spatial scales—those comparable to the width of the diffraction ring of the input point spread function—that have the largest impact on the contrast. We can also use the 'grid of Gaussian bumps' model to derive a requirement on the power spectrum of the surface roughness, as illustrated in Figure 6. Each curve on the plot represents the power spectral density (PSD) of substrate surface height for one value of the bump width and spacing *F*. The cases with broad bumps contain no power at high spatial frequencies and roll off at a spatial frequency proportional to F^{-1}. Each of the 5 spatial scales (i.e., set of Gaussian bumps of a particular size) have been allocated a contrast contribution of 10^{-11}, and evaluated to determine the vertical normalization of each PSD. The sum of the PSDs gives the net requirement on the surface; a surface matching this PSD will give a contrast of 5×10^{-11}. The contrast is proportional to the PSD; lowering the PSD by a factor of 10 will give a contrast of 5×10^{-12}. The shape is somewhat flexible, since we are free to reallocate the desired total contrast to the separate spatial scales. There is no requirement on the surface for spatial frequencies higher than 1 cycle per $f\,\lambda/D$, since the scattered light is blocked entirely by the Lyot stop.

Table 1: RMS wave front error for 1e-11 contrast at various locations in the image plane.

		Spatial scale of mask error, F (units λ_c / D_{in})				
		0.125	0.25	0.5	1	2
$\lvert\mathbf{x}_0\rvert$ (units λ_c/D_{in})	3	22 pm	17 pm	15 pm	22 pm	68 pm
	4	27 pm	20 pm	17 pm	26 pm	81 pm
	5	31 pm	23 pm	20 pm	30 pm	91 pm
	10	52 pm	39 pm	33 pm	45 pm	136 pm

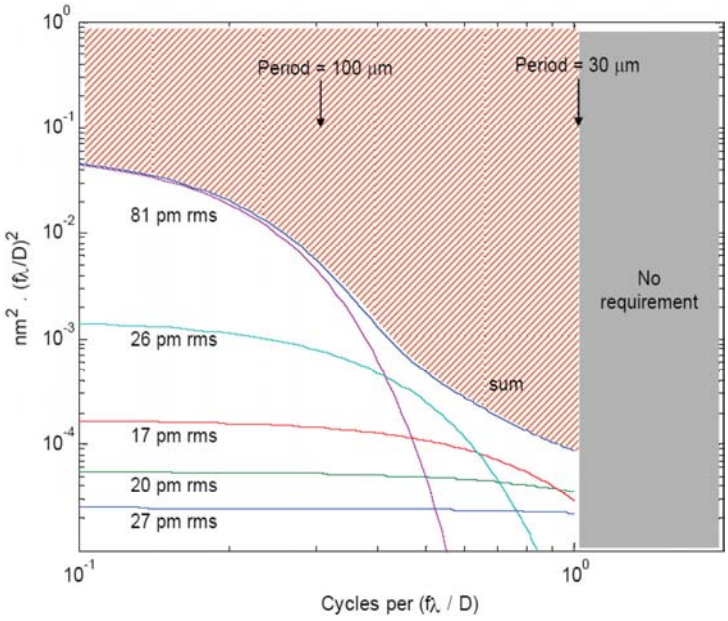


Figure 6: Power spectral densities for Gaussian mask errors with different spatial scales. The surface rms allocated to each curve results in a contrast contribution of 10^{-11} at an angular offset from the star of 4 λ/D (Table 1). Spatial frequencies higher than 1 cycle per f λ/D do not impact the contrast. The spatial periods shown at the top apply to the case where f / D = 0 and λ = 550 nm.

We now compare the derived requirement for the substrate surface rms to values that have been obtained in practice. We assume that the TPF-C optical system has $f / D = 60$, for which the spatial scales of interest have periods of 30 μm or more (Fig. 6). Using λ = 550 nm, the lower right corner of the hatched region in Figure 6 has a spatial frequency of 1 cycle per 30 μm or 0.03 μm^{-1}, and a PSD of 10^5 nm^4. Weis demonstrated superpolishing of sapphire with an rms of 70 pm for scales between 12.5 μm and 500 μm[18]. Duparre et al.[19] show power spectra for a fused silica substrate with a PSD of 10^5 nm^4 at 0.03 μm^{-1} (see their Fig. 18). We conclude that the substrate roughness requirement for a contrast contribution of 5×10^{-11} is within the current state-of-the-art.

We now have **lesson 3**: image plane masks require better than 1 Angstrom phase transmission for spatial scales up to 2 $f\lambda/D$. Mask errors look like planets because they do not scale with wavelength. While this is a disadvantage with respect to wave front control, it does simplify calibration of errors. Being fixed in the mask, the leakage spots can be rotated or translated with or independent of the telescope. This may allow a relaxation of mask requirements but requires detailed modeling to make the case.

V. Stability Requirements

Optical stability requirements have driven the design of TPF-C since its inception. The requirement to hold the shape of the optics led to the V-groove thermal shroud and the isothermal cavity. Jitter requirements call for an active vibration suppression system (though a passive system only violates a few requirements and could be made to work). Pointing control is achieved by few-milli-arcsecond rigid body stability combined with secondary mirror and a downstream fine-guiding mirror control. Metrology with a precision of ~ 25 nm is used to monitor the primary-secondary despace and shear. Dynamics requirements and an overview of the optical control systems and modeling approach are described by Shaklan et al.[20]

We have adopted a 'set and forget' approach to wave front control. In this approach, the wave front is controlled once at the beginning of an observation, and is required to remain stable for the duration T of the observation. This approach places demanding requirements on the system since observations can be long (~ 1 day) and stability requirements are extremely challenging, as will be shown below. However, a set-and-forget approach is straightforward to model—we can define clear engineering requirements for well-defined periods. The requirements are not blurred by various control bandwidths, and they are not relaxed by as-yet unproven high-dynamic range wave front control. In the future, once detailed end-to-end models of active wave front control demonstrate the ability to reduce speckles to a level of 1×10^{-10} while estimating their value to 2×10^{-11}, we will consider changing to an active approach and subsequently relaxing temporal stability requirements.

The dynamics requirements are driven by three factors: the sensitivity of the system to changes in low-order aberration content; the combination of mask errors with changing aberration content; and the sensitivity of the system to beam shear across imperfect optics. Three different masks have been considered: a 4th-order band-limited mask[21], an 8th-order band-limited mask,[22,23] and a shaped-pupil mask.[10,24] We note that the visible nuller is mathematically indetical to a 4th-order mask with the aberration sensitivity, throughput, and image plane response of a 1-cosine 4th-order mask.

Figure 7 compares the aberration sensitivity of the 4th- and 8th-order band-limited masks. The 8th-order mask has far superior aberration sensitivity, allowing a relaxation of several orders of magnitude in the wave front stability budget. The good aberration rejection comes at the expense of reduced throughput and a larger diffraction spot. The shaped-pupil sensitivities are generally better than the 4th-order masks (and worse than the 8th) while exhibiting the same 2nd-order slope and reduced throughput.

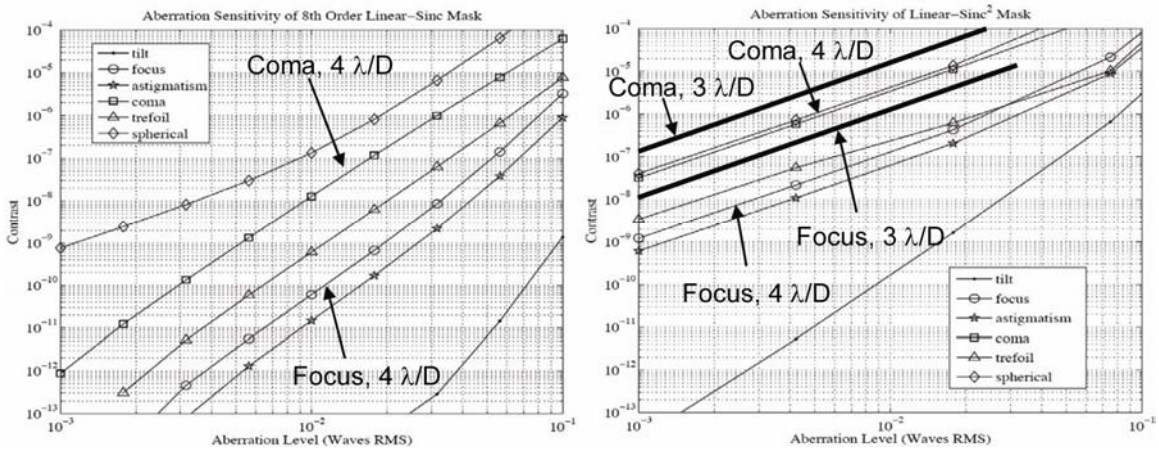


Figure 7. Aberration sensitivity of linear sinc^2 mask and linear eighth-order mask for low-order Zernike modes. For astigmatism, coma, and trefoil we display the more sensitive of the two orthogonal modes. Left: 8th-order mask. Right: 4th-order mask. Curves apply to a working angle of 4 λ/d unless otherwise indicated. The Visible Nuller has ~ the same aberration sensitivity as the 4th-order mask.

Mask errors allow aberrations to leak through the image plane and Lyot stop, and can destroy the aberration rejection properties of the band-limited designs. Even with the amplitude transmission error set to 5×10^{-4} at 4 λ/D (an intensity leakage of 1 part in 1000), mask errors combined with aberrations account for 11% of the TPF-C contrast error budget.[20]

Our models have shown that aberration sensitivity increases by an order of magnitude at 3 λ/D compared to 4 λ/D, while 2 λ/D is another 10× worse. Additionally, to achieve these small working angles, the Lyot stop diameter (or for the visible nuller, the transmitted pupil) decreases, which in turn decreases throughput and increases integration times. Thus wave front control system requirements become extraordinarily difficult due to the increased stability time frame and reduced flux. We see 3 λ/D as a practical wall for these coronagraph approaches, and have chosen to work at 4 λ/D where we find that our thermal and vibration isolation systems meet the requirements without requiring major new technological breakthroughs.

Beam walk is another important factor that limits coronagraph performance.[25] Beam walk is the shearing of the optical beam across the optics as a result of pointing errors or optical misalignments. Shear across imperfect optics changes the beam's spatial frequency content, resulting in changes to the speckle amplitudes and phases in the coronagraph dark hole. The assumed PSD of the optics[20] (which is substantially worse than the state of the art but high quality nonetheless) limits motions to a few nano-radians and a few nm, depending on the location in the optical train. Figure 8 shows the key stability requirements imposed by the dynamics error budget.

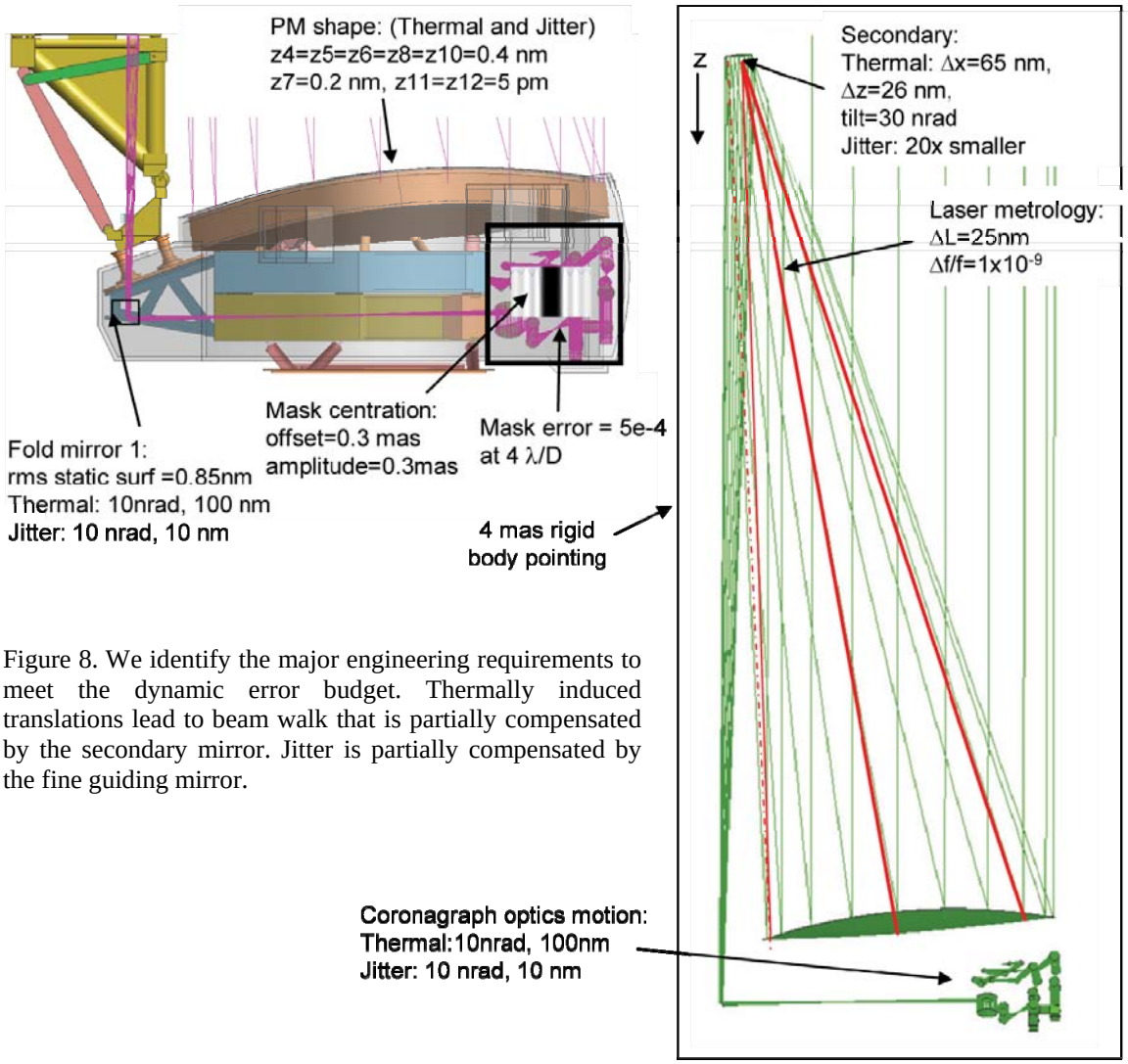


Figure 8. We identify the major engineering requirements to meet the dynamic error budget. Thermally induced translations lead to beam walk that is partially compensated by the secondary mirror. Jitter is partially compensated by the fine guiding mirror.

If the secondary mirror is not actively pointed, beam walk is driven by rigid body motions of the telescope. This places sub-milli-arcsecond (mas) requirements on the rigid body pointing. If the secondary mirror is used to compensate the pointing error (so that after the secondary mirror the beam is traveling parallel to the optical axis), then beam walk occurs only at the primary mirror where a large motion can be tolerated. With this architecture, the rigid-body pointing stability and offset requirement are ~ 4 mas (comparable to HST).

This approach does not work with 4^{th}-order masks, however. This is because tip-tilt compensation using the secondary mirror introduces wavefront aberrations that leak past the 4^{th}-order mask. With a steerable secondary, we are faced with sub-mas rigid body control and the additional requirement of sub-nm metrology between the primary and secondary. The same thing happens using 8^{th}-order mask working at 2 λ/D.

Finally we come to lesson 4: working at 2 or 3 λ/D is much, much harder than 4 λ/D. Breakthroughs in wave front control, optical surface stability, and a change in observing paradigm are needed. If an 8^{th}-order mask is not used (to achieve higher throughput, for example) then optical surface stability requirements move into the single-digit picometer regime for low-order aberrations and the sub-pm regime for spherical aberration and higher order terms. Rigid body pointing becomes extremely tight, and faster, more accurate wave front control is needed.

VI. Summary

The design reference mission modeling provides the flow down of science requirements to engineering requirements. To date, the DRM focus has been on observational completeness; requirements for orbit determination, disambiguation, and characterization have not been analyzed are not yet driving the engineering requirements. Things won't get easier! On the other hand, we have carried margin throughout the error budgeting process so it is possible that the new requirements can be absorbed by the margin.

We have a good handle on the surface height and reflectivity requirements through the system. We can now relate the requirements to optical bandwidth and several control options. We note that correction beyond ~ 25 cycles/aperture does is feasible at the 10^{-10} – 10^{-11} level in our current design because optical requirements exceed the performance of state-of-the art (< 0.3 nm rms) optics (see Fig. 3). The outer working angle grows with reduced optical bandpass, and may grow with improved designs that move optics away from intermediate image planes.

We also have a good handle on image plane mask requirements. These masks require superpolished surfaces (< 1 Å rms) over areas of a few hundred square microns. While this looks feasible, it has yet to be demonstrated. We also require the transmission to be good to a part in 1000 in the transition region from 3–5 λ/D.

Stability requirements have been studied in great detail and have been shown to meet requirements.[8,26] The requirements are met because our baseline approach is to work at angles >= 4 λ/D. It is tempting to design a more aggressive coronagraph, one that has higher resolving power to see planets closer to their stars, and planets around more distant stars. Indeed we are constantly looking for ways to do this. But our analysis clearly points to the virtual wall at smaller angles—scattered light increases dramatically at smaller working angles in large part because the wings of the Airy pattern are signficant in this region, and the Airy pattern itself grows when the Lyot stop is sized for small working angles. At smaller angles, the system architecture changes to one in which rigid body pointing is required with sub-mas precision.

Breakthroughs in wave front sensing and control could change these conclusions and change the observing paradigm. For example, the ability to measure speckles to 2×10^{-11} (while not confusing them with the planet) while controlling them to ~ 10^{-10} would relax the stability time frame requirement from many hours to the few minutes needed to collect the light at SNR = 1 and contrast = 10^{-10}. To date, model performance is a few orders of magnitude from predicting the speckle content at 10^{-11}; we will forge ahead with our set-and-forget observing paradigm while improving our modeling capabilities and searching for more optimal ways to perform planet detection.

REFERENCES

1. TPF-C Science Science and Technology Definition Team Final Report, JPL document number D-34923, 2006.
2. R. A. Brown, ApJ, 607, 1003 (2004)
3. R. A. Brown, ApJ, 624, 1010 (2005a)
4. R. A.Brown, Proc. IAU Colloquium No. 200, Proceedings Direct Imaging of Exoplanets: Science & Techniques, ed. C. Aime & F. Vakili, (France: Cambridge University Press, 2005b)
5. R. A. Brown, ApJ, 607, 1003 (2004)
6. R. A. Brown, this workshop
7. S. Hunyadi, & S. Shaklan, this workshop.
8. Terrestrial Planet Finder Coronagraph Flight Baseline 1 report, JPL 2005.
9. B. Kern, Band-limited masks, this workshop.
10. R. Belikov Shaped-pupil masks, this workshop.
11. M.G. Littman, M. Carr, J. Leighton, E. Burke, D. N. Spergel, and N. J. Kasdin, "Phase and amplitude control ability using spatial light modulation and zero path length difference Michelson interferometer," Proc. SPIE **4854**, 405-412 (2003).
12. S. B. Shaklan & J. J. Green, "Reflectivity and Optical Surface Height Requirements in a Broad-band Coronagraph 1: Contrast Floor Due to Controllable Spatial Frequencies," Appl. Opt, 45, 5143-5153 (2006).
13. S. B. Shaklan, J. J. Green, & D. Palacios, "The Terrestrial Planet Finder Optical Surface Requirements," Proc. SPIE **6265** (2006).
14. B. P. Mennesson et al, "Optical Planet Discoverer: How to Turn a 1.5m Class Space Telescope into a Powerful Exo-planetary Systems Imager," Proc. SPIE **4860**, 32-44 (2003). Also, see M. Shao and B. M. Levine, this conference.
15. A. Give'on et al, "On Representing and Correcting Wavefront Errors in High-Contrast Imaging Systems," JOSA-A **23**, 1063-1073 (2006).
16. P. J. Borde & W. A. Traub, "High-Contrast Imaging from Space: Speckle Nulling in a Low-Aberration Regime," ApJ **638**, 488-498 (2006).
17. O. Lay, J. J. Green, D. J. Hoppe, & S. B. Shaklan, "Coronagraph Mask Tolerance for Exo-Earth Detection," Proc. SPIE 5905, 148-161 (2005).
18. O. Weiss, "Direct contact superpolishing of sapphire", Appl. Opt. 31, 4355 (1992).
19. A. . Duparre, et al., "Surface Characterization Techniques for Determining the Root-Mean-Square Roughness and Power Spectral Densities of Optical Components", Appl. Opt. 41, 154 (2002).
20. S. B. Shaklan, L. Marchen, J. J. Green, & O. P. Lay, "Terrestrial Planet Finder Coronagraph Dynamics Error Budget," Proc. SPIE 5905, 110-121 (2005).
21. M. J. Kuchner & W. A. Traub, Ap. J. 570 900 (2002).
22. M. J. Kuchner, J. Crepp, & J. Ge, "Eighth-order Masks for Terrestrial Planet Finding," Ap. J. 628, 466-473 (2005).
23. S. B. Shaklan & J. J. Green, "Low-order Aberration Sensitivity for Eighth-order Coronagraph Masks," Ap. J., 628, 474-477 (2005).
24. N. J. Kasdin et al, "The Shaped Pupil Coronagraph For Planet Finding Coroangraphy: Optimization, Sensitivity and Laboratory Testing," Proc. SPIE 5487, 1312-1321 (2004).
25. C. Noecker, "TPF-C Wavefront Changes Due to Beam Walk," Proc. SPIE 5905, 122-127 (2005).
26. C. Blaurock et al, "Passive Isolator Designs for Terrestrial Planet Finder Coronagraph," Proc. SPIE 567, 282-293 (2005).

A Review of Coronagraph Laboratory Results

N. Jeremy Kasdin
Mechanical and Aerospace Engineering, Princeton University

ABSTRACT

This brief paper, and the attached figures, is a summary of a talk given at the 4th International TPF/Darwin Workshop, held on November 8–10, 2006 in Pasadena, CA. The workshop brought together researchers from around the world to discuss progress in developing a joint program for searching for and characterizing exosolar earthlike planets from space. The talk upon which this paper is based summarized the current status of experimental work to develop and test various coronagraph concepts for achieving high contrast in a filled aperture telescope. It is based on work at a myriad of laboratories and the author is indebted to the researchers who generously cooperated and provided the information and images for the paper.

1. Introduction

There are two approaches for creating high contrast in order to image extrasolar planets: coronagraphy, where a single aperture telescope is combined with some method for blocking the starlight, either internally or externally, and interferometry, where the light from multiple satellites is combined to null the starlight. Both are being pursued, though current expectations are that the first mission to fly will involve some type of coronagraph. There is much still to do, and much to learn about the way to design and implement coronagraphs and how to combine with a wavefront control system. Nevertheless, progress has been tremendous in the past few years. In this brief paper I will summarize progress and results in roughly a dozen laboratories around the world, each studying and experimenting with a different type of coronagraph and/or method of correcting the wavefront. It is thus by necessity a broad overview with only minimal explanation and detail; I also make no claims to be exhaustive. I will not try and present the theoretical basis or a performance analysis of each method; for that there are many useful papers. I will also not try and compare different coronagraphs. Instead, I focus only on the laboratory results to give a snapshot in time of our successes and failures, of actual measured results, at achieving high contrast.

I have many objectives. Of course, first is simply to give a summary and overview of laboratory activities in coronagraphs for both ground and space; this should provide the reader with both a sense of the broad base of research and the great strides being made toward TPF level performance. I will also try, however, to place these various laboratory results into the larger context and give a sense of how they support a pathway to a space telescope for finding planets. I will highlight the various limitations of ground experiments to show the difficult problems being tackled and the steps that still need to be taken. I also hope to provide a flavor of the diversity of coronagraph approaches being proposed and worked on and the various implementation issues for each. And, finally, I hope to get the reader thinking about the difficult questions that need to be asked and answered and how close (or far) we are to a space ready system.

2. Why Experimental Work?

While this may seem like an obvious question, it is useful to reflect briefly on the importance of performing laboratory research and experimental verification of the various concepts. At its most basic, successfully achieving high contrast in the laboratory provides of proof of principle that we are pursuing an achievable goal. Any TPF design, or progress of the TPF program, must be preceded by a clear an unambiguous demonstration in the laboratory that high contrast in a coronagraph is possible. As we'll see, the most recent results at JPL are remarkably close.

Experimental work is more than that, however. By moving from the notebook to the laboratory we discover what the most difficult problems actually are and uncover a myriad of surprises. No matter how careful we are in our analyses, nothing replaces laboratory work and there is always something new to learn. Additionally, the experience we gain in manufacturing and in testing will prove indispensable when we begin an actual program.

The act of performing experimental tests in an of itself is also a source of frequent inspiriation for new ideas and approaches. Finally, the facilities we are developing today will become the test labs for tomorrow for demonstrating the feasibility of a flight ready system.

What is our overall goal? Simply stated, it is to demonstrate a contrast of at least 10^{-10} in a realistic lab environment with a full, integrated, flight like system. This is an ambitious goal and not one that will be achieved quickly or precipitously. The distributed effort I will describe across multiple institutions reflects a measured and incremental approach, slowly reaching for deeper contrast and attempting to understand the limitations and roadblocks at each step. These limitations fall into five broad categories, and articulating them here will help us understand the status of many of the experimental results to come:

1. Design Limitations
2. Practical Limitations
3. Manufacturing Limitations
4. Algorithmic Limitations
5. Experimental Limitations

Design limitations refer to inherent characteristic of a given coronagraph that might prevent it from achieving the broadband contrast we seek (a good example are chromatic limitations of certain designs). Practical limitations refer to the realities of implementing a coronagraph that aren't captured by the idealized model (a good example are polarization and edge effects). Manufacturing limitations refer to the difficulty of making masks, stops, and the like (examples include chromatic image plane masks, fine structures in shaped pupils, surface grinding of pupil mapping mirrors, and others). Algorithmic limitations refer to the performance of various wavefront sensing and control algorithms when integrated with various coronagraphs (an example here is the interaction of classical speckle nulling with errors induced by shaped pupil truncation). Finally, experimental limitations refer to the fidelity of the test environment (detector performance, thermal variations, vibrations, etc.).

Interestingly, the first four of these issues apply both to the ground testing as well as to a final space mission, so understanding and overcoming them is fundamental to the eventual design of a TPF coronagraph. The fifth, on the other hand, is unique to the laboratory environment and simply limits our ability to measure and assess coronagraph performance. These either need to be well understand, and corrected for, or reduced to below the level where they influence performance. One of the distinctions among the various laboratories is the degree to which they have reduced these interfering affects (the High Contrast Imagine Testbed (HCIT) at JPL is by far the highest fidelity test environment available).

3. Uncorrected Performance

This section and the next encompass the heart of the presentation. In these two sections I make a distinction between the achieved performance in what I call the "raw" coronagraph, that is, the coronagraph alone without a wavefront sensing and control system, and corrected performance where the coronagraph is integrated with one or more deformable mirrors (or other wavefront correction device). This separation of the coronagraph from the correction scheme is typical of all the design teams and sensible for this stage of research. In other words, we want to understand the coronagraph as well as we can and find its limits of performance and then augment it by a WFSC system to determine what level of correction is possible. However, in my final remarks, I'll discuss briefly the need to start viewing these as an integrated high contrast system and understanding how to optimize design and performance of both together.

Figure 1 is a graphical depiction of what has been accomplished by coronagraphs alone. The first two arrows give a sense of what needs to be achieve at the smallest inner working angles (here I assuming an inner working angle of roughly 4 λ/D). At that angle, the Airy function has a contrast of roughly 10^{-3} compared to the needed contrast of 10^{-10}. Note that all coronagraphs, by design, theoretically achieve this contrast in monochromatic light (for some, the optical vortex in particular, there is no accepted design approach yet that maintains contrast in broadband light).

The inability to reach the high contrast results from a variety of sources. As shown by the right arrow, the newest coronagraph, the optical vortex, has the worst uncorrected performance, not much better than the Airy function. This comes mostly for the difficulty manufacturing the vortex mask. The PIAA tests at Subaru have achieved roughly a factor of 10 improvement over the Airy function without correction. The error here arises from the difficulty manufacturing the two mirrors, M1 and M2, performing the mapping. The poor optical quality of those mirrors, and the use of only a single post-apodizer, results in a severe degradation in contrast. The remaining coronagraphs all cluster around roughly 10^{-5}. This seems to be a canonical uncorrected contrast level for most facilities. There is general agreement that this residual contrast is due to phase errors on the optical components. In fact, the goal of every laboratory is to reach the point where performance is limited by the optical quality. Then, wavefront control is necessary.

Two facilities have achieved contrast quite a bit better, reaching down to as low as 10^{-7} without correction, the ExAO testbed at Santa Cruz and the experiments for SPICA at ISAS in Japan. This improved performance is due entirely to the much smaller size of the optics (and masks); SPICA, for instance, is using 1 mm masks in their shaped pupil tests. These smaller masks are much easier to super polish and have dramatically reduced phase errors. The third high contrast results occurred at the University of Colorado at Boulder in tests of the Occulter. The only laboratory experiments to date for the occulter have also achieved a contrast level of 10^{-7}. This improvement is due to the insenstivity of the occulter to wavefront error. The primary limitations in the occulter experiment were due to light scattering of the tips and edges of the mask.

Figures 2 to 19 show the various laboratory facilities and summarize the uncorrected performance at each.

4. Corrected Performance

Figure 20 is a reprise of Figure 1 but now with the addition of corrected results. These should be used with care, however, as different approaches and algorithms have been used for correction with the various coronagraphs and many are transportable to other concepts. However, it does provide a useful snapshot of the progress being made in wavefront control. There are also few systems displayed since only a small subset of the facilities have deformable mirrors and the capability of performing high fidelity wavefront control. Details of the wavefront control implementation and the various results are shown in Figures 21 to 33.

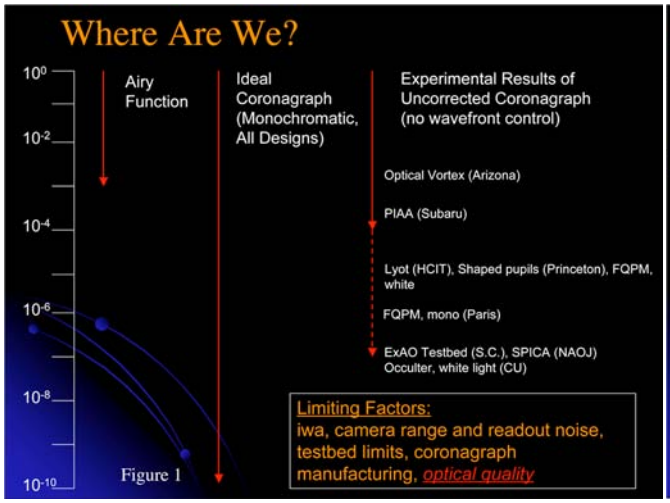
At the high end are the shaped pupil experiments at Princeton, pupil mapping at Subaru, and the visible nuller operating in a 15% band at JPL, all of which are operating at roughly 10^{-6} contrast. Here, there are a variety of limitations and no consistent reason for this level of contrast; it is mostly coincidental. This following figures go into more detail.

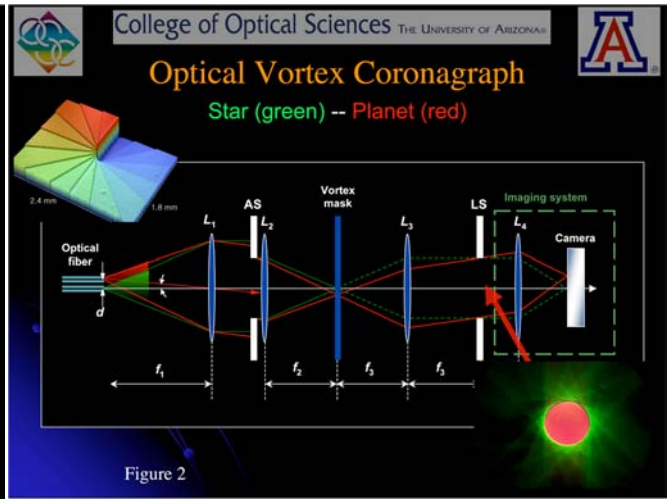
The ExAO testbed at Santa Cruz and the Visible Nuller group at JPL have both achieved a contrast of 10^{-7} in monochromatic light. The primary limitation in the ExAO results arises from amplitude errors on the optics, which were not being corrected, as well as scattered light (this is consistent with results at Princeton and JPL, where amplitude errors seem to kick in at roughly a contrast level of 10^{-7}). The primary limitation in the Visible Nuller was from errors in the optical lengths. It should also be noted that stability limitations limit to contrast to short time windows of a few seconds. It should also be noted that these results are for only a single subaperture using a translation stage rather than a DM; no results have been reported on a full pupil experiment.

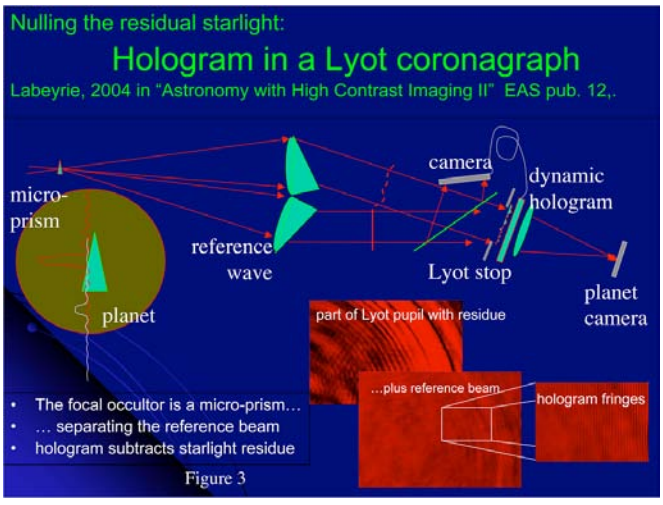
The only facility to achieve higher contrast has been the HCIT. There, contrasts of 10^{-8} were achieved with shaped pupils using classical speckle nulling last winter. More recently, a contrast of 10^{-9} was reached in monochromatic light with a bandimited Lyot coronagraph and classical speckle nulling. Unfortunately, contrast degraded by roughly a factor of 10 when a 10% bandwidth was tried. It is believed that the main polychromatic distortion was caused by chromatic phase shifts in the image plane masks. Current work is directed at new materials and manufacturing approaches to improve the wideband performance. Development of a larger format DM is also proceeding to increase the size of the dark field. Finally, a number of groups, including Princeton, Subaru, and JPL have begun experiments with optimal dark field algorithms to improve convergence speed.

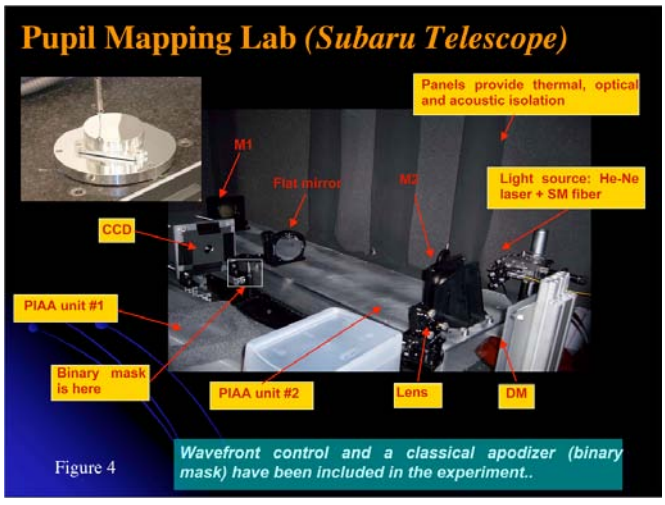
5. Acknowledgements

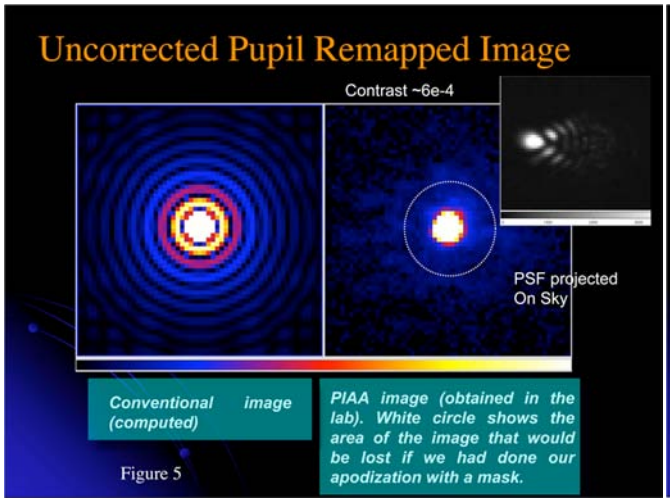
This presentation would not have been possible without the generous help and cooperation from many researchers around the world. They provided me with numerous images and patiently responded to all of my questions, no matter how trivial. By name they are O. Guyon of the Subaru Telescope, NAOJ, J. Evans and B. Macintosh of the Lawrence Livermore National Laboratory and the Laboratory for Adaptive Optics, S. Shaklan, J. Trauger, and M. Levine of the Jet Propulsion Laboratory (JPL), W. Cash of the University of Colorado, J. Schneider, A. Boccaletti, and P. Baudoz of the Paris Observatory, A. Labeyrie of the Collège de France, J. Codona of the Steward Observatory at the University of Arizona, G. Swartzlander, also of the University of Arizona, K. Enya of the Institute of Space and Astronautical Science of the Japan Aerospace Exploration Agency (ISAS/JAXA), and R. Belikov of Princeton University. While I extend my thanks for their contributions, all errors, omissions, misrepresentations, or blunders are solely my own.

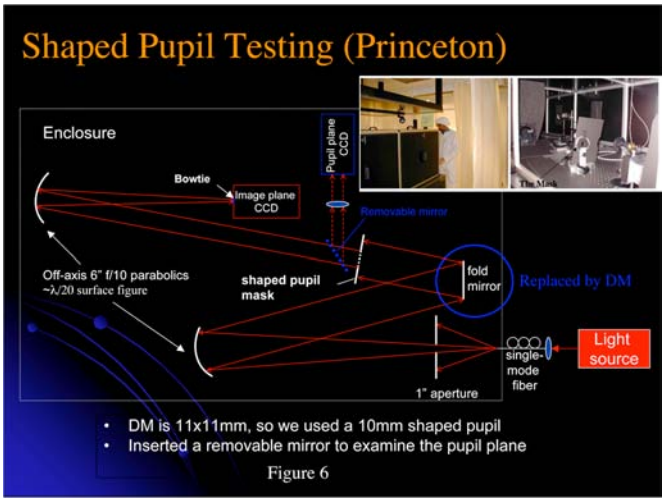


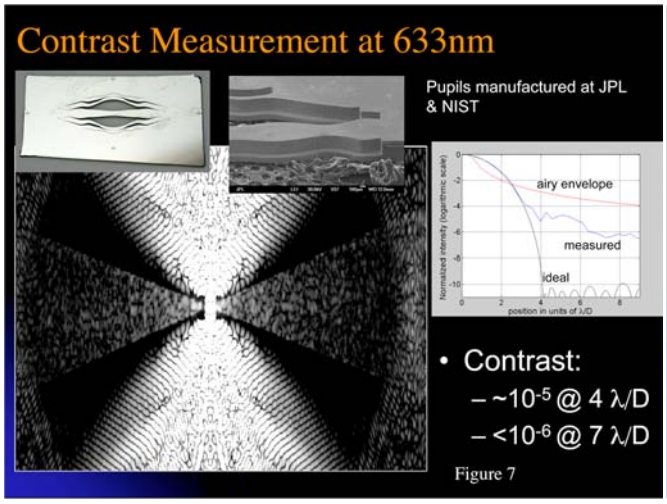


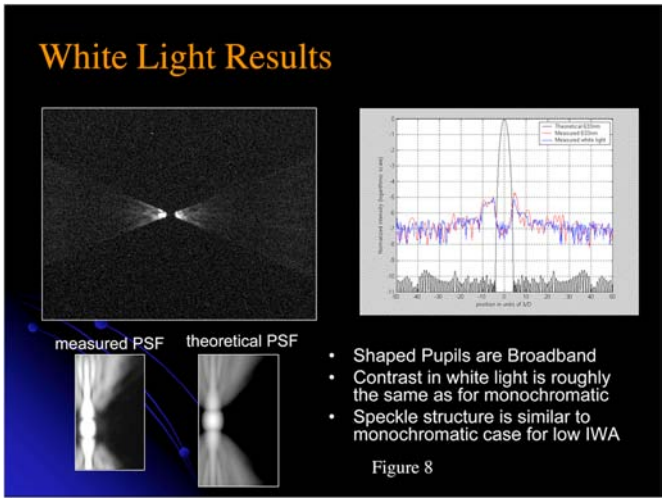


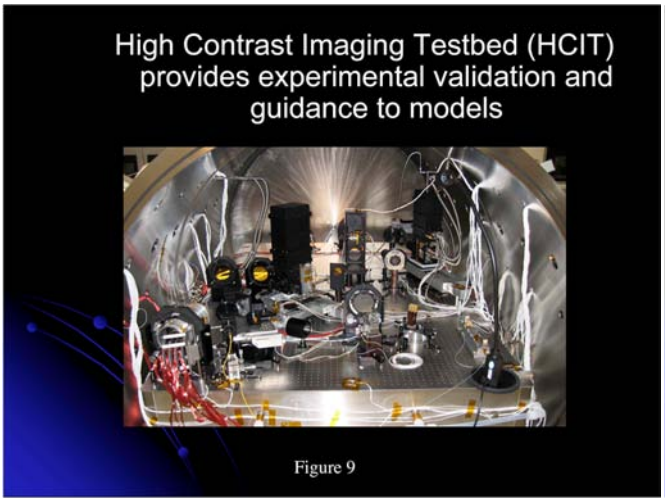


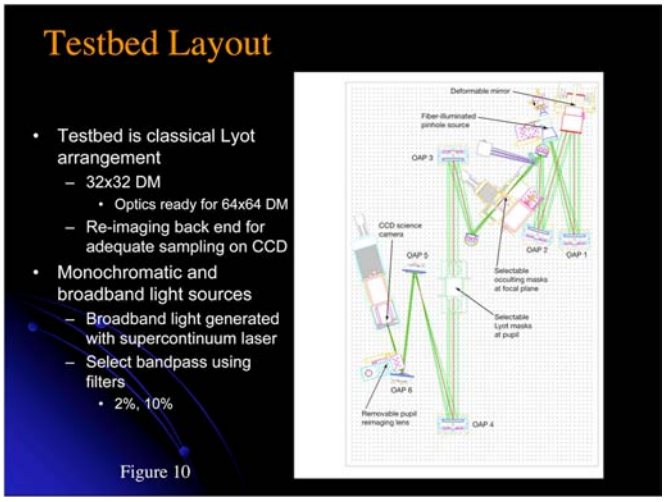

Uncorrected Pupil Remapped Image
Contrast ~6e-4
PSF projected On Sky
Conventional image (computed)
PIAA image (obtained in the lab). White circle shows the area of the image that would be lost if we had done our apodization with a mask.
Figure 5

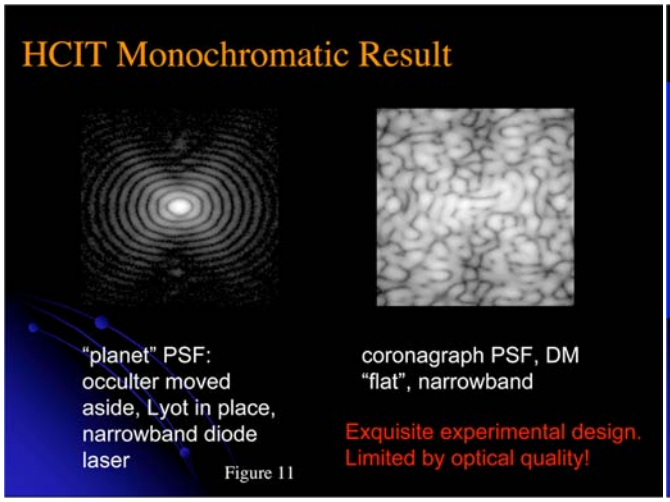

Shaped Pupil Testing (Princeton)
Enclosure
Pupil plane CCD
Bowtie
Image plane CCD
Removable mirror
Off-axis 6" f/10 parabolics ~λ/20 surface figure
shaped pupil mask
fold mirror
Replaced by DM
Light source
single-mode fiber
1" aperture
• DM is 11x11mm, so we used a 10mm shaped pupil
• Inserted a removable mirror to examine the pupil plane
Figure 6

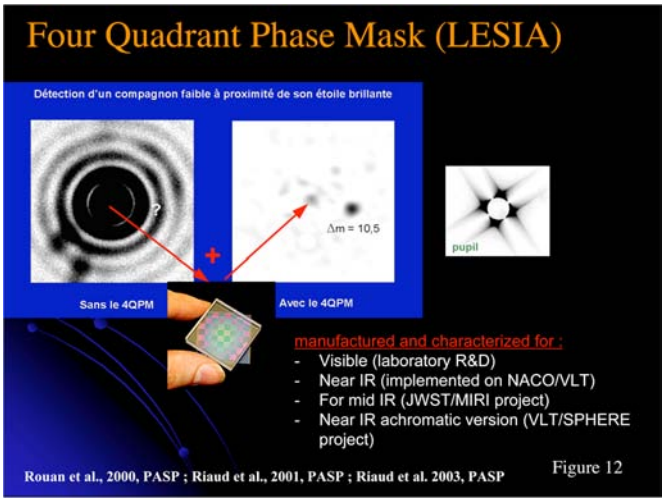

Contrast Measurement at 633nm
Pupils manufactured at JPL & NIST
airy envelope
measured
ideal
• Contrast:
– ~10-5 @ 4 λ/D
– <10-6 @ 7 λ/D
Figure 7

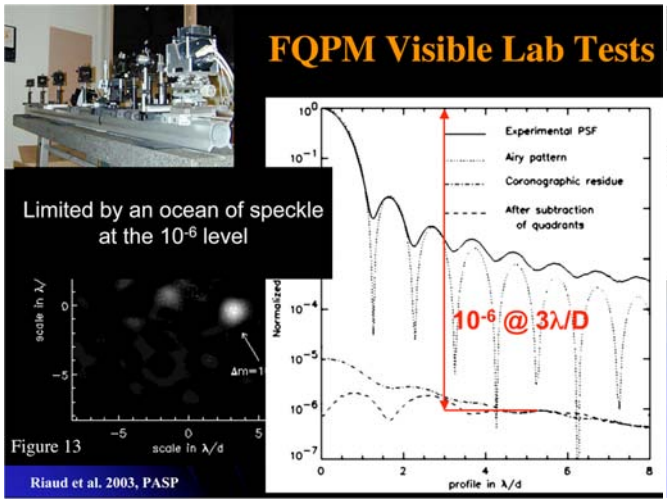

White Light Results
measured PSF
theoretical PSF
• Shaped Pupils are Broadband
• Contrast in white light is roughly the same as for monochromatic
• Speckle structure is similar to monochromatic case for low IWA
Figure 8

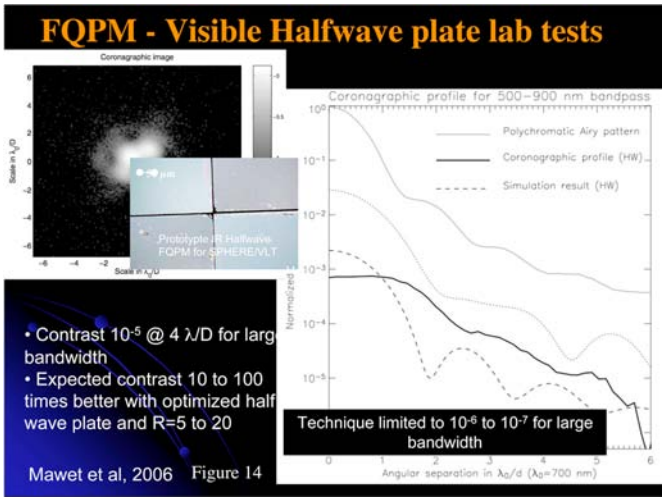

High Contrast Imaging Testbed (HCIT) provides experimental validation and guidance to models
Figure 9

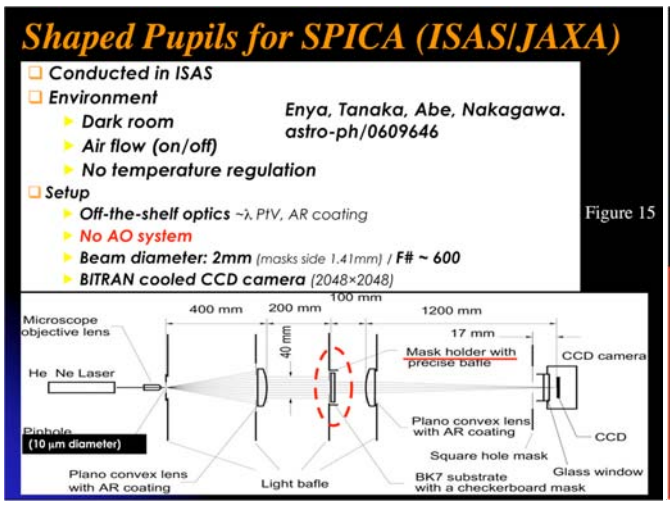

Testbed Layout
• Testbed is classical Lyot arrangement
– 32x32 DM
• Optics ready for 64x64 DM
– Re-imaging back end for adequate sampling on CCD
• Monochromatic and broadband light sources
– Broadband light generated with supercontinuum laser
– Select bandpass using filters
• 2%, 10%
Figure 10

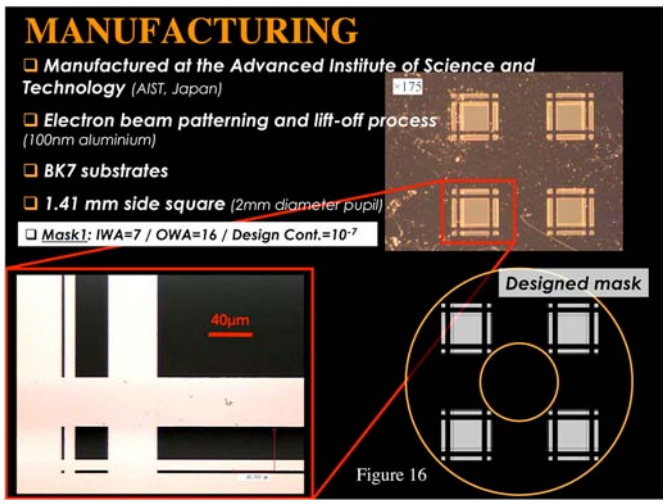

HCIT Monochromatic Result
"planet" PSF: occulter moved aside, Lyot in place, narrowband diode laser
coronagraph PSF, DM "flat", narrowband
Exquisite experimental design. Limited by optical quality!
Figure 11


Four Quadrant Phase Mask (LESIA)
Détection d'un compagnon faible à proximité de son étoile brillante
Δm = 10,5
pupil
Sans le 4QPM
Avec le 4QPM
manufactured and characterized for :
- Visible (laboratory R&D)
- Near IR (implemented on NACO/VLT)
- For mid IR (JWST/MIRI project)
- Near IR achromatic version (VLT/SPHERE project)
Rouan et al., 2000, PASP ; Riaud et al., 2001, PASP ; Riaud et al. 2003, PASP
Figure 12


FQPM Visible Lab Tests
Limited by an ocean of speckle at the 10-6 level
Experimental PSF
Airy pattern
Coronographic residue
After subtraction of quadrants
10-6 @ 3λ/D
Normalized
profile in λ/d
scale in λ/d
Δm=1
Figure 13
Riaud et al. 2003, PASP


FQPM - Visible Halfwave plate lab tests
Coronagraphic image
Coronagraphic profile for 500–900 nm bandpass
Polychromatic Airy pattern
Coronographic profile (HW)
Simulation result (HW)
• Contrast 10-5 @ 4 λ/D for larg bandwidth
• Expected contrast 10 to 100 times better with optimized half wave plate and R=5 to 20
Technique limited to 10-6 to 10-7 for large bandwidth
Angular separation in λ0/d (λ0=700 nm)
Mawet et al, 2006 Figure 14


Shaped Pupils for SPICA (ISAS/JAXA)
Conducted in ISAS
Environment
Dark room
Air flow (on/off)
No temperature regulation
Enya, Tanaka, Abe, Nakagawa. astro-ph/0609646
Setup
Off-the-shelf optics ~λ PtV, AR coating
No AO system
Beam diameter: 2mm (masks side 1.41mm) / F# ~ 600
BITRAN cooled CCD camera (2048×2048)
Figure 15
Microscope objective lens
He Ne Laser
Pinhole (10 μm diameter)
Plano convex lens with AR coating
Light bafle
400 mm
200 mm
100 mm
40 mm
1200 mm
17 mm
Mask holder with precise batle
CCD camera
Plano convex lens with AR coating
CCD
Square hole mask
BK7 substrate with a checkerboard mask
Glass window


MANUFACTURING
Manufactured at the Advanced Institute of Science and Technology (AIST, Japan)
Electron beam patterning and lift-off process (100nm aluminium)
BK7 substrates
1.41 mm side square (2mm diameter pupil)
Mask1: IWA=7 / OWA=16 / Design Cont.=10-7
40μm
Designed mask
Figure 16

MANUFACTURING

- No central obstruction design
- IWA=3 / OWA=30 / Design Cont.=10^{-7}

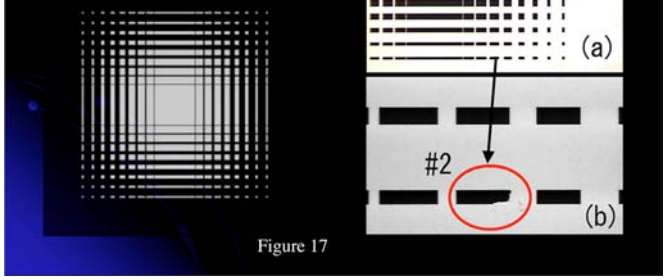
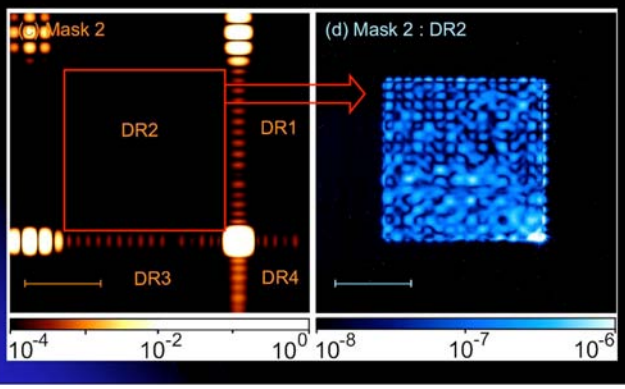


Figure 17

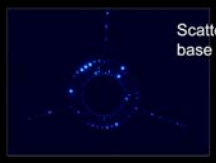
Uncorrected Performance

Figure 18

- **Mask2:** IWA=3 / OWA=30 / Design Cont.=10^{-7}/ Throughput=24%



Starshade Experiments (Colorado)



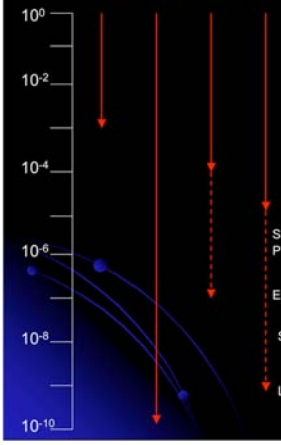
Scattered light from base and tips of petals

White light from heliostat illuminating 35 mm shade.

Average contrast in center of shadow of 5 x 10^{-7} broadband

Figure 19

Adding Wavefront Control



Experimental Results of Corrected Coronagraph (phase and amplitude control)

Limiting Factors: Manufacturing limitations, electronics, algorithm limits, testbed limits (alignments, vacuum), DM quilting, amplitude errors

Shaped pupils (Princeton)
PIAA (Subaru), Visible Nuller, 15% band (JPL)

ExAO Testbed (S.C.), Visible Nuller, monochromatic (JPL)

Shaped Pupil (HCIT), Lyot, 10% bandwidth (HCIT)

Lyot, monochromatic (HCIT)

Figure 20

Control Approaches / Algorithms

- Conventional AO (pupil based; non-common path errors)
- Speckle Nulling (iterative)
- Single Step (phase diversity, dark hole optimization)
- Interferometric (AHA)

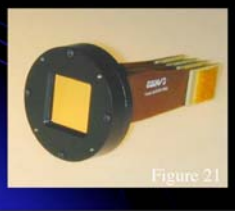
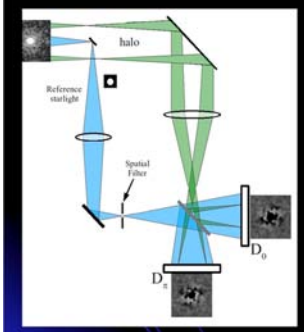


Figure 21

Focal Plane Interferometer: FPI (Arizona)

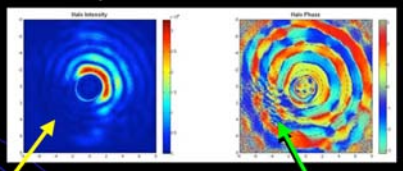


- **Uses discarded starlight**
- Core forms reference beam to probe halo.
- Halo processed as usual.
- Must capture both 0° and 180° phases for photon efficiency.
- **Gives an estimate of the halo phase from minimum number of photons.**
- **Allows measurement of halo phase in presence of bright incoherent background by using a much brighter reference beam.**

Figure 22

Using the FPI to drive a Pupil DM (PDM)

Using a simple speckle-steering algorithm, we used the FPI complex amplitude measurements to suppress the halo in closed loop.



Suppressed 90° sector

When fully suppressed, phase will become completely random. This simple system could have gone another 1-1.5 decades fainter.

Figure 23

Anti-Halo Apodization for 360 degree Suppression - AHA

Use AHA complex spatial light modulator to directly imprint a detailed "anti-halo" which interferes destructively with the star's halo

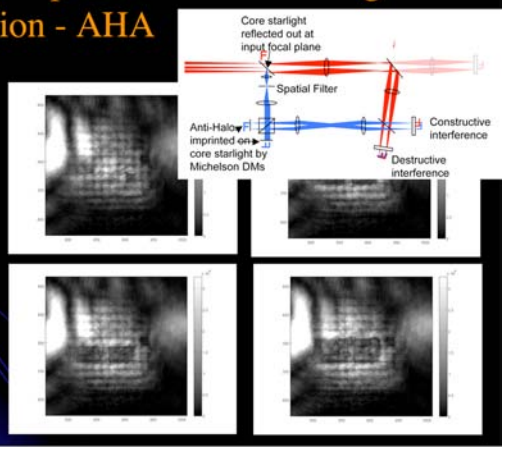
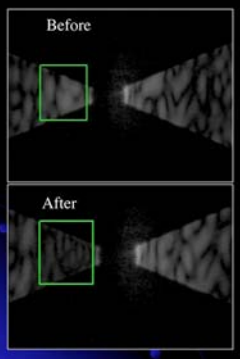
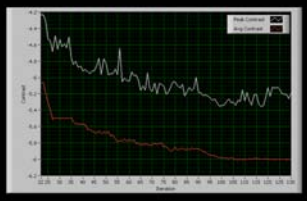


Figure 24

Speckle Nulling Results (Princeton)

Using single 32 x 32 MEMS DM from Boston Micromachines





- Dark hole: 4.5 to 12 l/D
- Peak contrast of 1e-5.2
- Average contrast of 1e-6
- Repeatable and stable
- Limited by mask accuracy and speckle nulling algorithm

Figure 25

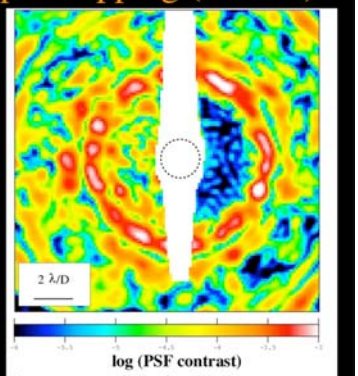
Dark Hole with Pupil Mapping (Subaru)

32x32 DM

Direct algorithm using DM induced diversity

Currently limited by variable alignment between beam and binary mask (binary mask closer to focus than pupil)

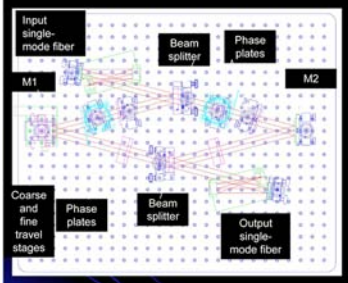
No mask block starlight currently subject to camera dynamical range



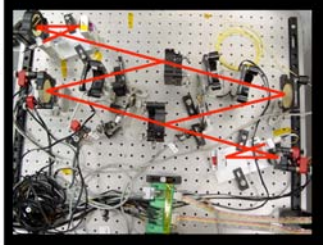
Average contrast in dark region = 2.5e-6
Vertical white feature is due to CCD saturation
Central circle is 1 lambda/D radius
Note: all distances are measured using the system chief ray

Figure 26

Visible Light Shearing Nuller



(Symmetric) Nuller Layout

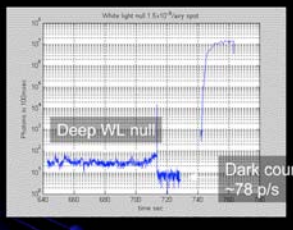


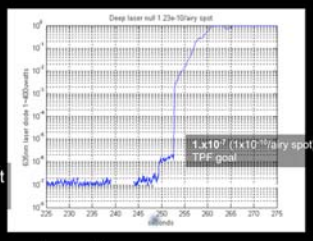
Experimental set up

- **All experiments have been run in vacuum chamber at 1 atm, with the door shut**

Figure 27

Monochromatic and White Light Contrast





- White light data
  - null over 60 sec ~1.06x10^{-6}
  - ~15% bandwidth around 650nm.
- Better control of dispersion correction will deepen null to same limit as the laser null
- Laser data:
  - Optical path error of 90 picometer will cause 2x10^{-7}
  - rms vibration and drift over ~15 sec is ~60pm

Figure 28

ExAO Testbed (Santa Cruz)

"Conventional" AO approach

- Phase Shifting Diffraction Interferometer (PSDI) for metrology
- Low wavefront-error optical design
- MEMS DM controlled in closed loop via:
  - PSDI or
  - SF Shack-Hartmann WFS
- Shaped pupils to suppress diffraction
- Science Camera in far-field for contrast measurements

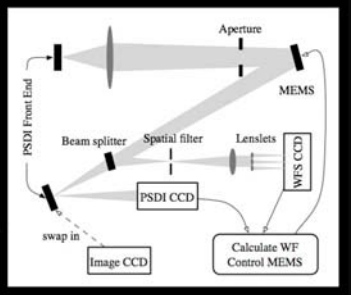
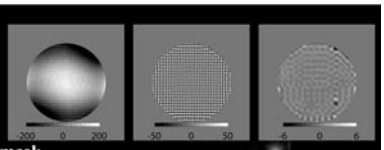
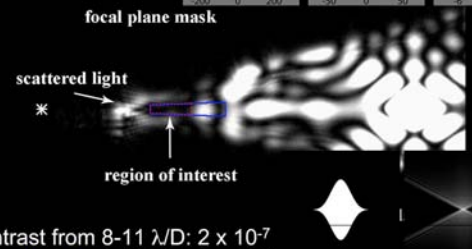


Figure 29

Contrast with MEMS

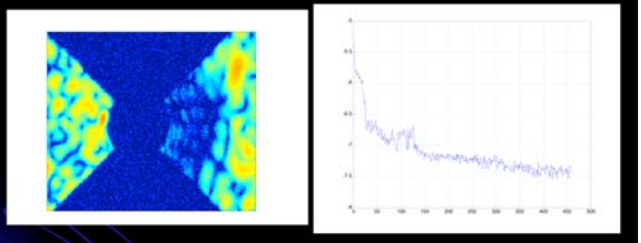




- Contrast from 8-11 λ/D: 2×10^{-7}
- Contrast from 8-13 λ/D: 6×10^{-7}
- Region of interest is limited by scattered light from the focal plane mask and controllable spatial frequencies
- Contrast is limited by phase and amplitude errors

Figure 30

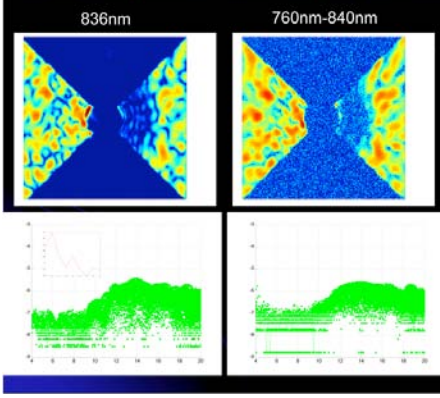
Shaped Pupil Speckle Nulling at HCIT 785nm with Ripple 3



- 457 iterations over a course of 3 days
- Went down from 10^{-5} contrast to 4×10^{-8} between 4 and 9 λ/D
- Some wavefront correction on the opposite side, until we hit amplitude errors

Figure 31

Shaped Pupils at the HCIT Monochromatic and Broadband Performance

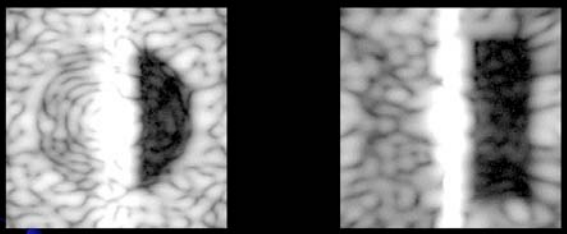


- After switching from 785 to 836nm
  - Contrast degrades very slightly
  - After a few speckle nulling iterations, contrast dips below 4×10^{-8}
  - longer wavelengths see lower aberrations

After switching to broadband light (760-840nm), contrast remained about the same

Figure 32

HCIT Monochromatic Result



Better than 1e-9 Contrast
4-5 Airy rings.
5e-10, 4-10 Airy rings.

10% bandwidth, Xenon lamp,
4-10 λ/D contrast = 6e-9,
Same spatial scale as narrowband images

Figure 33

Coronagraph Critique

C. Noecker[1], J. Kasdin[2], S. Shaklan[3]
[1] *Ball Aerospace,* [2]*Princeton University,*
[3]*Jet Propulsion Laboratory, California Instiute of Technology*

1. INTRODUCTION

This short report is a summary of the presentations and discussions at the coronagraph workshop held in Pasadena, CA in September, 2006. The respondents and attendees heard from numerous researchers about the technology and status of various coronagraphic approaches to achieving high contrast imaging very close to the image of a star. We use the term coronagraph here to refer to both internal coronagraphs, those that create high contrast around the image formed by a single telescope (with, of course, the aid of wavefront sensing and control), and external coronagraphs that employ a large occulter to block and diffract starlight before it can enter a distant telescope. This is in contrast to a multiple satellite interferometer such as TPF-I.

Our goals in this report are multifold. We not only summarize the presentations and current status of these various concepts, but also place them into the larger context of TPF-C readiness and, to some degree, make comparative assessments. While all agree that it is far too early to make final judgments on readiness and appropriateness of each concept, it is important to highlight our current understanding of the strengths and weaknesses of various designs, both to gain insight on possible TPF-C performance and to allocate resources and directions for current and future research. We also try to give a general picture of technical feasibility and readiness with regard to the basic question of achieving high contrast—in short, do we have an architecture that works? Finally, we lay out the broader questions of performance and readiness for TPF-C to help build a structure for making the important design decisions and to articulate the important open issues that cross all concepts and need to be addressed in the coming few years.

As to basic feasibility, arguably the most important outcome of the workshop is that we, as a community, are tantalizingly close to achieving the fundamental goal of 10^{-10} contrast in the laboratory. Some argued that, when combined with software manipulations such as simulated observatory roll and subtraction, we have achieved it. Experiments in the HCIT, with a bandlimited coronagraph, have achieved better than 10^{-9} contrast in laser light and demonstrated the needed stability of that contrast. **There is a consensus that there is no fundamental physical obstacle to achieving the required contrast through the proper combination of wavefront control and coronagraphy.**

This is not to minimize the challenges ahead. In the above experiments, contrast degraded significantly in broadband. In addition, the algorithm for correction—classical speckle nulling—while the simplest and most direct of all proposed algorithms, also takes many iterations and is impractical for use in a space mission. Nevertheless, these results point to the tremendous progress made in the past few years and provide a proof of principal for high contrast imaging. We also expect to see high contrast achieved with new and faster algorithms in the months ahead.

2. MAJOR ISSUES

All proposed coronagraph architectures achieve the needed high contrast by design. We must then look to other criteria for choosing the appropriate architecture for TPF-C and assessing overall progress. The important issues to consider fall into the following four broad categories: (1) Performance, (2) Wavefront Sensing and Control, (3) Manufacturability, Sensitivity, Risk and Cost, and (4) Laboratory Verification. Of course, these four areas are intimately intertwined; for example, high performance is often accompanied by unmanufacturable designs. However, this division provides us with a convenient framework for discussing the issues and challenges ahead. We discuss each of these briefly next with regard to all possible architectures. This is followed by a short summary of the five leading approaches.

3. PERFORMANCE

Since all coronagraphs achieve the needed contrast, the first category of discriminators relates to performance, that is, the amount of science that can be achieved. Up to now, all discussion has focused on a small collection of important performance metrics: inner working angle (IWA), throughput, and discovery space. This is a sensible thing to do and these will continue to be important tools in our discussion as they can, through analysis, be shown to directly relate to science outcome. However, one of the outcomes of the workshop was to re-emphasize the importance of examining coronagraph performance in the context of an overall science program. The work of Bob Brown at STScI and Sarah Hunyadi at JPL is a large step forward in this direction, but more needs to be done. It is important we reach a consensus on the proper metrics for "science harvest" and the proper structure and approach for an overall simulation. We then need to encourage all groups, within the framework of the Navigator Program, to begin system level simulations of the various architectures. In our various discussions we seem to be suffering a bit from the lack of a consistent, apples-to-apples method for evaluation and comparison.

Arguably the most important performance metric is inner working angle; it has certainly received the most attention recently. For most of the past decade it has been assumed that the best a coronagraph could achieve was 4 λ/D. This led to the baseline design that included an 8 meter primary mirror in order to see the complete habitable zone in the nearest 150 stars. There is concern, however, regarding the manufacturing difficulties (and risk) associated with such a large mirror. In the last few years a number of designs have been proposed (and are discussed later) that can operate at smaller inner working angles (as low as 2 λ/D). This could presumably allow a TPF-C with a smaller mirror (perhaps 4 meters) or with a wider science space. While either of these outcomes is appealing, the vulnerability to wavefront drift appears to be dramatically more severe at the smaller inner working angle and the ability to counteract it with rapid wavefront sensing and control is not yet well understood. These issues are discussed more below.

4. WAVEFRONT SENSING AND CONTROL

While one could argue that this subject belongs in the performance category, as it is integrally tied to the overall architecture, we list it separately because of its importance. It is widely recognized that the essential ingredient of a successful high-contrast imaging system for TPF-C is an effective wavefront control system. No internal architecture will achieve the theoretical performance in practice because of errors in the optics (an external occulter does not, in principle, need a wavefront control system as the starlight is prevented from entering the telescope). Fortunately, all of the research groups are working intensely in this area as well. There are numerous algorithms and architectures being both analyzed and tested in the laboratory. This should continue.

What seems to be lacking, however, is an integrated study of the WFSC with the coronagraph. There is a danger of making architecture decisions based solely upon performance and then assuming that the WFSC can simply correct and restore that performance. It is not at all clear that an integrated coronagraph and WFSC system has the same performance, that the corrector is performance neutral, or that the same coronagraph would be designed in the presence of a specific WFSC system as in a theoretically ideal telescope. These issues are only now beginning to be addressed. As we discuss below, many of the gains being made in this area cut across architectures and the next year should see an integration of the best of both areas.

All internal designs so far seem to be limited to between 10^{-5} and 10^{-6} contrast in the absence of control. The remaining contrast must be achieved through a combination of amplitude and phase correction. Amplitude is key as once we reach the limit induced by amplitude error (expected to be on the order of 10^{-7}), the distinction between the control system and the coronagraph begins to blur. It is the performance of this combined system that needs to be examined. For example, some architectures claim very small inner working angle (better than 2 λ/D). To actually achieve this inner working angle then requires phase and amplitude control of very low order aberrations; this has yet to be demonstrated. This also requires an analysis and understanding of the stability of the observatory and whether the WFSC can keep up with the low order distortions. If so, then the distinctions between some designs become less pronounced; for example, a shaped pupil/amplitude control combination may be possible with inner working angle comparable to pupil mapping.

One of the major open issues is achieving corrected high-contrast performance in broadband light (typically, broadband in this applications is considered to be a roughly 10% bandwidth). There appears to be a significant difference among architectures in their ability to reach the needed contrast in broadband. We try to highlight that below. This is intimately associated with the need to understand the coupling between coronagraph design, control architecture, and the wavelength dependence of the various errors.

5. MANUFACTURABILITY, SENSITIVITY, RISK, AND COST

As with WFSC, coronagraph performance cannot be viewed in isolation from the ability to make it. Some designs are far easier and cheaper to manufacture than others and in some cases this difficulty translates directly into requirements on the WFSC system. We try to highlight this in our discussion of specific architectures below. Sensitivity of various designs to changes in the optical system is also critical as there is limited signal-to-noise available for tracking these errors. This is also discussed below. The FB1 design effort did an excellent job of attempting to predict expected distortions and lay requirements on the coronagraph, though more could be done. There seemed to be wide disagreement at the workshop with regard to the significance of these issues. This becomes particularly contentious when discussing the tradeoffs in mirror size and inner working angle. Part of achieving a uniform basis for comparison is reaching consensus on the requirements imposed on the design by these practical engineering considerations. Soon it is going to be necessary to show that we have a practical technology development and risk reduction program with credible verification plans. In other words, we need to show that (a) we know how to build it, (b) we achieve adequate stability on orbit, and (c) we know how to test and verify on the ground. This is particularly true for the external occulter where the challenges are far more associated with the construction and test of the large space structures than with the optics.

6. LABORATORY DEMONSTRATION

Though last in our list, this is by far the most important category. It is essential that we have experimental verification of any concept seriously being considered for a flight program. Not only is this demonstration a key step in showing viability, but the process of performing experimental work informs the design and uncovers any number of issues that were overlooked. We can't emphasize enough the importance of a broad range of experimental work. Fortunately, we are rather far along in this area and we expect continued progress. We point out experimental progress where appropriate in our summaries below.

7. SUMMARY OF CORONAGRAPH ARCHITECTURES

We list below the six concepts considered most promising at the workshop. Their desirability derives from a combination of maturity, feasibility, performance, manufacturability, and sensitivity, the balance among these competing criteria differing for each one. In each of the following sections we provide a summary of the basic optical concept, the strengths and weaknesses, highlighting any important limitations, and the current developmental/experimental status. We try to also point out the most important areas for future focus and the open questions. In short, we address the following three questions for each architecture:

1) What is the concept and what makes it unique?
2) What is the current status in our theoretical understanding, in the technology, and in laboratory demonstration?
3) What are the steps needed to bring it to fruition?

Many of the descriptions include a specific wavefront sensing and control approach. For the most part, this is because the researchers involved have also been experimenting with different techniques; it is not necessarily because a specific approach only works with a certain coronagraph. In fact, in many cases, these techniques cut across multiple architectures; we try to point that out. As we discussed above, it is important that we make progress in understanding the general framework of wavefront control and how it interacts with the requirements on the coronagraph and observatory stability.

8. BAND-LIMITED CORONAGRAPH

The idea behind band-limited masks is to place a variable-transmission filter in the image plane of a telescope, followed by a Lyot stop that removes, potentially, all of the diffracted light. If presented with an ideal Airy pattern, and assuming the mask transmission follows exactly the prescribed shape, the diffraction from an on-axis point source is completely removed and only light from an off-axis planet (or other background or exozodiacal light) is transmitted. Band-limited coronagraphs (BLCs) can be designed to have a 4^{th}, 8^{th}, or higher order null. The 8^{th} order null is very effective at rejecting low-order aberrations and light from the small but finite-diameter star, but it suffers lower throughput and produces a larger Airy spot than 4^{th}-order masks. The aberration rejection—insensitivity to tilt, focus, etc., is so effective that this was the driving factor in selecting 8^{th}-order band-limited masks as the baseline for TPF-C.

These masks have been used in the laboratory to achieve the best dynamic range imaging results to date. With contrast better than 10^{-9} at 4 λ/D and λ=780 nm, these masks (in combination with a wavefront control system) have been used to achieve contrast about 50× deeper than shaped pupil masks and 1000× deeper than other approaches. In 10% bandwidth, results have approached 10^{-9} but were still degraded by roughly a factor of 5.

The wave front control algorithm used to achieve these results is the simple but exceedingly slow speckle nulling algorithm. While it takes the initial uncorrected 10^{-6} contrast to 10^{-8} in a few iterations, it requires thousands of iterations (each one requiring 3 new images) to reach 10^{-9}. Speckle nulling has served the purpose of reaching unprecedented contrast levels but it is not suitable for TPF-C. New algorithms are under development but have not been proven in the laboratory.

The modeling effort includes end-to-end diffraction analysis and incorporates measured values of mirror surface properties, source amplitude profile, and mask spatial and dispersive properties. The modeling is in good agreement with broad-band experimental results and is being used to derive requirements for broad-band masks.

The results obtained in JPL's HCIT testbed are tantalizingly close to meeting TPF-C requirements. The testbed contains almost all the complexity of the TPF-C back-end: a large format (32 × 32 actuators, with 64 × 64 in preparation) deformable mirror (DM), reimaging optics to provide the Lyot and intermediate image planes, filter wheels, and a CCD camera. It is presently missing a second DM that will enable broad-band control of both phase and amplitude non-uniformities. The results are obtained over 4–10 λ/D, adequate for most TPF-C stars. The stability of the nulls over time scales of several hours is already good enough to detect earth-like planets in the habitable zone.

The most challenging aspect of the BLC is achromatizing the spatially variable (over 10s of um) mask transmission function over ~10% bandwidth. The aforementioned results were obtained using electron-beam bombarded photosensitive HEBS glass. Characterization of this glass reveals that it is dispersive both in amplitude and in phase. It does not appear to be suitable for TPF-C broadband work and JPL is now pursuing the manufacture of multi-layer metallic/dielectric combinations produced using scanning slits and/or multi-step lithography. Additionally, even if the right combination of materials is found, the mask and substrate wavefront transmission errors must be controlled to Angstrom levels over a spatial extent of ~ 100 um.

The focus for future work is clearly in two areas: mask development and wave front sensing and control. The former is unique to band-limited masks: development and production of new broad-band materials with micron-scale features. The latter is shared by the other approaches (except the external occulter) in one way or another.

9. VISIBLE NULLER

The visible nuller (VN) consists of two cascaded shearing, nulling interferometers that create a 4^{th}-order null. The baseline approach (compatible with the 8-m elliptical TPF-C aperture) produces a transmission function identical to a 4^{th}-order band-limited 1-cosine mask in monochromatic light, hence it has higher throughput but is much more sensitive to aberrations than the 8^{th}-order mask now favored for the BLC. The VN includes an array of lenslets matched to an array of single-mode spatial fibers placed at a Lyot plane. The fibers remove high-order wavefront corrugations so that they cannot leak into the dark hole. In principal, they also enable a method of controlling the amplitude balance of the four beams that combine to form a null there but this has yet to be

demonstrated. At the output side, an array of pinholes matched to the fiber core positions eliminates any residual core modes. In its originally proposed configuration, the wave front phase and amplitude were to be controlled with a segmented facesheet DM. A redesign is under study that may use a second DM with a continuous face sheet for broad-band amplitude control. A single-nuller VN called PICTURE is being built for a sounding rocket experiment to fly in 2007.

The VN is designed to achieve a high-contrast null without placing a mask in the image plane. A Mach-Zehnder interferometer configuration using balanced coatings and tilted glass for dispersion compensation nulls the starlight in a collimated beam. The shear of one beam relative to the other allows planet light to be transmitted over certain angles. Because the null is achieved in a pupil, the VN does not exhibit chromatic pupil shearing (except a small amount due to the prisms) and is well-suited to nulling complex, segmented apertures. It is also designed with a calibration channel that interferes light from the bright (non-nulled) interferometer output with the nulled output to measure the complex wavefront. While band-limited and shaped-pupil coronagraphs can be set up in a similar way using light reflected from the mask or transmitted through a pinhole, the calibration channel is a central part of the VN design and is currently under study in the laboratory.

A single-channel VN configuration has been used to demonstrate ~10^{-7} nulls in laser light for ~ 10 s duration and 10^{-6} nulls in a 100 nm wide bandpass. The experiment is meant to demonstrate a single channel of what is ultimately a > 1000 channel system. This demonstration system includes the beamsplitters and compensating flats but lacks DMs, reimaging optics, and the lenslet and fiber arrays. For random errors, i.e., residual fields emerging from each fiber which are uncorrelated from one to the next, the final contrast per pixel in the focal plane is predicted to be ~ 1000× deeper, since the residual starlight leakage is distributed evenly over the entire focal plane. In this way a null depth of 10^{-7} in each fiber might ultimately achieve 10^{-10} starlight leakage per detector pixel. But some effects, such as polarization- and color-dependent nulling imbalances, will limit the null in a correlated way across all fibers; this correlation allows the leaked light to concentrate in a small portion of the focal plane, limiting the starlight leakage there to levels more like 10^{-7} per pixel. This will necessitate either much tighter control of leakage variations or a smaller tolerance for correlated leakage fields per fiber. Presently the effects limiting broad-band contrast beyond 10^{-6} have not been explored.

Fiber arrays have been constructed with several hundred SM fibers coupled to lenslet arrays. Results in several papers have shown that after polishing the fiber length uniformity is consistent with diffraction-limited imaging at the back-end of the system. Coupling to lenslets is non-uniform but this can be partially offset by the wave front control system. Results of these tests were not presented at this workshop.

An end-to-end model including diffraction effects has been developed but was not presented at the workshop and is not presently in the design loop. It is not known whether the modeling is consistent with laboratory observations. Detailed models of SM fiber spatial filtering have shown that a few cm is adequate for obtaining a pure mode when a pinhole at the output limits emerging light to the region of the core.

The VN has been touted as a system capable of functioning at 2 λ/D. However during the presentation it became clear that the error budgets so far established are for operation at 4 λ/D. Achieving an inner working angle (IWA) of 2 λ/D is much more challenging (10^{-100}× tighter control of wavefront variations and large spatial scale chromatic and polarization terms). Although the calibration channel (above) could enable relatively rapid measurement and control of wavefront errors, the issue of measurement biases in that wavefront sensor has not been addressed. Given these uncertainties, it is not at all clear that such a small IWA is feasible.

While the laboratory work has been impressive, the first order of business is to show that there is a path to obtaining high contrast at levels 1000× beyond what has been demonstrated. The focus should be on identifying, budgeting, and studying the correlated or partially correlated nulling errors—i.e., the effects that *do not* benefit from evenly spreading leakage over the detector plane. Then for robustness in the worst case, a single-channel experiment like the existing system (but placed in vacuum and under improved vibration isolation) should be able to demonstrate broad-band control at 10^{-10}. Further end-to-end modeling, and in particular validation of the models via laboratory experiments, is also crucial. The calibration channel could then be added and demonstrated. The goal is to demonstrate control and calibration to levels consistent with detection as required

by TPF-C. It is also important to demonstrate that the fiber array behaves as predicted, but this is secondary to understanding the limitations of the optics in the nuller.

10. SHAPED PUPIL CORONAGRAPH

The shaped pupil coronagraph (SPC) is a binary version of a continuous-tone apodized aperture. By rolling off the pupil transmission (generally higher in the middle, lower near the edges), the image plane diffraction pattern is redistributed into a function that has almost all its energy within a prescribed core area (e.g., a 4 λ/D radius) and diffracted energy in the 'dark hole' region (e.g., 4–30 λ/D) reduced to below 10^{-10}. When low order aberrations are present, they scatter light in and near the core but do not exhibit significant ringing as is found with the BL and VN designs. If one allows a small guard-band around the IWA, e.g., from 4–4.3 λ/D, the SPC is quite insensitive to changes in low-order aberration content.

The binary SPC is much simpler to fabricate than a continuous-transmission aperture. Patterns are etched into silicon wafers using precision lithography. No substrate is required, eliminating ghost reflections, surface height, and transmission non-uniformities. Because the masks are in pupil-conjugates, the mask features are large (minimum scale of a few μm) and tolerancing can be met with current technology. Vector diffraction modeling (discussed below) has shown that the mask designs perform well in unpolarized, broad-band light.

The SPC is the simplest coronagraph to implement. The mask is placed at a pupil plane, and at the image plane a simple sharp-edge mask is used to remove the starlight. A Lyot plane is not required but may prove to be useful for wavefront control. The masks can be designed for any aperture shape.

The masks have low throughput. Typically they pass only ~10–20% of a planet's received flux to the image plane. They have discovery space that can be traded against throughput: masks can be designed to optimize characterization (high throughput, low discovery space in the shape of wedges) or to broaden discovery space at the expense of integration time. In addition, it is possible to design masks with a diffraction core smaller than 4λ/D by reducing either discovery space, contrast, or outer working angle, or combinations of all three.

Laboratory results have been very promising. The SPC 'ripple' mask was used to achieve 4×10^{-8} contrast from 4–10 λ/D in 10% broadband light on HCIT. Models have shown that the results were limited by the particular mask design. Improved designs capable of achieving better than 10^{-9} contrast are now going into fabrication. The masks are also undergoing testing at the Princeton facility (which lacks a vacuum tank) where wavefront control algorithms and model validation are taking place. The setup employs a 32×32 continuous facesheet MEMS DM and has achieved ~ 10^{-6} contrast. The wave front sensing and control approaches include: pupil and image plane sensing, Gerchberg-Saxton algorithms, and reference beams generated by holes in the image-plane mask. The algorithms can be transferred to HCIT where they can be used to achieve TPF-C contrast levels.

An extensive modeling effort has led to a detailed understanding of the mask behavior. Vector diffraction modeling of fields passing through the masks has shown how the field behaves at the mask edges, and has been used to analyze real manufacturing defects such as undercutting, passivation layering, and edge roughness. The vector modeling is combined with propagation models of representative optical systems to show that the masks will produce high contrast, broad-band images in unpolarized light. The vector models have also shown that it is advantageous to build masks using tapered walls through the silicon. These are presently under development.

SPC masks appear to be well understood and technologically feasible. Development is underway to learn how to use the SPC in an efficient, broad-band wave front control approach, one that requires few iterations and could be implemented on TPF-C. In spite of their low throughput, and potential inner working angle limitations, shaped pupil coronagraphs represent a promising, well understood approach to high-contrast imaging.

11. OPTICAL VORTEX

The optical vortex coronagraph (OVC) is similar to the BLC in that it requires an image plane mask, a Lyot stop, and it can remove all the light from an on-axis point source. Because the Lyot stop is large (equal in diameter to the pupil image), the OVC has higher throughput and a tighter point-spread function than the BLC, VN, and

SPC. High order OVCs (charge = 4 or 6) have good aberration rejection. However, the OVC is the least mature approach and has not been used in a high-contrast experiment. Only ~ 10^{-2} contrast was obtained in a first, simple laboratory proof-of-concept experiment.

The OVC image plane mask is an azimuthal phase ramp that accumulates an integer number of wavelengths of phase over 2π of azimuth. The phase-only mask is related to the four-quadrant phase mask and phase groove mask, but by increasing the depth of the ramp to 4 or 6 waves (OVC4, OVC6), the coronagraph null goes to 4^{th} and 6^{th} order respectively. The center of the mask is undefined—it is a mathematical singularity—that in practice has a finite size and shape fixed by the width of the lithographic process used in manufacture. This scatters light but in a way that is understood through modeling to be non-fatal and controllable in part by slightly resizing the Lyot stop.

The major problem with the conventional OVC approach is that material dispersion limits the bandwidth to impractically small values. Achromatizing the masks using multiple substrate materials has proven to be infeasible because the masks become hundreds of microns thick. A novel approach has been suggested: the OVC is encoded in the first order diffraction of a blazed phase hologram (depth ~ 1 wave). All diffracted light (in first order) has the same phase ramp, but the diffraction angle scales with wavelength. A prism is added (forming a grism arrangement) that reassembles the diffracted light into a collimated beam that then propagates to the Lyot plane. The phase hologram approach has good efficiency over ~ 10% bandpass and is much easier to manufacture than the phase ramp. A first mask has been fabricated (without the prism) and is under study at JPL.

There is still much to be learned about the OVC. Like the BLC, the OVC has angstrom-level requirements on the wavefront transmitted by the mask. There is presently no detailed design for the prism and other optics, and there has been no study of effects such as residual lateral chromatic shear. The prism correction is only approximate and will result in laterally sheared pupils that may be problematic for wave front control.

12. PUPIL MAPPING

The coronagraph approach that promises the highest throughput and smallest inner working angle is the pupil mapping (PM) coronagraph, also known as phase induced amplitude apodization (PIAA). Like optical systems that reshape diode laser beams from elliptical to circular or other shapes, the PM is based on two optics: the first one reshapes the beam, concentrating light at the center of the beam so that it has a prescribed amplitude taper; the second optic compensates the phase creating a planar output wavefront. When this re-profiled beam of light is focused, it forms a sharp core with low wings, similar to the SPC.

While a ray-optics analysis suggests that pupil mapping could match the PSF shaping ability of a purely apodized system but with 100% throughput, a careful diffraction analysis has shown that pure pupil mapping (that is, just two mirrors to shape the amplitude) cannot achieve better than 10^{-5} contrast. To gain the high contrast performance, the system requires two edge apodizers, one on or near the first optic, and one following the output optic. These apodizers reduce edge ringing and can be made using the binary approach of the SPC. By modifying the profile of the pre- and post-apodizers, the broadband performance of the pupil mapping system can also be adjusted. Current designs show that the apodizers absorb ~ 20% of the incident light and have an effective diameter of ~ 0.9 of the pupil. When the light loss from the apodizers and mirror coatings is accounted for, the Lyot throughput (the light transmitted by the central core of the PSF) will be on the order of 65%, still a significant improvement over the other coronagraph approaches.

Pupil mapping also has the potential to achieve smaller inner working angle. A two mirror pupil mapping system creates large distortions of off-axis sources. These distortions create an effective magnification that allows close in targets to move outside the main lobe of the PSF. If followed by a blocking mask and a reversed pupil mapper to restore the target sharpness, images as close as 2 λ/D can be potentially imaged with no impact on throughput or discovery space. The reverse pupil mapper also restores the Airy pattern shape of the planet's PSF. It thus has a full-width at half maximum (FWHM) of ~1.1 λ/D, making it 2× sharper than the BLC, VN, or SPC point spread function, and contain at least 2× as much energy. This renders the PM system potentially much more efficient at detecting planets than the others. Unfortunately, this same magnification property makes pupil

mapping extremely sensitive to aberrations at these angles, requiring single-digit picometer stability of low order terms such as focus, astigmatism, and coma.

PM has been the subject of extensive modeling including both ray tracing and diffraction. The modeling effort has been used to optimize design parameters: throughput, edge radius of curvature, inner working angle, and bandpass. Diffraction models have been used to derive aberration sensitivity as well. However, the modeling does not yet include the wave front control (WFC) system. Without WFC, the requirements on the optical surface quality can not be derived. Further, the useful bandpass is not yet known (it may be limited by the large pupil distortion) and multi-DM configurations have not been modeled, though models have been developed to show that starlight reflected from the image plane mask can be used to quickly measure low order aberrations and derive a control signal. For instance, current manufacturing experience show that it is unlikely that the pupil mapping mirrors can be made much better than roughly $\lambda/100$; this limits contrast to 10^{-5} or so. It is not yet known how a WFC system would be best integrated and what its performance implications are.

Two sets of PM optics have been fabricated and are under test in the laboratory in Hilo, HI. The test setup includes a 32×32 DM of the same design as Princeton's (but using different electronics). A contrast approaching 10^{-6} has been obtained at 2 λ/D and out to ~ 10 λ/D.

The major gaps in the current state of PM work are: 1) The requirements on the fabrication of PM optics are not known; these are linked to the design of the WFC system. 2) The WFC system has not been defined; where should DMs be placed, how many are needed, and can they work with highly-distorted pupils? 3) What is the practical limit to the inner working angle? The last question applies to all the concepts, but is particularly relevant for PIAA because it maintains higher throughput well inside of 4 λ/D.

13. EXTERNAL OCCULTER

The external occulter (EO) is a two (or more) spacecraft approach that has a completely different set of issues from the internal occulters described above. The EO places a large (~25–50 m diameter) mask at 30,000–80,000 km from a 4 m telescope. The telescope can be segmented and should provide a good quality image but does not have any particularly challenging stability requirements. The EO mask is in the shape of a flower with perhaps dozens of petals. The target star, mask, and telescope form a line and the shadow of the mask creates a dark hole around the telescope. The hole is a few meters larger than the telescope, setting the relative alignment accuracy requirement to ~ 1 m. Bandwidths of > 1 octave are possible. At the long wavelength limit, light diffracts into the dark hole limiting performance. The mask moves around the sky and the telescope is repointed to study the various TPF-C targets.

The mask size sets the IWA. Inner working angles of ~60 mas are possible, consistent with the baseline TPF-C design (8 m at 4 λ/D at λ=600 nm). There is no outer working angle; high contrast is obtained at all angles > IWA. Also, there are no low- and mid-spatial frequencies to scatter light outside the IWA. As long as the mask is properly made and positioned, and sunlight reflections are at an acceptable level, the discovery space is devoid of scattered light. When the mask deforms, however, the edge of the mask or center of the field gets bright, forming a 'ring of fire' that can mask planets near the IWA.

Some mask tolerancing has been studied through analysis and modeling though more needs to be done. The mask is required to be deployed to within ~ 1 mm of the design shape. Holes (due, e.g., to micrometeoroids) can total a few square cm as can the integrated size of edge defects. Starshade tip/tilt and distance requirements are easily met. The tolerancing appears to have been done conservatively—the values quoted represent 10^{-10} scattered light at the center or edge of the mask. The scatter beyond the IWA will be substantially less. Modeling studies are challenging: the Fresnel propagation requires huge arrays and results so far appear to be limited by modeling grid error.

An alignment system involving a small telescope that transmits an image of the star to the telescope has been conceptualized but requires detailed study. Other means of alignment are under study. This issue is not considered to be a show stopper.

A laboratory demonstration using a small mask placed in a 40 m long tunnel and illuminated with sunlight is said to have achieved a contrast of $10^{-7.}$ The mask was scaled to cover the same number of Fresnel zones as an orbiting occulter. This experiment was not discussed during the workshop and details beyond those published in SPIE vol 6265 (Cash et al, 2006) are not available.

The number of targets that can be observed is limited not only by the time it takes to move the occulter tens of thousands of km from target to target, but also the solar view angle restrictions that limit the observing season. Models of both high thrust (hydrazine) and low thrust (electric propulsion) show that given reasonable assumptions of occulter mass and initial wet mass, transit times are a minimum of 5–7 days for 15° slews. Operation at either L2 or Earth-trailing orbits appears feasible. Using a single occulter, the EO may be limited to ~ 100 visits in a 3-year time frame. The solar restriction angle is quoted as a cone 45°–95° from the sun (90° is a right angle to the sun). This results in a ~ 3 month observing season for targets near the ecliptic. Initial mission modeling has shown that the science return is about 50% of the baseline TPF-C mission. Detection-confirming observations and additional orbit-determination observations will be limited. The impact of this on the science return is under study. For these reasons, a second occulter (thus a third spacecraft) is proposed. One occulter would be smaller and nimbler, located closer to the telescope, but would have a larger IWA. This occulter could be used for discovery and for follow up observations when planets are expected to be at larger separations from their stars. The other occulter is larger and further, allowing deeper observations at smaller IWA for discovery and characterization.

The most important issues to be addressed in future work are: 1) Modeling: construct computer models that agree with analysis and can be used to accurately tolerance the system; 2) Demonstrate high-contrast in the laboratory and validate the models; 3) Detail the deployment approach—how can 1-mm deployment be achieved? 4) Continue to study the mission observation scenario and optimize the trade between science return, mass, and mission complexity.

14. CONCLUSIONS AND RECOMMENDATIONS

There is general agreement in the astronomical community that the science behind searching for terrestrial planets is compelling and important. It is gratifying that there seem to be numerous paths toward a system for imaging such planets. It is also extremely important to note again that the fundamental problem of achieving high contrast has been demonstrated in the laboratory, though only monochromatically. The community believes that a credible plan for a space based, visible light system for terrestrial planet finding in the 2010–2020 decade will become possible within the next few years. All six of the approaches described above provide a feasible avenue to a space mission, though each has its own tall poles and concerns to be addressed. While new ideas continue to be generated, we do feel that narrowing our attention to these six is necessary to best utilize the available resources and converge on an architecture in the near future.

Progress needs to be made on two fronts, and appropriate resources need to be dedicated to achieving these goals. First, experimental work must continue and each design must demonstrate, in the laboratory, that high contrast is achievable, in broadband, with an implementable wavefront control system. Second, detailed, and common, performance simulations must be completed for each candidate architecture, utilizing common assumptions and metrics, to allow fair and comprehensive comparisons on science yield, complexity, robustness, and cost. It is of the utmost importance to validate the models with the laboratory experiments. These two activities will provide the information necessary to both convince the broader astronomical community that the technological readiness for TPF-C is there and to make the eventual architecture decisions for a future space telescope.

Mirror Technology Assessment

Gary Matthews
ITT Space Systems Division

MIRROR TECHNOLOGY ASSESSMENT ABSTRACT

A mirror technology assessment was provided to ground the TPF science community regarding the current state-of-art in space mirror capabilities within ITAR constraints. A generic assessment regarding the various mirror configurations (active vs passive, on-axis vs off-axis, round vs elliptical, coronagraphic quality vs visible quality) was discussed. The maturity level of each of the various concepts was provided in order to help guide future mission planners regarding programmatic risk and the need for early technology development.

The fundamental conclusion was that any of the concepts are certainly possible from an engineering standpoint. The major variable was the cost to reduce the risk to an acceptable level such that a robust schedule could be proposed and managed on the program. This assessment and long term program plan in turn would drive the technology development process that would be required prior to baselining the program architecture.

TOP LEVEL PRIMARY MIRROR PARAMETERS

The top-level primary mirror parameters are shown on the chart shown on the previous page. The color scale reflects the proven capability with respect to demonstrated TRL and other technology development completed by ITT.

A Top Level Requirements Comparison Provides Some Keys to Required Technology Development

	TPF-C (Baseline)	New World Observer (NWO)	Phase-Induced Amplitude Apodization (PIAA)	Visible Nulling Interferometer (VNI)
PM Size Range	3m x 6m 3.5m x 8m	2m --- 4m	4m	4m
PM Configuration	Off-axis	On-axis	Off-axis	On-axis
PM Actuation	Low Authority	Passive	High Authority	Passive
Quality	Corona-graphic	Visible	Visible	Visible

— No Technology Development (TRL 6 Minimum)
— Engineering Process Development (TRL 5)
— Technology Development Needed (TRL 4 or Less)

ITT

4

PRELIMINARY MIRROR DESIGN CONFIGURATIONS

Mirror Configurations for the Various Sizes of Primary Mirrors Under Consideration

Configurations Characteristics	2.4m Passive f/1.25	3.2m Passive f/1.25	4.0m Passive f/1.25	4.0m Set & Forget f/1.25	3.5m x 8.0m F/3.8
Mirror					
Outside Diameter (m)	2.43	3.23	4.03	4.03	3.5 x 8.0
Inside Diameter (m)	0.44	0.62	0.77	0.77	N/A
Pocketmilled Facesheets	Yes	Yes	Yes	Yes	Yes
Core Cell Shape	Hexagonal	Hexagonal	Hexagonal	Triangular	Hexagonal/Segmented
Force Actuators	N/A	N/A	N/A	21	20-50
Material	ULE	ULE	ULE	ULE/Composite	ULE/Composite
Total Weight (kg)	176.4	304.1	601.5	450.9	1100.0
PM Mounts					
Material	MP35N/Invar	MP35N/Invar	MP35N/Invar	MP35N/Invar	MP35N/Invar
Total Weight (kg)	4.5	13.6	26.3	26.3	TBD
AMS					
Material	Composite	Composite	Composite	Composite	Composite
Total Weight (kg)	38.1	67.7	106.1	276.1	750.0
Mirror Area (m^2)	4.5	7.9	12.3	12.3	44.0
Total PMA Mass (kg)	219.0	385.4	733.9	753.3	1850*
PMA Areal Density (kg/m^2)	48.8	48.8	59.7	61.3	42.1*

* Does not include PM Mount Mass

Note: Design details are available
No contingency is included in these estimates

ITT

5

ITT has completed a very preliminary assessment of the mirror design parameters for the various sized mirrors required for the TPF-C program. As can be seen, reasonable areal densities can be achieved for all of the potential telescope designs.

FACILITY ASSESSMENT

Facility Break Points have been identified

	2m - 2.5m	2.5m – 3.0m	3.0m – 4.0m	3.5mx8m
ULE® Glass Manufacturing	No issues. Raw glass manufacturing has been demonstrated to 8m diameter at Corning			
Mirror Blank Manufacturing	Facilities in place at Corning	Facilities in place between Corning and ITT	New furnace required	Segmented manufacturing process requires development
Mirror Processing and Coating	Facilities in place	Minor upgrades required for final finishing and test. Coating chamber required.	Some modifications required for finishing and test. Coating chamber required	Major facility upgrades required.
Recommendation	No issues. Pathfinder will reduce schedule risk.	Pathfinder suggested to harden processes	Qualification model recommended to verify processes	Subscale pathfinder and full scale pathfinder recommended

ITT

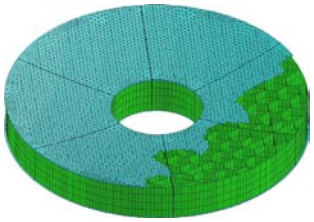
9

Facility limitations have also been evaluated and recommendations provided regarding engineering models. From our assessment, all of the primary mirror systems in our study are possible to build. Some configurations will require more engineering development than others.

For example, a Hubble-like 2.4m primary mirror is considered to be very low risk. In that configuration, no pathfinders would be needed. As the primary mirror systems become larger and more complicated, pathfinders or qualification units are recommended. In the case of the baseline TPF-C 3.5m × 8m concept, a subscale pathfinder would be needed to work out the critical engineering details prior to making a full scale pathfinder model. This would all have to occur prior to building the flight model off-axis primary mirror. Each of the mirrors fabricated would demonstrate key technology and demonstrate a progressively higher TRL as the development program progressed. The demonstration units and pathfinder program would provide high confidence that the flight primary mirror would be successful and the program cost and schedule could be maintained.

MIRROR FABRICATION TECHNOLOGY

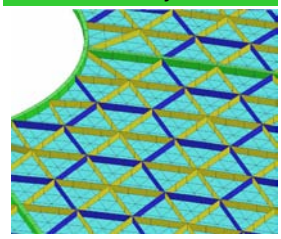
Passive Primary Mirror Model



PM Model

- In order to achieve acceptable areal density goals for very large primary mirrors, pocket milled face sheets will most likely by required
- This technique has been demonstrated on demonstration mirrors

ITT has done detailed modeling of large passive and active systems



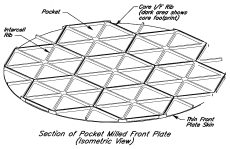
Front Plate Model with pocket milling included

ITT 10

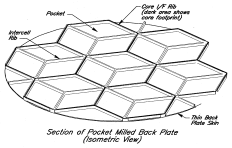
Mirror Pocketmilling Details

Typical of 2.4m, 3.2m and 4.0m

- **Front plate pocket milling**
  - **68%-72% plate lightweighting**
- **Back plate pocket milling**
  - **72%-74% plate lightweighting**
- **Reduces faceplate and core mass while minimizing processing induced quilting**



Section of Pocket Milled Front Plate (Isometric View)



Section of Pocket Milled Back Plate (Isometric View)

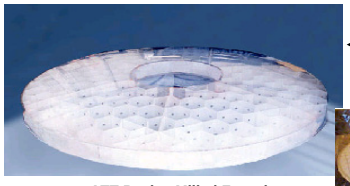
Glass Mirror Construction

Pocket milled detail

ITT 11

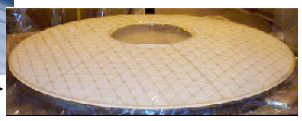
Ultra Lightweight 2m class Passive Primary Mirror Technology has been Demonstrated

- Abrasive waterjet lightweight segmented core reduces risk
- Pocket milled facesheets reduces weight to about 1/3 the areal density of HST with comparable stiffness
- Low Temperature Fusion (LTF) process eliminates the effects of Frit-bonding
- Directly scalable to 4m size



Advanced Technology Testbed (ATT) 2.4-meter Primary Mirror

ATT Pocket Milled Facesheet



ITT 12

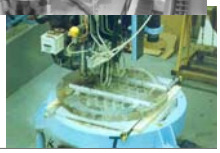
CORNING

Corning Has the Capacity to Produce Any of the TPF Mirror Blanks With Some Facilitization

8m Furnace Facility with 4m mirror

Very large water jet cutting facility

Deep Core Cutting for Lightweight Mirrors

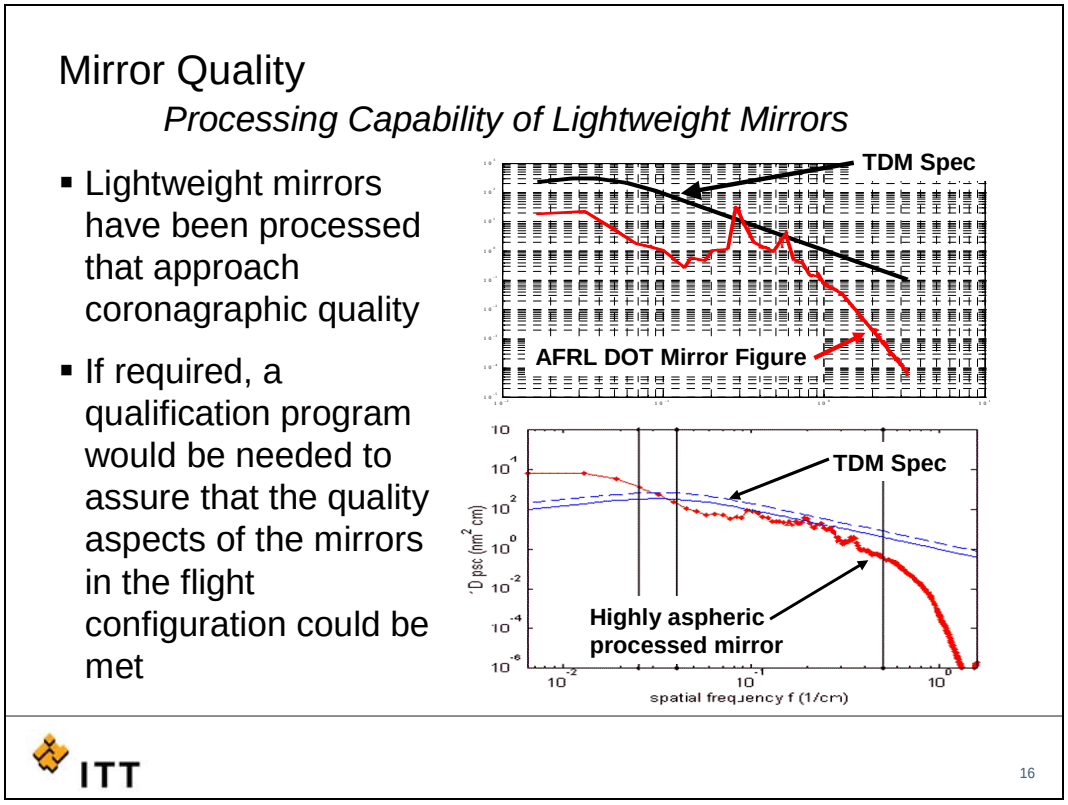


ITT 13

Mirror technology has been developed by Corning and ITT over the last 40 years. These technologies have been demonstrated on various programs and are available and ready for use on the TPF program. These technologies include faceplate pocket milling which provides faceplate support to minimize polishing quilting effects during processing. This added faceplate support allows mirrors with lower areal densities to be fabricated. Using this type of construction, large mirrors can be fabricated with areal densities as shown on page 48.

Facilities at Corning typically fabricate large ground based mirrors up to 8m as shown above. In addition, Corning has large waterjet cutting machinery that allows very deep cores to be fabricated. These facilities and those at ITT allow large mirrors to be successfully produced.

MIRROR QUALITY PARAMETERS



ITT has produced lightweight mirrors that demonstrate that coronagraphic-type quality can be achieved. Two example power spectral density (PSD) results are shown above. In each case, the quality achieved far exceeded the required specification for the mirror and was only compared to the coronagraphic specification for the Technology Demonstration Mirror (TDM) being developed by ITT for JPL. This provides evidence that the very difficult specifications required for the baseline TPF-C mirror are achievable in lightweight mirror systems.

SUMMARY

All TPF Primary Mirror concepts reviewed to date appear feasible to manufacture, process, and test. Some design configurations will require technology development in order to achieve a finished primary mirror with a predictable cost and schedule. Pathfinder/Back-up mirrors are recommended for almost all the configurations in order to reduce schedule risk on the program. For modest sized systems (<3.0m), the primary mirror will not be the critical path of the observatory based on ITT's experience. For larger systems, a rigorous and well-planned development program will provide sufficient risk reduction to allow the primary mirror to be successfully delivered within the cost and schedule constraints of a flight program.

Large Optical Systems Integration, Test and Verification

Robert Egerman
ITT Space Systems Division

Fundamentally the process of integration and test of large optical space borne systems is no different than the process which needs to be used on any space system. Because we are unable to perfectly reproduce space conditions on the ground, we must rely on a combination of techniques to verify and "buy-off" any space payload before it can be launched. Testing is a significant part of the verification process and deserves critical attention. However, other verification techniques, which per the NASA Systems Engineering Handbook (SP-610S) include analysis, demonstration, inspection, simulation and records validation are all important methods to be considered. The main point here is that the early design work of any space system should not focus on how the system will be tested prior to flight, but rather how it will be **verified** prior to flight.

Given that large optical systems might be considered by some to be more complicated than many other types of systems (See Figure 1), a disciplined engineering approach must be applied to each subsystem in terms of design and verification. It is especially critical to give attention to the verification plan early on in the design process to reduce technical and programmatic risk. It is certainly preferable to have an end-to-end ground test to demonstrate optical performance of payloads on the ground. However, if the uncertainties in the test and in ground-to-orbit effects are not accommodated by design and/or operationally adjustable parameters (such as focus shift and temperature control), then such a test is meaningless. In other words, it is possible to demonstrate performance on the ground, and then not have the required performance on-orbit.

An end-to-end test does not guarantee mission success
If done improperly, it adds no value and can give a false sense of security
If done properly it can be: Expensive, Time consuming, and
More complicated than the mission itself

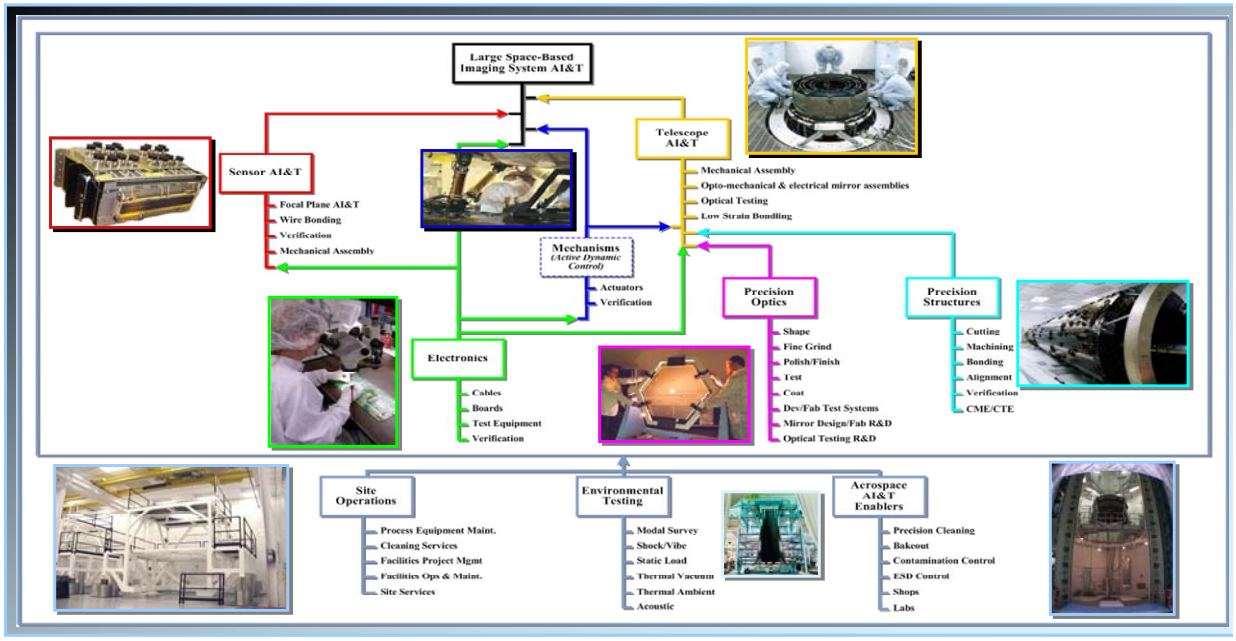


Figure 1: *Large Optical Systems Hardware and Verification Tree*

There have been many mission successes that have not included full up end-to-end testing of full observatories. The CHANDRA X-Ray observatory is a good example of this. Chandra underwent simplified end-to-end testing of the main telescope optics over a period of several months utilizing an engineering model of the science

instruments. Final verification testing after the integration of the flight instruments was abbreviated lasting only a couple of weeks and optical performance was demonstrated to within 10% of the flight requirement. To mitigate test uncertainty risks, Chandra utilized two basic engineering principles:

1. The Chandra verification plan was developed to ensure that system would perform within capture range of flight system adjustable parameters (The temperature set point of HRMA was adjustable and system focus was adjustable)
2. Performance error budgets had adequate margin

The success of the verification plan implemented on the CHANDRA program has been demonstrated through the extraordinary imagery and science that has been generated from the observatory.

Future systems such as TPF and JWST face similar verification challenges that were addressed on CHANDRA. Though technically more challenging, the utilization of a disciplined systems engineering approach which focuses on system verification early in the design process can help ensure that these large programs are executed in an efficient, low risk manner.

The challenge in verifying large space borne systems is being recognized in the community. A good summary of this issue was documented in the JWST Science Assessment Team Interim Report (26 July 2005):

"JWST is a space telescope of unprecedented aperture, but it is also cryogenic and passively cooled, a challenging combination. This means that ground-test verification that JWST will perform as planned on-orbit involves not only accommodation of the difference between 1-g and 0-g, but also a thermal environment that is very difficult to replicate. Demanding a full-up demonstration of JWST's performance on-orbit could involve test equipment that rivals the sophistication and complexity of JWST itself, potentially adding substantially to the capital cost and the time line to become one of the major expenses in the mission. While such a demonstration is desirable in this or any project, it is unlikely that it will be affordable.

This raises the question, to what extent is a full-up, end-to-end, simulated performance on the ground necessary? The experience with the Hubble, where the primary mirror was figured to the errant radius-of-curvature, which resulted in serious spherical aberration, has naturally led to a cautious attitude about performance verification before launch. However, it is easy to take the wrong lesson from the HST experience."

Science Operations and the Minimum Scale of TPF

Robert A. Brown
Space Telescope Science Institute

Stuart B. Shaklan, and Sarah L. Hunyadi
Jet Propulsion Laboratory, California Institute of Technology

1. INTRODUCTION

For any mission, the most stringent requirement is the one that determines its minimum scale—its minimum size, complexity, and cost. In recent years, it was widely assumed that *search completeness*[1] was *TPF*'s most stringent requirement, and therefore its greatest technical challenge was ensuring that the instrument design reliably detects and resolves extrasolar planets in the presence of unsuppressed starlight and zodiacal emission. These assumptions were held for both the thermal-infrared (interferometric; *TPF-I*) version and reflected-light (coronagraphic; *TPF-C*) version of the *TPF* mission. Recently, however, mission studies performed for the *TPF-C* Design Reference Mission (DRM)[2] have cast doubt on these assumptions. It now appears that the most stringent requirements of both *TPF-C* and *TPF-I* are to acquire critical astrometric information about the motion of a candidate planet in the epoch of its discovery. These new requirements are *science operational*: (R1) to ensure that a candidate planet is promptly discriminated from possible background confusion[3] and (R2) to ensure a planet's recovery at later observing epochs.[4] Until the implications of requirements R1 and R2 are understood, the minimum scale of both versions of *TPF* will remain unknown.[5] The analysis is now underway.

The most critical aspects of technical performance for meeting the new requirements are astrometric accuracy and the pointing restrictions—notably the solar avoidance angle. A timed series of measurements of a source's apparent position relative to the star must be obtained, and the data quality—accuracy and orbital coverage—must be sufficiently high to discriminate background sources and provide orbital solutions useful for predicting the future observability of the planet. For *TPF*, obtaining high-quality data is especially challenging, because (1) much of the typical planetary orbit is not available to observation,[6] (2) the intervals of time over which any star can be viewed are limited to a fraction of a year by solar avoidance, and (3) aliasing occurs for planetary periods approximately resonant with one Earth year. Because of these challenges, the astrometric accuracy of each positional measurement must be even higher than otherwise—but *how* high has not yet been determined. Because higher astrometric accuracy can only be achieved by reducing the size of the point-spread function—i.e., increasing the size of the instrument, which is the design parameter with the steepest cost slope for an instrument in space—it is important to resolve this issue decisively and correctly as soon as possible.

If the astrometric accuracy of *TPF* were insufficient to satisfy R1 and R2, then errors and inefficiencies would be introduced into science operations. Mission time would be wasted on void or misguided observations, and the

---

[1] Brown, R. A. 2005, "Single-Visit Photometric and Obscurational Completeness," *ApJ* 624: 1010–1024. Hunyadi, S. L., Shaklan, S. B., and Brown, R. A. 2007 "Single Visit Completeness Optimization," submitted to *ApJ*.

[2] Brown, R. A., Hunyadi, S. L., and Shaklan, S. B., 2006, "A DRM for *TPF-C*: A Design Reference Mission for the Coronagraphic T*errestrial Planet Finder*." http://sco.stsci.edu/tpf_downloads/TPF-C_DRM.pdf/.

[3] Brown, R. A., 2007, "Differentiating Extrasolar Planets from Background Confusion by Apparent Motion," submitted to *ApJ*, http://sco.stsci.edu/tpf_downloads/background_confusion.pdf/.

[4] DRM §5.2, p. 49.

[5] A related, existing requirement is in the *TPF-C* Science Requirements Document (SRD): "For at least 50 percent of detected planets whose projected semimajor axes lie within the habitable zone of their parent star, *TPF-C* shall measure the actual semimajor axes to within 10%. *TPF-C* shall also measure their orbital eccentricities to an absolute accuracy of ±0.3." This requirement is scientific and not time critical. It is less demanding than #2 and would be subsumed by it.

[6] DRM §5.3, p. 51.

science goals would be at risk, and possibly unobtainable. While the impacts are not yet fully and quantitatively known, they seem likely to strongly affect the design process for *TPF-I* and *TPF-C*.

In the remainder of this paper, we illustrate the science-operational issues for *TPF* with particular reference to the FB-1 design and associated DRM research for *TPF-C*. For *TPF-I*, the results should be qualitatively similar to those for *TPF-C*, involving a different planetary phase function, but similar instrumental parameters.

We show illustrative results for three typical target stars; the same results for all 136 top *TPF-C* target stars are available at *StarVault*.[7]

Two relevant discussions are available online: (1) the roles of technical performance in selecting target stars[8] and (2) the considerations for determining the photometric orbit in reflected starlight.[9]

2. THE SCIENCE-OPERATIONAL VIEWPOINT & MISSION MODELING

In the future science operations center for *TPF*, decisions about what star to observe next, and for what purpose—searching, background discrimination, orbit determination, spectroscopy, etc.—will be made continually. *Making and implementing sound scheduling decisions is the essence of science operations.* For the possible newly discovered planets, these decisions will be made with limited information available—but many possibilities about the nature of an object and how it might move and change in brightness. Because each imaging observation typically demands days, or even weeks, of precious observing time (spectroscopic observations may take ten times as long), scheduling decisions must be of the highest possible quality, making the best use of available knowledge (both observational and theoretical), while seeking more information.

Consider the aftermath of the very first *TPF* search observation of the very first target star. Assume it finds a source with a brightness and separation consistent with being a planet of interest. What's next? An immediate follow-up observation with a roll of the telescope aperture around the line of sight will confirm the external reality of the source—i.e., that it is not a speckle. But it could still be a background confusion source: very little is known about background sources as faint as those sought by *TPF*. To discriminate a possible planet from background, another observation must be made after a delay that is sufficiently long for the planetary movement (probabilistic) to be measurably different from the apparent motion of a background source induced by the parallactic and proper motion of the star (deterministic). Yet, this delay must also be sufficiently short to (1) minimize the probability that the planet has become undetectable, and (2) ensure the star is still viewable according to solar-avoidance constraints (see reference in footnote #3 for a full discussion.) Assuming the discriminating observation has confirmed the companionship of the source, additional astrometric observations must be made—in the discovery epoch, before the annual viewability window closes—to determine the orbit well enough to ensure future recovery. If we don't know the orbit at the time the viewability window closes, then at least six months will pass as the solar-avoidance restriction lapses, and we can only guess about when the companion might be observable again. The odds of a successful guess will be different for each star, but rarely high, as we will see §3.4. And if the planet is *not* recovered, we can only guess again and again, potentially wasting much more time, or give up, declaring the planet lost.

If this logic for the very first observation seems complex and risk riddled, consider how much more daunting the situation will be later in the mission, with many stars and possible planets in play at various levels of knowledge, including the possibility of multiple planetary systems, and with many possible observations all competing to be the next observations. The stakes are high, because mission time is limited and exquisitely valuable. Currently, NASA plans for this mission to last only five years.

---

[7] http://sco.stsci.edu/starvault/.

[8] http://sco.stsci.edu/tpf_downloads/on_target_stars.pdf/.

[9] Brown, R., 2006, "Determination of the Photometric Orbit in Reflected Starlight," http://sco.stsci.edu/tpf_downloads/photometric_orbit.pdf/.

Fortunately, it is possible to achieve the science-operational viewpoint and test a mission design by Monte Carlo simulation and mission modeling. Using double blind experimental formats, we can test the ability of a design to perform the necessary observations in the restricted informational environment of the mission.[10] In particular, we can use Monte Carlo simulations to estimate mission success as a function of astrometric accuracy. Mission modeling is also useful for testing the decision-making process itself. For example, modeling can quantify (in terms of probabilities) the consequences of *not* discriminating background, or of simply guessing about future observability. The ultimate goal of mission modeling is confirming that an optimized *combination* of two designs—one for the instrument and one for science operations—meets the science requirements for the mission.

3. EXEMPLIFICATION OF SCIENCE-OPERATIONAL ISSUES FOR *TPF-C*

The *TPF-C* DRM defines particular models of stellar qualification,[11] planets of interest (habitable zone),[12] and instrumental performance,[13] which have been used for the calculations reported here.

In assessing the suitability of target stars, an important quantity is the discovery rate R, for which convenient units are milliplanets per hour. For the initial search observation, $R_1 = C_1/T_{EXP}$, where C_1 is the search completeness (the fraction of Earth-like planets in habitable-zone orbits that are detectable at the time of the first observation), and T_{EXP} is the exposure time to detect a source at the limiting delta magnitude (Δmag_0). C_1 is always less than one, because planets can hide behind the obscuration or show the dark hemisphere to the observer. (C_∞ is ultimate completeness, achieved by exhaustive searching, using many searching observations at staggered times. C_∞ may equal one, but also may be less, as some possible planets may be permanently hidden behind the coronagraphic mask.)

Table 1. Three Typical *TPF-C* Target Stars

Star	HIP 23311	HIP 79672	HIP 61174
Ranking	71st of 136	68th of 136	59th of 136
Spectral Type	K4III	G5V	F2V
Apparent magnitude m_v	6.22	5.49	4.29
Absolute magnitude M_v	6.50	4.75	2.99
Color B–V	1.05	0.66	0.38
Parallax Π (mas)	113.5	71.3	54.9
Luminosity (relative to Sun)	0.14	0.95	5.42
Mass (relative to Sun)	0.86	1.10	1.46
Δmag_0 (optimized)	23.5	25.0	25.0
Exposure time T_{EXP} (hrs)	28	114	17
C_1	7%	35%	6%
R_1 (milliplanets hr^{-1})	2.59	3.11	3.34
Total discovery fraction C_∞	36%	94%	39%
$T_{10\% LOST}$ (days)	2.5	9.5	7.9
F_{DISC}	49%	45%	40%
F_{SOLAVD}	47%	47%	47%
$F_{SOLAVD \cap DISC}$	12%	1%	40%

[10] For one example of Monte Carlo mission simulation and testing, see: Brown, R. A. 2006a, "Expectations for the Early *TPF-C* Mission," in *Direct Imaging of Exoplanets: Science & Techniques*, Proceedings IAU Colloquium No. 200, (C. Aime & F. Vakili eds.), http://sco.stsci.edu/tpf_downloads/mission_studies_appendices/MS-1_Brown_Mission.pdf/.

[11] DRM §2.1, p. 5.

[12] DRM §2.2, p. 11.

[13] DRM §2.3, p. 13.

We focus on three typical stars with similar rankings (similar discovery rates R_1), but which offer quite different science opportunities and science-operational challenges. Various relevant stellar parameters are shown in Table 1. We use these three stars to exemplify four features relevant to science operations for *TPF*: (1) completeness and discovery rate; (2) properties of the discoverable subset of planets; (3) discrimination of background confusion by testing for common proper motion; and (4) planetary recovery at future observing epochs.

3.1 Initial Search Completeness, C_1, and Discovery Rate, R_1

We assume the maximum detectable magnitude difference for *TPF-C* is $\Delta mag_0 = 25$ (a factor of 10 billion), set by uncontrolled speckles of unsuppressed starlight, which can masquerade as planets. The angular radius of the central obscuration is assumed to be 57 milliarcseconds (mas), which means the minimum apparent separation of an observable planet is $s_0 = 57/\Pi$au, where Π is the annual parallax of a star in mas.

We introduce our Monte Carlo techniques for mission modeling by discussing the estimate of C_1. As illustrated in Figure. 1, we prepared random samples of 10,000 planets for each star, drawing them one at a time from the planetary population of interest. (In the computer, we use random deviates that emulate the specified probability distributions of orbital and planetary parameters to produce particular values of the randomized orbit for each planet.) Then, we counted the number of these random planets that can be detected—i.e., tested and found those that satisfy the constraints $s > s_0$ and $\Delta mag < \Delta mag_0$, where s is the apparent separation between star and random planet, and Δmag is the ratio of planetary to stellar flux expressed in magnitudes. Our estimate of C_1 was then the number detectable, divided by 10,000. (We estimated C_∞ by repeating this "observation" in the computer many times at random epochs—allowing the unfound planets to revolve in their orbits and evolve in their brightness—and counting planets as they were found.)

In mission *simulations*, the number of planets actually found in an initial search, $n_p = 0$ or 1, is a Bernoulli random variable with probability C_1.

Table 1 shows that the rank ordering of the three typical target stars tracks the value of discovery rate R_1, as expected.[14] The values of C_1 differ by a factor 5–6, which is compensated by a similar variation in (optimized) exposure time T_{EXP} (0.7 to 4.7 days).

3.2 Discoverable Subset of Planets

Each dot in Figure 1 is a particular planet of interest, possessing unique—and known—orbital elements: period, semimajor axis, three orientational (Euler) angles, and orbital phase (mean anomaly) at the initial observing epoch. Each *red* dot is one such planet that is also *detectable*, and therefore a member and exemplar of the discoverable subset. Knowing the orbital elements for each Monte Carlo planet, we can compute the true planetary position, the apparent planetary position (apparent separation and position angle), and apparent brightness of the planet *at any epoch*.

We can explore the behavior of the discoverable subset in various ways, for various purposes. We can study them like *real* planets, even adding noise to measurements of their position and brightness. We can use them to simulate actual observations to test our science-operational concepts, or to calculate properties of the ensemble. In particular, we can use them to test whether or not requirements R1 and R2 can be satisfied by a mission design. (See the reference in footnote #3 for a detailed approach to R1.)

[14] See *StarVault* or Table 6, DRM p. 19, for results on all 136 top *TPF-C* target stars.

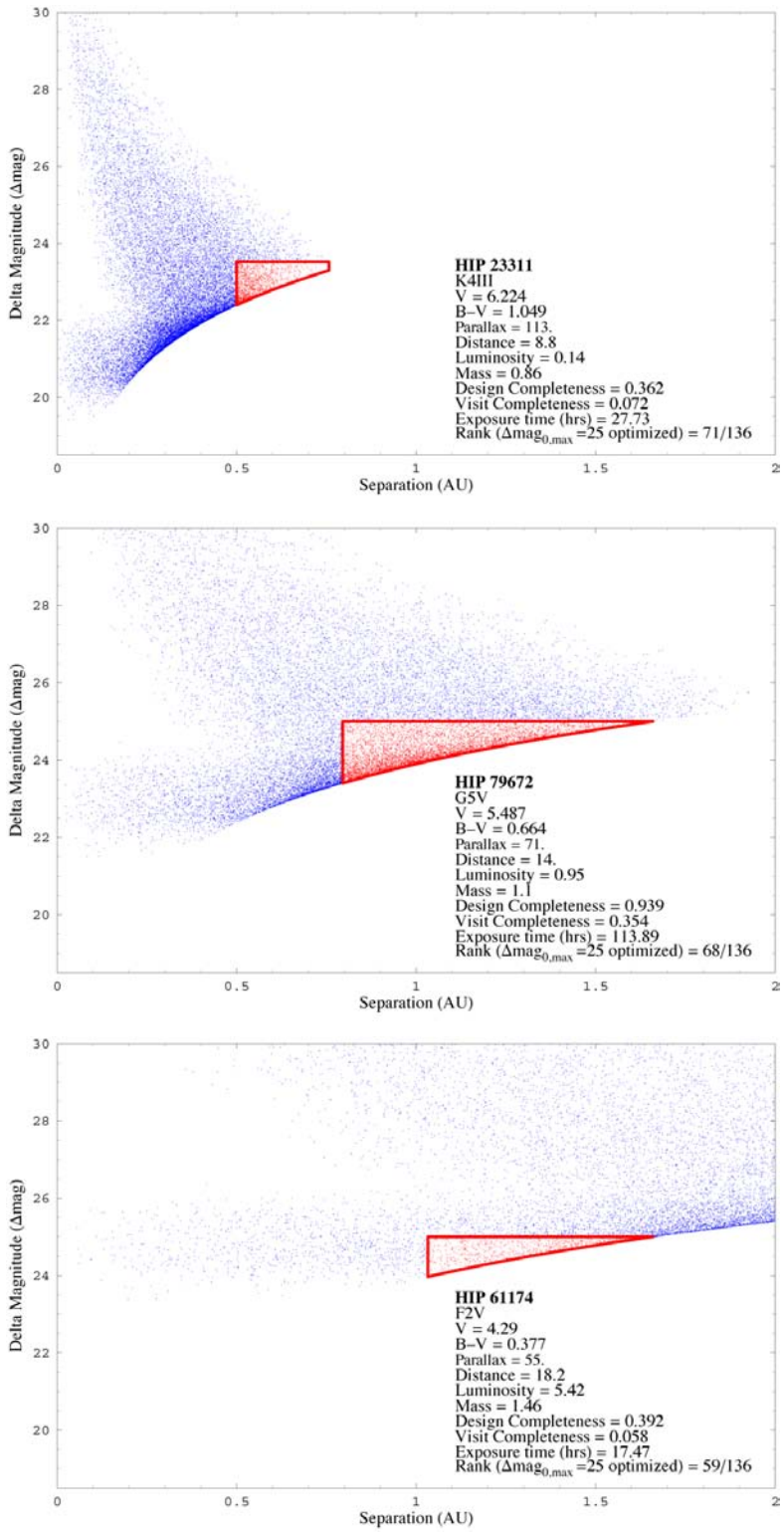


Figure 1: *Estimating C_1, the fraction of planets discovered. Out of 10,000 random planets from the population of interest, the red dots are members of the "discoverable subset" and the blue dots are not detectable. Instrumental (s_0, Δmag_0) and stellar (L, Π) factors select different discoverable subsets for each star.*

In simulations of observations, for example, we might proceed as follows. Let η_{Earth} be the probability that a star has a planet of interest. The number of planets and confusion sources actually found in the next search, $n_p = 0$ or 1, is a Bernoulli random variable with probabilities η_{Earth} C. In each trial, we would generate a sample value of n_p from a Bernoulli random deviate. If $n_p = 0$ in a trial, the experiment would proceed to consider another star (after a delay to allow the remaining possible planets to evolve in brightness and position, this star could be searched again). If $n_p = 1$ in a trial, we would select one "red-dot" planet at random to be the sample planet we had found. Only the computer would know the orbital elements, of course, but it can produce simulated observations of this planet as requested for double blind simulations of the science-operational decision-making process. Also, a new sample of planets consistent with an observation could be generated—the "found subset"—as a refinement over the discoverable subset.

The general behavior of the discoverable subset (or alternatively the found subset) can be studied as follows. For example, we want to know their "depletion time." That is, if a planet of interest were found around a given star, we want to know how long could a follow-up observation be delayed before the risk is greater than, say, 10% that the planet will have moved or changed in brightness enough to become undetectable? Call this time $T_{10\%LOST}$. To address this question, we prepared 10,000 discoverable planets for each of the 136 top TPF-C target stars, and tracked them until they became undetectable, either by disappearing behind the central obscuration or by becoming fainter than Δmag_0. The results are shown in Figure 2, and the values of $T_{10\%LOST}$ for our example stars are given in Table 1 (and in *StarVault* for all 136 stars).

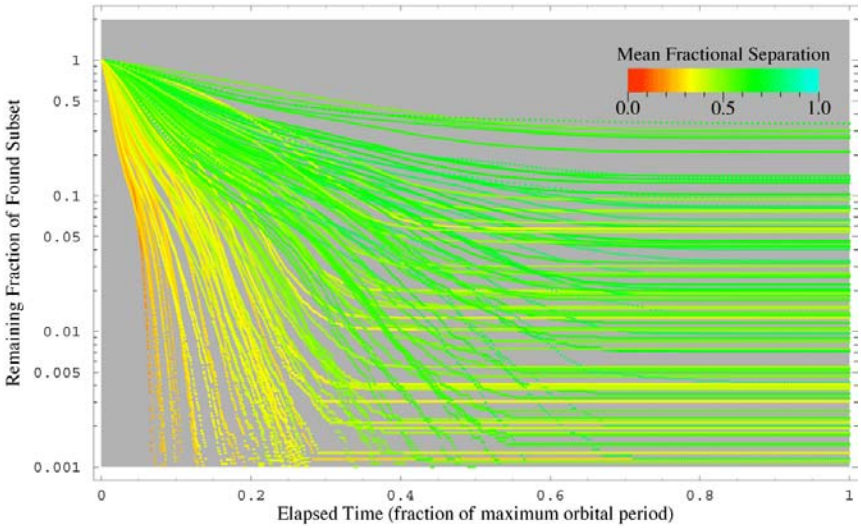


Figure 2. *The depletion with time of the discoverable subset of planets. Each curve is one star on the target list, color coded by the mean separation of the discoverable subset at the time of discovery expressed as a fraction of the maximum separation for that star (semimajor axis times one plus the eccentricity). The asymptotic value is the permanently observable fraction of the discoverable subset. The color trend is due to greater apparent transverse speed when smaller separations are preferentially sampled. The depletion curves for individual stars can be found at StarVault. $T_{10\%LOST}$ is the time for a depletion curve to drop to 0.90. Values of $T_{10\%LOST}$ are given for the three sample stars in Table 1, and for all 136 target stars in StarVault.*

In another example, we want to compute the apparent displacements of planets in the discoverable subset after some delay time, say $T_{10\%LOST}$, from the epoch of their discovery. We want this information to assess the ability of given astrometric accuracy to discriminate background confusion. This involves comparing the probability distribution of planetary displacements with the apparent displacement of a background source induced by proper motion and annual parallax.

3.3 Discriminating Background Confusion Sources[15]

It is critical to quickly confirm that a source is a planet and not background confusion. The timing of the confirming observation is a compromise between waiting too long, which risks losing the planet by "depletion," and waiting too little, which risks not achieving planetary displacements that are usefully different from the induced apparent displacement of a background source. A good compromise is $T_{10\%LOST}$. Because we know the orbital elements for each planet in the discoverable (or found) subset, we can compute the probability distribution of the displacements of discoverable (or found) planets in time $T_{10\%LOST}$.

By comparing the background and planetary displacements to the astrometric error of the instrument, we can estimate the probability, $1–P_{CONF}$, that an observation will allow us to reject the null hypothesis, that the putative planet is a background source. For the common proper motion test to be useful, P_{CONF} must be small—less than 0.0027, say, for three standard deviations of confidence.

The solar-avoidance restriction means that stars cannot be viewed if they lie within some specified angle of the Sun (95° for *TPF-C* FB-1). This angle dictates F_{SOLAVD}, the fraction of the year that a given star can be viewed. ($F_{SOLAVD} \leq 0.47$ for *TPF-C* FB-1.)

The intersection of the two conditions—stellar viewability and background discrimination utility—is of special interest. Searching observations must occur during the annual fraction when both are possible, which will be true for a fraction of the year we call $F_{SOLAVD \cap DISC}$. Note in Table 1 the huge difference in this quantity for the three target stars being used as examples. Figure 3 shows the details of how viewability and background-discrimination utility vary through the year for the example stars.

StarVault gives the same numerical and graphical results for all 136 *TPF-C* target stars assuming *TPF-C* FB-1. Background discrimination is *never* useful for 10% of the 136 stars, due to inadequate parallactic and proper-motion displacements, independent of solar avoidance ($F_{DISC} = 0$). It is not useful for 33% of the 136 stars at the same time the star is viewable according to solar avoidance ($F_{SOLAVD \cap DISC} = 0$). Furthermore, $F_{SOLAVD \cap DISC}$ >20%, >30%, and >40% for only 50%, 41%, and 20% of the stars, respectively. Because these 136 stars are the most productive qualified stars for *TPF-C*, replacement stars will be less productive than the least productive star currently on the list. The 137th star is HIP 88175 (zeta Ser), which offers the measly productivity of 0.72 milliplanets per hour, assuming it has one planet of interest. Smaller astrometric error (larger aperture) would improve the situation.

3.4 Planetary Recovery after the Viewability Epoch of Discovery

To expose issues of planetary recoverability, we prepared 10,000 planets in the discoverable subset for each of the 136 *TPF-C* target stars. We assigned each planet a day of discovery in the calendar year that was uniformly distributed over the first 25% of the stellar viewability window for that star, with the rationale that search observations should occur as early as possible, to allow time for follow-up observations before the viewability window closes. Next, for each planet, we ran time forward, after the closing of the viewability window of the discovery epoch, until the planet became observable again (i.e., the star was viewable and the planet was detectable); this time delay, measured from the time of discovery, is the *x*-coordinate of the plots in Figure 4. Next, the time was run forward again, to the point where the planet became unobservable again, either because the star was no longer viewable or the planet was no longer detectable; this duration, measured from the epoch of the first possible recovery, is the *y*-coordinate of the plots in Figure 4.

A planet's recovery window begins at the *x*-coordinate of a dot and extends to the right for a length in days equal to the y-coordinate of the dot, as illustrated by the red lines in the middle plane of Figure 4. The extended caption to Figure 4 is a tutorial. The following is a limited commentary on the results for the three example stars.

[15] See the reference in footnote #3 for a detailed discussion of this topic.

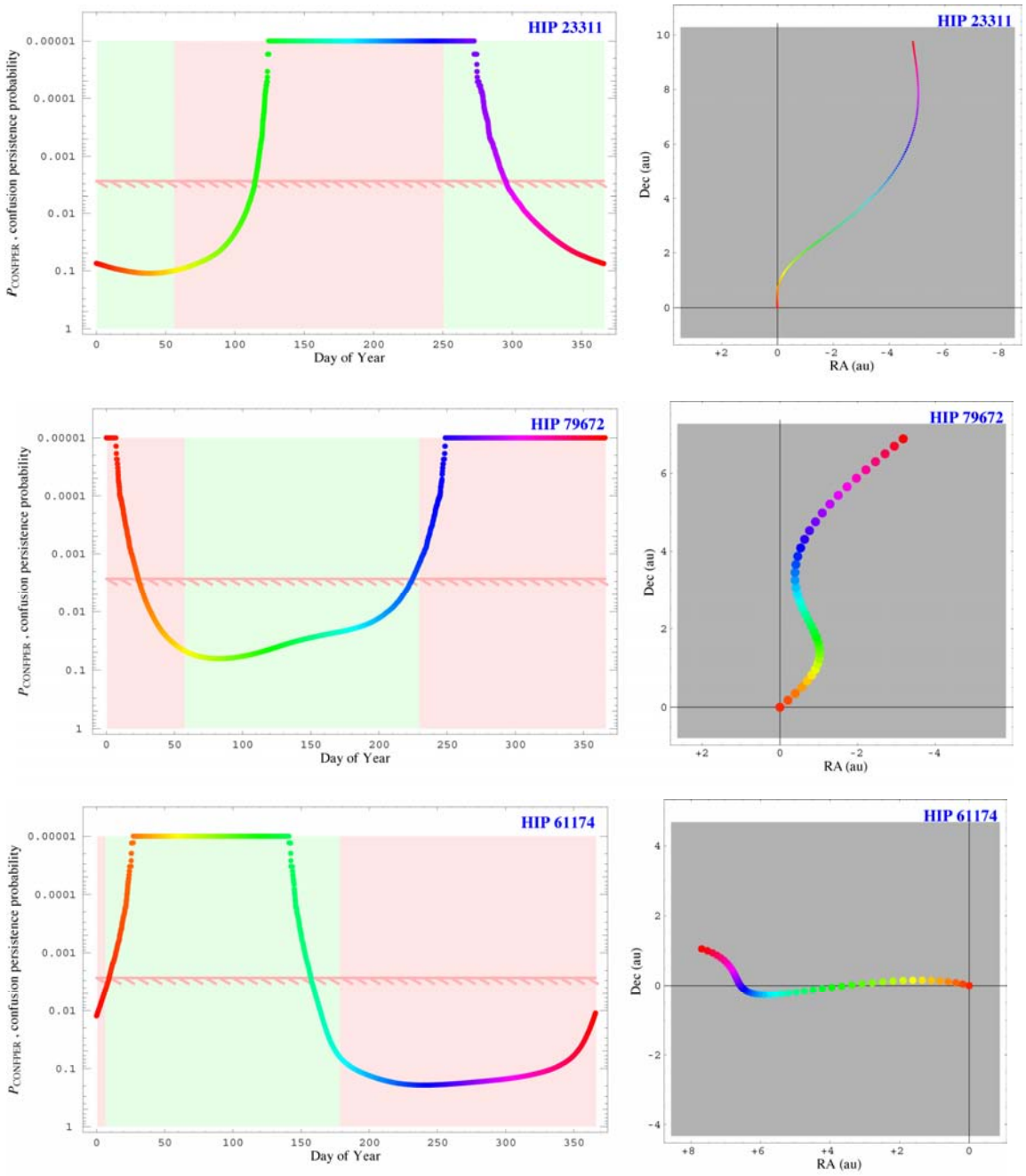


Figure 3. *Stellar viewability and background confusion differentiation. Left, light green: the days of the year when the star is more than 95° from the Sun and viewable by TPF-C FB-1. Rose: not viewable. Testing for common proper motion is useful when the curve is above the red, hatched line, which signifies three standard deviations of confidence that confusion will not persist after the second, differentiating observation. Right: displacement of background sources due to parallax and proper motion through the solar year. Points occur at intervals of $T_{10\%LOST}$, and the point radius is the mean amplitude of the planetary displacements in that time interval. The more widely separated the points on the right, the more useful is the confusion test. Rainbow color codes are synchronized: the larger the gap between points (right), the lower the probability of confusion persisting, $P_{CONFPER}$ (left). See the reference in footnote #3 for the details behind these calculations.*

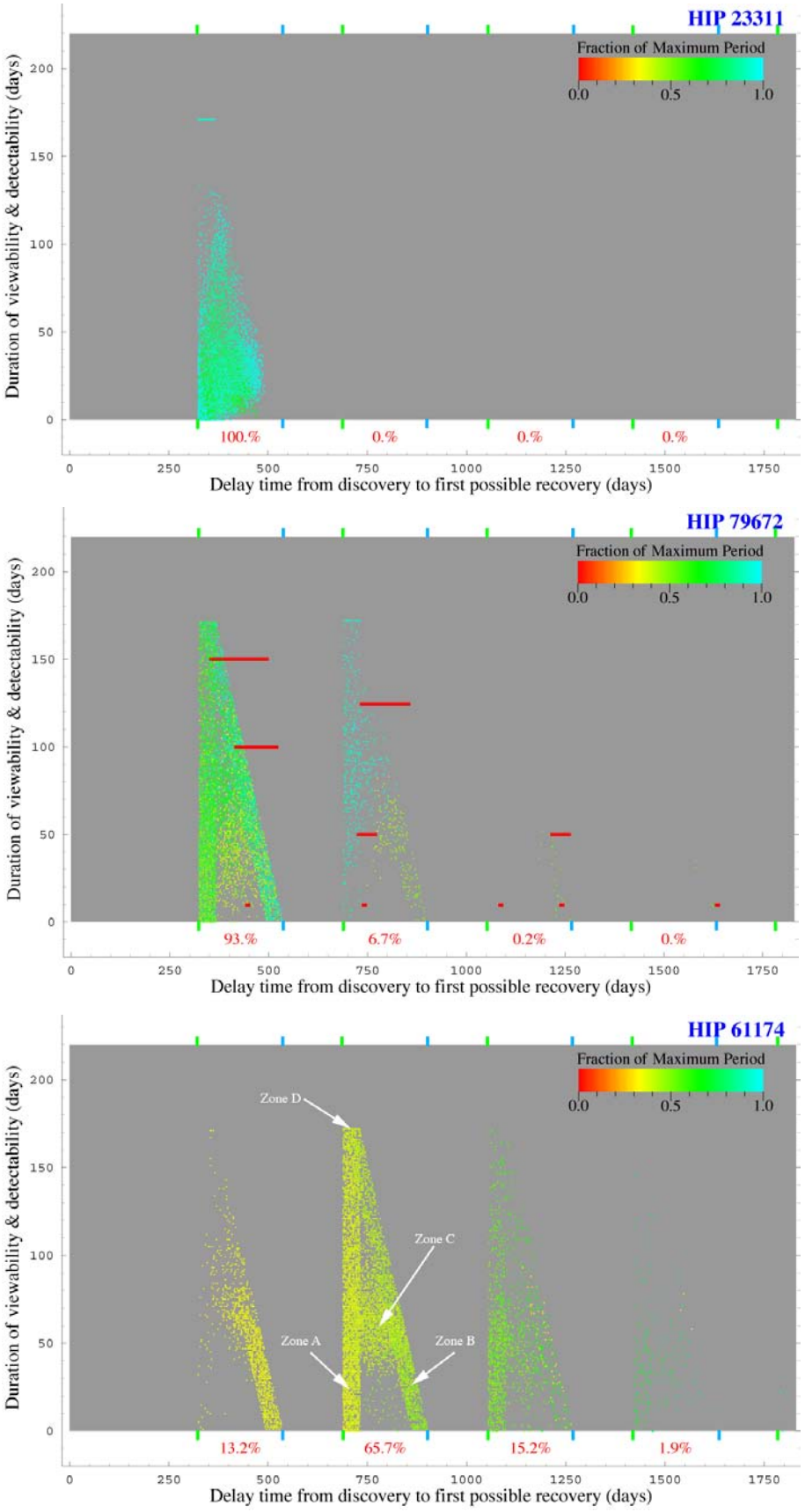


Figure 4. *(Extended figure caption follows.)*

HIP 23311 is the low-luminosity K star. From Figure 2, we see the discoverable subset consist largely of wider orbits (longer periods, as confirmed by the color code in Fig. 4), with the planets discovered near elongation. The maximum orbital period for planets of interest is 166 days, which is less than the stellar viewability window. Therefore, the entire discoverable subset is recoverable in the first stellar viewability window after the epoch of discovery. The duration of observability tends to be short, however, except for the permanently detectable planets, in orbits viewed nearly face on (the bar of dots at $y = F_{SOLAVD}$ 365.25 = 172 days).

HIP 79672 is a G star, like the Sun. From Figure 2, we see the discoverable subset consist of a wider range of orbits, with the planets viewed from half to gibbous phase. The maximum orbital period for planets of interest is 616 days, which is much longer than the stellar viewability window. Some 93% of the discoverable subset is recoverable in the first stellar viewability window, and the remainder are aliased out to subsequent stellar viewability windows. The duration of observability covers the whole range, from one day to the full width of the viewability window. There are some permanently detectable planets.

HIP 61174 is a high luminosity F star. From Figure 2, we see the discoverable subset consist of a wide range of orbits viewed mostly side-on, with the planets discovered rushing to or from full phase at superior conjunction. The maximum orbital period for planets of interest is 1973 days, which is much longer than the stellar viewability window. Only 13% of the discoverable subset is recoverable in the first stellar viewability window, and these planet tend to appear late in the stellar viewability window. Some 66% and 15% can only be recovered in the second and third windows, respectively.

Planetary recoverability. *X*-axis: delay time from planetary *discovery* to the first possible *recovery* in a subsequent observability period (star viewable and planet detectable). *Y*-axis: duration of the observability period of recovery. Each dot is one of 10,000 planets representing the discoverable subset. Dot color gives the planetary period as a fraction of the longest period (365.25 days times (a_{max} = 1.5 $\sqrt{luminosity}$)$^{1.5}$ ($mass$)$^{-0.5}$, where stellar luminosity and mass are given in Table 1. The red lines on plot for HIP 79672 show the observability periods of ten particular planets explicitly. For this experiment, the time of planet discovery was confined to the first quarter of the stellar viewability period. Numbers in red give the percentage of planets that are first recoverable in each stellar viewability period.

Each stellar visibility window comprises four zones, identified in the plot for HIP 61174.

Zone A is the vertical strip that starts at x = 365.25 (n–0.25 x F_{SOLAVD}) days, for n = 1, 2, 3... (green marks). This value of x is the minimum possible delay in the n^{th} viewability window. It implies a planet that was discovered at the last possible moment in the first quarter of the initial viewability window and is first observable again at the very first possible moment of the n^{th} subsequent viewability window. Hence the minimum possible delay of 0.75 x 365.25 days. The end of the Zone A strip is 0.25 x F_{SOLAVD} x 365.25 days later, at 365.25 x n days. It implies a planet that was discovered at the very first possible moment in the initial viewabiliity window and is observable again at the very first possible moment of the n^{th} viewability window.

In Zone A, the observability period for the recovery of any planet is punctuated at the start by the stellar viewability window opening, and at the end by the planet becoming undetectable—by the planet going behind the central obscuration or becoming fainter than the sensitivity limit—not by the end of the stellar viewability window.

Zone B is the diagonal strip that starts at the top of Zone A and extends down to the right. At y = 0, Zone B ends at x = 365.25 (n + F_{SOLAVD}) days (blue marks). This value of x is the maximum possible delay in the n^{th} viewability window. It implies a planet that was discovered at the first possible moment in the first quarter of the initial viewability window and is first observable again at the very last possible moment of the n^{th} subsequent viewability window. Hence the maximum possible delay in the first subsequent viewability window of (1 + F_{SOLAVD}) x 365.25 days. The start of the Zone B strip at y = 0 is 0.25 x F_{SOLAVD} x 365.25 days earlier, at x = 365.25 (n + 0.75 x F_{SOLAVD}) days. It implies a planet that was discovered at the very last possible moment in the first quarter of the initial viewability window and is observable again at the very last possible moment in the n^{th} observability window.

In Zone B, the observability period for the recovery of any planet is punctuated at the start by the planet becoming detectable, and at the end by the stellar viewability window closing.

Zone C is the region under Zones A and B. Here, the observability period for the recovery of any planet is punctuated at the start by the planet becoming detectable and at the end by the planet becoming undetectable.

Zone D is the line at the top of Zones A and B, where the observability period for the recovery of any planet is punctuated at the start by the stellar viewability window opening, and at the end by the stellar viewability window closing. Zone D includes the permanently detectable planets, in face-on orbits.

S*tarVault* contains these plots for all 136 stars on the target list.

4. SUMMARY, COMMENTS AND CONCLUSIONS

We have studied the requirements for planetary discovery using *TPF*. Although the focus has been on *TPF-C*, the general conclusions apply to *TPF-I*, as well. Our science-operational concerns do not apply to *TPF* observing planets found by other means—the *Space Interferometry Mission* (*SIM*) or ground-based radial-velocity searches, for example.

We have found ample indication that the science operations of *TPF-I* and *TPF-C* will face significant challenges eliminating background confusion and avoiding wasted observations in the attempt to recover previously discovered planets.

In the case of *TPF-C*, brightness will vary strongly in the detectable portion of the orbit for some planets. In principle, such brightness variations should be helpful for discriminating background sources and gaining orbital information useful for planetary recovery. We expect this phase effect to be more helpful for stars with better resolved habitable zones. Studies in progress should clarify and quantify the benefits. We expect to find a smaller, less helpful phase effect for *TPF-I*, at least for planets that efficiently redistribute the energy deposited by starlight, by rotation or atmospheric circulation.

Adequate astrometric accuracy addresses these challenges. Studies are underway to determine what astrometric accuracy is required to make science operations tolerably efficient and robust against the risks posed by undifferentiated confusion and guesses about recovery based on poor orbital information. These studies will use Monte Carlo methods to model aspects of the issues as well as the mission itself. The results will determine the minimum scale of the *TPF* mission.

Instrument designers are encouraged to study and report the expected astrometric performance of their *TPF-I* and *TPF-C* instrument concepts when measuring the star-planet separation and position angle.

Modeling similar to that shown here for the *discoverable* subset, could be performed on the *found* subset of planets—those consistent with the *measured* apparent separation and brightness of a particular discovered planet. This would more closely emulate the science-operational situation.

5. ACKNOWLEDGEMENTS

We salute Christian Lallo for his expertise in aiding the computations, graphics, and website. We thank Sharon Toolan with her help preparing the manuscript. JPL contract 1254081 with the Space Telescope Science Institute provided essential support.

TPF-C: Size and Completeness

Sarah Hunyadi and Stuart Shaklan
Jet Propulsion Laboratory, California Institute of Technology

Robert Brown
Space Telescope Science Institute

The Terrestrial Planet Finder Coronagraph (TPF-C) is required to find 30 terrestrial planets, provided that each star has such a planet. This requirement may be evaluated using program completeness [1,2], the cumulative sum of habitable zone search space observed over the duration of the mission. Beyond meeting mission goals, program completeness is used to compare different types of coronagraph designs.

In this study we compare the program completeness of different coronagraph designs:

1) TPF-C FB-1 design,
2) Phase induced amplitude apodization (PIAA),
3) Optical Vortex Coronagraph (Vortex), and
4) New Worlds Observer external occulter (NWO).

For all cases we assumed a habitable zone populated with 10,000 Earth-like pseudo planets uniformly distributed in semi-major axis and having an orbital ellipticity in the range [0, 0.1]. The exozodi level was chosen to be equivalent to the mean Solar System zodi level over the entire habisphere. The central wavelength of interest was 550 nm with a bandpass of 110nm. The throughput for the different mask designs is shown in Table 1.

The TPF-C FB-1 mission design baseline is an 8×3.5m elliptical primary mirror with a band-limited 8^{th} order occulting mask. An elliptical mirror allows larger habitable zone coverage in one direction, but requires additional rolls and integration time to circularize the inner working angle. The PIAA, Vortex and NWO designs may be able to utilize a 4m telescope, though this is still under study. Therefore, we simulated program completeness with this reduced-size concept.

For FB-1, PIAA and Vortex, potential targets were required to be at least 95 degrees from the sun. We also assumed 1 hour of overhead per target to account for slew time between targets.

For NWO, we assume a single occulter with the overhead shown in Table 1 and up to 1 full day of integration time per target. We assume that NWO can detect planets at a contrast ratio of 26 magnitudes, while the other designs achieve up to delta magnitude 25.5. We also restricted the potential observable region to a 15 degree latitudinal slew from the previously viewed star. This restriction was combined with an additional constraint on sun-angle that limited the potential targets to be between 45 and 95 degrees from the sun. The results for all cases are shown in the table below.

---

[1] Brown, R. A., Hunyadi, S. L., and Shaklan, S. B., A DRM for *TPF-C*: A Design Reference Mission for the Coronagraphic *Terrestrial Planet Finder*.
http://sco.stsci.edu/tpf_downloads/TPF-C_DRM.pdf

[2] Hunyadi, S. L., Shaklan, S. B., and Brown, R. A. Single Visit Completeness Optimization *(submitted to ApJ, April 2006)*.

Table 1: Completeness simulation results.

	Long Axis (m)	Ellipticity	IWA (mas)	IWA (λ/D)	Overhead (hrs)	Throughput	Completeness	# Targets
FB-1	8	Elliptical	65.5	4	1	11%	31.97	66
PIAA	4	Circular	56.7	2	1	28%	70.14	140
PIAA	4	Circular	70.9	2.5	1	28%	50.59	104
PIAA	4	Circular	85.1	3	1	28%	35.25	75
PIAA	4	Circular	99.3	3.5	1	28%	25.02	56
PIAA	4	Circular	113.4	4	1	28%	17.34	33
Vortex	4	Circular	56.7	2	1	23%	65.73	133
Vortex	4	Circular	70.9	2.5	1	23%	48.01	99
Vortex	4	Circular	85.1	3	1	23%	34.12	71
Vortex	4	Circular	99.3	3.5	1	23%	24.46	54
Vortex	4	Circular	113.4	4	1	23%	17.33	33
NW0	4	Circular	65.5	*	312	80%	23.01	58
NW0	4	Circular	65.5	*	480	80%	18.45	41

* The IWA for NW0 is based on the mask size and distance from the telescope, not λ/d

The Vortex and PIAA methods at 3 λ/D achieve scientific gain greater than or equal to that of our baseline FB-1 design (at 4 λ/D). These results depend on our ability to demonstrate reliable reduction of the inner working angle (IWA) for these devices. The NWO design may also be able to achieve similar results through the use of multiple or larger external occulters. Further completeness studies will implement variable throughput over the coronagraph field of view to account for different throughput levels at different working angle locations.

The Super-Earth Explorer

J. Schneider, P. Riaud, G. Tinetti, H. M. Schmid, D. Stam, S. Udry, P. Baudoz, A. Boccaletti, O. Grasset, D. Mawet, J. Surdej, and the SEE-COAST Team

LUTH—Paris Observatory, LESIA—Observatory of Paris-Meudon, HOLOLAB, Department of Physics, University of Liège,

ABSTRACT

The Super-Earth Explorer Coronagraphic Off Axis Space Telescope (SEE-COAST) is a space mission concept to be submitted to ESA call for proposals. It is devoted to the direct imaging, in the visible, of the stellar light reflected by its planetary companion. Here we present the type of planetary characterization we may accomplish with SEE-COAST, and we include a brief description of the spacecraft and mission strategy.

1. SCIENTIFIC OBJECTIVES

The main objective of the Super-Earth Explorer Coronagraphic Off Axis Space Telescope (SEE-COAST) is the characterization of a selected ensemble of exoplanets previously discovered by radial velocity and astrometric surveys. The project will be submitted to ESA call for proposals for small size missions to be launched in 2016.

Several questions can be raised about exoplanets' properties, some of them can be addressed by the detection of the stellar light reflected by the planet or of the thermal photons emitted by the planet. Both approaches are valid and may provide a complementary information. Of course, all questions may not be answered by a single mission. For some aspects we will have to wait for the second generation missions.

The planetary properties we are interested in observing and constraining are: the size (mass and radius), the atmosphere (chemical composition, clouds, seasonal variations, thermal inertia), the surface (type -rocks, ice, water, "vegetation"-, inhomogeneities), rotation (period, atmospheric dynamics) and environment (rings). Reflected light and/or thermal emission may be used to study these planetary characteristics. The former approach relies on the information that can be extracted from the stellar light reflected by the planet as a function of different parameters:

Wavelength λ, polarization $\vec{P}$, time t. The planetary reflected flux $F_{pl}(\lambda, t, \vec{P})$ is $\sim 3 \cdot 10^8$ smaller than the stellar for a Jupiter at 1 AU with albedo A = 0.5. This high contrast ratio imposes an efficient stellar light suppression mechanism (see section 2).

The planet flux allows to derive only the product $A \times R^2_{pl}$ from the observed flux at a given time. Nevertheless the information embedded in (1.1) is very rich and can be used to disentangle the planetary albedo from the radius. The planetary characteristics may be retrieved from the dependence of the planetary flux to each of these parameters: λ(atmospheric composition, Rayleigh scattering, surface albedo, biosignatures) $\vec{P}$ (nature of the surface, Rayleigh scattering by a clear atmosphere) t (seasonal (climate), random (variation in cloud coverage), surface inhomogeneities, planetary rotation period, rings, thermal inertia...). Some of the planetary features, such as the albedo, can be constrained only using the combined information provided by the dependence on λ, $\vec{P}$ and *t*: a complete planetary characterization needs a 3-D fit of the function to the data. In addition, the cartography of the whole planetary system may give hints on the planet-disk connection. Further details on these topics are given in the next sections.

1.1 Wavelength Dependence

A spectral resolution R ~ 40 may provide information on essentially three aspects of the planetary atmosphere: chemical composition, pressure and optical properties:

Atmospheric species

Table 1 shows the most important species accessible to the reflected light approach.

Table 1: *List of the most important molecular species absorbing in the spectral range accessible to SEE-COAST. Successful detection of some of those lines will depend on molecular abundance, S/N obtainable for that particular target and planetary atmosphere's conditions.*

Species	Spectral Lines (nm)
CH_4	486, 543, 576, 595, 619, 668, 681, 703, 790, 840, 864, 727, 889
CO_2	1.05, 1.25
H_2O	514,575, 610, 650, 730, 830, 940, 1150
NH_3	552, 647, 929
O_3	Chappuis band centered at 602
O_2	626, 688, 762

Rayleigh scattering

When clouds or aerosols are not present in the atmosphere, the planetary spectrum may show an increased re-flectance in the blue due to Rayleigh scattering. From Rayleigh scattering one can infer constrains on pressure and transparency of the atmosphere and surface reflectance properties. In the Solar System, we can observe Rayleigh scattering on six planets: Earth, Titan, Uranus, Neptune, Jupiter and Saturn.

Clouds

We will observe the planet at different phases during its orbital period. The scattering properties of clouds, when averaged over the visible disk, are very sensitive to phases (Tinetti et al. 2006). This information, combined with the expected effects due to polarization (see sec. 1.3) will help to constrain the cloud coverage and the absolute albedo of the planet.

1.2 Time Dependence

The planetary flux may vary on different time scales, for our purposes we are interested in monthly and daily variations. Periodic fluctuations on short time scales are possibly related to the rotation of the planet around its axis and might be detected using surface inhomogeneities and surface reflectance properties (Ford et al. 2001, Tinetti et al. 2006). If clouds do not entirely cover the surface, this method can be used to estimate the rotation period. Note that this modulation contrast and the period do not require the knowledge of the absolute value of the albedo; the ambiguity given by the product $A \times R^2_{pl}$ is not a problem here. Such investigation requires an extensive monitoring of the planet flux which can be limited to a few days or weeks.

Short term random variations may be explained by fluctuations in cloud coverage.

Long term patterns might be correlated with the orbital motion, and might have different explanations: seasonal variability, due to a combination of tilt and orbital eccentricity effects (e.g., the case of Neptune (Sromovsky et al. 2003)), or effects induced by the planetary rings (Arnold and Schneider 2004).

Eccentric orbits

Half of the planets detected have eccentricities larger than 0.35. The origin of this large eccentricity is unclear. It may or may not be the case for terrestrial planets. If it is related to the planets' migration and terrestrial planets are not affected by this process, they could have essentially circular orbits. But we must be prepared for this possibility: we expect in this case noticeable oscillations in the planetary temperature as a function of orbital parameters (Sertorio and Tinetti, 2002). As a first order approximation, the temperature gradient between the periastron and the apoastron can be estimated (for a relatively small eccentricity e) by $\Delta T = e\ T$. For instance, if $e = 0.35$ and $T = 300$ K, $\Delta T = 100$ K. The effect is thus not negligible.

Rings

Due to their negligible mass, we may suppose that rings are extremely difficult to detect. But this very small mass is compensated by a large area, as large as the planet itself in the case of Saturn. Since they may have a high albedo (~0.5), rings can be as easy to detect by reflected light imaging as the planet itself. The master formula (1) is valid only for spherical bodies and not for disk structures like rings. Planets with and without rings produce different light curves as a function of orbital angle, due to mutual shadow and occultation between the

ring and the planet itself 2 (Arnold and Schneider 2004). We can use this property to discriminate whether the planet is ringless or not. From a planetology point of view, rings are not only a curiosity. Their existence reveals the presence of satellites (which are continuous suppliers of dust by their collisions) and their radius R_R gives an indication on the density $\rho_R = M_{pl}/R^3_R$ of their building blocks.

1.3 Polarization

The degree of polarization of reflected light provides extra information about the scattering properties of the atmosphere under study. Polarization may help to estimate the degree of cloudiness of a planetary atmosphere, to retrieve the size of the scattering particles compared to the wavelength of the impinging light and to constrain the albedo because of its sensitivity to different phase angles (Schmid et al. 2006, Stam et al. 2005).

Multiple diffusion in fact destroys polarization. Therefore the degree of polarization is a measurement of the thickness of clouds. A high degree of polarization is a signature of single Rayleigh scattering (clear sky) or of a solid surface. The combination of the degree of polarization with a λ^{-4} dependence is a strong sign of Rayleigh scattering. Extrasolar planets may be observed at various angular phases (with respect to the parent star) from 0 to 90°. The angular phase dependence is thus correlated with the orbital motion of the planet.

2. SPACECRAFT AND MISSION STRATEGY

These scientific objectives can be achieved with the 1.5 m coronagraphic telescope of the SEE-COAST mission. Its 2.2 m^2 collecting area is sufficient to detect $1R_{Jup}$ planets up to 15 pc and $3R_\oplus$ planets up to 5 pc. The high star to planet contrast of 10^8 - 10^9 imposes the use of a stellar light rejection mechanism. We have chosen a high performance coronagraphic phase mask with a suppression factor of 60,000 on the stellar peak (Riaud et al), associated with a superpolished prime mirror. The focal instrument is a multiple coronagraphic device (Mawet et al. 2005) to attenuate the star and a spectropolarimeter with a low spectral resolution of R = 40 in the band 0.4 – 1.25 micron. An extension toward the UV is under study.

One possibility is to place the spacecraft on an L2 orbit (i.e., around the Earth-Sun Lagrange point L2), allowing full accessibility of about half of the sky at any epoch. But other possibilities are considered.

The mission philosophy is not to search for new planets, but to characterize planets already known by radial velocity or other techniques. At the end of 2006, about 220 planets were discovered: among them, already a sub-set of 30 can be detected with SEE-COAST. We anticipate that by the time of launch (2016), more planets will be accessible to SEE-COAST. As ancillary science, SEE-COAST will also observe circumstellar disks and AGNs.

REFERENCES

Arnold L. & Schneider J., 2004. Astron. Astrophys., 420, 1153

Ford E., Seager S. & Turner E., 2001. Nature, 412, 9885

Mawet D. et al.. 2005. ApJ. 633, 1191

Schmid H. M. et al., 2006. Search and investigation of extra-solar planets with polarimetry. in Direct Imaging of Exoplanets: Science & Techniques. Proceedings of the IAU Colloquium 200, Edited by C. Aime and F. Vakili. Cambridge University Press, p.166

Sertorio L. and Tinetti G., Constraints in the coupling star-life, 2002, Nuovo Cim. 025C, p. 457-500; astroph/0110031.

Stam D., et al., 2006. Integrating polarized light over a planetary disk applied to starlight reflected by extrasolar planets. Astron. & Astrophys., 452, 669

Sromovsky I., Fry P., Limaye S. Baines K., 2003, Icarus, 163, 256

Tinetti G., Meadows V., Crisp D., Kiang N., Khan B., Fishbein E., Velusamy T., Bosc E., Turnbull M., Detectability of Planetary Characteristics in Disk-averaged Spectra II: Synthetic Spectra & Light-curves of Earth, Astrobiology, 6, 6, 2006

Achromatic Optical Vortex Coronagraph with Subwavelength Gratings

Dimitri Mawet[1], Pierre Riaud[1,2], Jacques Baudrand[2], Anthony Boccaletti[2], Pierre Baudoz[2], Daniel Rouan[2], Cédric Lenaerts[3], and Jean Surdej[1]

[1] *Institut d'Astrophysique et de Géophysique de Liège (IAGL), University of Liège,*
[2] *LESIA, Observatory of Paris-Meudon,*
[3] *HOLOLAB, Department of physics, University of Liège*

ABSTRACT

We present in this short paper an efficient and powerful solution for making achromatic optical vortex coronagraphs. We propose to use the unique properties of subwavelength grating integrated meta-materials to induce achromatic phase shifts that can be implemented to yield vortices of any topological charge. TPF-C specifications are very challenging and require pushing this concept to its limits.

1. INTRODUCTION

Optical vortices have recently gained interest in the coronagraphists community (see the papers from Palacios and Swartzlander, these proceedings). Such an attraction is comprehensible since they present many advantages over other concepts of coronagraphs like for example, the conjunction of high rejection ratios, small inner working distance, full discovery space, high throughput, etc. In the framework of ground-based coronagraphy for second-generation instrumentation, they are foreseen to advantageously replace classical Lyot coronagraphs, and even more recent coronagraphs such as the four-quadrant phase-mask coronagraph which is already online at the VLT (Boccaletti et al. 2004, Riaud et al. 2006). For TPF-C, where the detectivity constraints are dramatically emphasized, optical vortices represent a competitive approach. Nevertheless, in order to be compliant with the daunting specifications of an Earth-like planet-finding mission like TPF-C, one has to push the concept to its limits. Two major problems are to be solved before seriously considering the implementation of optical vortices into this ambitious space observatory. First of all, their sensitivity to low-order aberrations has to be significantly kept below an acceptable level for the telescope pointing stability constraint to remain in comfortable margins. This issue also relates to the sensitivity to stellar angular diameter which is known to dramatically affect small inner working distance coronagraphs such as the FQPM, for instance (Guyon et al. 2006). Finally, as optical vortices are phase masks, one has to find a way to efficiently achromatize them, i.e., making them keeping their high rejection ratios over wavelength ranges equivalent to the usual astrophysical filters (bandwidth of ~ 20 %).

2. SUBWAVELENGTH GRATINGS

When the period of a grating is smaller than the wavelength of the incident light, it does not diffract as a classical spectroscopic grating. All the incident energy is enforced to propagate only in the zeroth order, leaving incident wavefronts free from any further aberrations. The subwavelength gratings are therefore often called Zeroth Order Gratings (ZOGs). This type of gratings behaves like homogeneous media with unique characteristics, which can be used to synthesize artificial birefringent achromatic waveplates. Quarterwave or halfwave plates are extensively used in astrophysics for polarimetric studies. Subwavelength gratings constitute an elegant, flexible and integrated solution to produce them. The key point is that by carefully controlling the geometry of the grating structure, one can tune the so-called form birefringence and induce a precise phase shift between the orthogonal polarization components of the incident light. The phase shifting optical function can then be integrated at the surface of a substrate with any kind of geometry (Figure 1), leading to an efficient and flexible solution for implementing the phase shifting spatial distribution of phase-mask coronagraph such as the FQPM (Mawet et al. 2005a) or the AGPM (Mawet et al. 2005b).

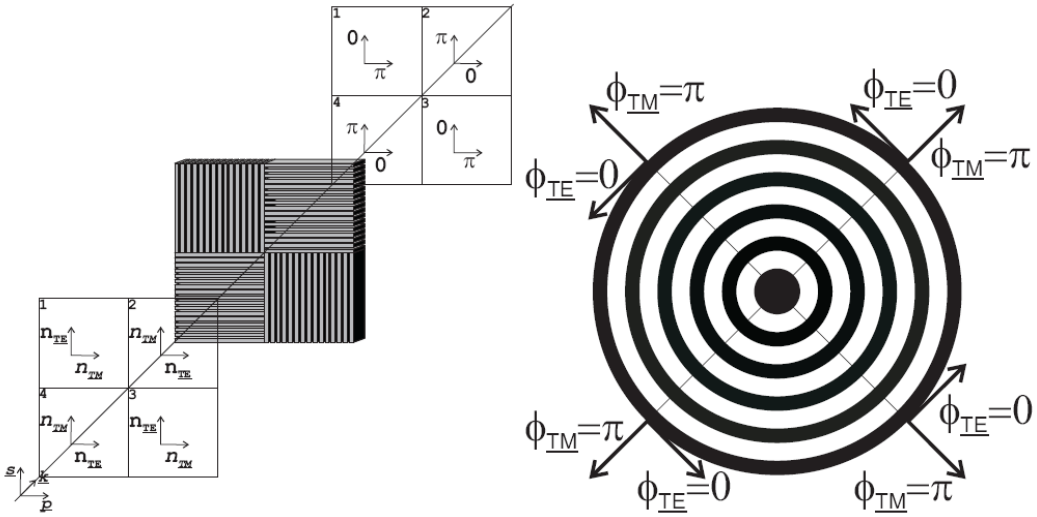


Figure 1: *Left: 4QZOG implementation (Mawet et al. 2005a), where the four subwavelength gratings engraved on a unique substrate are strictly identical and implemented in the following way: 2 of them in 2 quadrants along 1 diagonal are rotated by 90° around their normals with respect to the two others. Right: AGPM implementation (Mawet et al. 2005b). The AGPM consists of a concentric circular surface-relief grating with rectangular grooves with depth h and a periodicity Λ. It can be seen as a circularly symmetric FQPM in polarization.*

3. LIMITATIONS OF THE ANNULAR GROOVE PHASE MASK CORONAGRAPH

The AGPM is an optical vortex of topological charge 2 (Mawet et al. 2005b). It can be demonstrated that the sensitivity to a low-order optical aberration x depends on the topological charge m as x^m (Mawet et al. 2005b). This sensitivity to low-order aberrations is the harmful counterpart of the property of small inner working distance of this coronagraph and others like the FQPM or the AIC, for instance. This side-effect is too penalizing to consider the AGPM directly as such for TPF-C. It is however worth noting that, by adding a tiny Lyot dot at the center, this effect can be reduced together with the sensitivity to the stellar size (Mawet et al. 2005b). As far as the achromatization is concerned, one can expect total nulling depths of about 5×10^{-6} over 20% bandwidths (without the central Lyot dot). This means that the contrast at an angular separation of 4 λ/D can reach about 5×10^{-8} in the unaberrated case. This is still more than 2 orders of magnitude above the required 10^{-10} specification imposed by the contrast between an Earth-like planet in the habitable zone and its host star. The 5×10^{-8} limit is not fundamental but design-dependent which also means technology-dependent.

4. HIGHER-ORDER MASKS

Higher-order masks must be investigated for TPF-C. A compromise between the IWD and aberration sensitivity must be found. According to Guyon et al. 2006, an m=6 vortex mask would be the better compromise for TPF-C. Achieving such a topological charge with subwavelength gratings only depends on our ability to control the orientation of the grating lines to be imprinted (Figure 2). In this context, state-of-the-art realizations (see e.g., Niv et al. 2006) have already achieved m=4 masks without any difficulty in the infrared (10.6 microns) and on top of AsGa substrates (Figure 2).

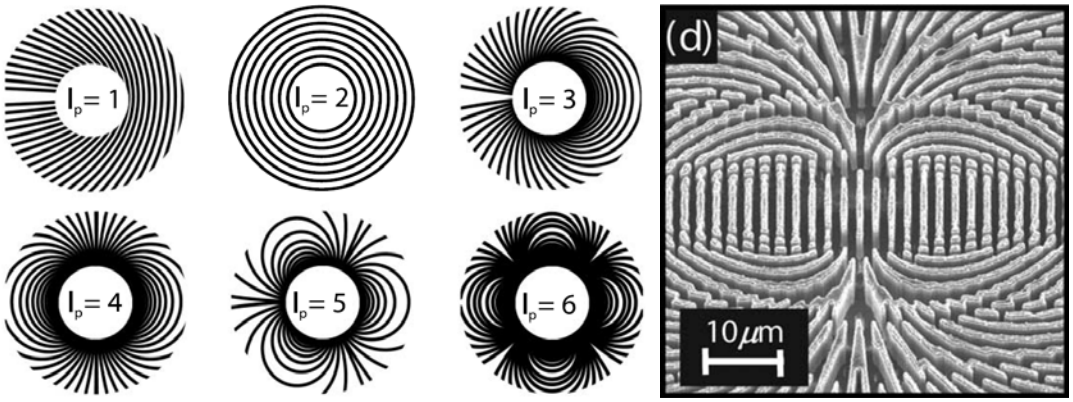


Figure 2: *Left: grating geometry for different topological charges (1-6). Right: SEM picture of a realization by Erez Hasman's group (Israel Institute of Technology). This picture shows an m=4 optical vortex engraved in an AsGa substrate for infrared applications (10.6 microns).*

5. VISIBLE OPERATION

Making subwavelength gratings for the visible wavelength range is not trivial. Subwavelength means that the period of the grating must be smaller than the wavelength of the incident light. In the visible, this means that the period of the grating must absolutely be smaller than λ/n, n being the index of refraction of the chosen substrate material. We therefore have to deal with periods in the 300 to 400 nm range according to the material. Controlling the micro-structure at such scales is challenging but feasible using for instance nano-imprint lithography techniques. In this context, let us cite the work of Deng et al. (2005) for optical pickup units. They indeed proposed, manufactured and tested (now commercialize) achromatic subwavelength grating waveplates for the visible. Converting the performance of their manufactured components (Figure 3) in terms of nulling efficiency would lead to a mere 1000 rejection ratio over a bandwidth of more than 20 %. There is still room for improvement. An intensive research in the grating optimization for the visible is nevertheless needed.

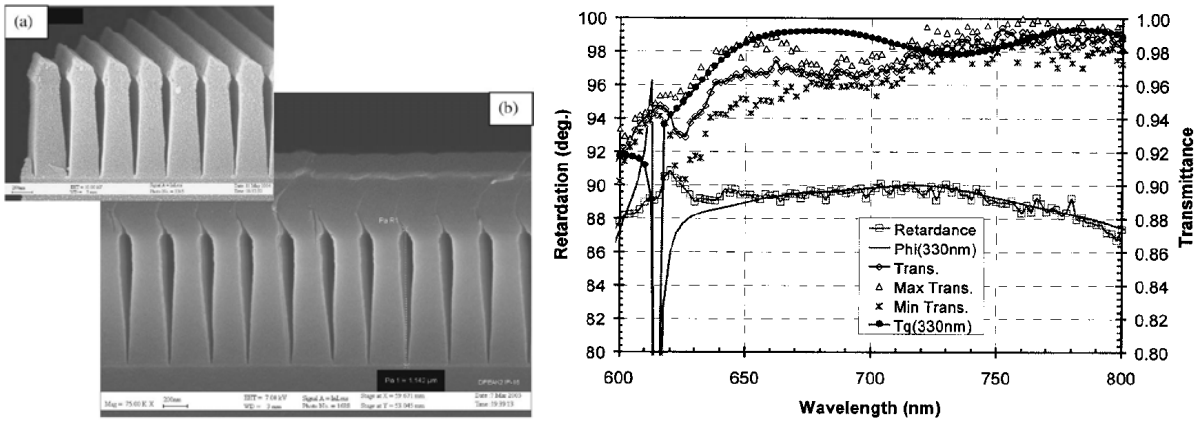


Figure 3: *Left: SEM picture of a subwavelength grating achromatic phase shifter for the visible. Right: measured performance of the micro-structure in terms of phase and throughput over the visible (Deng et al. 2005)*

6. CHROMATIC RESIDUALS FILTERING

Assuming a state-of-the-art nulling performance of 5×10^{-6} over a 100 nm bandwidth in the visible from the subwavelength grating technology, i.e., a contrast of about 5×10^{-8} at 4 λ/D in the unaberrated case, this imposes us to gain an additional factor of 1000 to reach the required level of 10^{-10}. If we consider the vectorial nature of

the vortex induced by the space-variant grating (see Niv et al. 2006, for instance), we can use the Jones formalism that gives the following matrix for the transmission through the component, in the helical basis:

$$J_{vortex}(x,y) = \frac{1}{2}\left(\eta_{TE} + \eta_{TM} e^{i\Delta\phi}\right)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{1}{2}\left(\eta_{TE} - \eta_{TM} e^{i\Delta\phi}\right)\begin{bmatrix} 0 & e^{i2\theta} \\ e^{-i2\theta} & 0 \end{bmatrix}$$

This equation indicates that the field emerging from a space-variant subwavelength grating comprises three components. The first (left term) maintains the original polarization state and phase of the incoming beam. The second (right term, top right component) is right-handed circularly polarized and has a phase modification of $2\theta(x,y)$, inducing a vortex of topological charge m according to the adopted geometry ($\theta(x,y)$ defines the orientation of the grating grooves, see Figure 2) . The third (right term, down left component) has an orthogonal polarization direction and opposite phase modification with respect to the second component. Note that the magnitude of the different components is determined by the local birefringent parameters of the subwavelength gratings: the diffraction efficiencies in transmission for the polarization components TE (or s) and TM (or p) η_{TE}, η_{TM} and the phase shift between them $\Delta\Phi$. The transmission of subwavelength dielectric gratings is relatively high and the retardation $\Delta\Phi$ is primarily a function of the subwavelength groove geometry (filling factor) and etching depth. The ideal situation is for η_{TE}, η_{TM}=1 and $\Delta\Phi=\pi$. Unfortunately, this idyllic behaviour cannot be maintained over large bandwidths and lead in practice to the 5×10^{-6} over 20 % bandpass limitation mentioned here above. However, any departure from the perfect conditions will result into a *transfer of energy* between the circularly polarization states with the embedded optical vortex and the first term that conserves the original polarization state (free of the optical singularity). This crucial property of vectorial vortices, and vectorial vortices only, allows us to consider splitting and filtering the polarization. This splitter/filter should induce an extinction ratio of 1000 or more, corresponding to the additional factor needed to reach the 10^{-10} contrast at 4 λ/D. A proper data post processing with both polarizations could also help increasing the contrast to reach 10^{-10} at 4 λ/D by calibrating the residual speckles. In practice, the polarization filter/splitter only consists in quarterwave plates and linear polarizing filters/beamsplitters. Note that the subwavelength grating technology could synthesize both components, and with the advantage that the whole optical function could be integrated on top of a single substrate (Deguzman & Nordin 1999).

7. CONCLUSIONS

Optical vortices such as the AGPM are promising coronagraphs for ground-based applications of extrasolar planet-finding instruments. In particular, we are currently manufacturing prototypes for the VLT-Planet Finder second-generation XAO instrument for the VLT (SPHERE). Optical vortices of higher orders are well ranked for TPF-C. However, manufacturing them at the required specification for Earth-like planet detection is extremely challenging. In this paper, we have reviewed the current state of the art of the vectorial vortex technology. Vectorial vortices are induced by space-variant subwavelength gratings. Pushing this technology to its limits, we proposed some hints to reach the 10^{-10} over usable bandwidths. This however would require further studies to definitely conclude on the feasibility of the proposed concept for TPF-C.

REFERENCES

Boccaletti et al. 2004, PASP 116, 1061
Deguzman & Nordin 1999, Appl. Opt. 40, 5731
Deng et al. 2005, Optics Letters 30, 2614
Guyon et al. 2006, submitted to ApJ
Mawet et al. 2005a, Appl. Opt. 44, 7313
Mawet et al. 2005b, ApJ 633, 1191
Niv et al. 2006, Opt. Express 14, 4208
Riaud et al. 2006, A&A 458, 317

Japanese Terrestrial Planet Finder

Motohide Tamura
JTPF working group and the National Astronomical Observatory of Japan

ABSTRACT

A Japanese roadmap on direct extrasolar planet studies is presented, from the current ground-based telescope to future IR/Opt space-based telescopes. Several searches for young planets and disks have been conducted with the Subaru 8.2m telescope and its adaptive optics (AO) infrared coronagraph, CIAO. The instrument will be soon upgraded to a new AO and a coronagraph with simultaneous spectral and polarimetric differential imaging modes (HiCIAO), which improve the contrast performance by more than order of magnitudes. A sensitive unbiased survey for extrasolar zodiacal emission around nearby stars is being conducted with the recently launched AKARI space IR telescope (0.7m diameter, 2–200 micron). A successor space IR telescope, SPICA (3.5m diameter, 5–200 micron), is also planned; its high sensitivity will enable the detection and characterization of outer-most planets around nearby stars. For the studies of extrasolar terrestrial planets in Japan (JTPF), a high contrast space telescope (HCST; 3.5m, 0.3–2 micron) is under discussion. We are also seeking for collaborations with or joining to foreign missions.

A roadmap shown in Figure 1 is an approach to the explorations of exoplanets in Japan. Our ground-based direct explorations are concentrated on the fully-operated and successful Subaru 8.2m telescope at Mauna Kea, Hawaii. Making use of the good observing site and the excellent image quality, the infrared coronagraph CIAO (Coronagraphic Imager with Adaptive Optics) has been used for various kinds of surveys. CIAO is the first among IR coronagraphs equipped with a full cold coronagraph mode (with various cold occulting masks and Lyot stop optics) on the world 8-m class telescopes. Compared to previous coronagraphs, it has unique features of near-infrared operation and small occulting masks (down to 0.2 arcsec in diameter; Tamura et al. 1998, 2000). CIAO has been successful to directly reveal the morphological diversity of protoplanetary disks, such as rings, spirals, and banana-split, with a high spatial resolution of 0.1″ (Fukagawa et al. 2004, 2005). In addition, first 2 micron imaging polarimetry of beta Pic has demonstrated that the disk is composed of several planetesimal belts with a gap at r ~ 100 AU (Tamura et al. 2005). It has also detected young brown dwarf around a classical T Tauri stars DH Tau (Itoh et al. 2005).

HiCIAO is a new high-contrast instrument for the Subaru telescope, which is currently developed by NAOJ and University of Hawaii (Tamura et al. 2006; Hodapp et al. 2006). HiCIAO is used in conjunction with the new 188 actuators adaptive optics system at the Infrared Nasmyth platform. It is designed as a flexible camera comprising several modules that can be configured into different modes of operation. The main modules are the AO module with its near-future extreme AO capability, the warm coronagraph module, the high contrast optics module, and the cold infrared camera module. HiCIAO will be the first instrument on the 8-m class telescopes which can combine coronagraphic techniques with simultaneous polarization and spectral differential imaging modes which minimizes the common path errors.

Regarding space-based explorations, the MIR and FIR space telescope, AKARI (aka ASTRO-F), is developed by JAXA/ISAS and is successfully launched in 2006 February. This telescope has a similar size to the Spitzer telescope, but is optimized to all-sky survey observations. Although its low spatial resolution is not suitable for direct explorations, it will conduct an unbiased survey for faint disk emission. This enables us to make census of exo-zodi and disk evolution with a large sample of stars. As a successor of AKARI, the SPICA mission plans to launch as a cooled (4.5K) large (3.5m) single-mirror telescope into a halo orbit around the Sun-Earth L2 orbit around 2014. Its unprecedented sensitivity at MIR and FIR and its simple telescope pupil make SPICA to be one of the best platforms to implement coronagraph instrument for exoplanet studies.

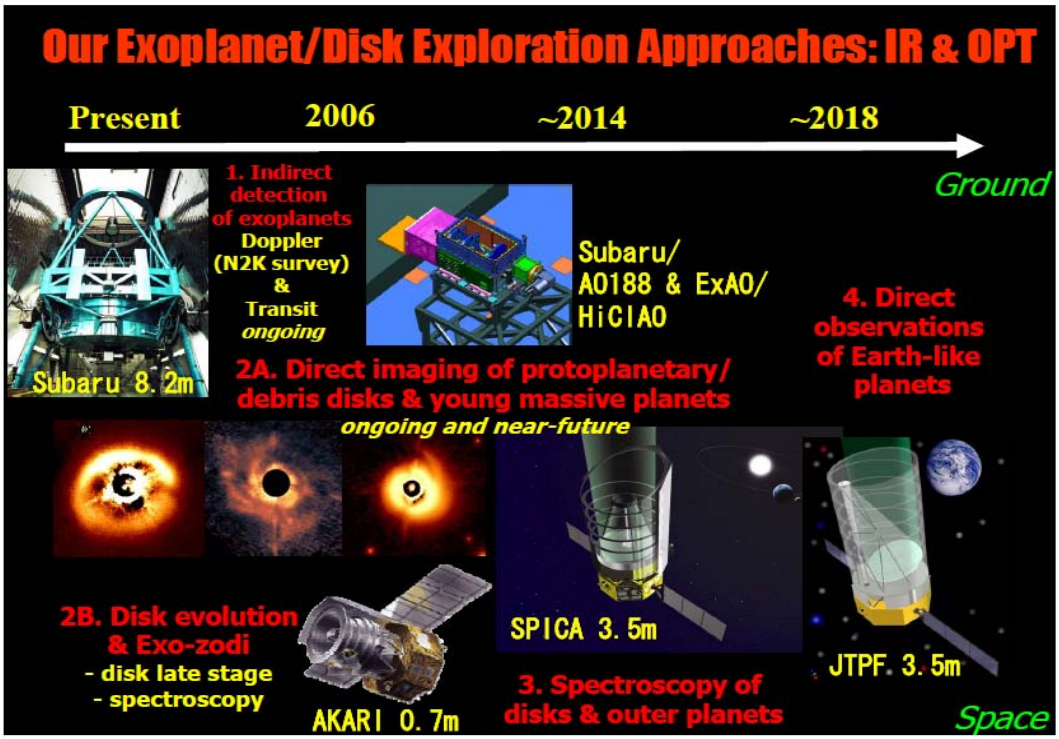


Figure 1. *A roadmap of exoplanet explorations in Japan.*

JTPF (Japanese Terrestrial Planet Finder) is a future space mission whose main science driver is the extrasolar terrestrial planet studies. The project is officially approved both by NAOJ and ISAS. In NAOJ, its activity is regarded as one of the main themes of the Phase A project which has started in April 2005. In JAXA/ISAS, the JTPF Working Group has been approved by the ISAS Science Steering Committee in 2002. Discussion has been made among both astronomy and planetary communities in Japan, and regarded as one of the important programs for optical and infrared astronomy. It could be either an independent mission or a joining collaboration to TPF/Darwin missions. At present, an 3.5m high contrast optical space telescope (HCST), which realizes clean and stable images, is under main discussions for the JTPF architecture among JTPF WG (see Figure 2). JTPF/HCST might make full use of the SPICA bus system with an optimization to shorter wavelengths.

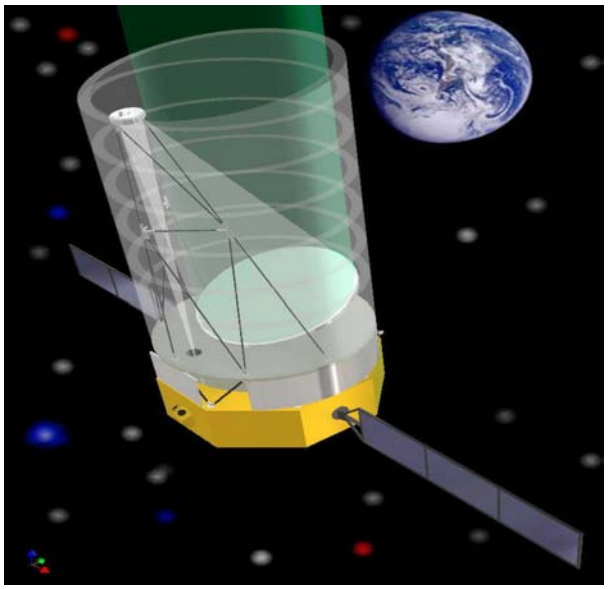


Figure 2. *JTPF/HCST -3.5m high contrast optical telescope.*

The SPICA Coronagraph Project

L. Abe[1], K. Enya[2], S. Tanaka[2], T. Nakagawa[2], N. Murakami[3], J. Nishikawa[3], M. Tamura[1], K. Fujita[4], Y. Itoh[4], H. Kataza[2], O. Guyon[5], and the Spica Working Group.

[1]Optical and Infrared Astronomy Division, National Astronomical Observatory of Japan, [2]Department of Infrared Astrophysics, Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, [3]MIRA Project, National Astronomical Observatory of Japan, [4]Graduate School of Science and Technology, Kobe University, [5]Subaru Telescope

ABSTRACT

We present the latest results of our coronagraphic instrumentation study for the Japanese 3.5m SPICA telescope. Our aim was to validate some of the candidate concepts, amongst which the binary checkerboard mask (Kasdin et al. 2003). This solution benefits from theoretical intrinsic achromaticity, simplicity of its optical implementation which is a key issue since the SPICA telescope is designed to work at cryogenic (4.5K) temperatures. The huge drawback of this solution is its poor so-called inner working angle (IWA) and its overall transmission. This solution, although a clear compromise between performance and complexity, has proven to be a secured backup solution for the SPICA coronagraph.

SPICA MISSION SUMMARY

The Space Infrared telescope for Cosmology and Astrophysics (SPICA) is the next generation mission for infrared astronomy led by Japan (Nakagawa et al. 2004). The SPICA telescope uses on-axis Ritchey-Chretien optics with a 3.5m diameter monolithic primary mirror. The whole telescope will be cooled to 4.5K and infrared observations will be made at wavelengths within the 5–200μm range. A coronagraphic instrument is currently being considered for the SPICA mission (Enya et al. 2006, Abe et al. 2006). The primary target of this instrument is the direct detection and observation of extra-solar Jovian outer planets (typically beyond 5AUs). The baseline requirements for the SPICA coronagraph are as follows: 10^{-6} peak-to-peak contrast, smallest possible IWA, and a core wavelength within the range 5–20μm.

LABORATORY EXPERIMENT

Optical configuration

We chose to test 2 masks: one optimized for a pupil with a central obstruction and another for a clear pupil. In the optimized solutions providing throughput higher than 15%, the design which had the smallest IWA was adopted for manufacture. The targeted contrast to be demonstrated in this work was 10^{6}, so that the mask was designed to achieve a contrast of 10^{-7}. The size chosen for the central obstruction of the first mask (Mask1, IWA of 7 λ/D) was 30% of the entrance pupil diameter. Subsequently, for comparison, a second mask (Mask2) was designed for a pupil without obstruction, and to probe the coronagraphic sensitivity at much lower IWA (3 λ/D). The OWAs for Masks1 and 2 were assumed to be 16 and 30 λ/D. In our experiment, λ=632.8nm and D=2mm (i.e., the diagonal of our square masks). The smaller OWA gave a solution where the rectangular patterns of the mask were not too narrow and therefore made the fabrication simple and robust. The larger OWA provided a wider dark region while it made the fabrication more complex. Table 1 summarizes the specifications of the masks. Figure 2 shows the shape of the masks, and the expected coronagraphic point spread function (PSF) for Mask1 and Mask2.

Figure 1 shows a schematic of our optical setup. All the optics were set in a dark room with an air cleaning system. Clean air flowed from the top of the room to the optical table during the settings and measurements. Although we plan to use an adaptive optics system in future experiments, it was not used in this work. A He-Ne laser was used as the light source. A spatial filter consisting of a microscope objective lens and a 10μm diameter

pinhole was also used. A commercially available cooled CCD camera (BITRAN) with 2048×2048 pixels was used to measure the PSF. The overall chip size was 15.16mm×15.16mm with a flat glass window in front. The CCD was cooled and stabilized at 0°C throughout the experiment. The camera was mounted on a linear motor drive stage in order to scan along the optical axis to find the best focus position.

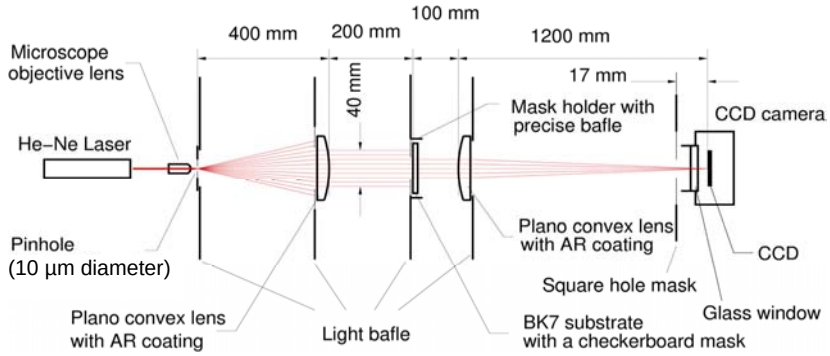


Figure 1: *Laboratory experiment optical setup.*

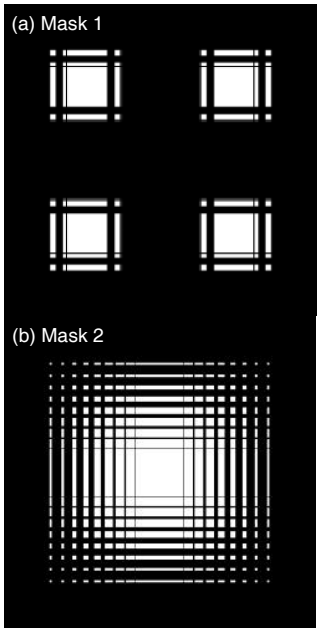


Figure 2*: Mask1 and Mask2 theoretical shapes. Both masks are 2mm in diagonal.*

MASK FABRICATION

Each checkerboard mask consists of an aluminum film on a BK7 substrate, and were manufactured using nano-fabrication technology at the National Institute of Advanced Industrial Science and Technology (AIST) in Japan. A 100-nm thick film aluminum was evaporated onto the surface of the sub-strate. After evaporation, electron beam patterning and a lift-off process were applied to produce a rectangular trans-missive area for the checkerboard pattern. The substrate was 2mm thick and had no wedge angle. The flatness, parallelism and scratch-and-dig of the substrate were less than $\lambda/10$, 5 arcsecond and 10^{-5} respectively. A narrow-band multiple-layer coating was applied to both sides of the substrate to reduce reflection at the surface after fabrication. The temper-ature of the mask during coating was kept less than 100°C to avoid deterioration of the residual resist around the mask pattern. The reduction in reflectivity of the BK7 substrate was confirmed to be less than 0.25% as a result of this coating. Small defects were found on Mask2, but we found no signi-ficant degradation in performance for this level of contrast.

DATA ACQUISITION AND RESULTS

The dynamic range of the CCD was insufficient to capture the full dynamic range of the PSF in one single exposure. Therefore, several exposure times were used: 0.03, 0.1, 1.0, 10, 100, and 1800 seconds. The 1800 seconds exposure was for measurements in the dark region. The brightest region of the PSF saturated the detector even with the shortest exposure time, so we used additional neutral density filters with an optical density of 2. We also carefully checked the CCD linearity by imaging the PSF core through a combination of neutral density filters with a total optical density of 6, and with an 1800 seconds exposure. We found that no significant correction was needed so that we could rely on the multi-exposure time method.

The majority of the area of the dark region in the PSF with Mask1 is less than 10^{-6} as shown in Figure3. On average, the observed contrast for each dark region on a linear scale is 3.4×10^{-7}, 2.1×10^{-7}, 1.8×10^{-7} and 3.5×10^{-7} for DR1, DR2, DR3 and DR4 respectively, where DR1–DR4 are the dark regions corresponding to the quadrants around the core shown in Figure4. The average contrast for all the dark regions is 2.7×10^{-7}. The 3σ detection

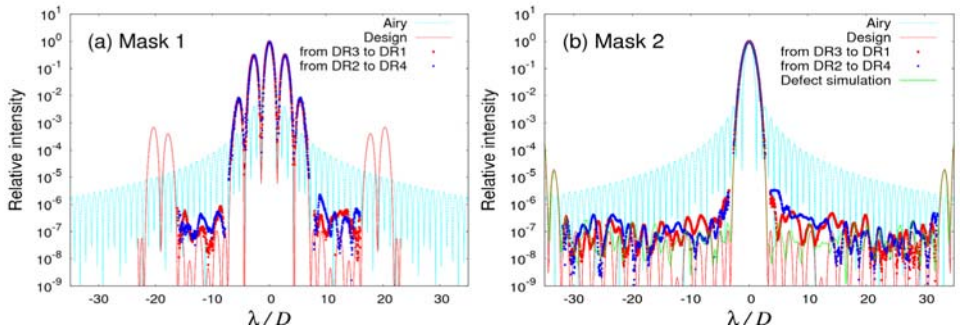


Figure 3: *Diagonal cuts across PSFs of both Mask1 (left) and Mask2 (right).*

limit computed within each observed dark region DR1–DR4 is 7.5×10^{-7}, 5.8×10^{-7}, 4.1×10^{-7}, and 9.1×10^{-7}, respectively. PSF subtraction or fitting has the potential to provide better performance, though such methods require assumptions about the repeatability of the pattern observed in the dark regions or the shape of the PSF. Therefore, the simple 3σ limit corresponds to a conservative estimate of the detection limit. Nevertheless, all the values are below our 10^{-6} limit.

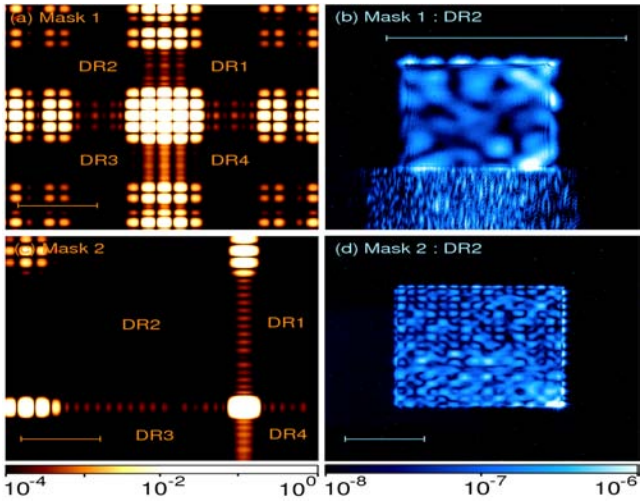


Figure 4: *Mask1 and Mask2 PSFs (left) and close-up of the respective dark regions (right). The scale bar corresponds to 10 λ/D.*

For the coronagraphic image with Mask2, the average contrast of DR1, DR2, DR3 and DR4 on a linear scale is 1.1×10^{-7}, 1.0×10^{-7}, 9.0×10^{-7} and 1.3×10^{-7}, respectively. The average contrast of the whole dark region is 1.1×10^{-7}. The 3σ limit is 3.8×10^{-7}, 3.0×10^{-7}, 2.7×10^{-7} and 3.8×10^{-7} for DR1, DR2, DR3 and DR4 respectively. In the bottom part of Figure 4(d), an irregular speckle pattern can be observed, getting brighter close to the optical axis (bottom right). This pattern was invariant in repeated measurements with a fixed setup. Change in the air flow and suspension conditions made no significant difference in the observed image. On the contrary, when we shifted and rotated the mask, this speckle pattern changed, suggesting the same limiting factors as for Mask1.

CONCLUSION

We reported the results from a laboratory experiment using a binary checkerboard-type pupil shaped mask coronagraph. Although this kind of mask presents several limitations (large IWA and low throughput), its manufacture and implementation are very robust. Two masks, optimized for either obstructed pupil or not, were fabricated and tested. Both masks achieved a higher performance than the 10^{-6} requirement: 2.3×10^{-7} and 1.1×10^{-7}. Therefore, we conclude that a binary checkerboard mask is a very attractive solution for the future MIR coronagraph of the SPICA space telescope, for which this study was carried out. Other coronagraphic solutions, more complex to implement are also considered, but the checkerboard mask can be regarded as the safest option.

A more exhaustive description of this experiment is available at astro-ph/0609646 (soon to appear in A&A).

REFERENCES

Abe, L. et al., 2005, proc. of IAU Colloquium **200**, 329

Enya, K. et al. 2006, proc. of SPIE, **6265**, 97

Kasdin, N. J. et al., 2003, ApJ, **582**, 1147

Nakagawa, T. and SPICA working group, 2004, Adv. Sp. Res., **34**, 645

The Coronagraphs of MIRI/JWST and SPHERE/VLT as Valuable Experiences for TPF-C

A. Boccaletti and P. Baudoz
LESIA, Observatoire de Paris

ABSTRACT

This paper presents two projects for which we are developing advanced coronagraphic system based on phase mask coronagraphs.

INTRODUCTION

For about 10 years, a lot of coronagraph concepts were proposed and studied theoretically by mean of numerical simulations. Some are tested in the lab in the context of TPF-C project but except the classical Lyot coronagraph only the Four Quadrant Phase Mask (FQPM, Rouan et al. 2000) was installed on a ground-based telescope (Boccaletti et al. 2004) in combination with Adaptive Optics system and has already provided scientific results (Gratadour et al. 2005, Boccaletti et al. 2007). The purpose of this paper is first of all to present some actual projects both in space and on the ground where FQPMs will be implemented to address the characterization of giant gaseous planets at near and mid IR wavelengths, and second of all to outline the importance of such developments for TPF-C.

1. MIR/JWST

By 2013, the James Webb Space Telescope (JWST, Gardner et al. 2006) will be the largest telescope in space to provide a large spectral coverage from visible to mid-IR wavelengths (0.6μ–25μ). JWST has a large scientific program defined in a Design Reference Mission. One of these programs is entitled “planetary systems and origins of life”. It includes the understanding of planet formation and evolution also in relation with circumstellar disks.

JWST is equipped with three instruments: NIRCAM, NIRSPEC and MIRI. The latter is the mid-IR facility of JWST and is developed in a NASA-led partnership with a European consortium sponsored by the European Space Agency (Wright et al. 2003). MIRI is composed with 2 modules one which is the imager under French responsibility (Dubreuil et al. 2003). At the Observatoire de Meudon, we are designing a coronagraph suite installed in the focal plane of the imager. The coronagraphic system is combining 3 monochromatic FQPMs for exoplanets research and characterization operating at 10.65μ, 11.40μ and 15.50μ, plus a standard Lyot coronagraph at longer wavelengths (23μ) for the study of silicates in cold circumstellar disks.

The selection of coronagraph filters was made upon the analysis of exoplanet spectral models (Allard et al. 2001) compared to the background noise. As illustrated in Figure 1 the best spectral region to search for exoplanets is located between 9μ and 15μ. At shorter wavelengths, MIRI is undersampled and at larger wavelengths the background is dominating the planets signal. The set of filter allows measuring the presence of ammonia in the planetary atmosphere, which is a good sensor of the temperature and of the sedimentation.

To address the performance of the coronagraphic system we developed numerical simulations fully described in Boccaletti et al. (2005) and Baudoz et al. (2006). We considered the JWST segmented pupil with 30μ gaps, phasing defects of about 84nm rms, mid frequencies aberrations of 58nm rms and high frequencies of 17nm rms. In our system analysis we found that most of the limitation in MIRI comes from the pupil position uncertainty (with respect to the Lyot stop) which may reached 4.5% of the pupil size, the defocus at the focal plane of JWST (±2μ) and the pointing accuracy (about 5mas). Expected performances are given in Table 1 in some particular condition. Unlike MIRI, NIRCAM is corrected for defocus and pupil position. Nevertheless, exoplanet contrasts

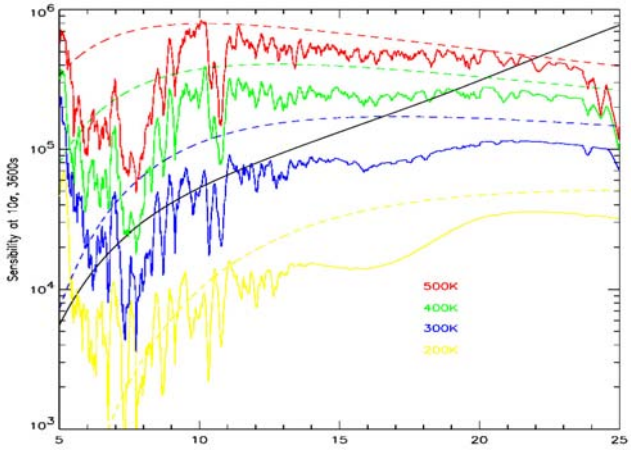


Figure 1: *Spectra of exoplanets modeled by Allard et al. (2001) for several temperatures and compared to the background model for MIRI/JWST. The range between 9μ and 15μ is clearly the best place to look at. The sensitivity will allow to detect planets with temperature larger than 300K.*

Table 1: *Raw contrast on the coronagraphic image for different cases (jitter is 7mas rms and pointing accuracy is 5mas).*

Observation conditions	Attenuation on the peak	Contrast		
		1 λ/D	3 λ/D	6 λ/D
1.5% shear + 1μ defocus	420	1392	4338	11940
3% shear + 2μ defocus	256	901	2912	9345
5% shear - 3μ defocus	115	386	1399	5016

are more favorable in the mid-IR, which makes MIRI and NIRCAM very complementary. In Figure 2, we give the impact of the pointing accuracy on the expected performance.

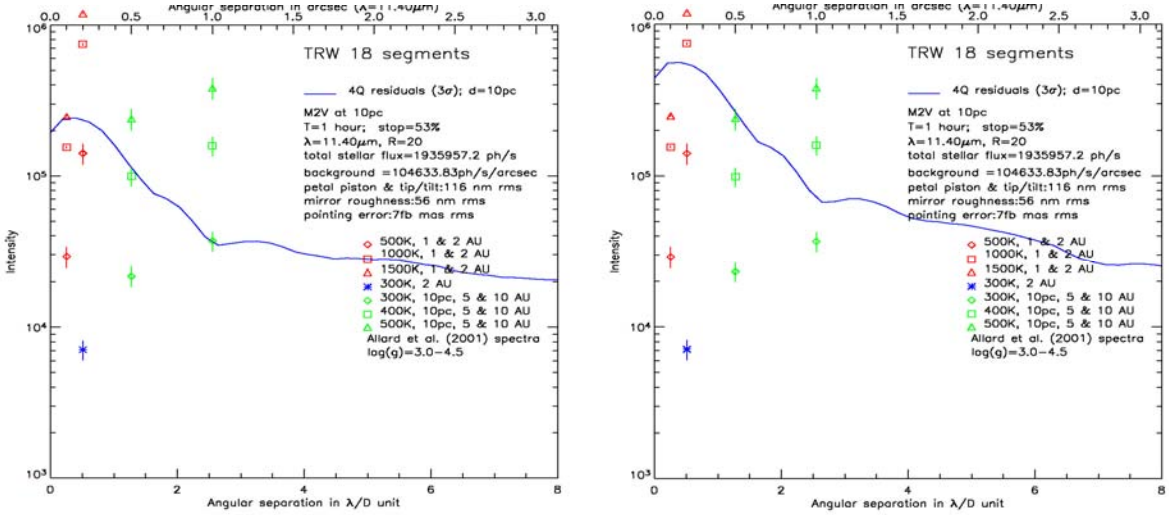


Figure 2: *Achievable contrast at 3 sigma versus angular distance for 5mas pointing accuracy (left) and 10mas (right). Colored symbols are corresponding to different types of planets as indicated on the plot. The presented performance implies the calibration of the coronagraphic image residue with either a reference star or observation of the target star at a different roll angle. The degradation is about linear in between 5mas and 10mas (for comparison PSF size is 360mas at 11.4μ.*

In the mid next decade, MIRI will be the only instrument in space equipped with a specific coronagraphic system. Although several parameters are imposed by the telescope, the system analysis we made has contributed to the optimization of the coronagraph design. In this respect, this study can benefit to TPF-C at several levels:

development of phase masks for use in space in specific condition (vibration, thermal cycling, etc.); evaluation of critical issues that reduce performance (defocus, pupil positioning, etc.); experience we get in simulations and manufacturing. The requirements we put on MIRI to achieve detection of giant planets is giving a very good idea of what TPF-C would have to do in terms of stability to achieve the detection of telluric planets. Several orders of magnitude will have to be gained at many levels. The error budget will have to be controlled much better than in any previous instruments.

2. SPHERE/VLT

Unlike MIRI, SPHERE is a 2nd generation ground-based instrument at the VLT (ESO). SPHERE stands for Spectro Polarimetric High contrast Exoplanet REsearch and is probably one of the first instruments to be fully optimized for the direct detection and characterization of exoplanets. SPHERE is studied among a large European consortium including 12 institutes in several countries (France, Germany, Italy, Switzerland, Netherland). Target objects will be giant gaseous planets down to 1 Jupiter mass and will be searched for inside young stellar associations (50–150pc), nearby stars (smaller orbits), stars with known planets (from radial velocity surveys with long term residuals) and stars of intermediate age (a few 100Myrs). Detailed specifications of SPHERE can be found in Beuzit et al. (2006). The fundamental principle in this instrument is to take advantage of differential imaging whether in several spectral bands or different polarization states. For that, an AO corrected wavefront is feeding an IR coronagraphic device, which distributes the light into 2 instruments IRDIS (an IR camera using dual band imaging) and IFS an integral Field Spectrograph adapted to high contrast imaging. A dual polarization camera ZIMPOL is installed in the visible channel of the instrument with its own coronagraph.

A realistic simulation of all observing modes of the instrument is developed under CAOS environment (Carbillet et al. 2004). This tool has several goals which are :

– to put specifications on the system in terms of :
  - AO related parameters
  - PSD and amplitude of phase aberrations upstream the coronagraph
  - pointing alignment (accuracy, reproducibility)
  - pupil centering (accuracy, reproducibility)
  - focus stability
  - chromatic aberrations upstream/downstream coronagraph
  - comparison of coronagraphs and optimization of each design
  - chromaticity
  - dual imaging filter parameters
  - detector related parameters (Flat Field)
  - to assess performance in terms of astrophysical parameters (Sp, distance, age, planet mass, atmospheric model)
  - to provide inputs to the Exposure Time Calculator

In Figure 3, we give an example of system analysis concerning the impact of offset pointing which demonstrates how accurate have to be the pointing in such a system (about 0.5mas). Figure 4 shows an example of detectability for an M0V star at 40pc. High order AO system combined with advanced coronagraphs, differential imaging and smart calibrations are needed to detect $1M_J$ young planets if specifications of system and sub-systems are met.

The originality of SPHERE is the combination of the 3 instruments to improve spectral coverage, and hence the ability to characterize exoplanets and also the chance of detection by reducing false alarms. Another peculiarity is the coronagraphic device, which actually includes several types of coronagraphs to take advantage of each concept. The main component will be an achromatic coronagraph to be used in the NIR survey (simultaneous observations with IRDIS in H and IFS in YJ). For that, we are developing prototypes of achromatic FQPM based on the use of half wave plates (Mawet et al. 2006) and Apodized Lyot coronagraph (Soummer 2005). The coronagraphic device will also include classical Lyot masks, and some achromatic coronagraphs optimized for

the H and K band, whether an Apodized Lyot or a FQPM/AGPM using Zero Order Gratings (Mawet et al. 2005a,b).

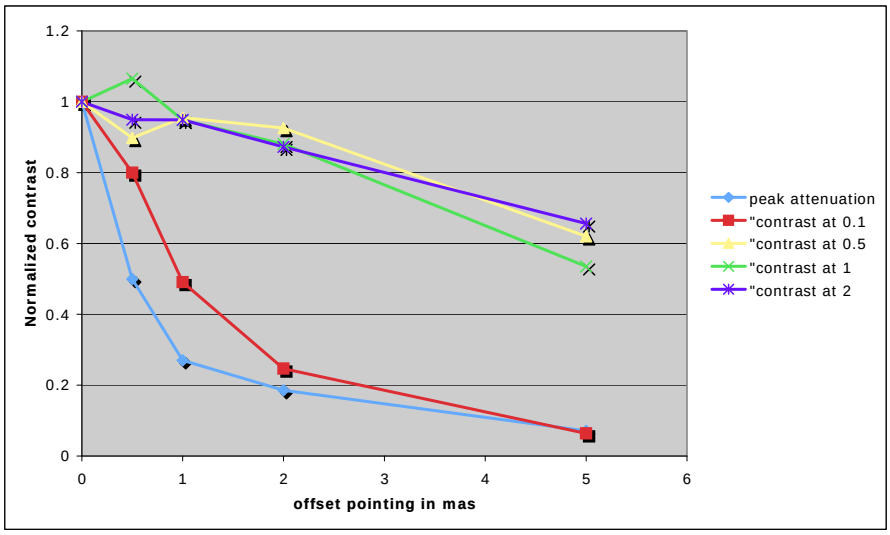


Figure 3: *contrast (with differential imaging) as a function of offset pointing (in mas) at different positions in the field.*

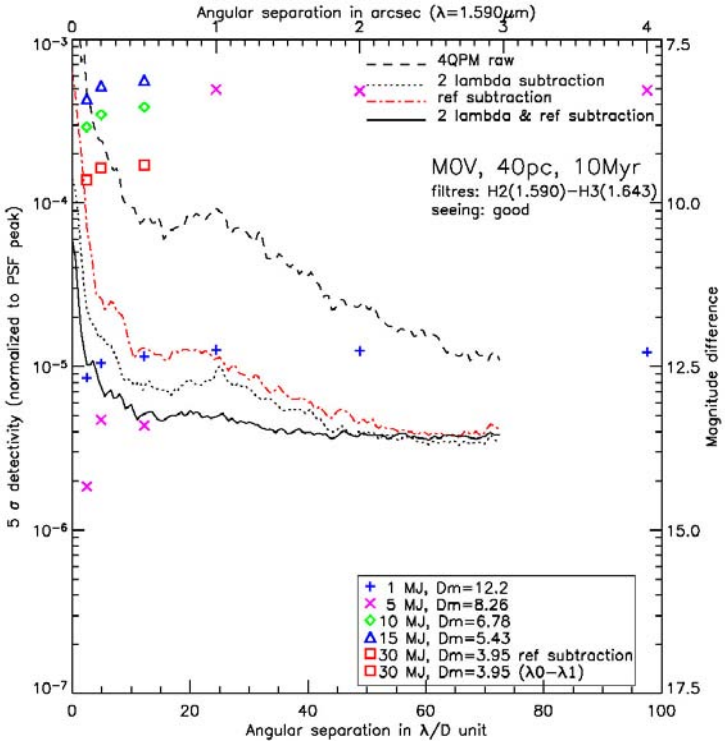


Figure 4: *Contrast at 5 sigma versus angular distance for the raw coronagraphic image (dash), the differential imaging (dot) and the double differential imaging (solid). Colored symbols are corresponding to different planetary mass.*

3. CONCLUSIONS

Although TPF-C has a much more challenging objective than MIRI and SPHERE, this project will benefit of the many developments we made on 1/ prototyping of coronagraphs, 2/ system analysis by numerical simulations and 3/ estimation of performance for ground-based as for space telescopes not only in terms of technique but also regarding scientific aspects (input catalog for instance). The FQPM is the only high contrast coronagraph manufactured and validated in the laboratory and on the sky at several spectral bands (see Baudoz et al. this proceeding). A new system will be implemented at the VLT by Feb. 2007 to take advantage of the Simultaneous Differential Imager (Lenzen et al. 2004, Biller et al. 2005).

In what concern estimation of performance or system analysis, we have developed a very complete model of SPHERE that could be easily turned to match TPF-C requirements. To carry on this study in a TPF context, a concrete collaboration between the TPF-C team and our group is desirable, not necessarily through ESA but rather with students and postdocs.

The selection of a coronagraph for TPF-C is definitely a puzzling issue. It is obvious that inner working angle and throughput are not sufficient parameters to do the choice in such a complex program. And, at some point, a thorough simulation including all limiting factors would be needed to perform the selection on the basis of signal to noise ratio estimate.

REFERENCES

Allard, F., et al. 2001, ApJ, 556, 357

Baudoz, P., Boccaletti, A., Riaud, P., Cavarroc, C., Baudrand, J., Reess, J.-M. & Rouan, D., 2006, PASP, 118, 765

Beuzit, J.-L., et al. 2006, The messenger 125, 29

Biller, B., Close, L., Lenzen, R., Brandner, W., McCarthy, D., & et al., 2005, IAU Colloq. 200

Boccaletti, A., Riaud, P., Baudoz, P., Baudrand, J., Rouan, D., Gratadour, D., Lacombe, F., 2004, PASP 116, 1061

Boccaletti, A., Baudoz, P., Baudrand, J., Reess, J. M., & Rouan, D., 2005, Adv. Space Res., 36, 1099

Boccaletti, A., Augereau, J.-C., Pantin, E., Baudoz, P., Riaud, P., 2007, in preparation

Carbillet, M., Verinaud, C., Guarracino, M., Fini, L., & et al., 2004, SPIE Proc. 5490, 637

Dubreuil, D., et al. 2003, Proc. SPIE, 4850, 564

Gardner et al. 2006, Space Science Reviews, in press

Gratadour, D., Rouan, D., Boccaletti, A., Riaud, P., Clénet, Y., 2005, A&A 429, 433

Lenzen, R., Close, L., Brandner, W., Hartung, M., & Biller, B., 2004, Proc. SPIE 5492

Mawet, D., Riaud, P., Surdej, J., & Baudrand, J., 2005,

Mawet, D., Riaud, P., Absil, O., & Surdej, J., 2005, ApJ 633, 1191

Mawet, D., Riaud, P., Baudrand, J., Baudoz, P., Boccaletti, A., Dupuis, O., & Rouan, D., 2006, A&A 448, 801

Rouan, D., Riaud, P., Boccaletti, A., Clénet, Y., & Labeyrie, A. 2000, PASP, 112, 1479

Soummer, R., 2005, ApJ 618, 161

Wright, G., et al. 2003, Proc. SPIE, 4850, 493

Coronagraphs on the Hubble Space Telescope

John Krist
Jet Propulsion Laboratory, California Institute of Technology

The Hubble Space Telescope (HST) is the most successful high-contrast astronomical imaging instrument that has yet been built (Krist 2004). It has provided numerous images of the scattered light from circumstellar disks and the glow from faint substellar companions. The key to its performance is the stability and high resolution provided by its placement above the atmosphere. Typically, on the ground, high-contrast observations are carried out using coronagraphs that suppress the diffracted light from the central source, some of which remains unsuppressed because of scatter introduced by the uncorrected turbulence of the atmosphere (which exists to some extent, even with adaptive optics, at levels equal to or greater than the scatter created by imperfections in the telescope optics). The residual light is difficult to subtract out because of time-dependent variations in the optics and atmosphere.

On Hubble, though, even without a coronagraph, the diffraction pattern and the scattered light from the optics created by a star can be largely subtracted out because the point spread function (PSF) does not vary greatly over time. In some cases, an image of another star taken months after the science observation can be subtracted out, for instance. Such direct (non-coronagraphic) HST observations have been made of many circumstellar disks and jets near stars.

Coronagraphs do, however, provide additional performance benefits to HST. Simply by occulting the star, scatter from optics after the occulter is reduced, as are detector artifacts such as saturated columns and electronic ghosts. With the addition of a Lyot stop, the diffraction pattern of a bright star can be suppressed to below the level of the scattered light from the optics (which cannot be suppressed optically without some sort of wavefront correction such as a deformable mirror). This reduces the effects of PSF instability caused by time-dependent aberration changes and color mismatches between the target and reference star PSFs, as well as reducing photon noise by diminishing the wings of the PSF.

HST has three coronagraphs that cover the near-UV to near-IR wavelength range, with varying levels of performance. While none of these coronagraphs could be considered to be anything close to optimal, they have provided high-contrast imaging that has not been matched by ground-based telescopes.

NICMOS (Near Infrared Camera and Multi-Object Spectrometer)

NICMOS provides imaging over the near-IR wavelength range (λ=0.95–2.4 μm) with three separate channels providing different resolutions (0.043″, 0.076″, and 0.22″ per pixel). Each channel uses a 256 × 256 element detector. The mid-resolution channel (NIC2) includes a coronagraphic capability (Schneider 2002) with an r=0.3″ (2.2 λ/D at λ=1.6 μm) occulting spot (actually a hole drilled in a field offset mirror) and Lyot stop. The occulting spot is located near the corner of the NIC2 field to minimize its impact on normal imaging (because it is always in place). The Lyot stop is actually the cold stop of the camera, which is intended to block emission from the warm obscurations of the telescope (primary edge, secondary mirror, spiders). Because the cold stop is always in place, throughput in normal imaging mode was optimized at the expense of diffraction suppression performance (i.e., the Lyot stop is not well matched to the rather small occulter). Unfortunately, deformation of the NICMOS dewar due to abnormal thermal expansion of the cryogen on-orbit resulted in a misalignment of the stop, so diffraction from the spiders and much of the primary and secondary mirror edges is not reduced. In addition, small instabilities in the Lyot stop alignment during the deformation period (the first couple of years on-orbit) caused significant time-dependent variations in the residual coronagraphic starlight pattern. Such variations are greatly reduced today (NICMOS now uses an external cooler that was installed in a later servicing mission).

The NICMOS coronagraph reduces the surface brightness of the PSF wings by a factor of 2–3 (a factor of two worse than if the Lyot stop were aligned). Because of scatter by the rough edge of the drilled occulter hole, the effective inner working radius is usually r>0.5″. Despite these limitations, the NICMOS coronagraph has been

used to view a number of circumstellar disks for the first time in scattered light (e.g., HD 141569, HD 32297). So far, ground-based coronagraphs operating in the near-IR have not been able to match its performance on such disks.

STIS (Space Telescope Imaging Spectrograph)

STIS is primarily a spectrograph with three separate channels: far-UV and near-UV using MAMA detectors and visible-light using a CCD (0.05″ per pixel). Its coronagraph was a (relatively) low-cost addition to the CCD channel (Grady et al. 2003). Its occulter consists of two crossed wedges (r>0.3″) and is moved into the beam on command (because of pointing and focus uncertainties, the effective inner working angle is r>0.7″ in most circumstances). The Lyot stop masks only the diffraction from the outer edge of the primary mirror and is always in place, reducing throughput by about 20%. Diffraction suppression was sacrificed for better throughput in non-coronagraphic observations, with the coronagraph providing a 2–6 times reduction in the PSF wing brightness. Much of this contrast improvement is gained by blocking the star image so that its light does not enter the STIS CCD, which internally scatters light to large angles at λ>0.7 μm (up to 30% at λ=1.0 μm). Besides the incomplete masking of the diffracting edges (secondary and spiders), the greatest limitation of the STIS spectrograph is that it can only be used in an unfiltered mode, covering the entire bandpass of the CCD (λ = 0.2–1.0 μm). This makes it particularly sensitive to color mismatches between target and reference PSFs.

The STIS coronagraph has been used to image a number of circumstellar disks, including Beta Pictoris, HD 141569, and HD 163296. Because of its higher resolution and more stable PSF, STIS has been able to view fainter disks than has been possible with the NICMOS coronagraph. At present, STIS is not functional due to an electronics failure, though it is scheduled to be repaired in the next HST servicing mission. For most coronagraphic imaging, however, STIS has been succeeded by the ACS camera.

ACS (Advanced Camera for Surveys)

ACS, the newest of HST's instruments, consists of a near-UV camera with a MAMA detector and wide-field (200″ × 200″, 0.05″/pix) and high-resolution, narrow-field (25″ × 25″, 0.025″/pix) visible-light CCD cameras. The high-res channel (HRC) includes a coronagraphic mode (Krist 2003) that allows imaging using a variety of filters over λ = 0.2–1.0 μm. The occulters and Lyot stop are located on a flap that flips into the beam on command. This allows the Lyot stop to be appropriately matched to the occulters without sacrificing throughput in the normal imaging mode. The Lyot stop thoroughly masks all of the diffracting edges in the system, reducing throughput by 52%.

The coronagraphic mode was added to the HRC after construction of the camera began, and it was not possible to place the occulters at an intermediate image plane after the spherically aberrated beam from HST was corrected. Thus, the occulters are located at a spherically-aberrated focus, and the beam is corrected afterwards (by a mirror at which the Lyot stop is located). This resulted in the need for large (hard-edged) spots of r=0.9″ and 1.8″ (17 and 35 λ/D at λ = 0.6 μm). While significantly impacting the inner working angle, the well-matched occulters and Lyot stop effectively reduce the brightness of the PSF wings by a factor of 6 – 7, to the point that the uncorrectable scattered light from the optics is greater than the residual diffraction pattern. This allows the ACS coronagraph to view the faintest disks yet seen in scattered light (e.g., Fomalhaut, HD 92945, and the outer regions of HD 141569). Despite its strong performance, however, it is not capable of imaging extrasolar planets in reflected light around even the nearest stars (Figure 2).

Performance Advantages and Limitations of the HST Coronagraphs

The HST coronagraphs have a number of advantages over ground-based systems. The primary one is that the stability of the HST allows subtraction of the residual light from the central source at greater levels than can be achieved on the ground. When HST is rolled between observations of the same star and the images from the two orientations are subtracted from each other (thus avoiding color differences between the target and reference PSFs), it is possible to reduce the residual light by up to factors of nearly 500. When another star is used as a reference PSF, improvements of 50× – 100× are often achieved. In comparison, a factor of 10 reduction is considered good on the ground.

HST also provides sharp images over the entire coronagraphic fields, while the corrected fields on ground-based systems are limited by the number of deformable mirror elements. Of course, HST also allows detailed high-contrast imaging at visible wavelengths where ground-based AO does not currently work.

While a powerful high-contrast tool, HST is still just a 2.4 meter telescope, and it can collect only so many photons from faint sources within a reasonable amount of time. In the near-IR, ground-based telescopes are beginning to approach the capabilities of NICMOS for substellar companion searches. The very large telescopes being proposed with extreme AO systems will be able to outperform HST in the near-IR on point sources (e.g., brown dwarf companions), but will still probably not achieve as high a contrast on disks due to optical instabilities in the regions around the star common to both space and ground.

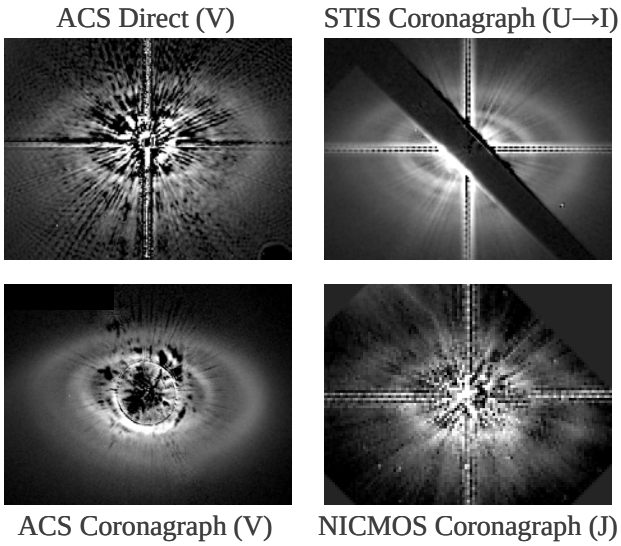


Figure 1: *HST images of the circumstellar disk around the Herbig Ae star HD 141569a. All of the images have been subtracted using images of reference PSF stars. The upper left image shows a non-coronagraphic image taken with ACS, and the others are coronagraphic images. The disk is ~7" across, and the total disk flux is about 0.02% of the total stellar flux. Subtractions by the author.*

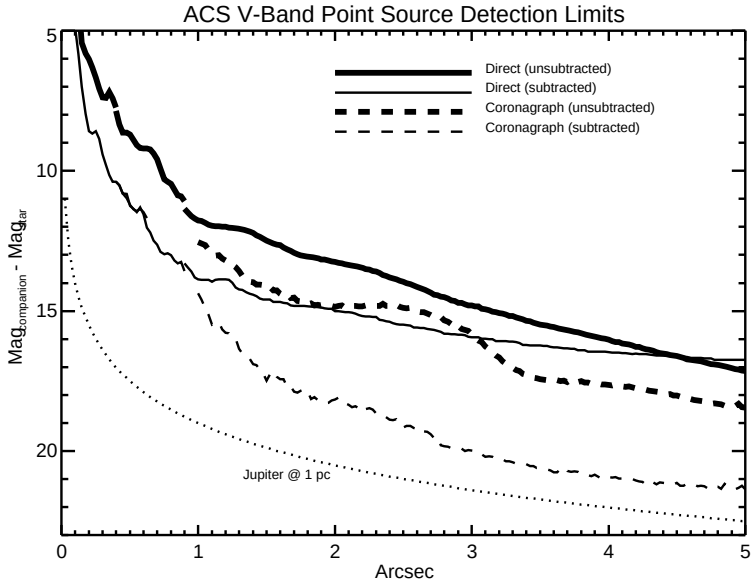


Figure 2: *Robust visual detection limits for a point source near a star in the ACS in V band, with and without the coronagraph and with and without subtraction of the stellar light (roll subtraction). Also plotted is the corresponding brightness of Jupiter as it would appear at various radii from a star located 1 parsec away.*

REFERENCES

Grady, C.A., et al., 2003, P.A.S.P., 115, 1036

Krist, J.E., et al., (2003) in Proceedings of the SPIE, 4860, 20

Krist, J.E., 2004, in Proceedings of the SPIE, 5487, 1284

Schneider, G., 2002, in Proceedings of the 2002 HST Calibration Workshop (available on-line at http://www.stsci.edu/hst/HST_overview/documents/calworkshop/workshop2002)

Recent Coronagraph Experimental Studies

Jun Nishikawa, Y. Sato, N. Murakami, N. Baba, Alexander Tavrov, Kaito Yokochi, Takashi Kurokawa, and Mitsuo Takeda

MIRA Project & Extra-Solar Planet Project Office, NAOJ, Hokkaido Univ., National Astronomical Observatory of Japan, Tokyo University of Agriculture and Technology, The University of Electro-Communications

ABSTRACT

Some activities on coronagraph experiments in Japan are reported. One is pre-optics for dynamic range absorption with nulling interferometer and deformable mirrors by Nishikawa et al. in NAOJ. Another is polarization differential imaging by Baba et al. in Hokkaido Univ. The third is common-path achromatic interferometer coronagraph by Tavrov et al. in NAOJ and TUAT.

Dynamic range absorption by nulling interferometer and AO

A new technique for very precise wavefront error reduction is now under study by Nishikawa et al.[1,2], and similar works are reported before[3-6]. The technique uses an unbalanced nulling interferometer (UNI) and phase and amplitude correction (PAC) by an AO system with a normal wavefront sensor and two deformable mirrors. The UNI+PAC is used as pre-optics of a coronagraph (Fig. 1). The pre-optics can absorb 10^2 dynamic range and the downstream coronagraph provides 10^8 for the total contrast which can be achieved by λ/1000 optics.

A brief description of the principle is as follows. The UNI for two beams provides some order of extinction of the central star light. Since the wavefront of the UNI output has still enough large amplitude without phase singularity and wavefront errors are artificially magnified by the extinction ratio, a normal wavefront sensor and two deformable mirrors can be applied for the PAC. After the wavefront error correction at the error-magnified stage, the wavefront error is the same level as the initial wavefront but the star light is reduced. In the image plane, after the UNI-PAC method is applied, the peak intensity of the central star is dimmed and speckle noise level is also reduced relative to off-axis planet intensity (Fig. 2).

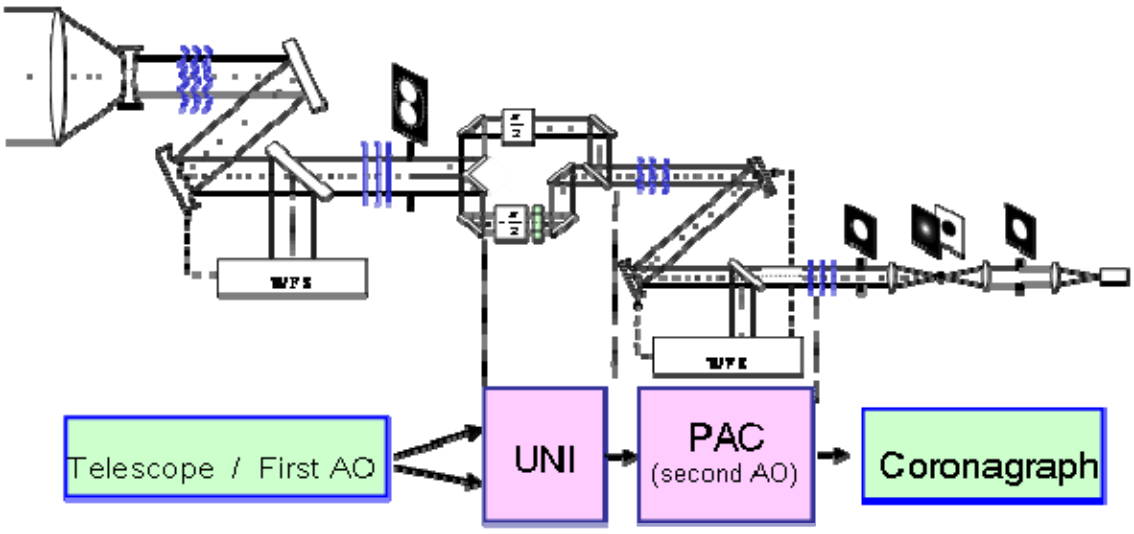


Figure 1. *A coronagraph configuration including the UNI+PAC method. The wavefront errors are magnified at UNI and corrected at PAC and its output light is led to a coronagraph.*

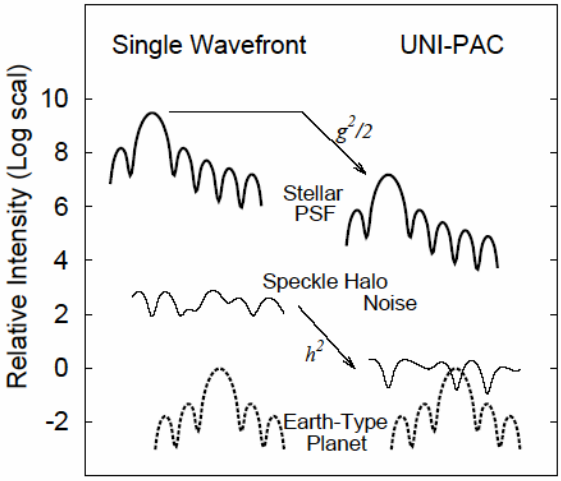
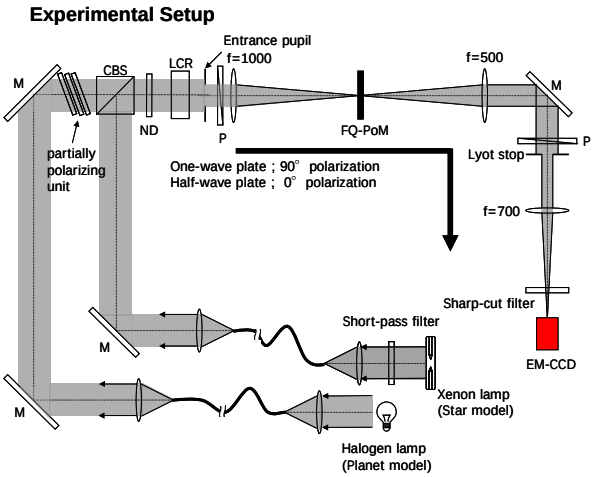


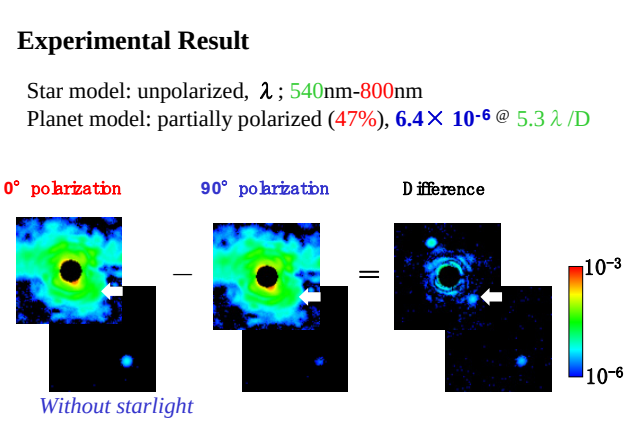
Figure 2. *Pre-optical effects by UNI+PAC in the image plane.*

The UNI-PAC method can be applied to two collimated beams at some styles of telescope configurations in front of various kinds of coronagraphs or interferometric nulling optics. Possible telescope architectures are: (a) Long Baseline Interferometer, (b) (Large) Single Telescope, Sub-aperture Interferometer, (c) Off-axis Single Telescope, Overlapped Sub-Aperture Interferometer, (d) Single Telescope, Rotation Shearing Interferometer. The multiple use of the UNI-PAC processes may be also possible for more than four telescope beams or VNC-like architecture. In this case the central star will be dimmed by 1E-4 but still the wavefront is compensated to the AO limit, where a coronagraph with a dynamic range of 1E6 would be used as a final stage. Note that we do not make detailed consideration on the use of various architectures, yet.

Polarization differential imaging with four-quadrant polarization mask

Baba et al. have been making experiments of some polarization applications with white light, a nulling interferometer using geometric (Pancharatnam) phase, a FQPM coronagraph with a liquid crystal, and so on[7–14]. A polarization differential imaging result is obtained by Sato, Murakami, and Baba, in which a partially polarized planet model light of 10^6 contrast at 5λ/D is observed using very wide band light of 540nm–800nm. The aim is to extract exoplanet image from speckle noise, and the method is a nulling stellar coronagraph with a four-quadrant polarization mask using a characteristics of the light from exoplanet is expected to be partially polarized, while speckle noises are caused by unpolarized stellar light.





Recent development of achromatic interfero-coronagraph for planet finder: practical and theoretical limits have become equal

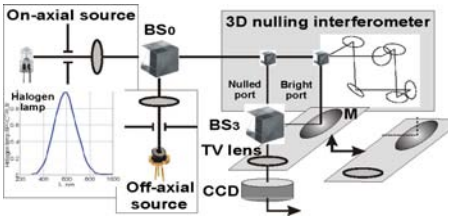


Figure 3. *General experiment schematic.*

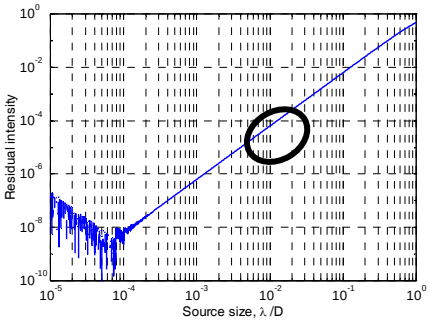


Figure 4. *Not nulled residual intensity. AIC scope in circle.*

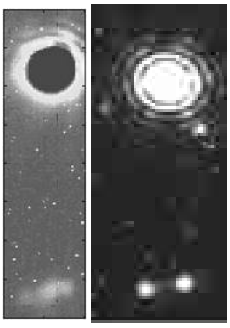


Figure 5. *Bright result up. Nulling results on bottom. (a) – 300 nm broadband light; (b) – He-Ne laser light.*

A three-dimensional common-path interferometer is proposed[15–17], which can achromatically null out an on-axis source, while maintaining the detectability of an off-axis source (Fig. 3). Geometric phase in the three-dimensional interferometer introduces an achromatic π-phase shift to the light from the on-axis source, such that destructive interference nulls out the axial light at one of the ports of the interferometer. Light from the off-axis source, which is exempt from the π -phase shift, comes out from both ports with equal intensity. The common-path scheme makes the system highly immune to environmental disturbances. In the described experiment, a 10^{-4} peak-to-peak nulling contrast was obtained for spatially unfiltered image and 10^{-5} with spatial filter.

The best approach to date is to interferometrically null the star's output while letting the planet's light shine through. We have developed an achromatic interferometer-coronagraph for extrasolar-planet detection that has a common path to allow stable operation under nonideal conditions[1]. The 180° rotational-shearing interferometer, which has a 3-D path and is a modified Sagnac interferometer, introduces a geometrical π phase shift for an on-axis source, nulling the source.

The stellar disc is resolved by a telescope as an incoherent extended source, and an incomplete nulling has a residual intensity versus the angular extend of the stellar disc, which dependence is shown in Figure 4.

By a laser light, Figure 5. it can be nulled out just perfectly being limited only by the instrument error, which we studied. In the *Nulled* area, shown in Figure 6, the residual light forms a partially developed speckle field. It has the peak-to-peak intensity of 9×10^{-5} in respect to to the *Bright* area intensity.

In Figure 6, the *Bright* signal is shown in 10^3 times attenuated by neutral density filter. We consider that the intensity modulation being observed as the partially developed speckles is originated from the wavefront modulations, which caused by the reflections on $_{BB}$ mirrors, whose actual surface figures depart from an ideal plane. Since the reflective surface figures depart from a perfect plane by the order of wavelength fractions, we have to consider this factor as the main instrumental error in the present demonstrator. Therefore, a spatial filter being implemented by a pinhole can sufficiently improve the nulling contrast, by our experiment down to 9×10^{-6}. The measured dependence of the nulling contrast on the optical axis tilt is shown in Figure 7. Hence all the optical methods are limited by the same kind aberrations caused by optics imperfection, new differential techniques improve the nulling contrast to fit the required range of 10^{-6} in infrared and to 10^{-9} in visible. These techniques have important advances if they are used with our achromatic coronagraph.

[1] A. Tavrov, Y. Kobayashi, Y. Tanaka, T. Shioda, Y. Otani, T. Kurokawa, and M. Takeda, "Common-path achromatic interferometer-coronagraph: nulling of polychromatic light" Opt. Lett. **30**, 2224, (2005).

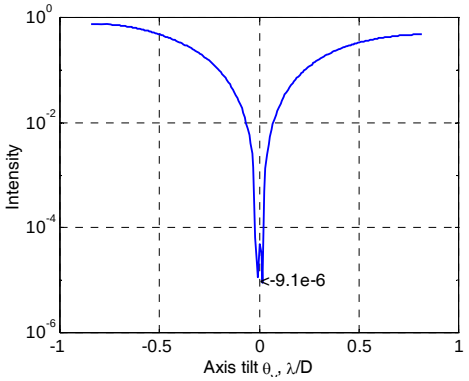


Figure 6. *Nulling in image plane.*

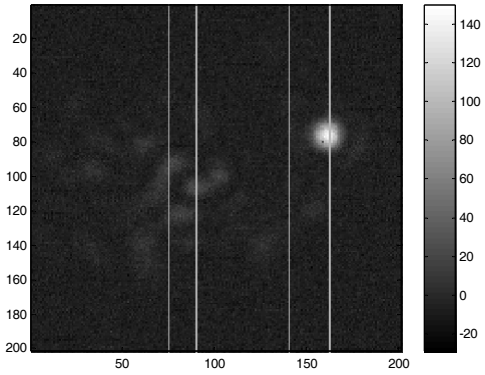


Figure 7. *Spatial filter used.*

REFERENCES

1. J. Nishikawa et al., A&A submitted (2006).
2. J. Nishikawa, N. Murakami, L. Abe, T. Kotani, and M. Tamura, "Nulling and adaptive optics for very high dynamic range coronagraph," (2006), Proc. SPIE, 6265, 3Q1-3Q5
3. L. Abe, J. Nishikawa, N. Murakami, and M. Tamura, "Removal of Central Obscuration and Spiders for Coronagraphy", (2006), Proc. SPIE, 6269, 501-5010
4. Abe, L., Murakami, N., Nishikawa, J., and Tamura, M., 2006, "Removal of central obscuration and spider arms effects with beam shaping coronagraphy", A&A , 451, p.363-373
5. J. Nishikawa, and N. Murakami, "Combination of nulling interferometer, nulling coronagraph, and modified pupil method", IAU Colloquium No.200, (2006) p.427-430
6. J. Nishikawa, T. Kotani, N. Murakami, N. Baba, Y. Itoh, and M. Tamura, 2005, "Combination of nulling interferometer and modified pupil for observations of exo-planets", A&A 435, 379–384.
7. N. Murakami, N. Baba, Y. Tate, Y. Sato, and M. Tamura, "Polarization Differential Objective Spectroscopy with a Nulling Coronagraph", (2006), PASP, 118, p.774-77
8. N. Murakami and N. Baba, "Polarization and Spectral Differential Imager using Channeled Spectrum", (2006), Proc. SPIE, 6265, 3T-1-3T-7
9. N. Murakami and N. Baba, "Pupil-remapping Mirrors for a Four-Quadrant Phase Mask Coronagraph", (2005), PASP, 117, p.295-299
10. N. Baba, N. Murakami, Y. Tate, Y. Sato, and M. Tamura, "Objective Spectrometer for Exoplanets based on Nulling Polarization Interferometry", (2005), Proc. SPIE, 5905, p.347-351
11. N. Murakami, T. Maeda, N. Baba, T. Ishigaki, N. Hashimoto, and M. Tamura, "Two-Channel Polarization Interferometric Coronagraph to Improve the Detectability of Extrasolar Planets," (2004), Proc. SPIE, Optical, Infrared, and Millimeter Space Telescopes, 5487, p.1368-1375
12. N. Baba and N. Murakami, "A Method to Image Extrasolar Planets with Polarized Light", (2003), PASP, 115, p.1363-1366
13. N. Murakami, Y. Kato, N. Baba, and T. Ishigaki, "Geometric Phase Modulation for the Separate Arms in Nulling Interferometer", (2004), Optics Communications, 237, p.9-15
14. N. Baba, N. Murakami, T. Ishigaki, and N. Hashimoto, "Polarization Interferometric Stellar Coronagraph," (2002), Optics Letters, 27, p.1373-1375
15. A. Tavrov, Y. Tanaka, T. Shioda, T. Kurokawa, and M. Takeda, 2005, "Achromatic coronagraph based on out-of-plane common-path nulling interferometer" SPIE 5905, 59051A-1-59051A-7.
16. A. V. Tavrov, Y. Tanaka, T. Shioda, T. Kurokawa, and M. Takeda, "Achromatic coronagraph based on out-of-plane common-path nulling interferometer," International Conference, New Frontiers in Stellar Interferometry, (Glasgow, UK) Ed. W. A. Traub, Proc. SPIE Vol. 5491, pp.824-830 (2004).
17. A. Tavrov, R. Bohr, M. Totzeck, H. Tiziani, and M. Takeda, "Achromatic nulling interferometer based on a geometric spin-redirection phase," Opt. Lett., Vol.27, No.23, pp.2070-2072 (2002,12).

Status of Development for the AIC and the FQPM

P. Baudoz and A. Boccaletti
LESIA, Observatoire de Paris

ABSTRACT

This paper presents the development of the Achromatic Interfero-Coronagraph (AIC) and the Four-Quadrant Phase Mask coronagraph (FQPM). Both coronagraphs have been tested in the lab and on ground-based telescope.

INTRODUCTION

Until about 10 years ago, only the classical Lyot coronagraph was used as a tool for high dynamical range imaging. In 1996, a new type of coronagraph, the Achromatic-Interfero Coronagraph (AIC), was proposed (Gay & Rabbia 1996). In 1997, Roddier & Roddier proposed to use focal phase plate as coronagraph. One of the most prolific phase mask coronagraph is the Four Quadrant Phase Mask (FQPM) proposed by Rouan et al. in 2000. Both AIC and FQPM have been tested on ground-based telescope in combination with Adaptive Optics system. The purpose of this paper is to recall the main results of these two coronagraphs, emphasizing on the laboratory and scientific results of the FQPM.

Achromatic Interfero Coronagraph

The Achromatic Interfero Coronagraph (AIC, Gay & Rabbia 1996, Baudoz et al, 2000a) was the first new type of coronagraph to be tested on a ground-based telescope (1.5 m telescope at the Observatoire de Haute Provence) combined with an adaptive optic (AO) system (Baudoz et al. 2000b). The test performed in the near Infrared (K band) demonstrated the principle of AIC (see Figure 1). The performance of the coronagraph was a attenuation of the star of about 50 mostly limited by residual atmospheric turbulence not corrected by AO. A detection of a faint component (difference of magnitude =3.5) was proven at a distance of only 1/3 of the telescope resolution (λ/D). This is the only coronagraph able to detect at such close angular distance to the star. Another prototype of AIC was installed on the 3.6 m Canada France Hawaii Telescope but the complexity of the optical train necessary to fit the CFHT beam, the mechanical instability of the Cassegrain focus, and the modest number of corrected mode of the AO system prevented the AIC to reach a better performance than at OHP. A new AIC that is less sensitive to the optical path difference is under development (Gay et al., private communication).

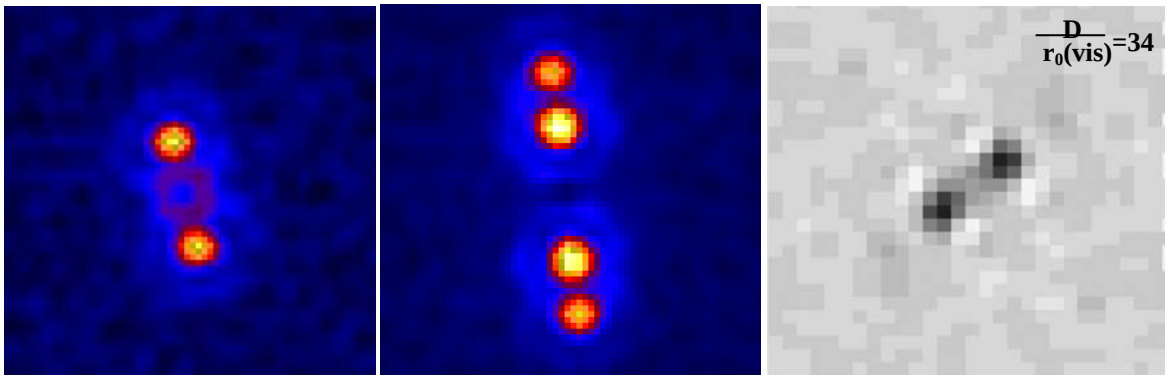


Figure 1: *Left: AIC image of a double star not centered on the coronagraph (low contrast between both star to show the effect of AIC). Middle: AIC image of the same double star centered on the coronagraph. Right: AIC image of a faint companion around a bright star. The difference of magnitude between the bright star and the faint component is 3.5 and the distance between the two star is only 1/3 of λ/D (see Baudoz et al. 2000b for more details).*

Four Quadrant Phase Mask Coronagraph

a. Visible

The first development and tests of the FQPM has been done in the visible range. Several techniques have been used to manufacture FQPM optimized for the visible wavelength. The simplest way to produce FQPM is to deposit an optical layer on a substrate of glass only on two opposite quadrants. The FQPM is then optimized for one wavelength. Riaud et al. 2003 have shown that contrast of 2.10^6 can be reached using a laser diode (wavelength adjustable between 630 and 640 nm, see Figure **2**). Several solutions have been studied to avoid the monochromaticity of a deposited layer. One of the solutions our group has implemented is the use of achromatic half-wave plates. Using 4 half-wave plates, 2 of them being rotating by 90° compared to the other is equivalent to get a FQPM achromatic (Mawet et al. 2006). The tests performed in the lab at visible wavelength have reached the performance (a contrast of 10^4 at 3 λ/D) expected from the theoretical point of view (see Figure **2**).

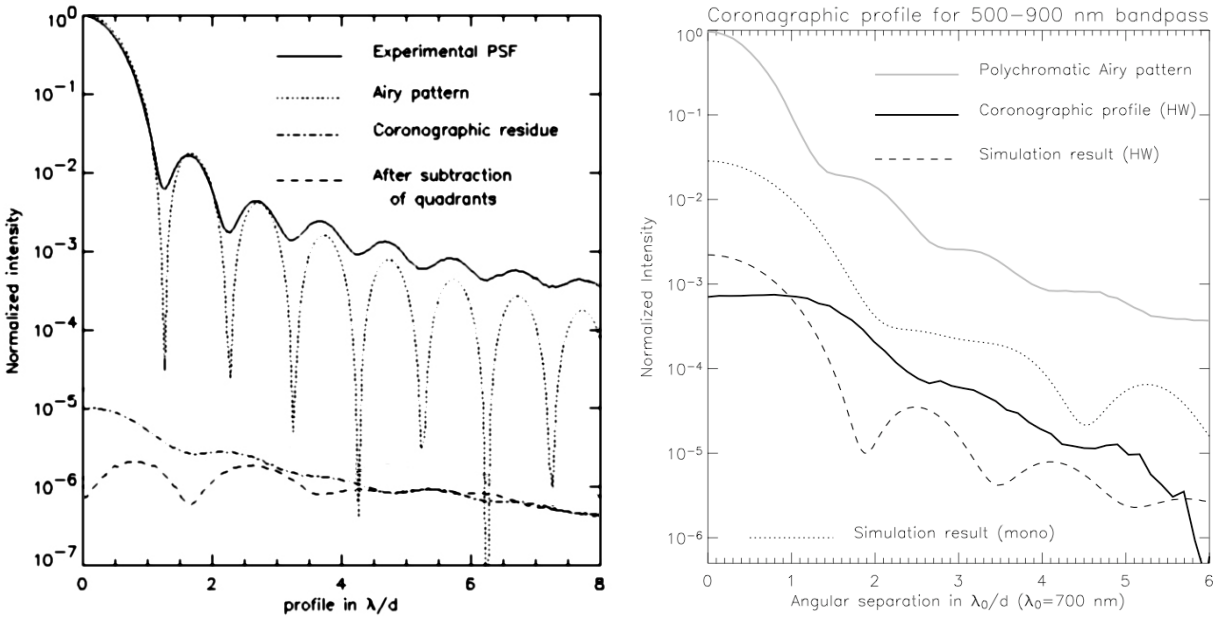


Figure 2: *Left: Laboratory tests of the FQPM in the visible and with monochromatic source (see Riaud et al. 2003 for more details). Right: Laboratory tests of achromatic FQPM with a 400 nm bandpass (see Mawet et al. 2006 for more details).*

b. Near-IR

While laboratory results in the near-IR reach the expected theoretical performance (Baudoz et al. 2004), the most interesting results in the near-IR are the scientific observation done on the VLT. Indeed, a FQPM has been installed inside the near-IR +AO camera NACO (Rousset et al. 2003) on UT4 at the VLT in 2003. Several scientific runs using this coronagraph have brought series of high dynamical range observation in different astrophysical areas. The tests performed on the sky have shown that a companion of a few 10^{-5} could be detected at 0.6″ (see Figure 3). Its high dynamical range combined with its small inner working angle enables it to reach information close to the star where no other instrument could observe at a given wavelength. This is the case for observation of the innermost part of the Beta Pictoris disk observed with the FQPM (Boccaletti et al. in preparation). This is also the case for the detection of the disk structure of the weak T Tauri star PDS 70 (Riaud et al. 2006). The FQPM also proved to be very useful for extragalactic observation because several structures that were never seen in the near-IR could be observed around NGC 1068 thanks to the FQPM (Gratadour et al. 2005, see Figure 4).

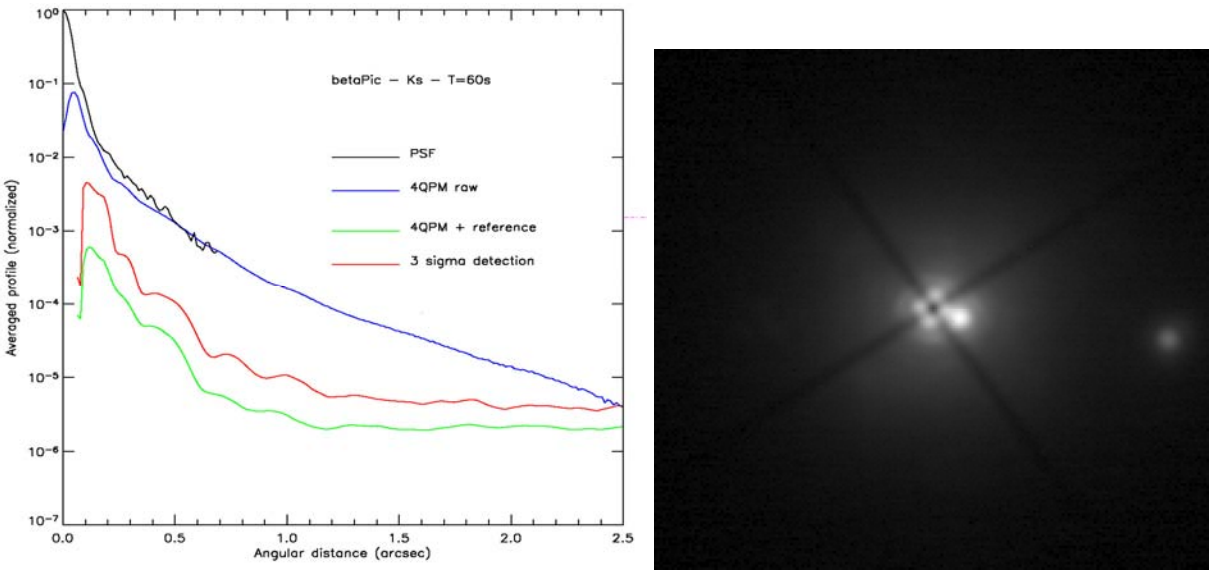


Figure 3: *Left: example of detection performance using the FQPM actually installed on VLT. The red curve shows the 3σ detection level (Boccaletti et al. in preparation). Right: Image of the visual triple star HIP 1306. The two faint components are respectively 1.6 and 3.5 magnitudes fainter than their star and located at distances of 0.13" and 1.0" (see Boccaletti et al. 2004)*

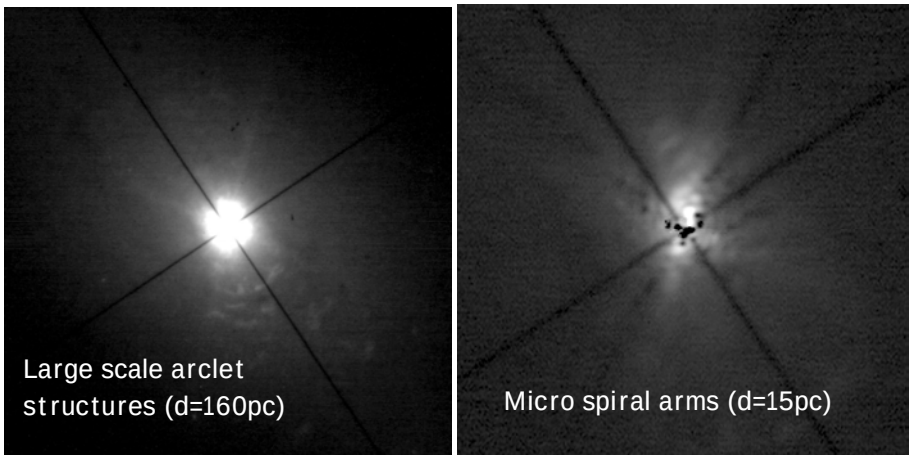


Figure 4: *The Active Galactic Nucleus NGC 1068 observed in Ks with FQPM on NACO. Left: FOV 11.7" where large scale arclet structure can be seen on the south side. Right: FOV 3.5" where we can see micro arms close to the center as well as aligned knots on the north side (see Gratadour et al. 2005 for more details).*

c. MID-IR

Our group is developing the three FQPM and the Lyot coronagraph that will take place in the Mid-IR Camera of the JWST (MIRI instrument). To validate the feasibility of such coronagraph, we manufactured coronagraph in the thermal infrared using different materials and fabrication techniques. We demonstrated that the behavior of the FQPM in the mid-IR and at very low temperature (12K) corresponds to the expected performance. The effective results of the FQPM in terms of contrast and attenuation (see Figure 5) are much better than what is specified for MIRI (Baudoz et al. 2006). In fact, the limitations from the James Webb Space Telescope (defocus, pupil shearing, jitter, aberrations, pupil geometry) are largely dominating over the intrinsic defects of the manufactured FQPM (see Figure 5). Thus, we can hope to reach the sensitivity estimated by simulations where the JWST limitations are taken into account. In that case, an Extrasolar Giant Planet orbiting at 10 AU around its parent star should be detected down to 400 K in one hour.

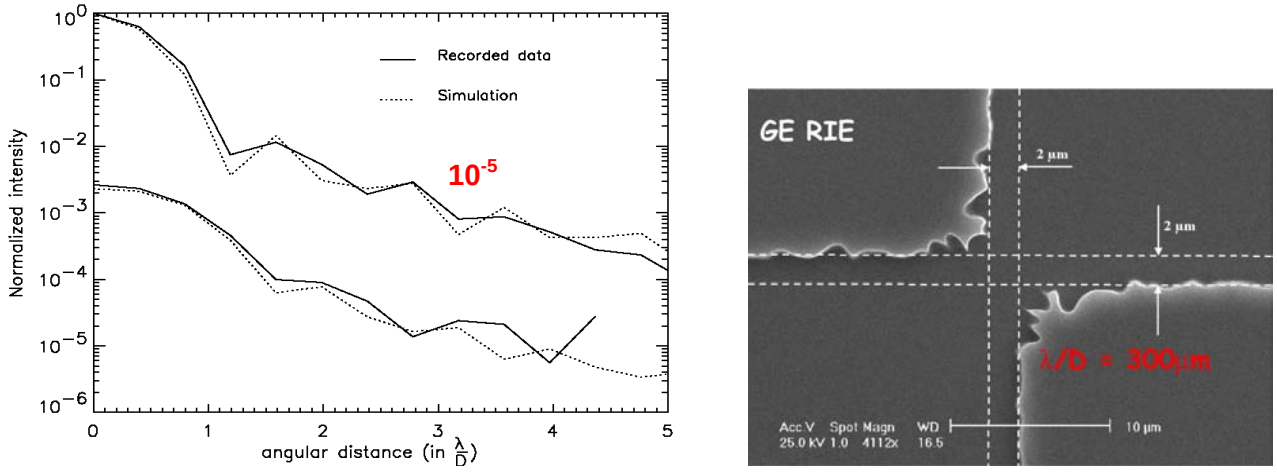


Figure 5: *Left: The solid lines show the radial profiles of the laboratory results with FQPM (lower curves) and without FQPM (upper curves). Dotted lines show the radial profiles calculated with the numerical code simulating the actual bench. Right: Electron microscope image of one of the FQPM manufactured out of Ge and using Reactive Ion-etching. The effect of the finite size of the transition between quadrants is negligible compared to other telescope limitations.*

CONCLUSIONS

The development of prototype of FQPM around the projects JWST and SPHERE has enabled the manufacturing of a number of different type of FQPM. The FQPM is the only coronagraph, aside of the classical Lyot coronagraph, that has been manufactured and validated in the laboratory and on the sky at so many spectral bands. The visible development of the FQPM shown that high rejection rate could be reached in monochromatic light. Achromatic FQPM were also tested in the visible using half-wave plates. In the mid-IR, the three FQPM that will take place in MIRI/JWST will mark a stone toward other coronagraph combined with space telescope. In the near-IR, the FQPM that has been installed on the VLT produced very promising results. A new system will be implemented at the VLT by Feb. 2007 and that will be combined to the Simultaneous Differential Imager (Biller et al. 2005). Even though the performance will be higher, it will still be limited by the quality of the correction of NAOS, the adaptive optics actually mounted on the VLT. To reach higher dynamical range, we are involved in the development of a 2nd generation ground-based instrument at the VLT. The instrument, called SPHERE (Beuzit et al. 2006), is probably the first high contrast imaging system to be fully optimized for the direct detection and characterization of exoplanets. Unlike all the coronagraphs that have been put on the sky before (AIC, FQPM), the design of this instrument is made altogether with the development and design of the coronagraph. Although TPF-C has a much challenging objective than SPHERE, this project could benefit of the many developments we made on the prototyping of coronagraphs and on the experience we gain at designing an instrument altogether with a coronagraph (system analysis, end-to-end simulation, …).

REFERENCES

Baudoz, P., Rabbia, Y., Gay, J., 2000a, A&AS, 141, 319

Baudoz, P., Rabbia, Y., Gay, J., Burg, R., Petro, L., et al., 2000b, A&AS, 145, 341

Baudoz, P., Boccaletti, A., Riaud, P., Rouan, D., Baudrand, J., Reess, J.-M., Parisot, J., 2004, SF2A : Scientific Highlights 2004, 175

Baudoz, P., Boccaletti, A., Riaud, P., Cavarroc, C., Baudrand, J., Reess, J.-M. & Rouan, D., 2006, PASP, 118, 765

Beuzit, J.-L., et al. 2006, The messenger 125, 29

Biller, B., Close, L., Lenzen, R., Brandner, W., McCarthy, D., & et al., 2005, IAU Colloq. 200

Boccaletti, A., Riaud, P., Baudoz, P., Baudrand, J., Rouan, D., Gratadour, D., Lacombe, F., 2004, PASP 116, 1061

Boccaletti, A., Baudoz, P., Baudrand, J., Reess, J. M., & Rouan, D., 2005, Adv. Space Res., 36, 1099

Boccaletti, A., Augereau, J.-C., Pantin, E., Baudoz, P., Riaud, P., 2007, in preparation

Gay, J., Rabbia, Y., Compte Rendus Academie des Sciences, 322, 625

Gratadour, D., Rouan, D., Boccaletti, A., Riaud, P., Clénet, Y., 2005, A&A 429, 433

Mawet, D., Riaud, P., Baudrand, J., Baudoz, P., Boccaletti, A., Dupuis, O., & Rouan, D., 2006, A&A 448, 801

Riaud, P., Mawet, D., Absil, O., Boccaletti, A., Baudoz, P., Herwats, E., Surdej, J., 2006, A&A, 458, 317

Rouan, D., Riaud, P., Boccaletti, A., Clénet, Y., & Labeyrie, A. 2000, PASP, 112, 1479

Rousset, G., et al. 2003, Proc. SPIE, 4839, 140

Lyot Project and Gemini Planet Imager

Rémi Soummer, Ben Oppenheimer, and the Lyot Project Team
American Museum of Natural History

Bruce Macintosh, and the GPI Team
Lawrence Livermore National Laboratory

In this communication we briefly present some results obtained with the Lyot Project Coronagraph (Fig.1), and analyze the current performance of this instrument. We present the Gemini Planet Imager (GPI) project which will be on sky in 2010.

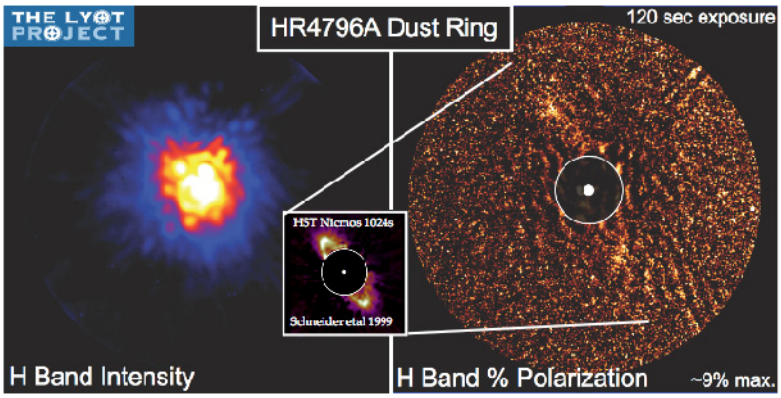


Figure 1: *Preliminary detection of the dust ring around HR 4796A. The left image shows the H-band coronagraphic image, and the right image shows the polarimetric image. For comparison, the HST image (Schneider et al. 1999) is shown. This detection was obtained in only 120 s, to be compared to the 1024 s of the HST exposure. This image was obtained during the commissioning of the polarimetric mode, with only 3 out of 4 modes operating. New data are being analyzed that utilize the full polarimetric mode and deeper exposures.*

With Extreme Adaptive Optics and Coronagraphy, the current on-sky performance limitations are due to quasi-static speckles with lifetimes up to one hour, which are not easily calibrated and do not average over long exposures (Figs. 2 and 3). Next-generation ground-based instruments will have to overcome this limitation using speckle reduction techniques. The Gemini Planet Imager will include an active calibration system, and a spectrograph (Integral Field Unit) to help push the dynamic range further using speckle nulling and multi-wavelength calibration techniques (Fig. 4).

Ground-based projects like GPI and its European counterpart VLT-Sphere will be capable of observing a large target sample in young associations and the solar neighborhood. Detection rates based on Monte-Carlo simulations predict a detection of about 100 planets for a survey of about 500 stars (Fig. 5). Imaging and spectroscopy of giant planets, and polarimetric imaging of circumstellar dust disks, will help the understanding of planetary formation and evolution.

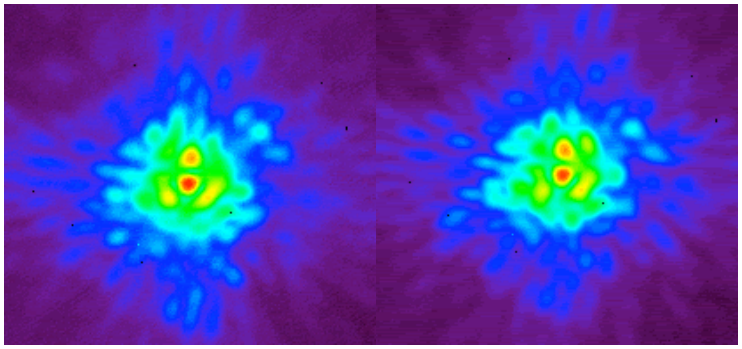


Figure 2: *Illustration of quasi-static speckles on H-band coronagraphic short exposure images. The two images are separated by a few minutes. Several quasi-static structures can be identified on both images.*

On the instrumentation side, the presence of these slowly varying speckles in Extreme AO coronagraphic images makes the case of ground-based coronagraphy relevant for TPF. Although the speckle parameters vary by orders of magnitude between ground and space, ground-based instruments like GPI will implement speckle nulling and multi-wavelength calibration techniques. This will provide useful information and results for the preparation of TPF, in particular on the speckle reduction and calibration aspects for increasing the dynamic range

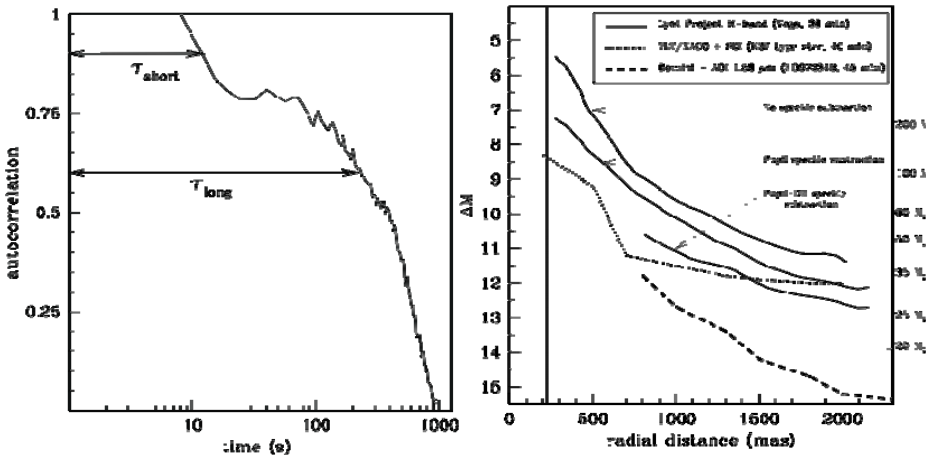


Figure 3: *Left: autocorrelation of the residual speckle in Lyot Project coronagraphic images, showing the short and long lifetimes. The long-lived speckles do not average over long integration times. Right: Lyot Project contrast performance obtained on Vega. Differential rotations at the Coudé focus have been used to remove some of the static speckles, resulting in an improvement of about 2 magnitudes in contrast.*

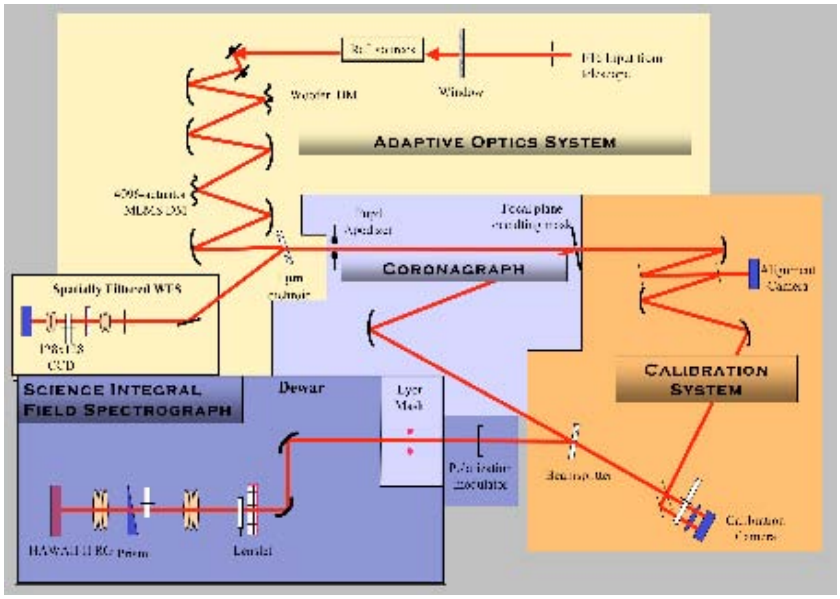


Figure 4: *Schematic layout of the Gemini Planet Imager. The system will include an Extreme Adaptive Optics system, an Apodized Pupil Lyot Coronagraph, a calibration system for quasi-static speckle reduction and an integral field spectrograph.*

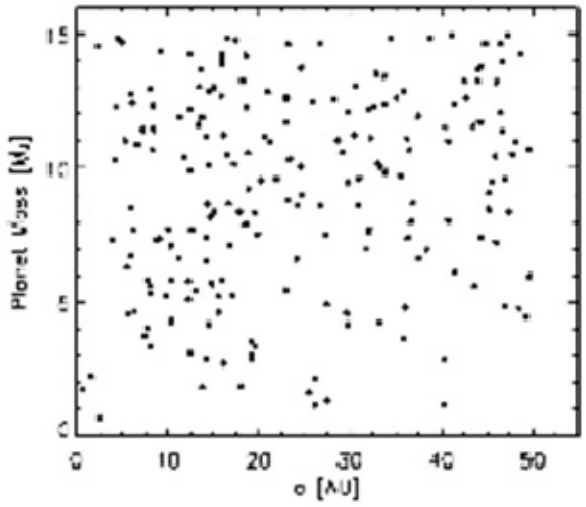


Figure 5: *The distribution of GPI-detected exoplanets in the semimajor axis/exoplanet mass plane. The detected planets are drawn from the field survey of nearby (<50 pc) stars (no age cut). This experiment samples semimajor axis and masses with uniformity*

The Near-Infrared Coronagraphic Imager

Christ Ftaclas, Mark Chun and Mike Liu
Institute for Astronomy, University of Hawaii

Douglas Toomey
Mauna Kea Infrared

INTRODUCTION

The Near Infrared Coronagraphic Imager (NICI) is a facility instrument built for Gemini South by Mauna Kea Infrared, in Hilo Hawaii with an 85-element curvature AO system provided by the Institute for Astronomy (IfA). The PI is Doug Toomey (MKIR), the author is the Project Scientist, and Mark Chun (IfA) has been the lead on the AO system. The dual-channel concept for NICI derives from experience gained at NASA's IRTF with CoCo, where we found that frequent beam switching between target and reference stars and imaging in and out of Methane bands were effective in detecting faint cold objects. The plan for NICI is to use spectral discrimination to detect young gas giant planets by accumulating images in two different spectral channels at the same time.

INSTRUMENT DESIGN

NICI is fundamentally a coronagraphic camera and it has been optimized to attack the circumstellar imaging problem. The instrument has a built-in, dedicated, 85-channel, curvature adaptive optics system. The basic layout is in Figure 1.

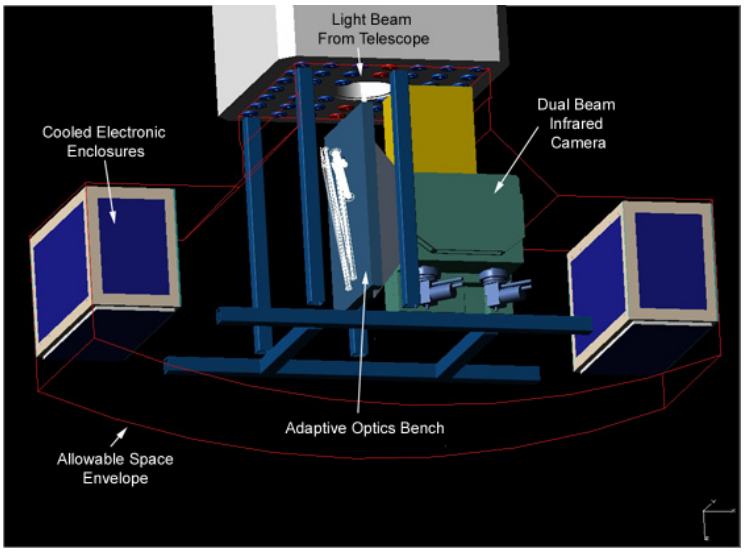


Figure 1: *NICI Instrument Layout. The telescope pupil is imaged on to the deformable mirror and then refocused. Prior to focus the beam is divided into science and wave front sensor light. The WFS light is sent through the optical bench to the WFS and the science light comes to focus on the warm occulting mask just ahead of the camera dewar window.*

Inside the dewar, the pupil is reimaged and a selection of Lyot stops is available. After the Lyot stop, the beam is divided between NICI's two channels. The design permits several options for doing this; at present only a 50-50 beam split has been implemented. The Lyot stop, beam splitting, and channel filter wheels are shown below. This resulted in detection capability plots like those shown in Figure 2.

Each channel is imaged onto an Aladdin 1KX1K array.

Features

In addition to its dedicated AO system and dual channel camera NICI includes several interesting features to either simplify design or enhance performance. These include:

a. A warm occulting mask. Experiments at the IRTF showed there was little penalty for doing this and it allow easy inclusion of user or specialty masks.
b. Halftone, semi-transparent apodized occulting masks. These masks have a central landing with neutral density ~3 to permit seeing the occulted star for alignment purposes.

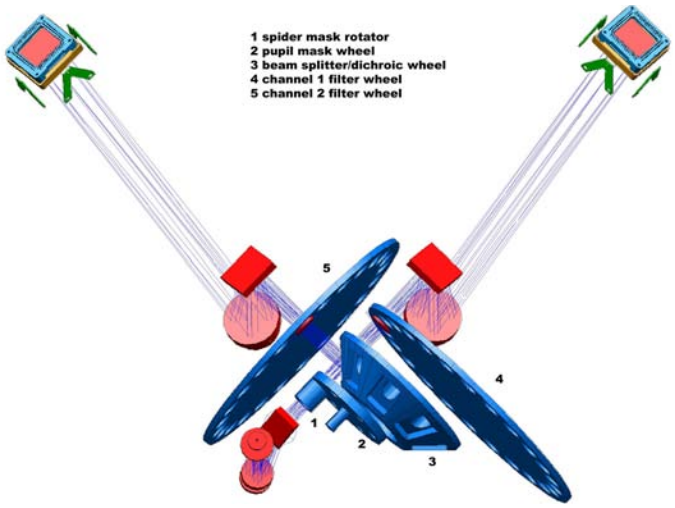


Figure 2: *The NICI Pupil Mask/Beam Division Assemblies. Options are available for a variety of pupil masks, beam splitting strategies and individual channel filters.*

c. A rotating spider mask. This mask is permanently in place and also includes a baseline inner Lyot stop.
d. An apodized pupil mask. This is used without an occulting mask to get low halo images in crowded fields.
e. Multiple channel filters. Each channel has a 25 position filter wheel which at present have IR band pass filters as well as 1% and 4% narrow band spectral differencing filters.
f. Detector mapping masks. Pinhole masks at the first focus permit imaging identical spot patterns onto both detectors to allow accurate mapping of the detectors onto one another.

The NICI optical system was designed and specified with two simple goals: that the planet image would have the highest possible Strehl, and that even for the best possible atmospheric conditions, the final focal plane would always be turbulence dominated. Meeting these goals required constraining both the magnitude and distribution of optical surface errors. We feel we have succeeded in meeting these goals as evidenced by the static image quality data shown in Figure 3.

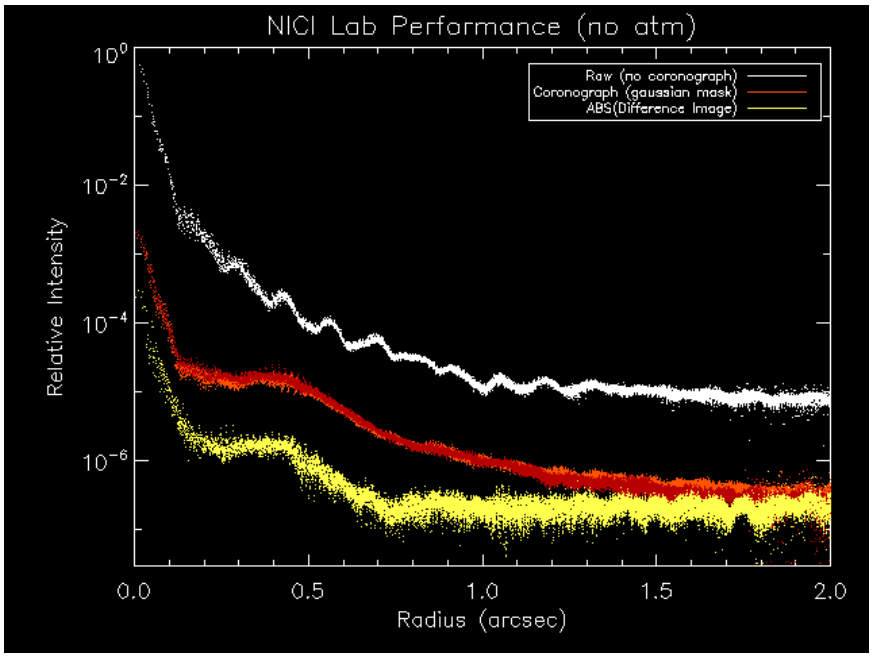


Figure 3: *NICI Static Image Quality. The plots above show the no coronagraph image profile, the post coronagraph image profile for each channel (red and orange traces) and finally the profile of the difference image. All images were taken with the Lyot stop in place and occulting mask was moved in and out of the camera. The difference images were clearly noise dominated and longer exposures will be required to get firm limits on the frame differencing limits.*

OBSERVATIONAL OUTLOOK.

Gemini has decided to use NICI initially in a campaign mode. To prepare for this campaign NICI performance was modeled by coupling planetary and brown dwarf spectra with AO and coronagraph simulations. This resulted in detection capability plots like those shown in Figure 4. They show that NICI imaging performance clearly compliments radial velocity detections. NICI has gone through acceptance testing and is currently en route to Gemini South. It will be commissioned in early 2007 and the planet detection campaign should begin later in the year.

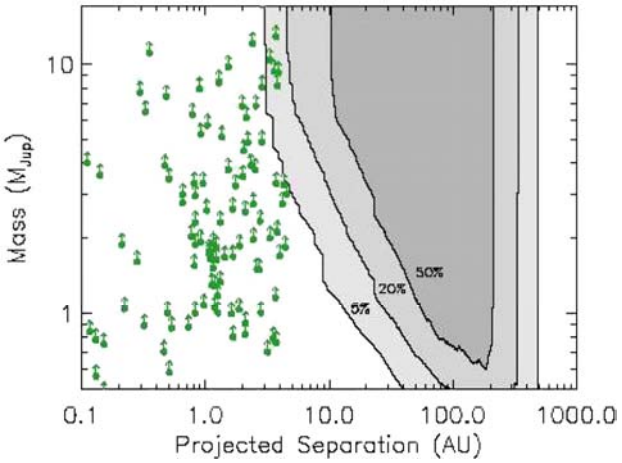


Figure 4: *NICI Planet Detection Performance (from M. Liu et al., 2005)—Sensitivity for our NICI survey compared with RV searches. The shaded areas of 5, 20, and 50% show the {mass, separation} regime imaged for our targets, i.e.,* ***at ≈10–30 AU projected separations, 20% of our sample will be surveyed to ≈1 M_{Jup} and most stars to 3 M_{Jup}.*** *RV planets around old Sun-like stars are shown as open circles with arrows. The inner edge of NICI detectability is due to its focal plane mask. (The detection floor at 0.5 M_{Jup} represents the low-mass limit of the theoretical models available for these calculations.)*

D.W. Toomey, C. Ftaclas, R.H. Brown, D. Trilling, *Proc. SPIE* Vol. **3354**, p. 782-790, IR Astron. Instrum., Albert M. Fowler; Ed.(1998)

High Contrast Coronagraphy and Extreme Adaptive Optics Experiments at Palomar

E. Serabyn, P. Haguenauer, B. Mennesson, J. K. Wallace, M. Troy, and E. E. Bloemhof

Jet Propulsion Laboratory, California Institute of Technology,

ABSTRACT

The next generation of adaptive optics (AO) systems, often referred to as extreme adaptive optics (ExAO), will use higher numbers of actuators to achieve wavefront correction levels better than 100 nm, and so enable a variety of new observations, such as high-contrast coronagraphy. However, the number of potential coronagraph types is increasing, and selection of the most promising coronagraph is subject to many factors. Here it is pointed out that experiments in the ExAO regime can be initiated with existing AO systems, by correcting a subaperture on an existing telescope rather than the full pupil. With a 1.5-m diameter well-corrected subaperture (WCS) on the Palomar Hale telescope, we have recently achieved stellar Strehl ratios as high as 92% to 94% in the infrared, corresponding to wavefront errors of 85–100 nm, and have carried out visible wavelength AO observations. With a WCS, a wide variety of ExAO experiments can thus be carried out immediately, including infrared ExAO imaging and performance optimization, a comparison of coronagraphic approaches in the ExAO regime, and visible wavelength AO.

INTRODUCTION

The direct detection of very faint companions close to much brighter stars requires very high-contrast observations at small angular separations, and a number of novel coronagraphs have recently been proposed to meet this requirement (e.g., Guyon, these proc.) The main limitation to high contrast coronagraphic observations on ground-based telescopes is the wavefront quality achievable after correction by the AO system, as scattering from both phase errors and pupil obstructions degrade the image quality. In particular, classical coronagraphs provide significant contrast improvement1 only for Strehl ratios above about 90%, a level not yet available with the current generation of AO systems. Thus, it had been thought that very high contrast coronagraphy might need to await the arrival of next-generation ExAO systems.

On the other hand, a more highly corrected wavefront can already be obtained by using an existing AO system to more finely correct a sub-aperture smaller than the full telescope pupil. Vignetting can also be eliminated by means of a clear off-axis sub-aperture. With such a clear off-axis, well-corrected sub-aperture (WCS) a dual image improvement is thus obtained: scattering due to both wavefront aberrations and pupil blockages is eliminated. Figure 1 shows the expected Strehl ratio improvement for the case of the Palomar 200-inch telescope, while Figure 2 shows potential Strehl ratio vs. wavelength. Evidently infrared Strehl ratios exceeding 90%, and visible Strehl ratios of order 50% are achievable with a WCS, implying that a WCS can be employed to immediately carry out ExAO-level experiments such as high-Strehl infrared imaging, high-contrast coronagraphy, and visible AO. Given this promise, we have recently implemented a WCS with a diameter of 1.5 m on the Hale telescope, using an appropriate set of relay optic[2, 3, 4].

OBSERVATIONAL RESULTS

Several initial ExAO experiments have now been carried out with our WCS at Palomar. A few examples are presented here.

1. **High Strehl infrared observations:** Figure 3 shows an image of the single star HD 121107 obtained with our WCS. The image closely resembles a diffraction-limited Airy pattern. Strehl ratios as good as 92–94% were obtained, corresponding to wavefront errors in the 85 – 100 nm range.
2. **Visible AO:** Figure 4 shows a B-band image, obtained with our WCS, of a binary star with a separation of 0.34 arc sec. Our WCS clearly provides good AO correction even at blue wavelengths.

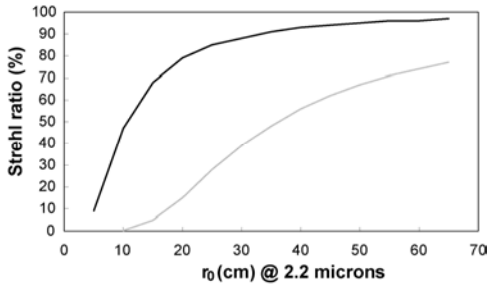


Figure 1: *Strehl ratios obtainable with the full Palomar 200 inch telescope (bottom curve) and a sub-aperture of diameter 1.5 m (top curve) vs. r_o, the atmospheric seeing cell size.*

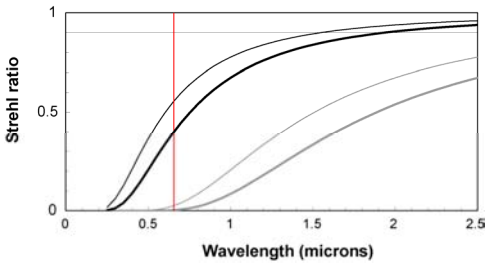


Figure 2: *Strehl ratios obtainable with a WCS for rms wavefront accuracies of 80, 100, 200 and 250 microns (curves, top to bottom). The Hα wavelength is marked by the vertical red line, and the dotted horizontal line is at a Strehl ratio of 90%, above which coronagraphic contrast improvement begins to become significant.*

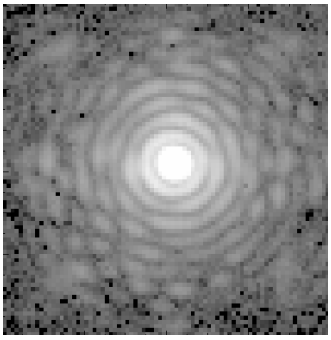


Figure 3: *Image of HD 121107 with our WCS in the 2.17 mu Br γ filter.*

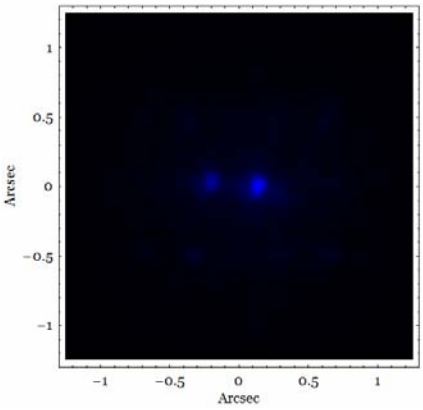


Figure 4: *B band AO-corrected image of SAO 37735 through our WCS.*

3. **Coronagraphy:** We also carried out initial infrared coronagraphic observations with our WCS using a four-quadrant phase mask (FQPM) coronagraph[5,6], because this type of coronagraph calls for an unobscured circular aperture for high rejection. Figure 5 shows our observations of the binary star HD 148112: the left-hand panel shows the normal "off-coronagraph" image, the central panel shows the "through-FQPM" image, and the right-hand panel shows the through-FQPM image after cross-diagonal subtraction. The coronagraph allows the fainter companion at ≈ 2λ/D to emerge as the brightest source in the field by a factor of 10.

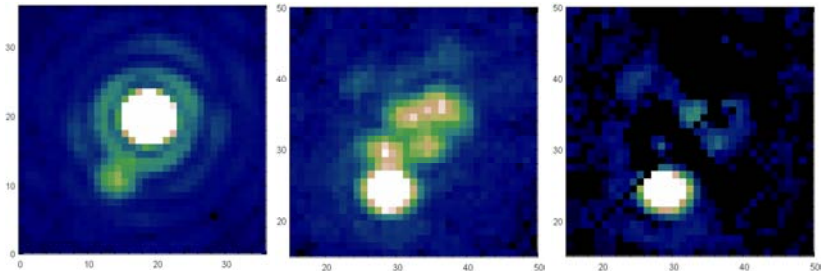


Figure 5: *WCS observations of the binary HD 148112 in the Br γ filter. Left: normal (no FQPM) image. Center: through-FQPM image. Right: cross-diagonal subtraction of FQPM image. The highest residual peak is down by 235:1.*

SUMMARY

With a WCS of diameter 1.5 m on Palomar's Hale telescope, we have been able to carry out several ExAO-level experiments, including high-Strehl IR imaging, high stellar rejec–tion with an infrared coronagraph, and visible AO. Many other experiments are now also possible, such as dark hole generation, implementation of other types of coronagraph, and wave front sensor optimization in the ExAO regime. Of course, implementing such a system on a larger telescope is also an option.

ACKNOWLEDGEMENTS

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Observations at the Palomar Observatory were made as part of a continuing collaborative agreement between Palomar Observatory and the Jet Propulsion Laboratory.

We wish to thank Rick Burruss and Jeff Hickey and the entire staff of the Palomar 200-inch telescope for their assistance at the Observatory, as well as the JPL Research and Technology Development program for supporting this work.

REFERENCES

1. A.Sivaramakrishnan, C. Koresko, R.B. Makidon, T. Berkefeld and M.J. Kuchner, *Ap. J.* 552, 397 (2001).
2. P. Haguenauer et al., *Proc. SPIE* 5905, 261 (2005).
3. P. Haguenauer et al., *Proc. SPIE* 6265, 62651G (2006).
4. E. Serabyn et al., Proc. SPIE 6272, 62722W (2006).
5. D. Rouan, P. Riaud, A. Bocaletti, Y. Clenet and A. Labeyrie, *PASP* 112, 1479 (2000).
6. P. Riaud, A. Boccaleti, D. Rouan, F. Lemarquis and A. Labeyrie, 2001, *PASP* 113, 1145 (2001).

High Contrast Imaging with Focal Plane Wavefront Sensing and PIAA for Subaru Telescope

Olivier Guyon, Hideki Takami, Eugene Pluzhnik, Shinishiro Tanaka, Basile Gallet, and Stephen Colley
Subaru Telescope

Motohide Tamura and Lyu Abe
National Astronomical Observatory of Japan

ABSTRACT

We describe in this paper the plans for a high contrast imaging system to be deployed on Subaru Telescope. This system makes efficient use of both the new curvature AO systems at Subaru which recently had first light, and the HiCIAO camera, which will be completed in 2007.

The proposed system will use “focal plane wavefront sensing” (FPWFS), a scheme where focal plane images are used to measure the wavefront. It will also use the highly efficient PIAA coronagraph. We show that these choices will result in a highly efficient high-contrast imaging instrument. Key techniques for this project are being demonstrated in our high contrast testbed at Subaru Telescope.

OVERALL DESIGN

The overall design for this project is shown in Figure 1. The proposed system takes full advantage of a very favorable environment for our high contrast imaging project:

– AO188 is a high performance AO system, delivering good quality images with approximately SR=0.5 in H band. It is a very good first stage AO system for our project.

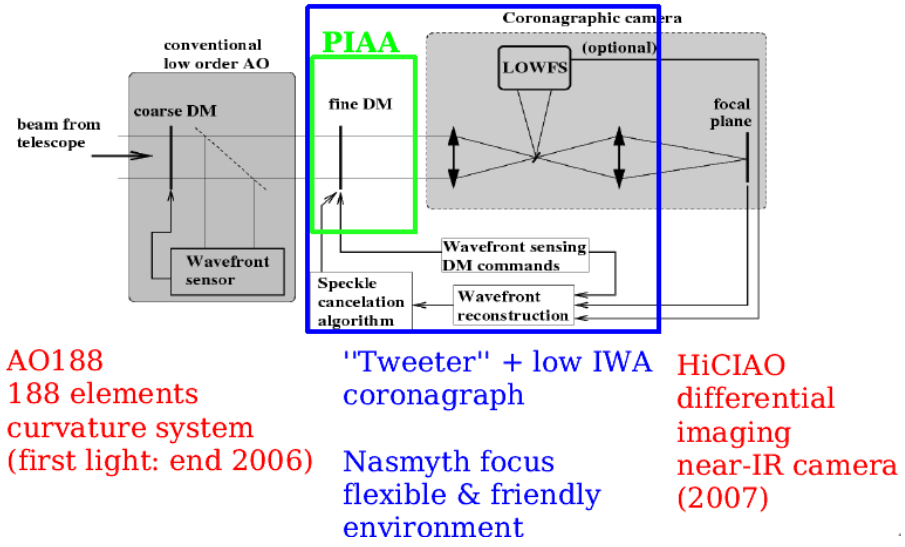


Figure 1: *Block diagram of the high contrast imaging system for Subaru Telescope. Both the AO188 system and HiCIAO are well advanced projects. Our goal is to include between HiCIAO and AO188 a PIAA coronagraph equipped with a deformable mirror (DM).*

– HiCIAO is an imaging camera specifically designed for high contrast coronagraphic imaging. It also offers simultaneous differential imaging around the methane absorption feature in H band. It has been designed to allow new coronagraph systems to be tested, and is therefore very flexible.

– Both HiCIAO and AO188 are at a Nasmyth focus, which offers the space and flexibility to easily deploy new instruments.
– We are acquiring unique experience in PIAA coronagraphy and focal plane wavefront control from our laboratory testbed at Subaru Telescope.

CORONAGRAPH: PHASE-INDUCED AMPLITUDE APODIZATION (PIAA)

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 is then yields a high contrast PSF in which starlight can be removed by a small focal plane occulting disk.

A schematic representation of the PIAAC is shown in Figure 2, 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).

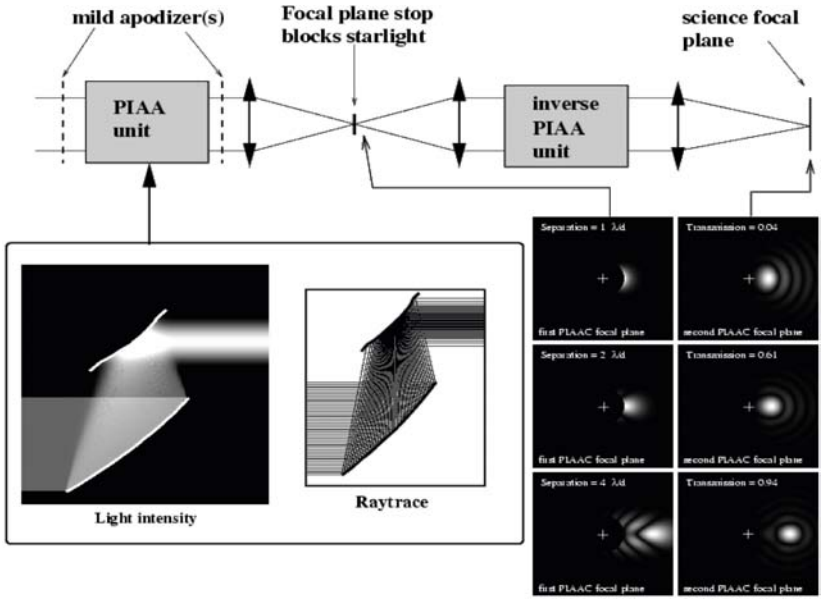


Figure 2: *Schematic representation of the PIAAC. The telescope light beam enters from the left and is first apodized by the PIAA unit. Mild apodizer(s) are used to perform a small part of the apodizations, and are essential to mitigate chromatic diffraction propagation effects and to allow for the design of "friendly" aspheric PIAA microns. An high contrast image is then formed, allowing starlight to be removed by a small occulter. An inverse PIAA unit is required to "sharpen" the image of off-axis sources.*

The PIAAC coronagraph performance is quantified and compared with other coronagraphs in the "Theoretical analysis of coronagraphs" paper in this volume. We summarize here the main characteristics of the coronagraph:

- nearly 100% throughput
- 1 λ/d angular resolution
- ~1 λ/d IWA
- full 360 deg search area
- good achromaticity

In the PIAAC design adopted for this project, the aspheric optics deliver the apodized entrance beam to a Lyot-type coronagraph. While some of the starlight is removed by the focal plane occulter, the Lyot pupil plane mask also contributes to starlight rejection. This design allows a smaller IWA (see Fig. 3) than for the more conventional PIAAC design considered for higher contrast space use (see the paper about PIAAC in this volume).

The PIAAC system optics are also designed to work with the Subaru Telescope pupil central obstruction and spider vanes. First, a **Spider Removal Plate** (SRP) removes the spider vanes from the pupil. This is achieved by translating 4 segments of the pupil closer together—this translation is obtained by a tilt in a transmissive glass plate, and the SRP is therefore a rooftop-shaped plate of constant thickness. The PIAA optics, chosen to be lenses for this design, can then remove the central obstruction while performing the apodization.

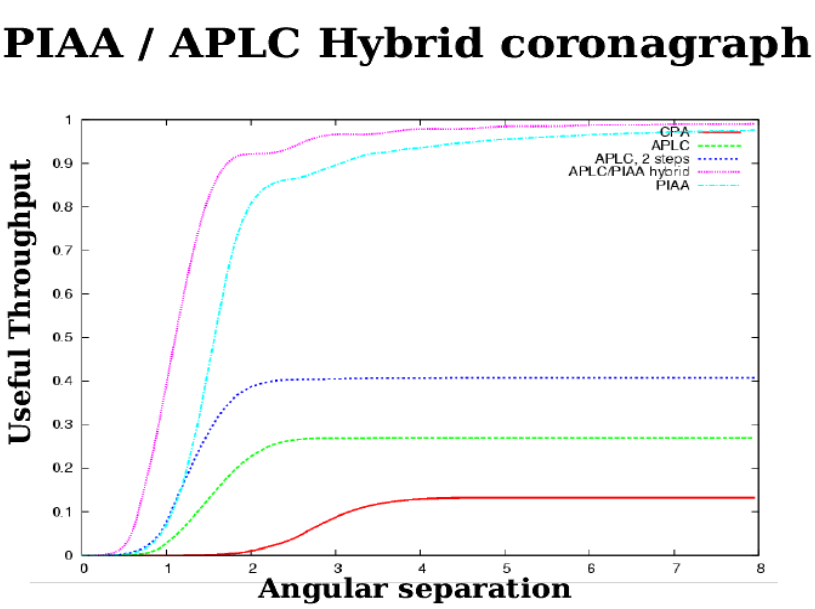


Figure 3: *Comparison between Conventional Pupil Apodization (CPA), Apodized Pupil Lyot Coronagraph (APLC), conventional PIAA (PIAA) and the hybrid solution adopted for the Subaru high contrast imaging system (APLC/PIAA hybrid). The APLC/PIAA hybrid solution offers improved IWA and throughput.*

WAVEFRONT CONTROL

The focal plane images acquired by the science camera (HiCIAO) will be used to estimate the wavefront in a phase-diversity sensing scheme. The DM will be used both to introduce the phase diversity and to correct for the wavefront aberrations. This scheme, which has been used to reach very high contrast on the HCIT bench, is also very attractive for ground-based systems:

- it is optically extremely simple, requiring no dedicated WFS optics
- it is extremely sensitive and makes optimal use of a small number of photons
- by design, it is not prone to non-common path errors
- The sensing wavelength is the same as the science wavelength, thus avoiding wavefront chromaticity effects

We have successfully demonstrated the technique in our laboratory testbed at Subaru Telesope (see Fig. 4).

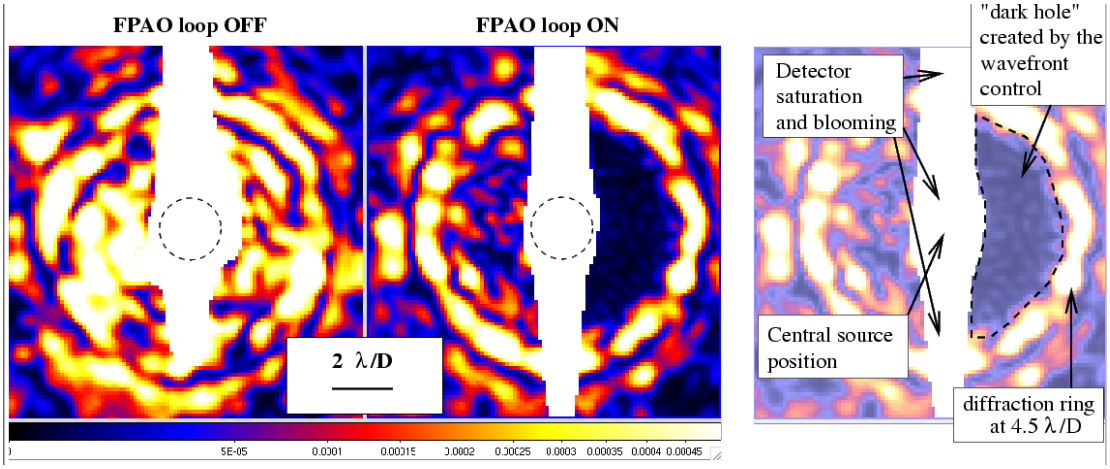


Figure 4: *Preliminary results from the PIAA coronagraph + Focal plane AO laboratory demonstration. A heavily stretched version of our laboratory PSF image (left) shows that almost all starlight is concentrated within 1.5 λ/D radius (the large vertical structure is due to charge bleeding on the detector).*

ACKNOWLEDGMENTS:

This work presented in this paper 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.

HiCIAO: High Contrast Instrument for the Subaru Next Generation Adaptive Optics

Motohide Tamura and Ryuji Suzuki
National Astronomical Observatory of Japan

Klaus Hodapp, HiCIAO Team
University of Hawaii

ABSTRACT

HiCIAO is a new high-contrast instrument for the Subaru telescope. HiCIAO will be used in conjunction with the new adaptive optics system (188 actuators and/or its laser guide star - AO188/LGSAO188) at the Subaru infrared Nasmyth platform. It is designed as a flexible camera comprising several modules that can be configured into different modes of operation. The main modules are the AO module with its future extreme AO capability, the warm coronagraph module, and the cold infrared camera module. HiCIAO can combine coronagraphic techniques with either polarization or spectral simultaneous differential imaging modes. The basic concept of such differential imaging is to split up the image into two or more images, and then use either different planes of polarization or different spectral filter band-passes to produce a signal that distinguishes faint objects near a bright central object from scattered halo or residual speckles.

In order for the Subaru telescope to be at the forefront of the ground-based high contrast science, we are currently developing a higher contrast instrument for the Subaru telescope. We call it HiCIAO (‘Hi’gh ‘C’ontrast ‘I’nstrument for the Subaru Next Generation ‘A’daptive ‘O’ptics). The HiCIAO camera is designed as a flexible camera that can be configured into different modes of operation in an optical bench environment at the Subaru telescope Nasmyth focus. HiCIAO will be the first instrument which can combine coronagraph with either simultaneous polarization or spectral differential imaging modes (PDI or SDI modes) with minimizing efforts for un-common path errors. The basic concept of such differential imaging is to split up the image into two or more images, and then use either different planes of polarization, or different spectral filter bandpasses, or both, to produce a signal that distinguishes objects near a star (dust clouds, planets, etc.) from scattered light artifacts (residual speckles).

HiCIAO consists of several modules; some of them can be easily upgraded. Those modules are as follows (from upstream):

- The Subaru Nasmyth 188-actuator AO system now under development, and its second MEM-based deformable mirror system to be added in near-future. The combination of these will eventually work as an extreme adaptive optics system.
- The coronagraph module in the room temperature environment. This will start from the classical Lyot coronagraph, but can be upgraded to more advanced coronagraphs in near-future.
- The high contrast optics module in the room temperature environment. This optics enables either polarization or spectral simultaneous differential imaging. Simple direct imaging with a high throughput is also supported.
- The infrared camera module in the cold environment. This is equipped with the state-of-the-art detector (2048 × 2048-pixel HgCdTe array - HAWAII-2RG) and its new array-controller (ASIC - “Sidecar”).

The instrument is under the assembling phase in Hilo, Hawaii and its first light at the telescope is expected in the spring of 2007.

As mentioned before, we plan HiCIAO to be upgraded in the future. We are considering using a second stage MEMS deformable mirror (a 1024 actuators MEMS DM from Boston Micromachine Company) which is

currently under test for another experiment at the Subaru telescope. This mirror, combined with advanced readout possibilities of the HAWAII-2RG ASICS controller, is expected to provide a solution for focal plane wavefront sensing (e.g., Guyon 2005). More advanced coronagraph solutions may also be considered for a future upgrade, especially to increase the throughput (since very high Strehl ratios are required to reach the coronagraph contrast capabilities).

Table 1: *HiCIAO overallspecifications.*

Platform	Subaru Nasmyth focus (IR-side)
Combined adaptive optics	188-elements curvature sensing AO With natural guide star (2007-) With laser guide star (2008-)
Wavelength coverage	zJHKs bands (optimized at H)
Detector	2048 × 2048 HgCdTe (Rockwell HAWAII-2RG #29)
Detector controller	ASIC (Rockwell "sidecar")
Pixel scale	0.01 arcsec/pixel (proper PSF sampling at J)
Spatial resolution	0.03 arcsec (J) , 0.04 arcsec (H), 0.055 arcsec (Ks) FWHM
Sensitivity	24.3 mag (J), 24.2 mag (H), 23.8 mag (Ks), ~23.3 mag (CH_4) (5σ, 1hr)
Contrast goal	10^4 at 0.1 arcsec, $10^{5.5}$ at 1 arcsec (with DI model + coronagraph)
DI mode	Direct Imaging 20 arcsec × 20 arcsec FOV, general purpose imaging High throughput (only 5 lenses), with or without coronagraph
PDI mode	Polarimetric (o-ray and e-ray) dual imaging with YLF Wollaston prism 20 arcsec × 10 arcsec FOV, polarized target imaging
SDI mode	Spectral simultaneous differential imaging with double Wollaston prism 6 arcsec × 6 arcsec FOV, highest contrast imaging
PI mode	Pupil imaging for optical alignment check in real-time
Filters	11 common filters (zJHKs and narrow band filters) 4 sets of differential filters (methane, [FeII], and others)
Coronagraph Occulting masks Lyot stops	Starting with the classical Lyot coronagraph. More sophisticated coronagraph will be implemented later. 3 sizes with real-time remote choice. Stop with spider/center-hole blocks & rotation compensation.

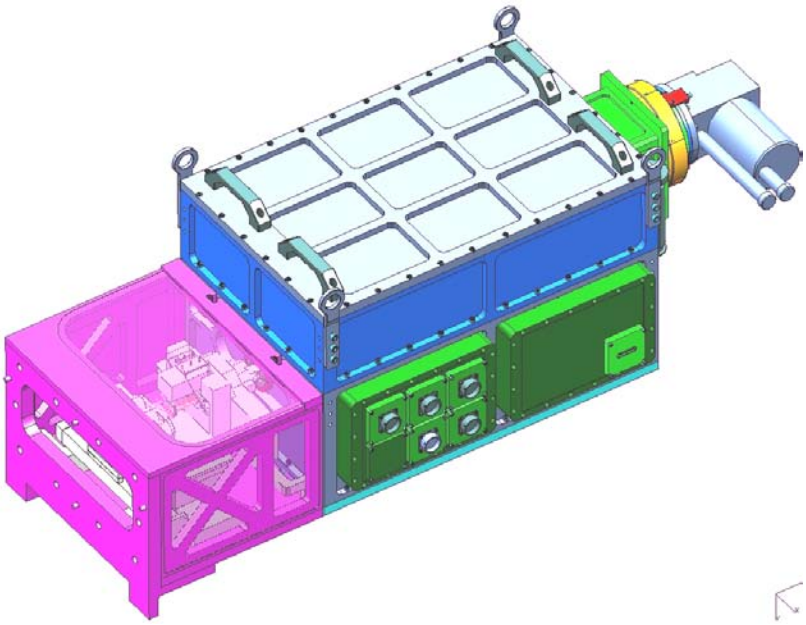


Figure1: *Mechanical design of HiCIAO. which is composed of warm foreoptics including coronagraph optics and cold IR camera including differential filters.*

First Laboratory Demonstration of Anti-Halo Apodization: A Coronagraph "Afterburner"

Johanan L. Codona
Steward Observatory, University of Arizona

ABSTRACT

A prototype of the AntiHalo Apodization (AHA) system has been implemented and initially demonstrated in the lab. This system uses the diffraction-limited core starlight, normally discarded by the focal plane stop in a coronagraph, to explicitly construct a negative copy of the halo directly in the focal plane. When coherently combined with the coronagraphic halo in the detector plane, the resultant halo is expected to be up to 100× fainter. The total system design includes measuring the complex sum of the halo and the antihalo with an interferometric focal plane wavefront sensor (FPWFS), which is also implemented in the lab. In this progress report, we show initial results from creating the antihalo with a 2-DM Michelson spatial light modulator, and the first modest halo suppression achieved by using an intensity-minimizing search algorithm in the 2-dimensional AHA actuator displacement space. The resulting attenuation of a factor of 6 to 8×, points the way to the expected 2 decades of attenuation when the system is fully functional.

INTRODUCTION

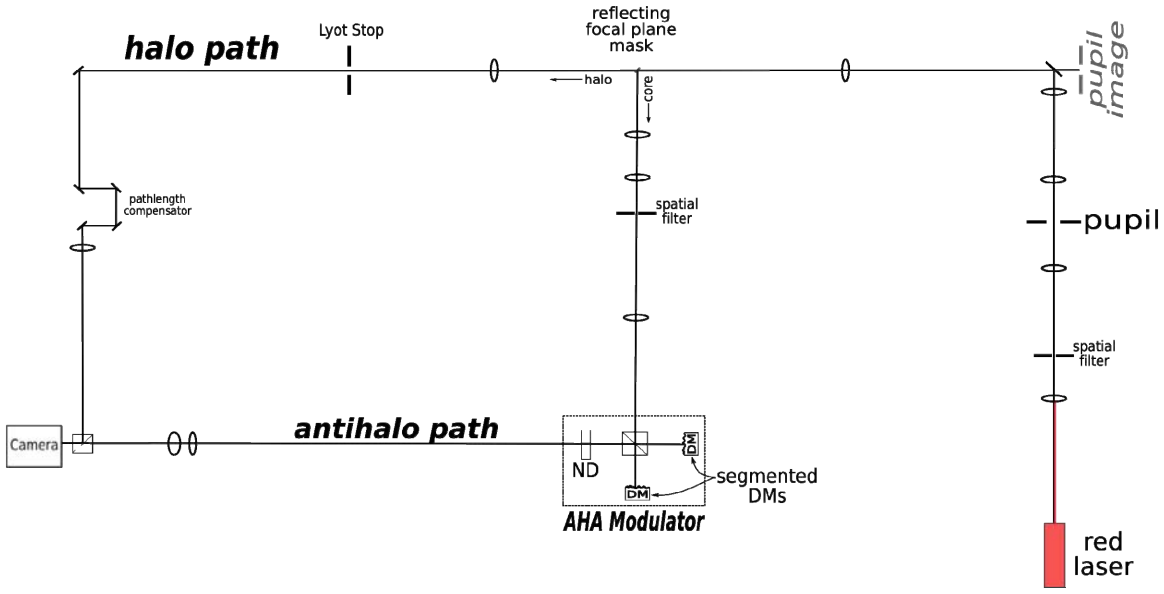


Figure 1: *Schematic of the Antihalo Testbed, showing only those paths and components used in this initial test.*

As this conference has clearly shown, great advances are being made in the design of stellar coronagraphs. To be able to see terrestrial exoplanets, the stellar halo needs to be suppressed on the order of ten decades below the diffraction peak of the star. This requirement is extremely difficult to meet, and indeed many promising techniques are falling 3 or 4 decades short of the mark. These problems are typically due to slight imperfections in manufacturing, or even very slight transmission variations, such as dust on the optical surfaces, which will cause halo light at levels much brighter than desired.

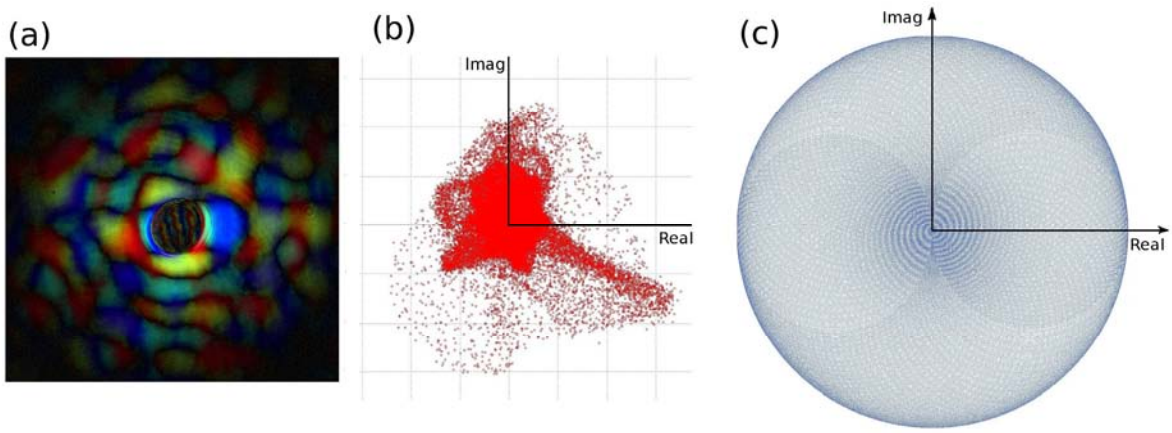


Figure 2: *Complex halo and addressable antihalo. (a) The complex halo can be measured using the FPWFS. This figure shows intensity as usual, but encodes phase as color. (b) The halo plotted as points in the complex plane. (c) A single wrap subset of addressable points in the complex plane using the MEMS DM and the high-resolution electronics.*

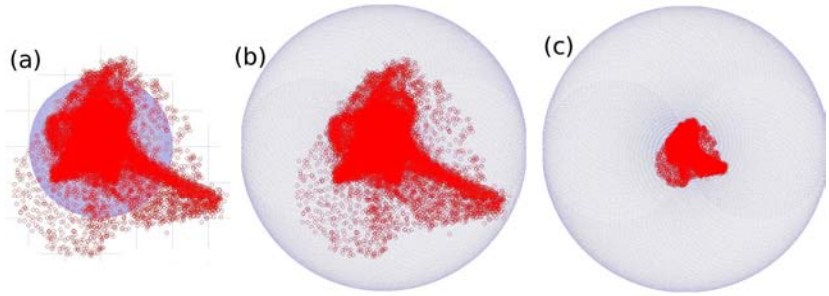


Figure 3: *Effect of AHA modulator illumination on addressable complex antihalo values. (a) AHA path too faint: can't reach or suppress all points in the halo. (b) AHA modulator illumination well matched to the halo. (c) AHA modulator illumination too bright: will work, but density of addressable points is lower limiting the maximum suppression.*

Since the scattered light halo is coherent with itself at various points in the focal plane, it is possible to mitigate the unwanted scattered light by measuring or estimating its phase and amplitude (i.e., its complex amplitude), and adjusting an upstream deformable mirror (DM) to intentionally create a speckle with the same amplitude but opposite sign, canceling the halo at that point. An iterative scheme can do this over a region, up to 180° around the star, and out to the control radius of the DM. The difficulty comes from the required level of control of the DM. To control a halo structure 10 decades below the star in visible light using a 1024-actuator DM would require each actuator to be controllable to about 10 pm.

The Anti-Halo Apodization (AHA) method is an active halo suppression technique that is intended to further suppress the residual halo achieved by a coronagraph (Codona and Angel, 2004). It accomplishes this by explicitly constructing a negative approximation of the residual halo directly in the focal plane and adding it to the halo, approximately canceling it out. To do this, the constructed antihalo must be temporally coherent with the residual coronagraph halo. In most coronagraph designs, the core starlight is blocked by a focal plane mask of some description, and is absorbed or otherwise prevented from further propagation though the coronagraph. We make the simple modification of diverting this discarded starlight and using it to create our temporally coherent antihalo. This method is similar to that originally proposed by Angel (2003) to build an interferometric focal plane wavefront sensor (FPWFS). Our system simultaneously implements both methods by splitting the core starlight: one part to build the antihalo, and the other part to interferometrically probe the sum of halo and antihalo.

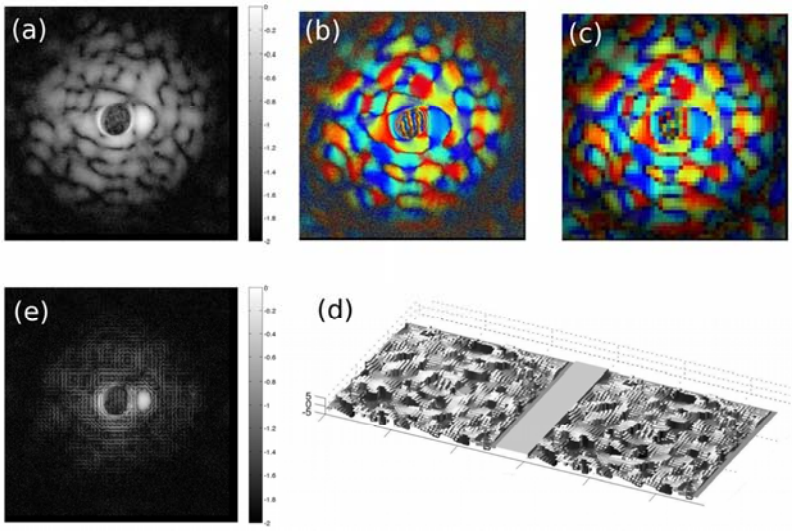


Figure 4: *Conceptual flow of the AHA method. (a) A measured intensity image of a speckled halo (high-gamma stretch). (b) The complex halo as measured with the FPWFS. (c) Pixelized version implemented on the AHA modulator. (d) The computed displacements for the two DMs in the AHA Michelson to create (c). (e) The intensity resulting from the difference between (b) and (c). Note the bright speckle at the 3:00 position that was too bright to be reached by the addressable complex values with the AHA modulator illumination. Increasing the modulator illumination would enable this speckle to also be suppressed.*

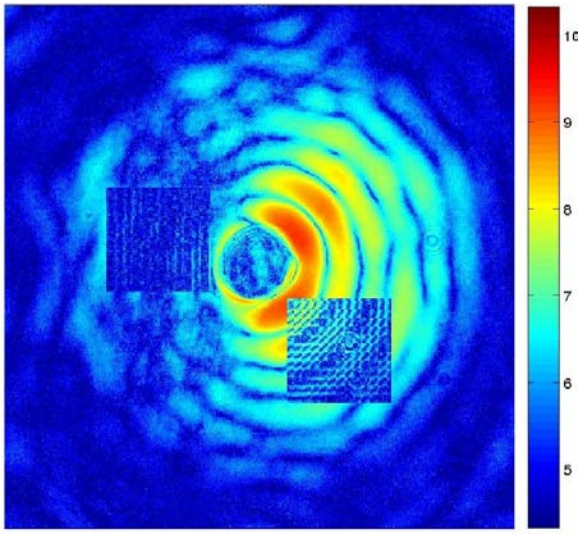


Figure 5: *With small DMs (say 12×12 actuators), only a small region of the focal plane may be suppressed. However, there are no restrictions on where it may be placed.*

The beauty of the AHA method is that the amount of starlight used to create the antihalo is controlled by some gross method, such as an ND filter, rather than by the incredibly fine control of a pupil DM. The additional suppression potentially achievable using AHA is inversely proportional to the density of addressable points in the antihalo complex plane. As discussed in Codona & Angel 2004, the suppression achievable by controlling the phase of each of the elements in the AHA SLM to 0.1 radians is about two decades. Using a MEMS DM with adequately high-resolution electronics should enable suppression at even higher levels, up to three decades. This assumes that the AHA "pixels" (the AHA spatial light modulator (SLM) actuators) are small compared with λ/D so that the uncontrolled shape of the antihalo over the actuator is not important. Since this technique works in conjunction with other coronagraph designs as an "afterburner" to get even more contrast out of the system, the addition of a relatively low-precision AHA system would significantly ease the accuracy and control requirements on the coronagraph and pupil DM, to levels such as a quarter of a nanometer, which is more robustly available.

At the time of this conference, we have just built the first laboratory implementation of an AHA modulator. As a first test, we used a single actuator search algorithm to perform a search in phase and amplitude to suppress the halo+antihalo. This process was repeated for a number of pixels over a rectangular test region. This first test, even with its algorithmic simplifications and the fact that it was search-based rather than using the FPWFS, successfully demonstrates and validates the AHA principle.

Lab Setup

In addition to the AHA technique, the UA Antihalo Testbed is intended to demonstrate and explore the use of the interferometric FPWFS, and a number of other adaptive halo suppression techniques and algorithms. Since the first test use of the AHA modulator did not involve these other components and subsystems, we will not describe them here. The relevant components are shown in Figure 1. Our light source was a simple red laser diode, which

was spatially filtered and collimated onto a pupil. This pupil is reimaged to another plane (to facilitate the use of a pupil DM not used here), and then imaged into the first focal plane of the coronagraph. Here there is a reflecting spot deposited on a glass substrate, with a 1:1.414 axial ratio, mounted on a 45° plane so that it appears to be a simple circular top hat mask when seen along the axis of the incoming light. The mask size used is comparable with the diameter of the first null in the Airy pattern. The light missing the mask is processed as usual for a normal Lyot coronagraph, being reimaged into a pupil plane where there is a Lyot stop, and finally being reimaged into a focal plane on the camera sensor. To enable the halo to remain in temporal coherence with the antihalo, the extra path length introduced by the AHA modulator must be added to the otherwise rectangular path. The diverted core light is spatially filtered and collimated onto the two segmented MEMS DMs of the AHA modulator. The light is split equally between the two mirrors with a 50/50 beam-splitter cube, which then remixes the reflected light with the phase modulation applied by moving the DM actuators. In the process, half the light entering the modulator is lost, but since it only needs to be a tiny fraction of the core starlight for an already faint halo, the loss is unimportant. Overall, the amplitude of the AHA modulator light is adjusted with a neutral density (ND) filter to have an input amplitude that is greater than that of any halo amplitude over the region to be suppressed. Finally, the newly imprinted antihalo is mixed with the residual halo from the coronagraph in another beam-splitting cube and is detected by the camera.

The AHA modulator is a fully complex spatial light modulator (SLM) built using a Michelson interferometer with a MEMS DM on each arm. The incoming light, having been spatially filtered and spread out over what will become the suppressed region of the halo, is intended to have no significant phase or intensity features. (In the full system, if it did, the FPWFS would automatically measure and compensate for them.) For optimal contrast and chromatic performance, the two arms of the modulator should be as close as possible to the same length, which should be similarly close to the path length of the halo path. The bandwidth and structure of the laser diode source used here is sufficiently broad to warrant careful alignment for best performance. However, for this first quick test, the only alignment done was to get high-contrast fringes in the AHA Michelson, and acceptable modulation between the antihalo and halo paths. Later, more careful adjustment will be necessary to achieve the best possible halo suppression. By examining the image of the illuminated AHA DMs from the final focal plane camera, the DMs were adjusted relative to each other so that their actuator images overlapped each other to within a small fraction of an actuator (roughly 30 microns on an actuator pitch of 0.3 mm).

By moving the DM actuators, the path length along each arm is altered, causing the phase shift along each arm to become $\phi arm = 4\pi zarm/\lambda$, where zarm is the path length to an actuator along one of the arms. After remixing the light in the beam splitter cube, the resulting complex amplitude applied to the featureless field derived from the diffraction core is proportional to

$$\exp\{i\phi_1(x,y)\} + \exp\{i\phi_2(x,y)\}.$$

This is sufficient to give the antihalo any complex value inside a circle of radius controlled by the ND filter (Figure 2c). The DM control electronics start with a digital-to-analog converter (DAC), which generates an equally spaced set of voltages, which are applied to the actuators. The actuators, in turn, feel a force that is proportional to the square of the applied voltage opposing a linear restoring force, and hence the displacement is roughly proportional to the square of the voltage. The resulting, non-uniform set of surface displacements determines the spacing of the AHA complex modulator points. Since the goal is to create an antihalo replica of the halo, which is to be subtracted from it, controllability residual field error is determined by the distance between the nearest addressable antihalo point and the actual halo value. This means that the mean error field is proportional to the distance between the antihalo complex field points, and the resulting intensity is inversely proportional to the density of addressable points in the complex plane. To achieve maximum control of the actuators, we worked with the company *IrisAO, Inc.*, to adapt their high-resolution *Smart Driver* electronics to drive 128 of the 140 actuators in two *Boston Micromachines* DMs. The DM electronics, like the rest of the experimental control and data acquisition systems were controlled from a Linux system.

To understand the role of the ND filter in adjusting the resolution of the AHA modulator, imagine measuring the complex halo over the region to be suppressed, and plotting all of the field points in the complex plane (Figure 2). The circular set of addressable antihalo field points has a maximum amplitude which is determined by the amount of light passed through the AHA modulator, and is attenuated further by the ND filter. As can be seen in Figure 3a, if the set of antihalo points is smaller than the set of halo points (i.e., the AHA light is too dim), it will not be possible to fully suppress the brightest portions of the halo. If the AHA light is too bright (Figure 3c), all points in the halo will be able to be suppressed, but since the mean distance between the addressable points is larger, the maximum suppression will be less. The optimum setting is when the antihalo set is larger than the set of halo field points to be suppressed, but not too much larger (Figure 3b). This is well defined when the halo is stable, but if it is fluctuating, the best setting will be to adjust the maximum antihalo to be larger than the largest expected halo values.

The loss of suppression caused by having too much light passing through the AHA modulator is not serious in practice, since the finite actuator size limits our ability to fit the shape of the antihalo to the halo. The smallest halo structure scales are on the order of the diffraction scale, λ/D. Over this scale, full amplitude swings and sign reversals are common, and must be reproduced in detail for full suppression. We use segmented DMs to avoid problems with the influence functions near points where the halo field changes sign. However, if the segmented actuators are flat, the best halo-antihalo difference will have a triangular error with a magnitude proportional to the angular gradient of the field over the scale of the actuator. This means that there is a halo-fitting error intensity that is proportional to the square of the ratio of AHA modulator actuator size to the diffraction scale. Using a zoom lens arrangement allows this to be modified, but it also has the effect of shrinking the overall field that can be suppressed. Larger DMs will ease this problem in the future. For now, we nominally use a magnification of about 3 to 4 actuators per λ/D. This should ultimately allow for greater than 2 to 3 decades of suppression, which is still much greater than the electronics will allow.

The operational concept is illustrated in Figure 4. The halo used in this figure was actually measured using the focal plane interferometer, but the AHA portion was simulated. In Figure (4a), we have the speckled halo of a star imaged through a weak phase screen. The Strehl ratio is still high enough to retain a substantial diffraction-limited core, which was reflected for use as the reference beam in the FPWFS, and available for use in creating the AHA antihalo. There is a bright speckle at the 3-o'clock position that we will intentionally miss with the AHA modulator illumination, just to show the effect. Figure (4b) shows the complex halo as measured with the FPWFS, and (4c) shows the pixelized version of the halo that can be created using the AHA modulator for some illumination strength. This is accomplished by adjusting the two DM surfaces to appropriate values, shown in Figure (4d) for a hypothetical huge DM. The two surfaces are very similar when the desired antihalo amplitude is much fainter than the light illuminating the modulator, and differ more where the antihalo is brightest. Careful examination of the two surfaces near the positions of the brightest speckles shows this effect. Finally, Figure (4e) shows the result of the complex sum of the halo and antihalo, which takes place in the final beam splitter. Since the light directly entering the camera is *halo-antihalo*, while the light exiting the beam splitter in the other direction is *antihalo-halo*, capturing both outputs and summing the detected photons means that no light need be lost in the mixing stage. Figure 5 illustrates the current situation, where we have only a relatively small format DM with which to build the AHA modulator. In that case, the suppressed region is small, but may be placed anywhere within the focal plane.

FIRST RESULTS

The first test of AHA was intended to demonstrate the mechanics of the halo interfering with a manufactured antihalo—with at least some attenuation. For this purpose, it was not necessary to use the FPWFS to measure the complex sum of the halo and antihalo. Instead, we used a simple search algorithm in the two-dimensional actuator displacement space to find the best suppression for a single actuator. Although this algorithm is fully independent of what the other actuators are doing, and is completely parallelizable, the first test was performed one actuator pair at a time. The result is shown in Figure 6. The first image (Fig. 6a) is an intensity image of the halo in the region to be suppressed. The second image (Fig. 6b) is the halo mixed with the antihalo, but set to

default initial positions. Finally, after sequentially running the AHA search algorithm on a single actuator pair (Fig. 6c) and a rectangular subset of the modulator actuators (Fig. 6d), the halo is suppressed. The maximum suppression shown here was only a factor of 6 to 8. But given that this was simply an initial demonstration, and the system was not optimally aligned for maximum visibility, nor was the search algorithm designed to find the best null near the white light null, this level of attenuation was reasonable. Analysis of these results suggests that the full expected suppression should be achieved when the system is properly adjusted.

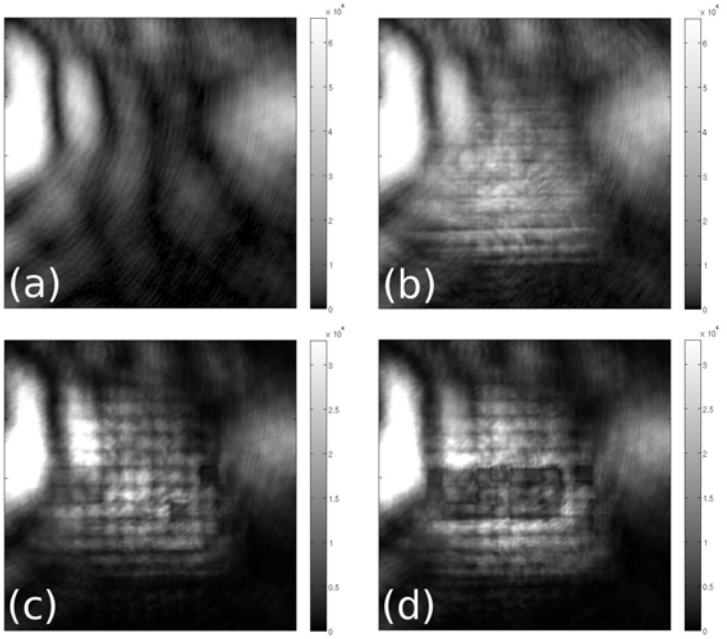


Figure 6: *Results of first AHA test. Bright ring on the left is the second Airy ring. (a) Initial halo. (b) Halo plus unset antihalo. (c) AHA effect after searching for a solution to a single AHA "pixel" value. (d) Repeated application of the single AHA "pixel" algorithm over a rectangular subset of the AHA region of the focal plane.*

SUMMARY

The AHA system is now working, and initial tests have validated the operational concept. Work continues in the lab to get the maximum suppression from the AHA modulator, along with simultaneous operation of the FPWFS, which will enable a closed-loop servo by measuring the complex field resulting from the coherent combination of the halo and the antihalo. We expect that when used with any coronagraph design where the discarded stellar core can be coherently preserved as a source for constructing the antihalo, the AHA method should permit 2 to 3 decades of improved halo suppression over that achieved by the coronagraph alone.

ACKNOWLEDGMENTS

Thanks to graduate student Peng Su for assistance in the lab. This work was supported by NASA grant Grant# NNG05GD28G.

REFERENCES

Angel, Roger, "Imaging Exoplanets from the Ground", in *Scientific Frontiers in Research on Extrasolar Planets*, eds. S. Seager and D. Deming, ASP Conf. Ser., 294, 543–556, Washington D.C. 2003.

Codona, J. L. and Roger Angel, "Imaging extrasolar planets by stellar halo suppression in separately-corrected color bands", ApJ **605**, L117–L120, 2004.

Laboratory Demonstrations of High-Contrast Coronagraph Imaging at JPL

John Trauger and Wesley Traub
Jet Propulsion Laboratory, California Institute of Technology

We report the laboratory demonstration of coronagraphic imaging at angular separations and contrast levels that could permit exoplanet detection at visible wavelengths; hence the existence of at least one viable solution for the "physics problem" of designing an instrument to detect Earth-like exoplanets, leading ultimately to the design of a TPF-C mission to detect and characterize planetary systems orbiting the nearby stars. The demonstration involves a collaboration of many individuals, including Chris Burrows, Brian Gordon, Brian Kern, John Krist, Andreas Kuhnert, Dwight Moody, Al Niessner, Fang Shi, Dan Wilson, and Marie Levine.

Coronagraphs are not new to astronomy, but only recently has the concept been considered for the imaging of Earth-like exoplanets from space. In space, free of the blurring effects of atmospheric turbulence, a coronagraph must further suppress the Airy rings diffracted from the edges of the primary mirror as well as the surrounding field of speckles due to irregularities in the surface figure of the optics. Diffracted light from the mirror edges can be removed by a variety of well-studied coronagraph configurations, each with its specific characteristics and limitations including efficiency, spectral bandwidth, and complexity. Here we demonstrate diffraction suppression with a Lyot-type coronagraph, and speckle suppression with a technique that is applicable to all coronagraph types.

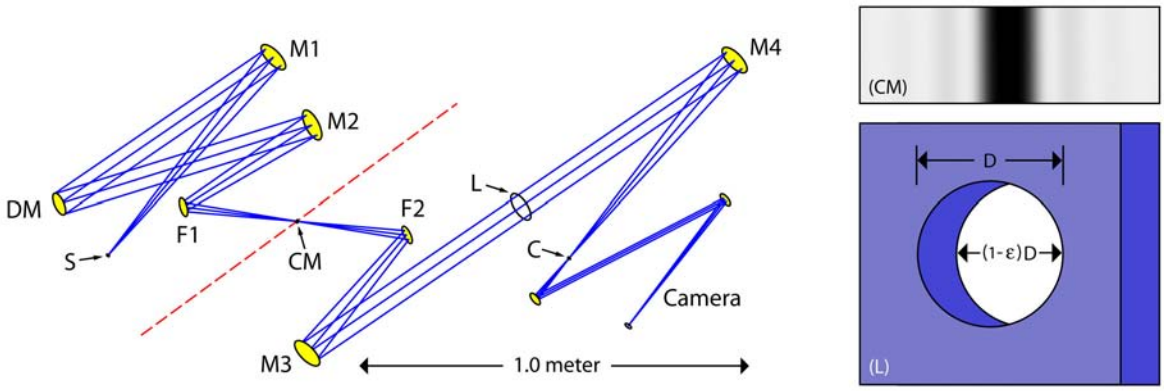


Figure 1. Laboratory coronagraph optical layout. *The essential optical elements, shown to scale, are the simulated star (S), four identical off-axis paraboloidal focusing mirrors (M1-4), a deformable mirror (DM) which also defines the pupil of the system, two flat fold mirrors (F1-2), focal-plane coronagraph mask (CM), Lyot stop (L), high-contrast coronagraph focal plane (C), and the science CCD camera. Detail at right shows the measured transmission profile of the coronagraph mask (CM) and the geometry of the Lyot mask (L), fashioned from a pair of circular stops, each with opening diameter D and offset by εD. A one-meter fiducial indicates the overall scale of the optical system. The pupil diameter D is 30 mm and the CM attenuation profile measures 82 μm (corresponding to 4 λ/D) from the center of the pattern to the half transmittance point.*

Our laboratory setup, the High Contrast Imaging Testbed (HCIT), is shown in Figure 1. Enclosed in a space-like environment inside a vibration-isolated, temperature-controlled vacuum chamber, this system captures the essential optical features of a space coronagraph. We report here the results of two experiments with the HCIT, a "snapshot" and a "movie". The snapshot experiment simulates a single exposure of a star and exoplanet system by a space coronagraph, allowing a view of a one-sided region near the star. The movie experiment simulates a

series of snapshots taken as the space coronagraph is rotated about the line of sight to the star, thereby allowing a search for exoplanets in an annular region around the star.

In the snapshot experiment, the image of the star is centered on the coronagraph mask and the deformable mirror is commanded, via iterations of a speckle nulling algorithm, to minimize the speckle intensity in the target field of view. This experiment used a simulated star, but no simulated planet. By offsetting the star to a clear part of the mask, we record what a planet would look like, shown in Figure 2a. With the star centered, a dark target field appeared (Figure 2b). If a simulated planet had been present in this D-shaped field, it would have appeared as a bright spot resembling Figure 2a. Quantitatively, the data plotted on the right in Figure 2 show: (a) the azimuth-averaged intensity of the star in the target field, without the focal plane mask present; (b) the azimuth-averaged intensity of the star in the target field, with the focal plane mask, and with the DM set to minimize the average intensity in this field; and (c) the same as (b) except prior to DM correction. As shown in (b) the average intensity is about 6×10^{-10} times the peak intensity of the star in a field between 4 and 10 λ/D on one side of the star. This snapshot experiment shows that the present apparatus, in a single exposure, is capable of suppressing both the diffracted and scattered light around a star, down to a contrast level that is slightly better than a Jupiter, but not quite as low as an Earth.

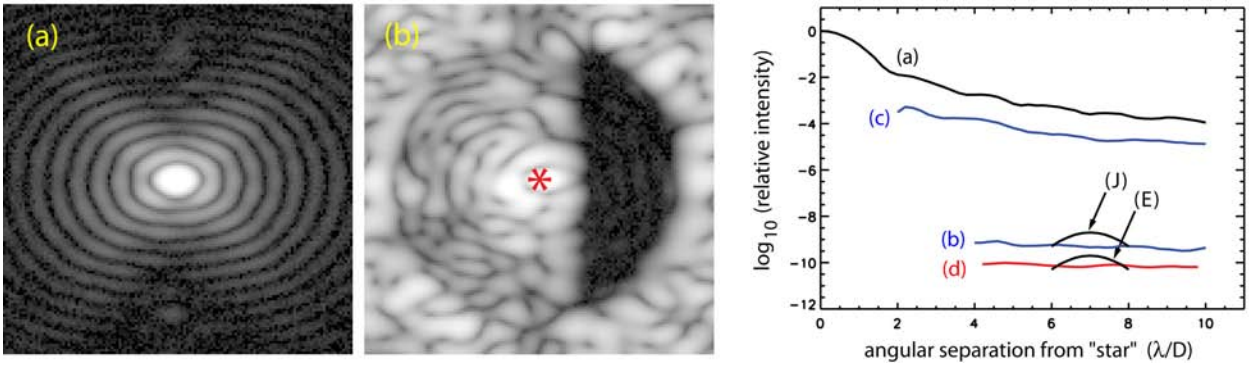


Figure 2. Representative coronagraph images and intensity profiles. *Image (a) and curve (a) illustrate the appearance of a "planet" offset from the star (and therefore not occulted by the coronagraph mask). The horizontal elongation of the diffraction rings is a result of the Lyot aperture in Figure 1. Image (b) and curve (b) show the high-contrast coronagraph "dark half-field" to the right of the masked central star (at the location of the red asterisk). Also shown are curves for (c) the coronagraph contrast prior to wavefront correction (with the DM nominally flat) and the improved contrast (d) obtained by roll deconvolution of a set of coronagraph images (cf., Figure 4c). Intensity profiles for a nominal Earth (E) and Jupiter (J) in reflected starlight are included for reference. Images and intensity curves are displayed on logarithmic scales.*

This experiment validates many, but not all, of the critical elements of an actively corrected space coronagraph for exoplanet imaging. It is an extremely simple, stable coronagraph configuration, operating in a space-like environment, with contrast performance that can be accurately modeled end-to-end using the known characteristics of the optical elements. It illustrates a robust method of optical wavefront sensing and control that requires only the DM and science camera to analyze the image of a star. It shows that current DMs are capable of suppressing scattered light to contrast levels and separations representative of a planet-finding mission and, as the movie experiment will show, that the precision DM settings remain stable over periods of hours or more without feedback. However, the experiment lacks a simulator for the dynamics of a large telescope structure in space, and the experiment was performed in polarized, narrowband laser light rather than unpolarized, continuum starlight filtered to a 10–20% ($\delta\lambda/\lambda$) bandwidth as would be required for photometric studies in astronomy. As verified by splitting the light into its component polarizations just prior to the CCD camera, our coronagraph is insensitive to polarization, working equally well in polarized and unpolarized light to the reported contrast levels.

While not reported here, we note that the question of spectral bandwidth is being addressed in stages. Early experiments with 2% bandwidth have demonstrated contrast of 1.5×10^{-9} using the same coronagraph masks and speckle nulling procedure as above. Our optical model indicates that an optimal choice of Lyot mask will improve this to the contrast levels seen in narrowband light. Initial experiments with 10% bandwidth light, again with the same coronagraph mask and speckle nulling procedure, produced a contrast of 6×10^{-9}. Our model predicts that this 10% contrast will be reduced by about an order of magnitude using a new coronagraph mask now being manufactured with standard techniques and common materials. This is an active area of development and a pathfinder for the TPF-C mission design.

In the movie experiment, to simulate a coronagraph operating in space, we continuously repeated the snapshot experiment 480 times over a period of 5 hours. The apparatus was very stable during this period; it was not adjusted in any way between exposures. The background speckle field evolved slowly, owing to room temperature changes and mechanical relaxation, much as might be expected in a coronagraph in space.

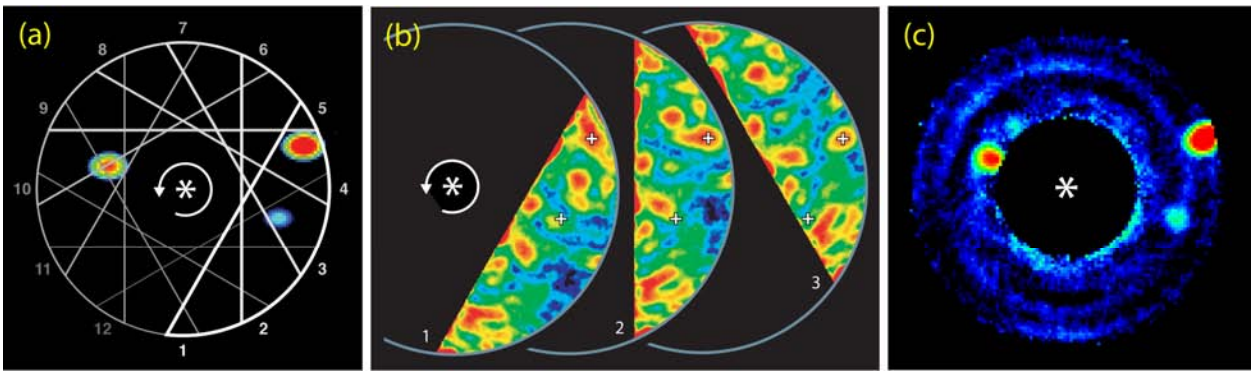


Figure 3. Laboratory images demonstrate contrast at levels required to detect an Earth-twin. *(a) Three planet images are shown on the sky. The planets are copies of the measured star but reduced in intensity by factors of (10, 5, and 1) $\times 10^{-10}$, corresponding to the typical intensities of Jupiter, half-Jupiter, and Earth, respectively. The Earth-twin is at about 4 o'clock, and Jupiter-twin at 2 o'clock. The D-shaped field of view rotates on the sky as the spacecraft is rotated about the line of sight to the star (asterisk). (b) Three sample images at different rotation angles illustrate the observing sequence. Note that the planets (white + signs) are fixed in inertial space, and just barely visible. The rotation sequence continues to fill a full annular field of view. (c) Roll deconvolution is applied to the data, removing the background speckles that rotate with the spacecraft, and keeping the part of the image (planets) fixed in the sky (see also Figure 2 for relative intensities). The planets stand out clearly against the residual background noise, which is the time-varying part of the speckles.*

The movie experiment demonstrates the process of planet discovery with a space coronagraph, as follows. As shown in Figure 3a, we hypothesize a space coronagraph that is aimed at a nearby star (the asterisk), with the background starlight suppressed in the target field, as in the snapshot experiment. We assume that the space coronagraph is rotated in small increments, so that the dark target field successively covers regions that ultimately fill the complete annular region between angular separations of 4 and 10 λ/D on the sky. For illustration we show twelve discrete 30° steps in Figure 3a, but in the example here we use 48 steps of 7.5° each, and in each step we co-add 10 sequential exposures, using a total of 480 exposures.

Lacking a simulated planet, we added attenuated copies of the star to each exposure, at the sky locations shown in Figure 3a. This procedure is valid because we have previously shown that the presence of a planet in the speckle field of a star has no effect whatsoever on our wavefront correction algorithm. The orbital positions were chosen artificially so that the projected planetary system could be captured in a single image here.

Three snapshots of the planets-plus-speckle field are shown in Figure 3b. With a relative speckle intensity of about 6×10^{-10} in each snapshot, the planets (centered under the white "+" signs) are barely visible. This combination of a slowly evolving instrumental background that is fixed with the rotating spacecraft, plus an astronomical object that is fixed in inertial space, has been encountered before with images from the HST. A roll deconvolution algorithm was applied to the sequence of these images of laboratory background plus superposed planets by John Krist. The result is shown in Figure 3c, where we see that all three planets stand out clearly against the remaining background. Quantitatively, the azimuth-averaged relative intensity is plotted in Figure 2 (right), where curve (d) is the roll-deconvolution background averaging about 0.9×10^{-10}, which is well below the single-snapshot intensity of 6×10^{-10}. The Earth and the Jupiter intensities are shown added to the background from roll-deconvolution, and superposed for clarity.

The movie experiment is an existence proof that it is possible to extract exoplanets close to a star, even when the residual speckle intensity is comparable the planet intensities. We believe, however, that for planet detection and characterization a space coronagraph should be designed to even stricter standards, with roughly a factor of 10 weaker speckles than achieved in the present experiments. The present work is a step toward this goal, but more work remains. Future laboratory work will focus on pushing the speckle background lower, broadening the spectral bandwidth, suppressing speckles simultaneously on both sides of the star with a pair of DMs, and increasing the radial field of view, both inward and outward. The present work lays the groundwork for the development of a TPF mission that will, for the first time, explore nearby Earth-like exoplanet systems by direct imaging and spectroscopy.

REFERENCE

J. Trauger and W. Traub, A laboratory demonstration of the capability to image an Earth-like extrasolar planet, Nature **445**, pp. 771–773, 2007.

Active Thermal Figure Control for Large, Lightweight Honeycomb Mirrors in Vacuum and Space

Roger Angel, Tae Kang and Brian Cuerden
Steward Observatory, University of Arizona

1. ABSTRACT

Active figure control of lightweight glass honeycomb mirrors will be valuable for making the on-orbit operation of large space optics more precise, lighter, more cost effective, and more thermally stable. Key applications will include very high contrast imaging of extrasolar planets and large scale vacuum test optics, whose optical quality must be controlled for better quality than the space systems under test.

The concept presented here relies on the low but finite thermal expansion of honeycomb mirrors made from fused silica, a material commonly used for precision lightweight space optics. The figure is controlled by varying the temperature of the faceplates and individual rib elements. Resistive heating is balanced in a servo control loop against radiative loss to cold fingers inserted in each honeycomb cell. Preliminary finite element models indicate that for a mirror with n cells up to n Zernike modes can be corrected to better than 90% fidelity, with still higher accuracy for the lower modes. An initial demonstration has been made with a honeycomb mirror of borosilicate glass. Interferometric measurements show a single cell influence function with 300 nm stroke and ~5 minute time constant.

2. RATIONALE

Glass honeycomb mirrors have long been the preferred choice for space optical systems operating at the optical diffraction limit, because of their high thermal and mechanical stability. Alternate primary mirror materials with higher elastic modulus can be used for smaller optical systems or for reduced quality. The beryllium segments for the JWST, for example, are aimed at diffraction-limited performance down to 2 μm wavelength. However, for large systems to be diffraction limited in the optical (λ=0.5 μm), glass honeycomb technology is likely to be preferred. But even glass has a performance limit—for example the HST primary has astigmatism that varies slowly over a range of 25 nm (Lallo et al, 2006). Active control of the wavefront would allow the highest optical quality to be maintained, even for systems with a much lighter build than HST, by compensating for gravity release effects and support force errors in space. It would also allow construction of "super diffraction-limited optical systems," as are needed for ultra-high contrast imaging to image extrasolar planets.

At present the only option for correcting primary aberrations is to add optics to relay a pupil image onto a separate deformable mirror and to recreate the original image. The concept we propose makes the correction directly at the glass-honeycomb primary mirror. This has several advantages

1) Higher throughput and simpler—no additional optics needed
2) Simple and lightweight translates to lower costs/lower mass
3) No cross coupling of phase into amplitude errors—limits spectral bandwidth for very high contrast imaging. This is very important for exoplanet detection
4) no increased field aberrations from the added relay

Many future astrophysics missions will be able to take advantage of precise optical wavefront prescription and/or the stability of thermally controlled honeycomb mirrors. Examples are exoplanet imaging, high spatial resolution UV and visible imaging (JDEM, Large Ultraviolet/Optical Telescope), or missions aimed at extreme astrometric stability (OBSS Origins Probe mission concept). The resistive heating control system described here has the potential to both lower the costs and improve the performance of these types of missions. It would also reduce the risk, since it works by combining two technologies—glass honeycomb mirrors and thermal control—that are already at very high TRL levels.

3. THERMAL ACTUATION CONCEPT

Conventional glass honeycomb mirrors for space obtain their thermal stability by use of very low or zero expansion glasses. Fused silica (coefficient 5×10^{-7}/C) and its titania doped form (ULE— coefficient $\leq 10^{-8}$/C) are both well-established and commonly used (Hobbs et al, 2003). Our strategy is to take advantage of silica's small coefficient to induce small wavefront corrections by differential thermal expansion. Accurate thermal controls will be added, on the scale of the honeycomb cells. This approach avoids the difficulties of primary mirror actuation by piezo or electrostrictive actuators, which arise because their relatively high expansion coefficients would compromise the natural stability and rigidity of glass honeycomb structure. Low order thermal control of a telescope primary has been tested by Lardiere et al. (2000), and of a silicon plate by Vdovin and Loktev (2002).

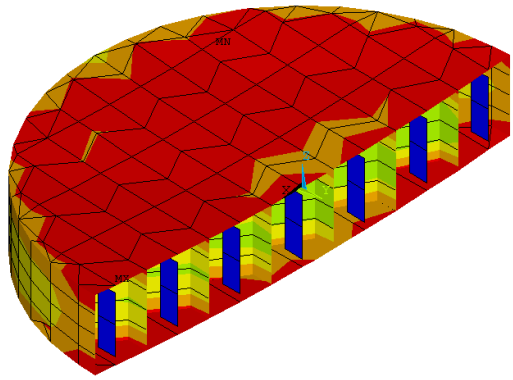


Figure 1*: Preliminary finite element model with 37 cells. Color coded for equilibrium temperatures; when the cold fingers are held isothermal and 3.5 W are dissipated across the mirror heaters, the face temperature (red) is 5°C warmer than the cold fingers (blue).*

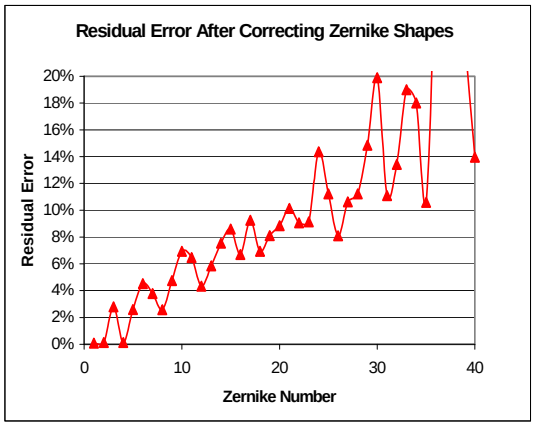


Figure 2: *Thermally induced fractional residual errors for Zernike terms.*

The control geometry we propose will use a combination of resistive heating and radiative cooling. Normally these would be balanced, but in order to achieve local expansion or contraction, the heater power will be altered. The resistive heaters will be deposited or glued to the internal rib and faceplate surfaces. Radiative cooling will flow to cold fingers inserted into each honeycomb cell and around the perimeter wall. The neutral state of the mirror will be one in which a steady state heat flow is established. To control edge effects, the back outer face will be insulated, and the front mirrored surface will have low emissivity and thermal coupling. The temperature of the different surface elements will be measured by thermistors and used as input to the thermal control servo.

A preliminary small-scale model of a 37-cell honeycomb mirror, using thermal finite element code, has been made (shown in Fig. 1). It is 30-cm in diameter and 5-cm thick overall. In the steady state the face sheets are 5°C warmer than the cold fingers for a total heat flow of 3.5 W. The mirror would be null figured in this state. When the heaters are adjusted to induce small distortions, wavefront amplitudes ≥ 300 nm peak-to-valley can be realized for up to the first 11 Zernike modes, and ≥100 nm for the first 30. The residual errors (shown in Fig. 2) range from a few percent at the lowest modes, and rise to 10% at mode 22. The calculated time constant is 30 minutes.

4. LABORATORY DEMONSTRATION

The principles of thermal actuation have been demonstrated with a single actuator implemented in a borosilicate honeycomb mirror. The 60-mm thick, 18″ diameter mirror (loaned to us by Richard Wortley of Hextek Corp) has 61 cells, is about 65 mm in diameter, and is polished as an optical flat. The backplate holes, in 4 cells near the center, were enlarged to 35-mm diameter for the experiment and radiative cooling fingers were installed to project into these holes, without touching the glass. These can be seen through the uncoated front face plate in Figure 3. The fingers are flat plates with the sides facing a Y shaped intersection of the hexagonal cell walls coated for high thermal emissivity; the outward facing sides are insulated to limit the cooling influence. Resistors were attached with thermally conducting silver epoxy

to the three ribs forming the central Y intersection. The temperature of the mirror surface was measured by a thermistor.

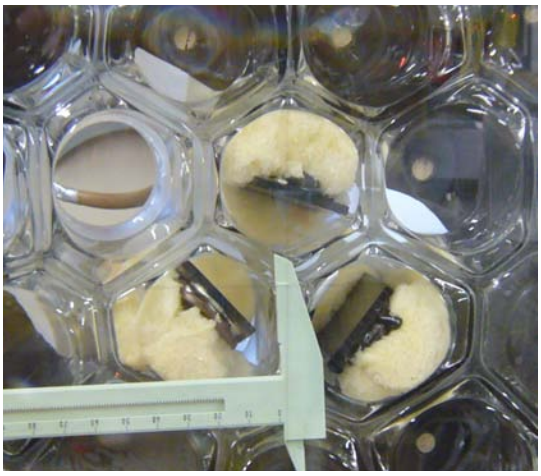


Figure 3. *Detailed view of honeycomb mirror cells with cooling fingers inserted through the hole in the back plate, seen through the face plate.*

In the experiment, the cooling mechanism consisted of ice water circulated through pipes soldered to the three cold fingers; heating was achieved by powering the resistors. The figure was measured with a 4D vision phase-shifting interferometer over a 145-mm diameter circular aperture, offset slightly from the test cell. Figure 4 shows a sequence of measurements of the surface figure. Initially the glass was in thermal equilibrium (all measurements are shown with reference to this equilibrium figure). The first frame of the equilibrium surface thus appears completely flat. The resistor power at this point was turned on at 200 mW, which produced a 150 nm peak surface actuation after 10 minutes (frame 5), with a surface temperature increase of 2.0°C. When ice water was circulated to the cooling fingers and the power dropped to 100 mW, the surface was restored to within 10 nm rms of its original equilibrium figure (frame 8). The power was then turned off, and after a further 6 minutes, a 130-nm depression caused by cold finger cooling was recorded in the last frame. From these measurements we were able to characterize the single cell influence function as having a full width at half maximum of 60 mm, a response time constant of around 6 minutes and an actuation rate of about 50 nm/°C. This rate can be compared with the expansion of a simple rod of borosilicate glass with the same expansion coefficient of 3.2×10^{-6}/°C and length of 60 mm (equal to the thickness of the mirror). This would expand ± 96 nm about its neutral center, if heated 1.0°C. The smaller expansion measured for the mirror faceplate arises because of the mechanical constraint of the neighboring unheated ribs and face sheets.

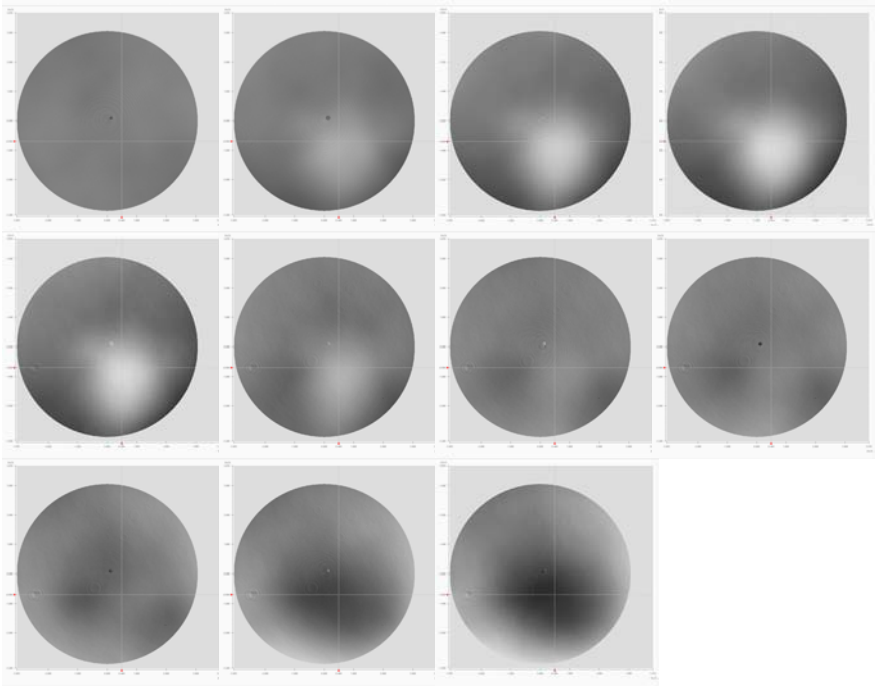


Figure 4. *Metrology of 145-mm diameter region covering one thermally controlled cell of a borosilicate mirror. The surface figure is shown by gray scale from -130 nm (black) to +150 nm (white). The actuator maximum corresponds to the 200 mW heating of three adjacent ribs, the minimum to the radiative cooling to the cold fingers at 0°C.*

5. FUTURE PLANS

We plan to validate the models and the control concept with a 0.4-m prototype honeycomb mirror with thermal control. The 61-cell blank will be made of borosilicate honeycomb by Hextek. It will be fitted with 250 heater elements and thermistors, before being figured as a 1.5-m radius sphere. The finned cold-fingers will be cooled individually with liquid from a refrigerated recirculator system (0.01°C control). Testing from the center of curvature will again be with a 4D Vision phase-shifting interferometer. Measurements will be taken of:

1) the control time constant
2) the power dissipation for initial steady flow state
3) the temporal stability of control
4) the power vs amplitude as a function of Zernike mode amplitude
5) the stability against external thermal perturbation

Good control up to the first 50 Zernike coefficients is projected, which (for the coronagraphic application) will allow the suppression of weak speckles out to ~ 4 λ/D in the focal plane. The larger system models will be refined on the basis of the prototype experience. Critical tests of the control accuracy for high contrast imaging will then be undertaken by combining the test mirror with coronagraphic optics. One test will be made with the AHA suppression system being developed by John Codona (these proceedings and Codona and Angel, 2006). Another will be with the PIAA optics recently demonstrated by Guyon and Pluzhnik (2006) to give suppression of the optical star halo to 10^{-6} at an angular separation of 2 λ/D. In this demonstration, the wavefront was corrected by a small MEMS deformable mirror. The prototype proposed here would be used to make a first test of achieving such contrast in a full system that includes a directly corrected primary mirror. At 40-cm diameter the primary would be far larger than the optics used in any previous high-contrast lab test, and the combined operation would make a major advance toward a practical coronagraphic space telescope.

This work is supported by NASA under grant NNG05GD28G.

REFERENCES

Codona, J. and Angel ((2006) ApJLetters, **604**, L117

Guyon, O and Pluzhnik, E (2006) these proceedings.

Hobbs, T. W., Edwards, M. & VanBrocklin, R. (2003) Current fabrication techniques for ULE and fused silical lightweight mirrors Proc SPIE **5179**, 1-11.

Lallo, M.D., Makidon, R.B. Casertano, S & Krist, J.E. (2006) *Temporal Optical behavior of JWST: focus, coma and astigmatism history* Proc SPIE **6270**, 54

Lardiere, O., Arnold, L. & DeJonghe, J. (2000) First light at the OVLA active mirror with its surface heating system Proc SPIE **4003**, 44

Vdovin, G. & Loktev, M. (2002) *Deformable mirror with thermal actuators* Optics Letters, **27**, 677–679

Thirty Meter Telescope Planet Formation Instrument[1]

B. Martin Levine[ac], Michael Shao[ac], Bruce Macintosh[ab], Mitchell Troy[ac], Rene Doyon[d], James Graham[ae], Kevin Baker[ab], Brian Bauman[ab], Christian Marois[ab], David Palmer[ab], Donald Phillion[ab], Lisa Poyneer[ab], Ian Crossfield[ac], Philip Dumont[ac], Gene Serabyn[ac], Chris Shelton[ac], Gautam Vasisht[ac], James K Wallace[ac], Jean-Francois Lavigne[d], Philippe Valee[d], Neil Rowlands[f], Ken Tam[f], and Daniel Hackett[f]

[a]NSF Center for Adaptive Optics
[b]Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551
[c]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109
[d]Universite de Montreal, Department de Physique, Montreal, QC, H3C 3J7, Canada
[e]Department of Astronomy, University of California at Berkeley, Berkeley, CA 94720
[f]Space Science Group, COM DEV, Suite 100, 303 Terry Fox Dr., Ottawa, ON, K2K 3J1, Canada

INTRODUCTION

Direct detection of extrasolar Jovian planets is a major scientific motivation for the construction of future extremely large telescopes such as the Thirty Meter Telescope (TMT). The instrument must be designed to meet specific scientific needs rather than a simple metric such as maximum Strehl ratio. The Planet Formation Imager (PFI) for TMT is a design for such an instrument. It has four key science missions: The first is the study of newly-formed planets on 5–10 AU scales in regions such as Taurus and Ophiucus—this requires very small inner working distances that are only possible with a 30 m or larger telescope. The second is a robust census of extrasolar giant planets orbiting mature nearby stars. The third is detailed spectral characterization of the brightest extrasolar planets. The final targets are circumstellar dust disks, including Zodiacal light analogs in the inner parts of other solar systems. To achieve these requirements, PFI combines advanced wavefront sensors, high-order MEMS deformable mirrors, a nulling coronagraph optimized for a segmented primary mirror, and an integral field spectrograph.[1]

SCIENCE MOTIVATION

Direct imaging of these planets, as opposed to detection of the effects of orbital motion on their parent star, is now feasible, and the first young planet in a wide orbit may have been detected using adaptive optics systems.[2]

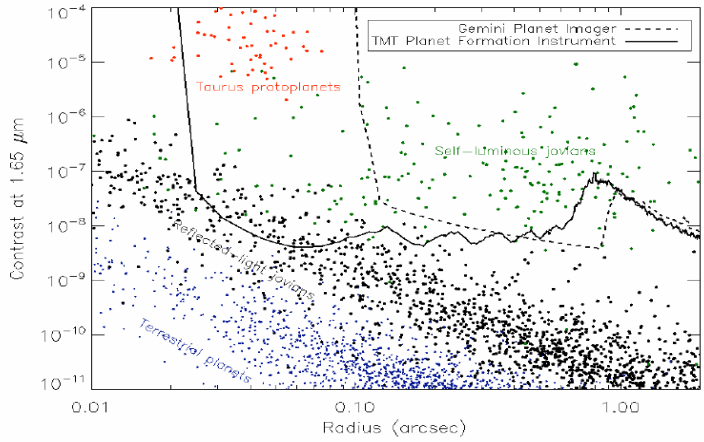


Figure 1: *Contrast-separation plot for a Monte Carlo simulation of a variety of targets in the solar neighborhood. Blue dots are rocky planets, beyond the reach of even TMT. Black dots are mature Jovian planets reflecting sunlight. Green dots are self-luminuous Jovian planets, typically those with masses of 3–10 Jupiter masses and ages < 1 Gyr. Red dots are extremely young planets, recently formed or still accreting, e.g., in the Taurus star-forming region. The expected sensitivity of PFI and the Gemini Planet Imager for a bright (4th magnitude) target are overlaid.*[1]

[1] This note is an extract taken from the publication: Bruce Macintosh, et al., "Extreme Adaptive Optics for the Thirty Meter Telescope", Proc. SPIE Vol. 6272, Advances in Adaptive Optics II, B. Ellerbroeck, ed. (2006). This paper was presented at the Coronagraph Workshop 2006, 28–29 September 2006, Pasadena, CA, JPL Publication 07-02, W.A.Traub, editor (2007).

Gemini and the VLT are building the first generation of high contrast "extreme" adaptive optics (ExAO) systems, which deliver planet-imaging performance at separations > 0.1 arcseconds. These systems will make the first surveys of the outer regions of solar systems by detecting the self-luminous radiation of young planets (1–10 MJ, 10–1000 MYr), but the 8-m ExAO systems cannot see close enough to the host stars to image Doppler-detected or other mature planets in reflected light, and they cannot reach the relatively distant, young clusters and associations where planets are likely to still be forming. The PFI will use the nearly four-fold improved angular resolution of TMT to peer into the inner solar systems of planet bearing stars to yield a unified sample of planets with known Keplerian orbital elements and atmospheric properties. In star formation regions, where T Tauri stars (young solar type stars) are found in abundance, PFI can see into the snow line, where the icy cores of planets like Jupiter must have formed. Thus, TMT could be the first facility to witness the formation of new planets directly. Because of the short lifetimes relative to the Galactic star formation rate, young planet-forming systems are rare and found in significant numbers only in distant (> 150 pc) star forming clouds. The inner working distance required to study planet formation in situ is therefore of order 35 milli arc seconds (5 AU at 150 pc). Since a typical coronagraph has an inner working distance >3–5 λ/D, it is evident that the TMT will be the first facility to enable direct observation of planets emerging from their parent discs. Likewise, to detect nearby mature planets in reflected light, a comparable angular resolution will be needed (30 mas = 0.3 AU at 10 pc). TMT will thus also be the first facility to be able to directly detect a sizable number of reflected light Jovian planets. The unique combination of angular resolution and sensitivity of TMT will thus enable direct images and spectra to be obtained for both young and old planets (Fig. 1). The main instrumental capability needed to take advantage of the TMT, in this regard, is high contrast imaging at an angular separation of a few λ/D from bright stars in the near infrared. This goal is the basic driver for the instrument described here, the Planetary Formation Instrument (PFI).[1]

OVERALL ARCHITECTURE

Current AO systems achieve contrasts on the order of 10^{-5} at angular separations of ~ 1 arcsecond; TMT PFI requires a three order of magnitude improvement in contrast and a factor of twenty in angular separation. The combined requirements of high dynamic range, a wide variety of target brightnesses, very high angular resolution and the need to minimize systematic errors, lead to a multi-stage integrated instrument. A high-speed front AO system, optimized for searching for planets, and achieving extremely high contrast ($>10^{-8}$) on bright targets, requires very high update rates (2–4 kHz) to minimize dynamic atmospheric errors. We have selected as our baseline a variant of the pyramid wavefront sensor run in a quasi-interferometric mode to take advantage of the high Strehl ratio at the wavefront sensing wavelength to achieve measurement errors a factor of 2–4 better than a conventional Shack-Hartmann sensor. Combined with a compact, high-order MEMS DM this system will produce H-band Strehl ratios above 0.9 on bright stars and 0.84 down to I = 9 mag. On dimmer stars, it will provide partial correction to enable the back (interferometric or calibration) wavefront sensor to provide the bulk of the wavefront correction. Unless controlled, light scattered by diffraction from the telescope pupil would completely swamp the signal from a planet. After considerable exploration of alternatives using simulations, we have selected a dual-stage shearing nulling interferometer or "nuller" as the diffraction suppression system (DSS). This combines four offset and phase-shifted copies of the telescope pupil to remove the uniform component of the electromagnetic field that causes diffraction. This has two major advantages over conventional coronagraph architectures. First, it allows for very small inner working angles (IWA)—as small as 3 λ/D—that are needed both to detect nearby planets in reflected starlight and to image young solar systems at distances as great as 150 pc. Second, it is robust against the large secondary obscurations likely to be found in any extremely large telescope. The large aperture of TMT allows PFI to exceed the contrast of 8-m ExAO systems by an order of magnitude, but does nothing to relax the fundamental requirement of sub-nanometer internal static optical errors; since the effect of internal errors is partially independent of telescope size, these requirements are even more exacting to reach the 10^{-8} contrast levels that we are targeting. To overcome this, we will use a dedicated, interferometric (calibration), wavefront sensor that is tightly integrated with the DSS, and controls a second MEMS DM located inside the nuller. Known as the back wavefront sensor, this system is a Mach-Zehnder interferometer combining the bright and dark outputs of the nulling DSS. It has two primary purposes. First, operating at the science wavelength and measuring the dark output of the DSS, it will provide sub-nanometer

absolute accuracy and correct these errors through feedback to the front AO or through its own MEMS. Second, especially on very red science targets, it will provide additional rejection of atmospheric turbulence, allowing PFI to reach contrast of 10^{-6} to 10^{-7} on H = 10 mag. young stars.[1]

PFI's science instrument is a dedicated Integral Field Spectrograph unit (IFS or IFU) optimized for high-contrast imaging, maximum scientific return, and for spectral follow-up of extrasolar planets. This combines moderate spectral resolution ($R = \lambda/\Delta\lambda \sim 70$) with a 2 × 2 arc second Nyquist-sampled field of view and spectral coverage from 1–5 μm. Carefully designed to minimize chromatic errors, this allows planets to be distinguished from artifact speckles by their differing behavior as a function of wavelength. The instrument also includes an R = 500 mode that can be used for followup of previously-detected planets. A dual-channel imaging polarimeter mode will be available to study circumstellar disks and distinguish disk structure from the planets embedded in them. Figure 2 shows a block diagram of the PFI system.[1]

PERFORMANCE SIMULATIONS

We ran a series of end-to-end simulations of PFI incorporating all the subsystems except the science instrument (only monochromatic PSFs were generated.) Typical simulation exposure times were 2–4 seconds, sufficient for 1–4 atmospheric speckle lifetimes[4]. These simulations correctly predict contrast due to atmospheric effects (including scintillation and propagation chromaticity) but are too short to show the effects of static wavefront

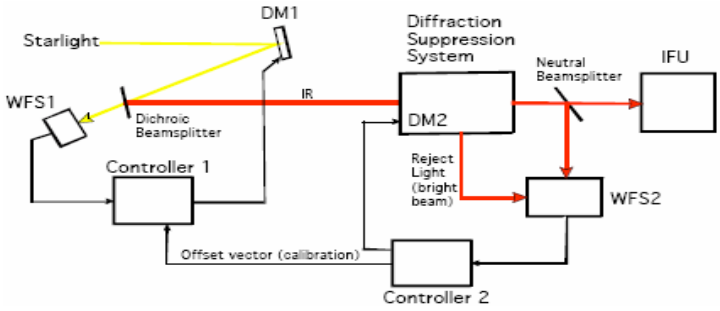


Figure 2: *Block diagram of the TMT Planet Formation Instrument*[1]

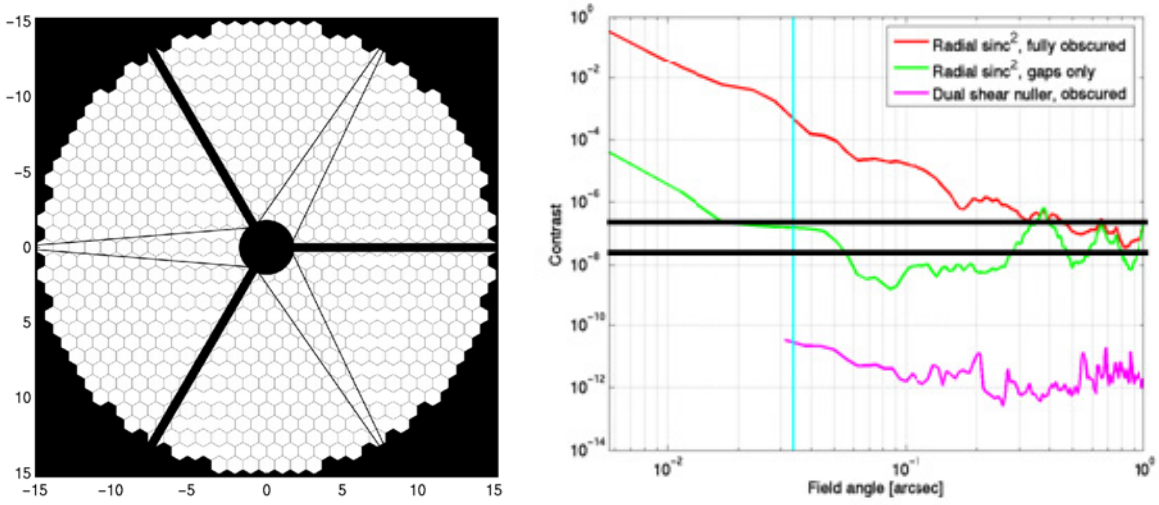


Figure 3: *Left, the segmented and obscured TMT pupil*[1]*. Right, results of achievable contrast from PFI due to the pupil as compared to other coronagraph types*[3]*. The segmented primary mirror does not adversely effects the resultant contrast, however, the obscuration grossly affects performance. The nulling coronagraph suffers no such loss of contrast.*

errors. Separate simulations of static effects were used to evaluate their magnitude and set instrument requirements. If these are met, sensitivity can then be extrapolated to 1–2 hour exposures. Additional post-processing suppression of speckle noise by a factor of 5–10 should be possible by the calibration wavefront sensor.[1]

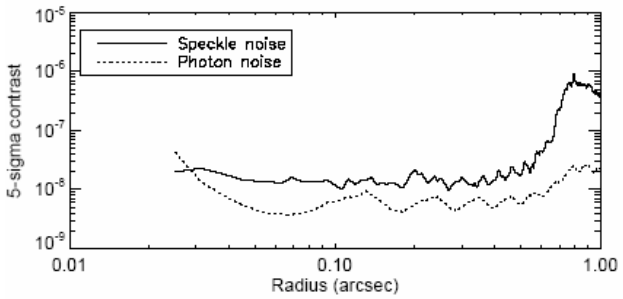


Figure 4: *The speckle and photon noise contrast is shown as a function of field angle for a G5 star at 10 pc. The simulation was run for 1.5 seconds and the results scaled to obtain the contrast in 2 hrs, assuming systematic errors are controlled.*[1]

SUMMARY

We have identified key science missions—the study of mature planets and of young protoplanets in starforming regions—in which future 30-m-class telescopes will have unique capabilities, in particular their small inner working angle. We have produced a science-driven design, practical using near-term technologies, that can carry out these science missions, with the emphasis on performance at 3–5 λ/D. Such an instrument will build upon the technological and scientific heritage of the 8-m class ExAO systems such as the Gemini Planet Imager, and will produce both a uniform census of extrasolar planets and the first images and spectra of planets in the process of formation.[1]

REFERENCES

1. Bruce Macintosh, et al., “Extreme Adaptive Optics for the Thirty Meter Telescope”, Proc. SPIE Vol. 6272, Advances in Adaptive Optics II, B. Ellerbroeck, ed. (2006).
2. G. Chauvin, et al., “A giant planet candidate near a young brown dwarf. Direct VLT/NACO observations using IR wavefront sensing”, A&A 425, L29 (2004)
3. M. Troy, et al., “Effects of diffraction and static wavefront errors on high-contrast imaging from the thirty meter telescope”, Proc. SPIE Vol. 6272, 62722C, Advances in Adaptive Optics II, B. Ellerbroeck, ed. (2006), and presentation viewgraphs therein.
4. B. Macintosh, et al., “Speckle lifetimes in high contrast optics”, Proc. SPIE, Vol. 5903, 170, (2005).
5. D. Phillion, et al., “Two-sided pyramid wavefront sensor in the direct phase mode”, Proc. SPIE Vol. 6272, 627228, Advances in Adaptive Optics II, B. Ellerbroeck, ed. (2006).
6. G. Vasisht, et al., “Post-coronagraph wavefront sensing for the TMT Planet Formation Imager”, Proc. SPIE Vol. 6272, 627253, Advances in Adaptive Optics II, B. Ellerbroeck, ed. (2006).

Shaped Pupil Coronagraph: State of the Art and Projections for TPF Performance and Readiness

N. J. Kasdin (PI)[1], R. Belikov[1], E. Cady[1], M. Carr[1], J. Kay[1], M. Littman[1], L. Pueyo[1], R. J. Vanderbei[1], J. T. Trauger (PI)[2], K. Balasubramanian[2], P. Echternach[2], A. Give'on[2], A. Kuhnert[2], S. Shaklan[2], F. Shi[2], A. R. Neureuther (PI)[3], D. Ceperley[3], M. Miller[3], T. Shih[3], S. Kilston[4], M. Lieber[4], and J. Beall[5]

[1]Princeton University, [2]Jet Propulsion Laboratory, California Instiute of Technology, [3]Univeristy of California, Berkeley, [4]Ball Aerospace, and [5]NIST

ABSTRACT

The Shaped Pupil Coronagraph (SPC) is a high-contrast imaging system pioneered at Princeton and designed for the TPF-C telescope. In this document, we summarize the work done to date on the SPC to date and evaluate its current and projected performance. What makes the SPC attractive for TPF is that it is very simple to make and set up, and it is inherently broadband. Owing to the simplicity of the SPC, it is quickly becoming a relatively mature technology with theoretical and experimental validations of its performance. Many shaped pupils have been designed to various specifications and tools are in place to quickly turn out more. Full vector-field simulations show that realistic shaped pupils can already achieve 10^{10} contrast in the absence of aberrations. A manufacturing process has been developed to make shaped pupils for as little as a few thousand dollars, at JPL and NIST. Shaped pupils have also been shown to be very insensitive to aberrations, and especially low order aberrations such as tilt and defocus. The SPC is undergoing extensive studies in the lab, and so far a suppression of 4×10^{-8} has been achieved in 10% broadband light (averaged across a region between 4 and 9 λ/D), after speckle-nulling-based wavefront correction. The limiting factor is now believed to be well-understood and is primarily the inability of the speckle nulling algorithm to correct for manufacturing errors in the mask. It was shown that this limitation can be overcome by using a more sophisticated estimation algorithm called peak-a-boo, or by using a shaped pupil design that is insensitive to manufacturing defects. The SPC lends itself well to many wavefront estimation and correction schemes. Simulations show that realistic shaped pupil manufacturing errors and realistic wavefront error can be corrected with a single DM at one wavelength, and 2 or 3 DMs in broadband. The main disadvantages of the SPC is throughput, sharpness, and working angle, but the throughput disadvantage may be counterbalanced to an extent by the fact that SPC requires very few optical components and the fact that the light blocked by the mask may still be used to sense aberrations.

INTRODUCTION

The Shaped Pupil Coronagraph (SPC) is a high-contrast imaging system under development at Princeton, designed for NASA's Terrestrial Planet Finder Coronagraph (TPF-C) mission[1]. The SPC is a type of an apodized-pupil coronagraph, where high contrast is achieved simply by apodizing the telescope pupil by a binary mask (the shaped pupil). Many shaped pupil designs can be found with high-contrast point spread functions (PSFs) that meet the TPF requirements of 10^{10} contrast at an inner working angle of 4 λ/D. Because the only required component in an SPC is a shaped pupil mask, the SPC is very simple and cheap to make, trivial to align, and is inherently broadband. Lab validations of the SPC are progressing and to date have shown it to be relatively close to the TPF requirements in broadband light. The main disadvantages of the SPC are its theoretically low throughput and limited discovery space. However, any considerations of throughput must also take into account practical issues such as number of required optical elements, bandwidth, and whether the "lost" photons can still be used. The SPC in principle only requires one optical element: the primary, it is inherently broadband, and the photons blocked by the shaped pupil can actually be utilized to help estimate aberrations.

Therefore, the SPC's low theoretical throughput may in fact be counterbalanced by gains in this practical throughput.

A rough schematic of the baseline option of the TPF-C telescope with an SPC and a wavefront correction system is shown in Figure 1. The system consists of the TPF telescope, followed by the wavefront correction system and the SPC. The wavefront correction system can be designed to a large extent independently of the SPC, so we do not show its details here. The SPC consists of the shaped pupil at a reimaged pupil plane, and the science camera with an optional star occulter if the camera itself is incapable of properly accepting light intensity with a 10^{10} dynamic range. The star occulter is a hard-edged mask designed to block the brightest portions of the PSF. If the star occulter cannot be placed very close to the chip, the science camera can be placed at a reimaged image plane downstream from the star occulter. Aside from imaging planets, the science camera will also be used to estimate aberrations and feed this estimation data back to the wavefront correction system (additional cameras can also be used for this purpose).

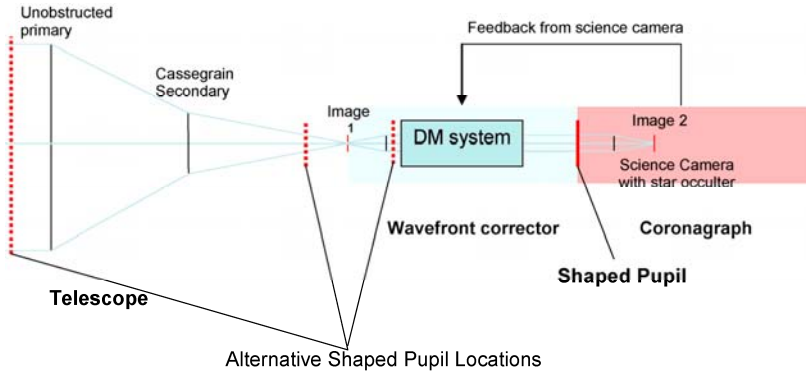


Figure 1: *Simplified schematic diagram of the TPF telescope with a shaped pupil coronagraph. The shaped pupil can be placed in virtually any "pupil" plane*

Even though our baseline position for the shaped pupil is as shown by the solid line in Figure 1, the shaped pupil can be located in many other planes, such as those indicated by the dotted lines. In fact, one attractive design is a single-mirror prime-focus telescope with a shaped pupil either in front of the primary or along the converging beam downstream. Such a prime-focus design is attractive not only because it is so simple, but also because it eliminates light losses and aberrations introduced by all optics except for the primary. (In fact, this design has no phase-induced amplitude or similar aberrations.) However, the prime-focus design may not be practical because it requires the ability to shape the primary to be free of aberrations and a few other technical challenges.

SHAPED PUPIL DESIGN AND MANUFACTURE

Shaped pupils are designed by specifying contrast as well as the areas in the PSF where this contrast is to be achieved (dark zones), and running nonlinear optimization techniques on the pupil shapes to optimize throughput[2]. Many pupil designs and design types were created over the years (see, e.g., Vanderbei et al. 2004[3]), but we have identified the so-called "ripple" design type as being the easiest to manufacture as a free-standing mask. Some of the designs we manufactured (by NIST and/or JPL[4]) are shown in Figure 2. The manufacturing process was deep reactive ion etching of silicon wafers. Mask were manufactured in both 10 and 25mm diameters. Ripple 1 was designed for the elliptical shape of TPF-C while ripple 3 was designed for a circular pupil. As could be seen by the horizontal cross-sections in the bottom row of Figure 2, both were designed to achieve better than 10^{-9} contrast at an inner working angle of $4\lambda/D$. Both have about 10% Airy throughput (i.e., the fraction of the input beam power that falls in the main lobe of the PSF). As could be seen from the PSFs in Figure 2, the dark zone for ripple 1 has 45° openings on two sides of the main lobe, and ripple 3 has 90° openings. As mentioned before, the theoretically low throughput of the shaped pupils can be compensated for to some extent by the fact that the SPC does not require any optics in addition to the shaped pupil, which eliminates losses due to reflections; and the fact that the SCP is inherently broadband, so that any photon losses due to narrow-band filtering are eliminated.

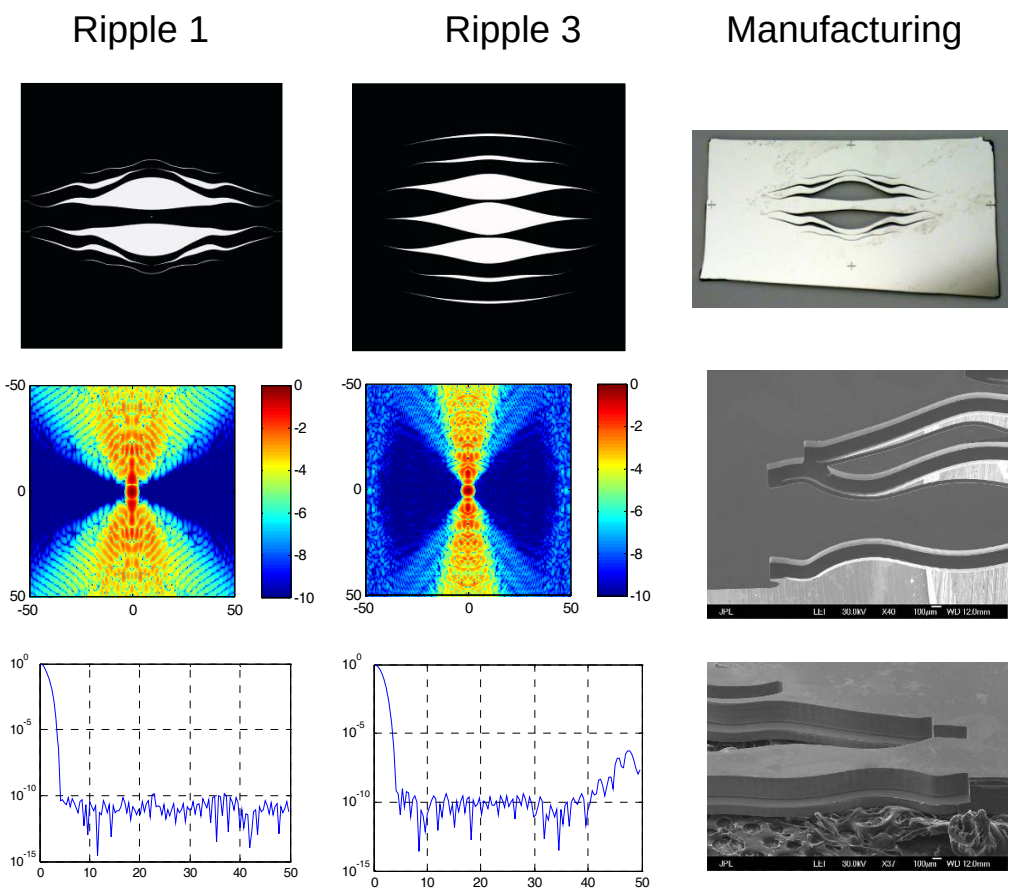


Figure 2: *Left two columns: examples of shaped pupil designs. Top row is the shaped pupil (white is transmissive); second row is the PSF, showing the high-contrast regions in blue; and third row is the horizontal slice through the PSF. All axes are in units of λ/D. Right column: microscope pictures of a manufactured mask.*

MASK SIMULATIONS

The shaped pupils are designed assuming that the masks will be ideal and that the image plane is an exact Fourier Transform of the pupil plane. One important question is whether these assumptions are valid in practice. In this section, we present the results of full vector simulations that show that in most cases of interest, these assumptions are indeed valid.

Ceperly et al.[5] have performed finite-difference-time-domain simulations of the interactions between the vector electro-magnetic fields and the edges of a realistic mask (i.e., real material, non-0 thickness) and compared the results to the assumption of an ideal mask (i.e., ideal apodization, 0 thickness). The main results are summarized in Figure 3. The left column specifies various cases that were simulated and the right column shows the magnitude of the difference between a real and an ideal mask (taking the worst case for the last 5 rows). By "vertical sidewalls", we mean the case where the edges of the mask have walls that form a 90 degree angle with respect to the substrate. By "20° undercut", we mean that this angle is 90°–20° (with the mask opening being narrower on the illuminated side of the mask). "Xμm opening" refers to the width of the mask opening in cross section. The value of the parameter on the right column (severity) corresponds to the width along the edge of the ideal mask that contains the same amount of energy as the perturbation due to the effect on the left column. The main conclusions of these simulations are that vertical sidewalls are to be avoided, and that small mask tilts, changes of polarization and wavelengths, and metal coatings are the least significant.

Mask Structure:	Severity:
Vertical sidewalls, Off-axis illumination, 10um Opening, 100um Thick	7 λ
Vertical sidewalls, Off-axis illumination, 10um Opening, 50um Thick	4 λ
Vertical sidewalls, On-axis illumination, 10um Opening, 100um Thick	3–4 λ
Vertical sidewalls, On-axis illumination, 10um Opening, 50um Thick	3 λ
Vertical sidewalls, On-axis illumination, 48um Opening, 50um Thick	3 λ
Vertical sidewalls, On-axis illumination, 48um Opening, 50um Thick	2–3 λ
20° undercut sidewalls, On-axis illumination, 48um Opening, 50um Thick	λ / 4
Physical effects on 50um thick Si masks with 48um openings at 630nm:	Worst case:
Undercut angle	3 λ
Small mask tilt	λ / 2 per degree tilt
Polarization	λ / 4
Wavelength (630nm to 785nm)	λ / 4
200nm Cr top-coat	λ / 100

Figure 3: *Extent of different effects on the fields around the edges.*

This data was used to study how vector interactions with the real mask affect the PSF of the real mask (Lieber et al[6]). The main conclusions were that severities of less than λ/4 do not degrade the 10^{10} contrast for mask sizes 10 cm and larger. Therefore, with a properly undercut mask, the mask can be assumed ideal and insensitive to polarization, wavelength changes, or coatings. For masks that are not undercut, the degradation depends on the mask type and size. In a 10 cm ripple mask, contrast is degraded to 10^{9}, while for a 10 cm checkerboard mask, contrast is not degraded. All such effects will be insignificant compared to wavefront error. One important corollary to this work is that manufacturing error or imprecision of the edges as large as λ/4 is insignificant (at least to the extent that manufacturing errors along the edges are similar to effective errors due to vector effects). Such an error is well within modern manufacturing tolerances. The results presented in this section indicate that real masks can be manufactured today which, in the absence of wavefront aberrations, will achieve the 10^{10} contrast required by TPF (except for the trimming limitation discussed later, which can be overcome by specialized mask designs).

SENSITIVITY TO WAVEFRONT ABERRATIONS

In order to assess the sensitivity of shaped pupils to wavefront aberrations, we simulated different Zernike aberrations at the pupil plane and studied their effects at the image plane[7,8]. Some of these results are summarized in Figure 4. For these simulations, we assumed a circular concentric ring shaped pupil. Each subplot corresponds to a different Zernike (starting from tilt, defocus, etc.), and shows the contrast in the image plane at a particular working angle (4 or 8 λ/D) as a function of RMS of aberration. Just as with many other coronagraphs, the sensitivity is more pronounced at lower inner working angles and higher Zernike orders. The main conclusion of this work, however, is that the SPC is very insensitive to aberrations (especially to pointing error or other forms of tilt, defocus, and astigmatism). The wavefront flatness requirement depends on the aberration, but assuming typical power spectral density fall-off such as $1/f^3$, the total flatness requirement needs to be on the order of λ/10,000.

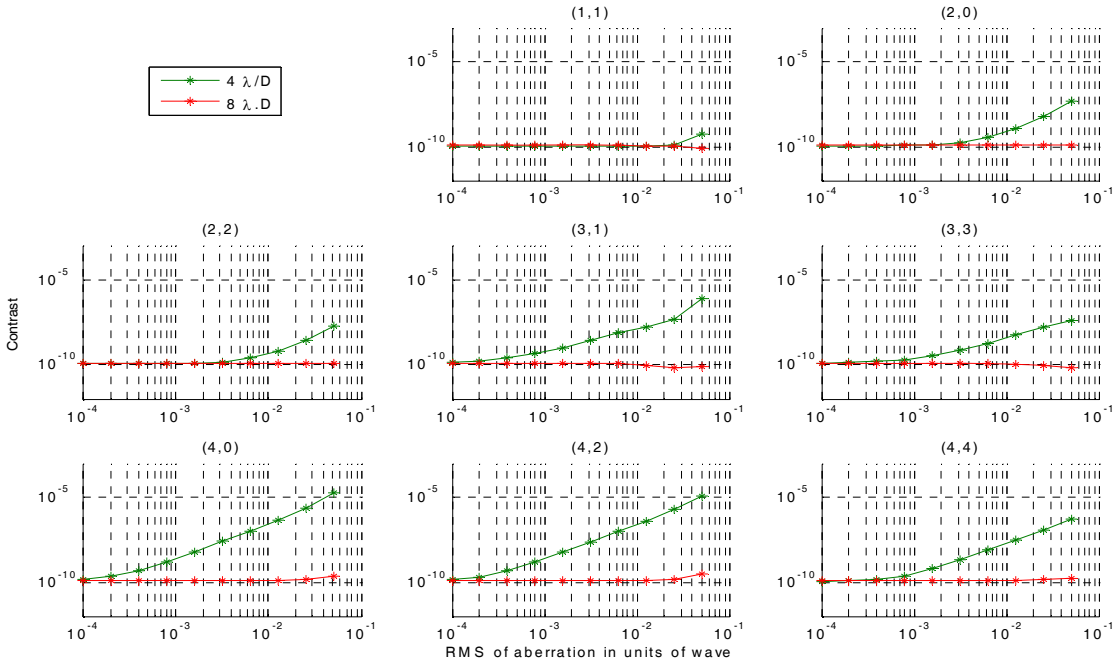


Figure 4: *Sensitivity to different Zernike aberrations for a concentric ring shaped pupil mask as a function of RMS of aberrations, at inner working angles of 4 and 8 λ/D. The title of each graph is the Zernike radial and azimuthal index.*

LABORATORY RESULTS

Experimental validation of shaped pupils is currently under way at Princeton and JPL[9,10]. The Princeton testbed is relatively new and low cost, consisting of an optical bench with an enclosure in a semi-clean room, and 2 recently acquired Boston Micromachines DMs. We are currently getting contrast levels of about 10^5 without wavefront correction and about 10^6 with wavefront correction. Both values hold for the largest bandwidth we tried, ~450–750 nm, demonstrating the broadband advantage of shaped pupils. Simulations indicate that the contrast-limiting factor for us is twofold. One, we are using a small mask (10 mm) because of the small size of our DM. Any manufacturing errors are proportionately greater for smaller masks, and in our case they appear at about 10^{-6} level. Two, the correction algorithm we are currently using, speckle nulling, is incapable of correcting for errors of this kind. Hence, we expect to be able to reach much better contrast if we either (a) adopt our setup to be able to accommodate a larger mask given the small DM without degrading contrast; (b) design a mask more tolerant to manufacturing limitations, or (c) use a more sophisticated correction algorithm such as peak-a-boo[11]. We are currently pursuing all three, with a focus on (c). Simulations show that peak-a-boo is capable of correcting for errors due to manufacturing defects which are beyond the ability of speckle nulling.

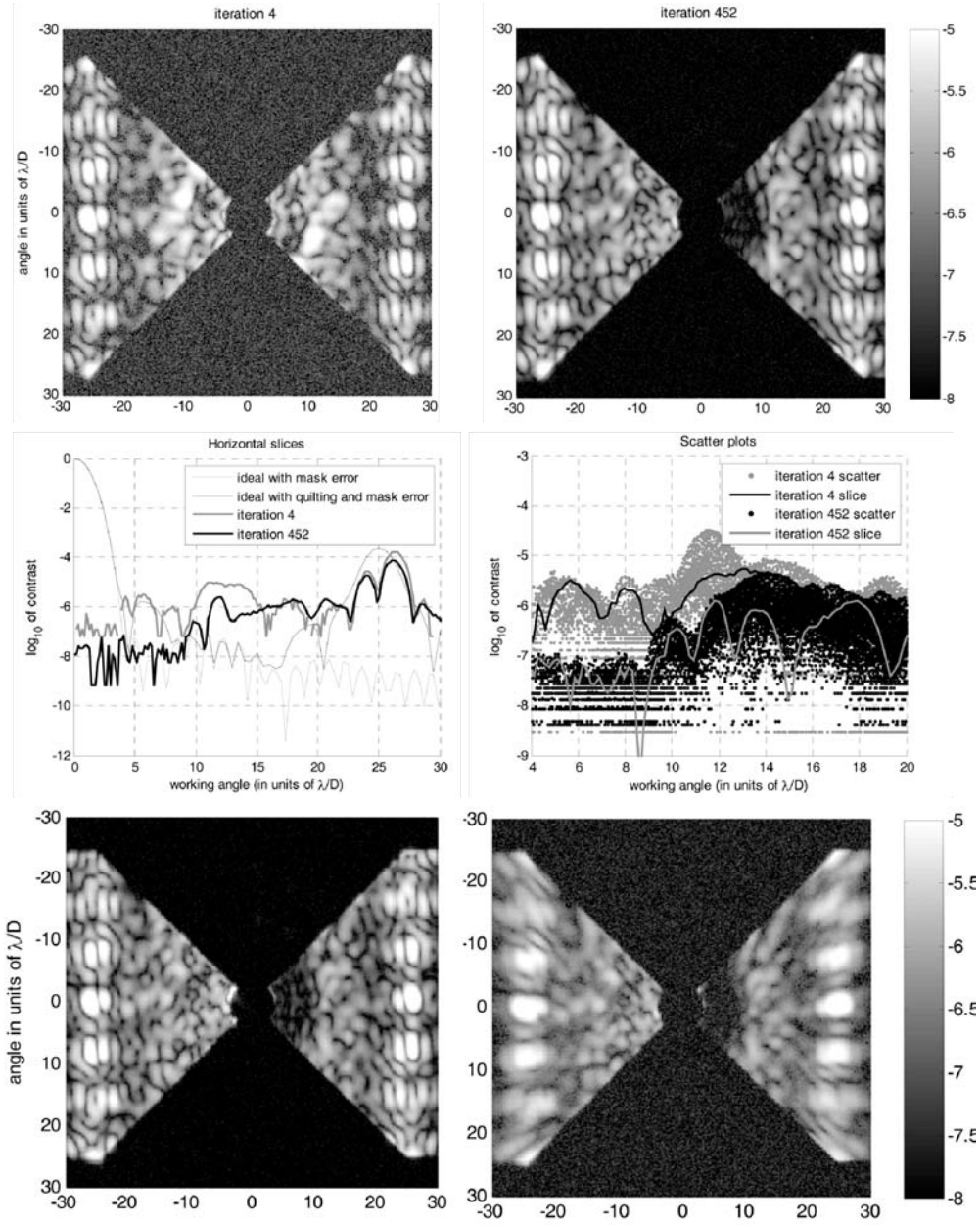


Figure 5: *Contrast plots from tests at the HCIT. Top left: one of the starting speckle nulling iterations (iteration 4) at 785 nm. Top right: one of the final iterations (iteration 452). The higher amount of noise in the image of iteration 4 is due to lower exposure time. Middle left: horizontal traces through the centers. Middle right: scatter plots in the dark region. Bottom left: after changing from 785- to 836-nm light. Bottom right: after changing to broadband light. For the λ/D scale, λ = 785 nm and D = 25 mm.*

JPL's testbed (HCIT), on the other hand, is somewhat more sophisticated than the one at Princeton, consisting of a vacuum chamber in a clean room and a having a large DM[12]. Owing to this, our best results to date are from an experiment performed at HCIT with a 25-mm shaped pupil, where we have achieved suppression of 4×10^{-8} with speckle nulling. Simulations show that this is about the level of contrast due to manufacturing limitations (trimming) for the particular shaped pupil we used. Work is progressing on a design and manufacturing process capable of yielding a mask with no significant manufacturing defects.

Our experimental setup at HCIT was designed to take an image of the PSF of the shaped pupil. A DM was present in a pupil plane upstream of the shaped pupil to correct for aberrations. A bowtie-shaped star occulter was used in an image plane upstream of the camera to prevent a lot of bright light hitting the CCD. Speckle nulling was performed at 785 nm and then the wavelength was switched to 836nm and to 760–840 nm broadband light. The results are shown in Figure 5. Without any corrections, the contrast on the HCIT testbed is similar to the Princeton testbed, or 10^5 at 4 λ/D (the HCIT system was already slightly corrected when we started, so that the starting point is closer to 10^6). After wavefront correction, the suppression is 4×10^{-8} between 4 and 9 λ/D (averaged across the dark region). As can be seen from Figure 5, this contrast is maintained after switching to either 836nm or broadband light, once again demonstrating achromaticity of shaped pupils.

WAVEFRONT SENSING AND CONTROL

Speckle nulling is a very robust and simple method of wavefront correction, but it has certain limitations, notably speed and inability to correct for certain kinds of aberrations. More sophisticated methods are necessary for use in space. Many correction methods are being developed for or can be adopted for use with the SPC. The hardware necessary for various methods is outlined in Figure 6, along with a few notes on each method. The simplest methods are those that involve the use of the science camera, in conjunction with algorithms such as Peak-a-boo[11] or Borde and Traub's method[13]. One disadvantage of these, however, is that speckles in the image plane are blurred in broadband light, and hence the performance suffers with bandwidth. An alternative that is not sensitive to bandwidth is to image the pupil plane after the star occulter. The price is additional complexity and possible slight loss of throughput. However, both of these methods work with light after the star occulter and, therefore, do not get a lot of photons. Guyon proposed that light blocked by the star occulter can be used to estimate low-order aberrations. A variation on this theme is to use the light reflected off the shaped pupil. Both of these schemes can be used in addition to the first two mentioned.

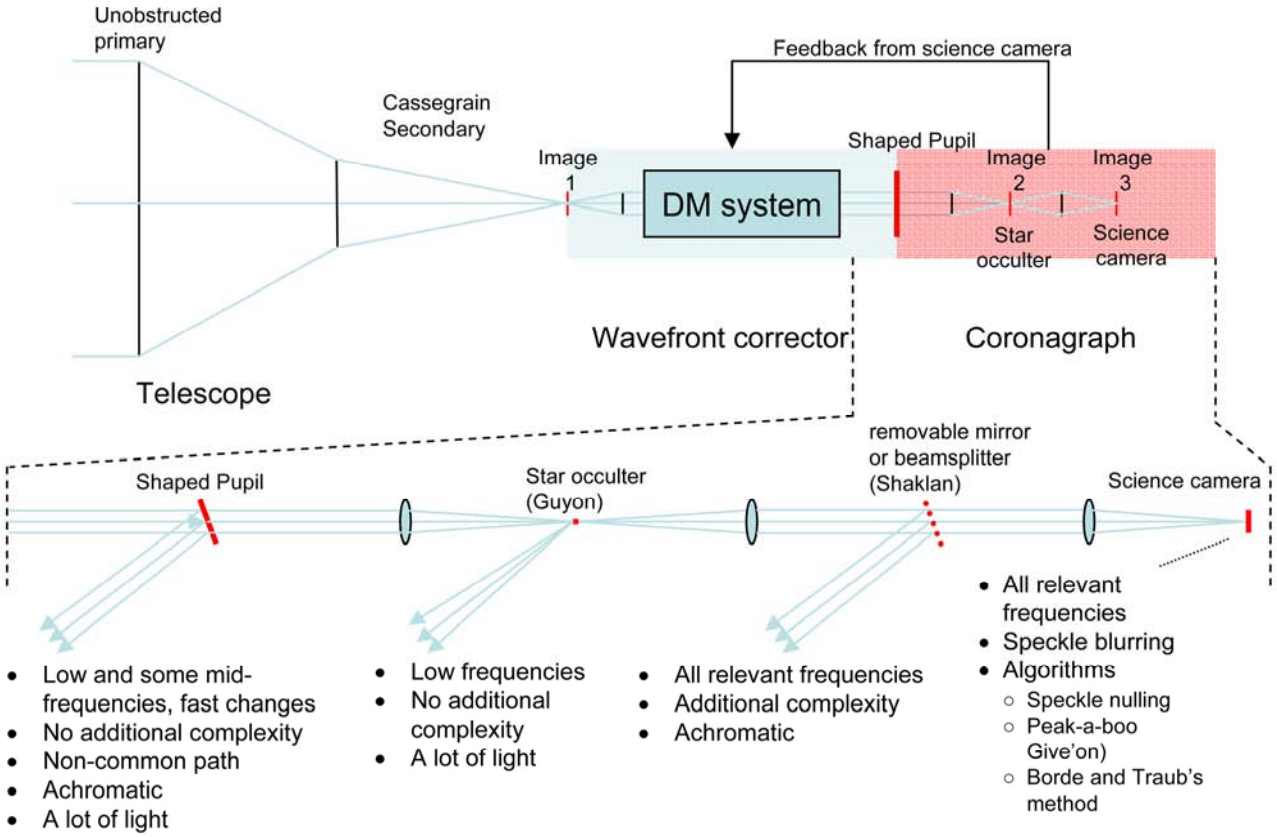


Figure 6: *Schematic diagram of wavefront sensing options.*

The ability to sense and correct for aberrations all the way to 10^{10} contrast in a SPC system that has real masks like those already manufactured today has already been shown theoretically in monochromatic light. Figure 7 shows a simulation of Peak-a-boo estimation and subsequent correction (assuming an ideal DM), in a system having phase and amplitude errors as well as the observed manufacturing errors (trimming) on the mask.

Such corrections work well for monochromatic light, or even in polychromatic light assuming no wavelength dependence of aberrations. However, in reality, both phase and amplitude error will have a non-negligible dependency on wavelength, due to a variety of mechanisms[15], severely limiting the bandwidth across which aberrations can be fully corrected. Nonetheless, in such cases, the required 10^{10} contrast can still be achieved with the SPC if one uses more than 1 DM for correction. Figure 8 shows some possible configurations of 2 and 3 DMs designed to correct for wavelength-dependent errors along with a simulation of correcting for such errors with a 3-DM correction system.

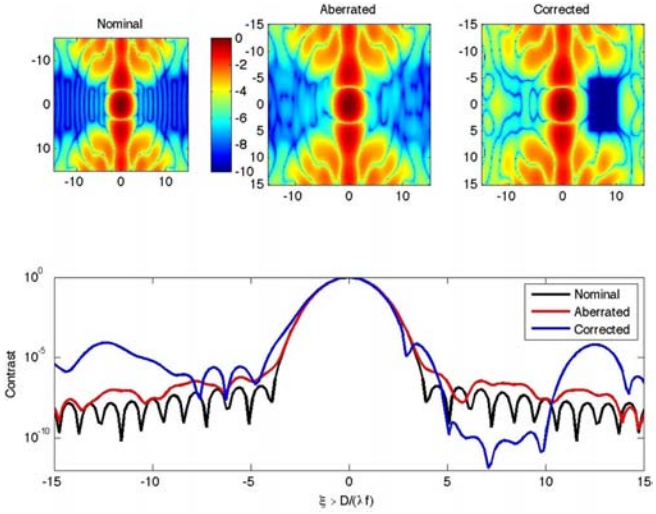
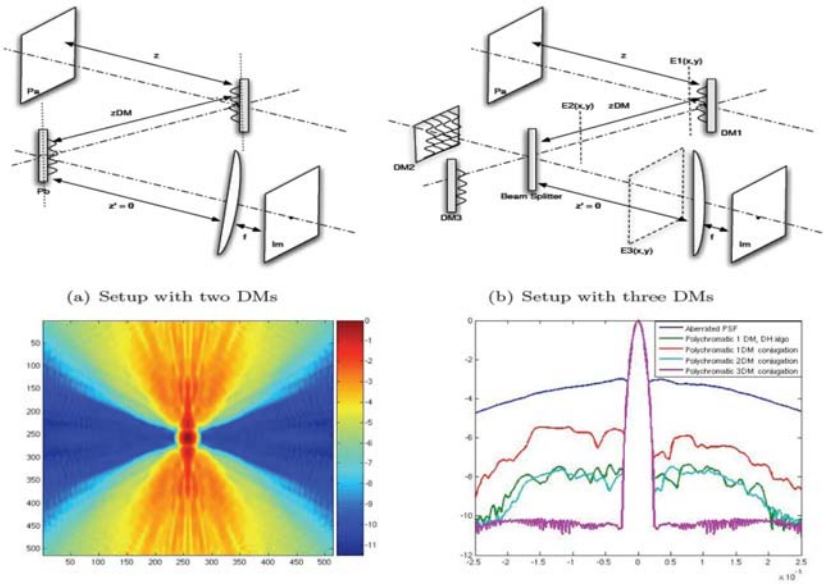


Figure 7: *Demonstration of Peak-a-boo sensing and correction with a realistic shaped pupil having manufacturing defects (trimming), in monochromatic light.*



(a) Setup with two DMs (b) Setup with three DMs

(a) Polychromatic correction with three DM $\Delta\lambda = 400$ nm (b) Slices of the several polychromatic PSFs presented above $\Delta\lambda = 400$ nm

Figure 8: *Top: Multi-DM configurations that allow corrections for realistic wavelength-dependent errors. Bottom: Even a severe case where wavelength-dependent aberrations limit the contrast to 10^3 can be fully corrected by a system of 3 DMs. In this simulation, perfect estimation of the aberrations is assumed.*

ACKNOWLEDGEMENTS

The Princeton group gratefully acknowledges the support of the National Aeronautics and Space Administration through the Jet Propulsion Laboratory, as well as the Michelson Science Center, California Institute of Technology for this work.

REFERENCES

1. Kasdin, N. J., Belikov, R., Beall, J., Vanderbei, R. J., Littman, M. G., Carr, M., & Give'on, A., "Shaped Pupil Coronagraphs for Planet Finding: Optimization, Manufacturing, and Experimental Results," *Proc. of SPIE* 5905, G1-G8, (2005).

2. Kasdin, N. J., Vanderbei, R. J., Spergel, D. N., & Littman, M. G., "Extrasolar Planet Finding via Optimal Apodized and Shaped Pupil Coronagraphs," *Astrophysical Journal* 582, 1147-1161, (2003).

3. Vanderbei, R. J., Kasdin, N. J., & Spergel, D. N., "Checkerboard-Mask Coronagraphs for High-Contrast Imaging," *Astrophysical Journal* 615(2004).

4. Balasubramanian, K., et al., "Fabrication and Characteristics of Free Standing Shaped Pupil Masks for TPF-Coronagraph," *Proc. of SPIE*, 6265 (2006).

5. Ceperley, Daniel; Neureuther, Andrew; Miller, Marshall; Lieber, Michael; Kasdin, Jeremy, "Stray-light sources from pupil mask edges and mitigation techniques for the TPF Coronagraph," *Proc. of SPIE*, 6271(2006).

6. Lieber, Mike; Neureuther, Andrew R.; Ceperley, Dan; Kasdin, Jeremy; Hoppe, Dan; Eisenman, Allan, " Evaluating the end-to-end performance of TPF-C with vector propagation models: Part I. Pupil mask effects," *Proc. of SPIE*, 5906 (2005).

7. Belikov, R., Kasdin, N. J., Vanderbei, R. J., "Diffraction-based Sensitivity Analysis of Apodized Pupil-mapping Systems," *Astrophysical Journal* 652 (2006).

8. Shaklan, Stuart B.; Green, Joseph J., "Low-Order Aberration Sensitivity of Eighth-Order Coronagraph Masks," *Astrophyciscal Journal* 628 (2005).

9. Belikov, Ruslan; Give'on, Amir; Trauger, John T.; Carr, Michael; Kasdin, N. J.; Vanderbei, Robert J.; Shi, Fang; Balasubramanian, Kunjithapatham; Kuhnert, Andreas, "Toward 10^{10} contrast for terrestrial exoplanet detection: demonstration of wavefront correction in a shaped-pupil coronagraph," *Proc. of SPIE*, 6265 (2006).

10. Belikov, Ruslan; Beall, James; Carr, Michael; Give'On, Amir; Kay, Jason; Kolade, Taofik; Littman, Michael; Mycroft, Frank; Pueyo, Laurent; Vanderbei, Robert J.; Kasdin, N. Jeremy, " Towards 10^{10} Contrast for NASA's Terrestrial Planet Finder Mission: Demonstration of High Contrast in a Shaped-Pupil Coronagraph at Princeton," Proceedings of the IAU Colloquium #200, pp. 415-420, 2006.

11. Give'On, Amir; Kasdin, N. Jeremy; Vanderbei, Robert J., "Closed-loop Wavefront Correction for High Contrast Imaging: The ``peak-a-boo" Algorithm," *Proceedings of the IAU Colloquium #200*, pp. 541-546, 2006.

12. Trauger, J., et al., "Coronagraph Contrast Demonstrations with the High Contrast Imaging Testbed," *Proc. of SPIE* 5487, 1330-1336, (2004).

13. Bordé, Pascal J.; Traub, Wesley A., "High-Contrast Imaging from Space: Speckle Nulling in a Low-Aberration Regime," *Astrophysical Journal*, 638 (2006).

14. Guyon, O.; Pluzhnik, E. A.; Kuchner, M. J.; Collins, B.; Ridgway, S. T., "Theoretical Limits on Extrasolar Terrestrial Planet Detection with Coronagraphs," *Astrophysical Journal Supplement Series*, 167, (2006).

15. Shaklan, Stuart B.; Green, Joseph J.; Palacios, David M., "The terrestrial planet finder coronagraph optical surface requirements," *Proc. of SPIE*, 6265 (2006).

A Nulling Coronagraph for TPF-C

B. M. Levine[1], M. Shao[1], R. Samuele[2], J.K. Wallace[1], S. Rao[1], R. Lyon[3], E. Schmidtlin[1], S.F. Fregoso[1], and F. Aguayo[1]

INTRODUCTION

This summary outlines a concept study for a high contrast instrument for the Terrestrial Planet Finder Coronagraph (TPF-C) mission (Levine 2006, Shao 2006). The objectives are to develop a nulling coronagraph based imager and spectrometer concept that will increase the number of planets TPF-C detects, and will expand the wavelength range of the spectrometer into the near-IR to enable detection of additional unique visible biomarkers. This instrument utilizes an alternative starlight suppression system (SSS) based on the principles of nulling interferometry, which allows inner working angles (IWA) within 2–3 λ/D to be obtained, and also to measure low resolution (R=80) spectra. Equally important, this concept contains a post starlight suppression wavefront sensor (or calibration wavefront sensor) to increase the achievable contrast level, and to substantially decrease stability requirements during integration. The search for planets will be conducted at short wavelengths, where the IWA is smaller. Extending spectroscopy to 1.7μm, the visible and near infra-red signatures of likely atmospheric constituents include oxygen, ozone, water, methane, and carbon dioxide, some combinations of which are considered to be biosignatures. Coverage of the 0.5 ~1.7 μm wavelength range is done in intervals of 25% bandwidth. Our latest experimental results show null depths using white light that are within an order of magnitude of the required value.

NULLING CONCEPT AND ARCHITECURE

The functional subsystems of a nulling coronagraph are shown in Figure 1. In this architecture, nuller #1 produces the deep null necessary to cancel the starlight. It passes this output to nuller #2, which broadens the angular extent of the null of starlight out to θ^4, and transmits the rejected starlight to the Calibration Wavefront sensor for further processing. After nuller #2, any remaining starlight is caused by scattering from imperfections in the nuller optics. The fiber array then filters the high spatial frequency errors that may cause speckles in the same field of view as the planet. A part of the light from the fiber array goes to the science camera (or spectrophotometer), and the remainder of the light is mixed coherently in the Calibration Wavefront Sensor to provide an error signal to correct upstream phase errors and also to estimate residual starlight that can be subtracted on a post-detection basis.

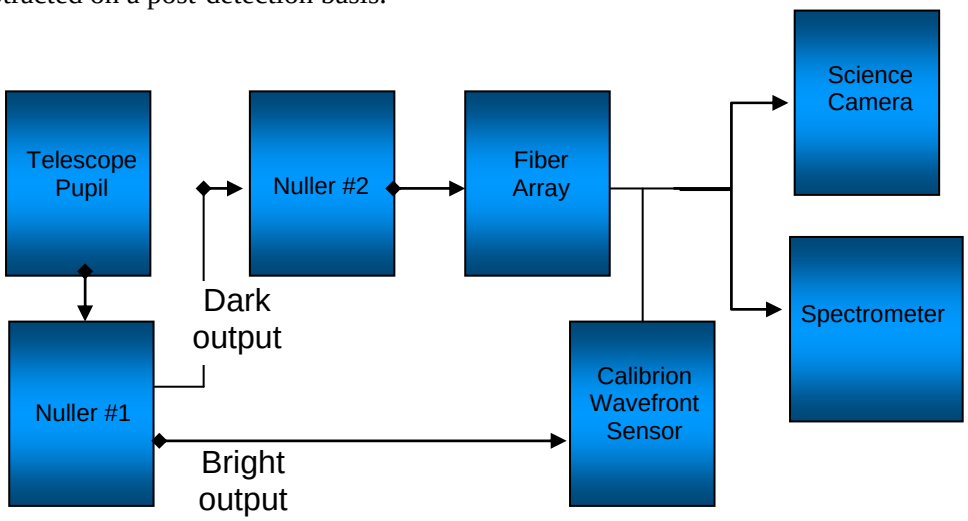


Figure 1: *Top level architecture of a Nulling Coronagraph*

[1] Jet Propulsion Laboratory, California Institute of Technology
[2] Northrup Grumman Aerospace Systems
[3] NASA Goddard Space Flight Center

USING A NULLING INTERFEROMETER INSTRUMENT WITH A SINGLE APERTURE TELESCOPE

In principle, a nulling interferometer effectively projects an interference pattern on the sky over the star-planet system to be imaged (Figure 2). The baseline, s, sets the fringe spacing. They attenuate the starlight and have 100% transmission for planet light when the optical path from the planet is λ/2 different from the star. For a modest sized aperture, about D=1m, a Jupiter-like planet can be resolved by synthesizing an interferometer with a 30-cm baseline, and at D=8m, an Earth-like planet can be resolved with a 50 cm baseline at λ=500 μm wavelength.

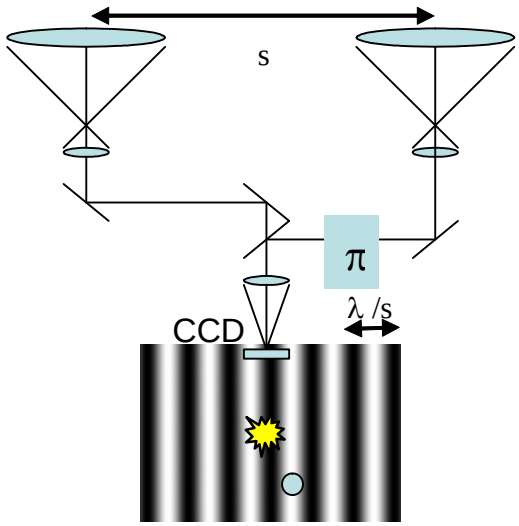


Figure 2: *Imaging with nulling interferometry. A fringe pattern is superimposed over the star and planet to be imaged, the star is placed at the bottom of a deep, achromatic destructive (null) fringe, while planet on the constructive fringe is transmitted.*

The implementation for TPF-C calls for two nulling interferometers in a series. This arrangement synthesizes a four element nulling interferometer (the first nulling interferometer produces a deep null fringe pattern, and by directing its output into a second nulling interferometer, the null fringe becomes wider, and minimizes light leakage from the finite diameter of the star, and also from pointing errors). After nulling, an array of coherent single-mode optical fibers, coupled on both sides with (commercially available) lens arrays is used to filter the effects of any residual stellar leakage (scattering) due to imperfections in the telescope optics and optical train. A simple imaging system after this array forms the final extra-solar planet image.

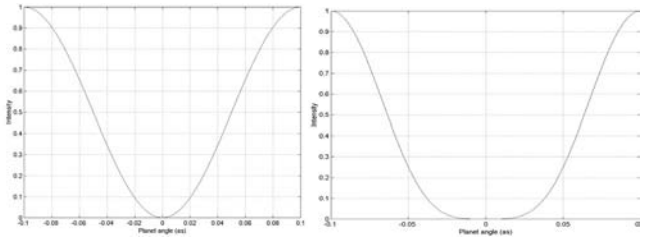


Figure 3: *Intensity response of a 2 and a 4 element nulling inteterferometer. Left, fringe pattern over a 0.1 as field of view. Right, nulling depth over same field but with a 4 element nuller*

To introduce an achromatic π-radian phase shift in a different fashion, we use pairs of dielectric plates of differing thicknesses. Solutions for achromatic π-radian phase shifts (to the needed accuracy) exist with two glasses (Figure 4, lower right). Thus, for this experiment, the final layout for the beam combiner consists of identical two-glass pairs of (rotatable) dielectric phase retarders in each leg of the interferometer. Note that with the combination of common BK7 and Fused Silica optical glasses the theoretical minimum of 10^{-7} can be realized over a 25% bandwidth. (Morgan 2001)

Another advantage of the nulling coronagraph architecture is the single mode fiber array component to provide the deep null required, and also to relax the optical figure quality requirements of the primary telescope optics. With this array of fibers, wavefront errors at high spatial frequency, on a spatial scale smaller than any deformable mirror (DM) actuator, will be filtered out by the single-mode fibers instead of propagating to the science focal plane. If the optical fibers all have the same length (to λ/20) the planet light from each fiber combines to form a coherent off-axis image. The phase of the residual star light exiting the fiber array is random, hence it is scattered evenly across the whole field of view. In summary, wavefront errors on a spatial scale larger than the DM actuator spacing are corrected by active control of the DM and wavefront errors smaller than the actuator spacing are filtered out by the fiber array.

With an 8-m telescope, the airy pattern of the telescope has dropped ~10^{-4} by the 5th airy ring, where we expect to find an earth at 10 pc. We need to suppress this by an additional factor of one million to achieve imaging. To keep scattered light below this 10^{-6} level, the optical system would need to be near-perfect (or have a Strehl ratio

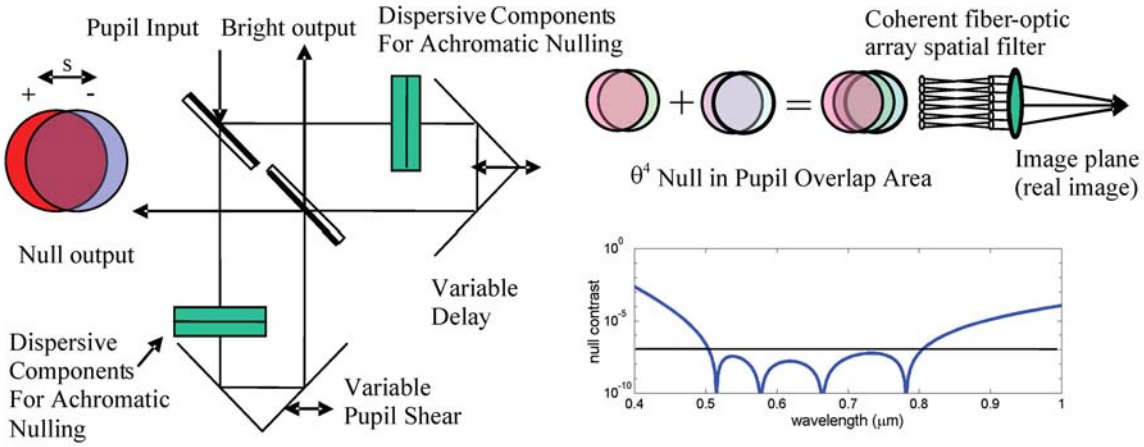


Figure 4: *Visible nulling coronagraph instrument concept. Left, a concept for single-input symmetric nulling interferometer. Note that the two pupils emerge from the null output displaced (or sheared) by a distance, s, proportional to the baseline of the instrument. Lower right, Calculation of "pseudo-achromatic" null using two dielectric plates. Upper right, the major components of nulling coronagraph. By coupling two nulling interferometers, a 4 pupil overlap results in an interference fringe pattern proportional to λ^4. The fiber-optic array is a coherent spatial filter to reject residual scattered light. An imaging system then projects the pupil into the far-field to form an image.*

of 99.9999%), or the scattered light due to any minute optical figure imperfections results in a residual diffraction pattern from a conventional Lyot coronagraph. A Strehl of 10^{-6}, in turn implies a full aperture optical figure that is almost perfect: ~λ/6000 rms, or ~1Å, hence the need for extreme wavefront control and Angstrom accuracy deformable mirrors. In our nulling coronagraph architecture, we plan to have a fiber array of approximately 1,000 fibers (arranged in an hexagonal array with a diameter of 39 fibers). Light from an 8m/39~ 20cm part of the primary falls on a lenslet which focuses the light into a single-mode fiber. At the output, this light is recollimated. Optically coupled to each fiber is a segmented deformable mirror in one arm of each nulling interferometer. This DM is used to control the amplitude and phase of the light. At the fiber input, if the starlight is suppressed by 10^{-7}, then at the output a total of 10^{-7} of the starlight will appear as scattered light. Because in general the nulling process will produce residual leaked light with random phase, this 10^{-7} starlight will be scattered uniformly over the airy spot output field of view, with an average scattered light level of 10^{-10}/airy spot.

CALIBRATION SYSTEM

Although starlight is suppressed to 10^{-10}, the resulting speckle pattern needs to be subtracted to a much lower level, 2×10^{-11} for a 5 sigma detection of a planet. The post-starlight suppression system (SSS) sensor has two functions. One is to provide real-time feedback to the adaptive optics (AO) system to correct the state of scattered light, and to make the wavefront more perfect. The second function is to provide post-detection information on the amplitude and phase of this post-SSS wavefront in order to generate an estimate of the post-coronagraph point spread function (PSF) for subsequent correction by software subtraction.

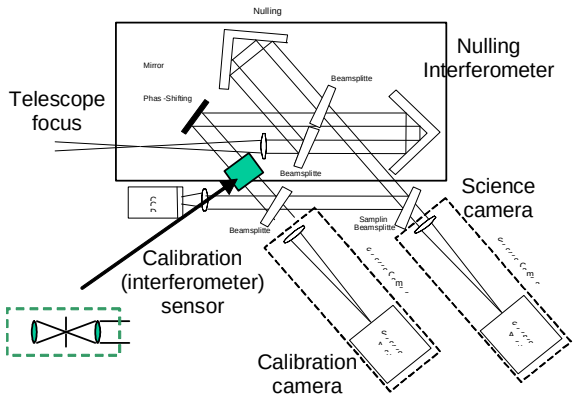


Figure 5: *Concept for a calibration wavefront sensor with a nullling coronagraph.*

The default TPF procedure for subtracting residual starlight is to roll the telescope around the line of sight after about 2–4 hours of integration. For this method to work, the null has to be stable to 2×10^{-11} for hours, which means that the wavefront must be stable to single digit picometers.

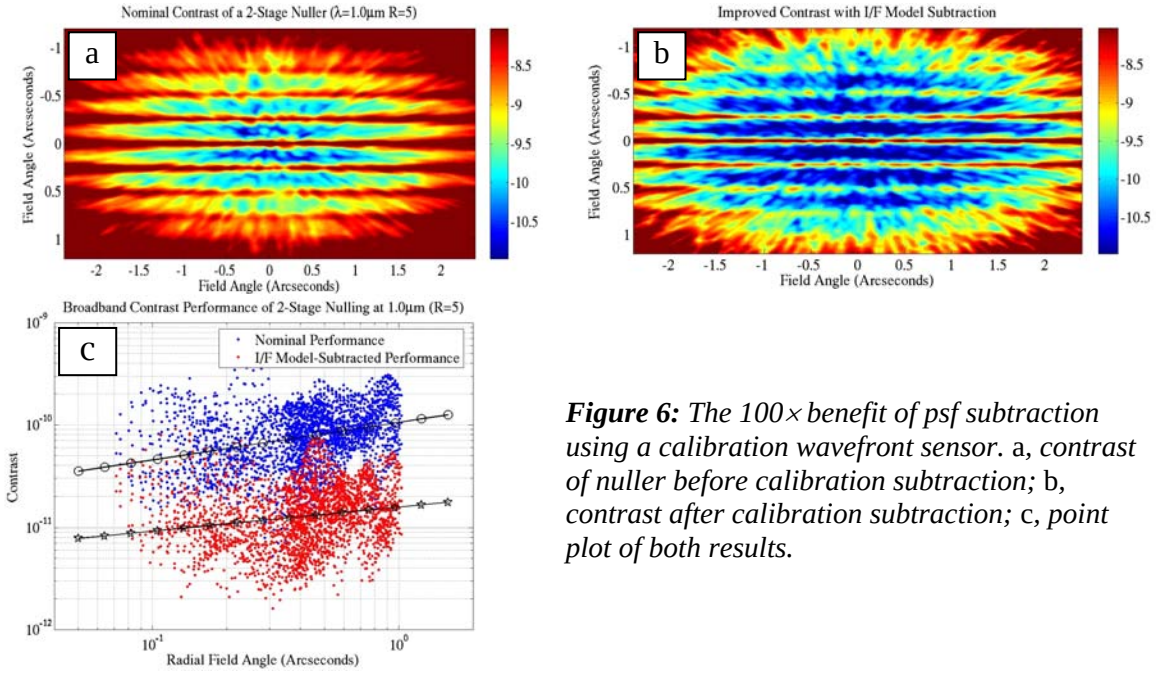


Figure 6: *The 100× benefit of psf subtraction using a calibration wavefront sensor.* a, *contrast of nuller before calibration subtraction;* b, *contrast after calibration subtraction;* c, *point plot of both results.*

With a calibration interferometer, we measure the residual wavefront concurrent during science integration. Having both the image and an estimate of the wavefront error, the wavefront stability requirements can be relaxed by 4 to 6 orders of magnitude (picometer (pm) level stability over hours of integration versus pm stability over a fraction of a second). Wavefront changes caused by nanometer motions of the telescope secondary with respect to the primary mirror generate errors (focus, astigmatism, coma, etc.) that result in scattered light at 2-3-4 λ/D. Making any coronagraph/nuller work at small λ/D is difficult, but the calibration wavefront sensor makes 2 λ/D coronagraphs feasible.

The basic idea of the post starlight suppression system (SSS) wavefront sensor is to use the light blocked by the SSS as a "reference" beam in an interferometer to measure the post SSS electric field, both amplitude and phase. It comprises two beams: a reference beam and an unknown. The reference beam is created from the light that passes from the bright output of the first nulling interferometer. The light that forms the second (unknown) part is the nulled light picked off after the fiber array and before the science camera. The two beams are combined in the calibration camera, and using modulation with a 4-bin algorithm, the amplitude and phase of the stellar leakage is computed. The leakage point spread function is estimated by propagating this field into the far-field.

The calibration procedure produces a two-order-of-magnitude reduction of the speckle pattern as shown in the following simulation of the 8 m × 3.5 m TPF telescope (Fig. 6). The upper and lower figure shows the contrast level in the image before and after removal of the residual starlight psf. The point plot shows the individual contrast of each field point—somewhat like a contrast histogram with field angle.

STATUS OF KEY EXPERIMENTS

Recent results from our nulling interferometer testbed yield contrast ratios at the 1.05×10^{-6} level, with a 15% visible bandpass, which is equivalent to 10^{-9}per airy spot. This result is at 65% of our final bandpass requirement, although limitations of our current configuration make major hardware changes essential to broadening the bandpass. We make the argument that broadening the bandpass should not necessarily adversely affect the null depth until beyond the 20% visible light level. Using the same setup we are able to reach monochromatic null depths of 1.11×10^{-7} (λ = 638 nm) averaged over three seconds (Samuele, 2007). This section describes our experimental approach for achieving deep broadband nulls, as well as error considerations and limitations, and the most recent results for our nulling coronagraph testbed.

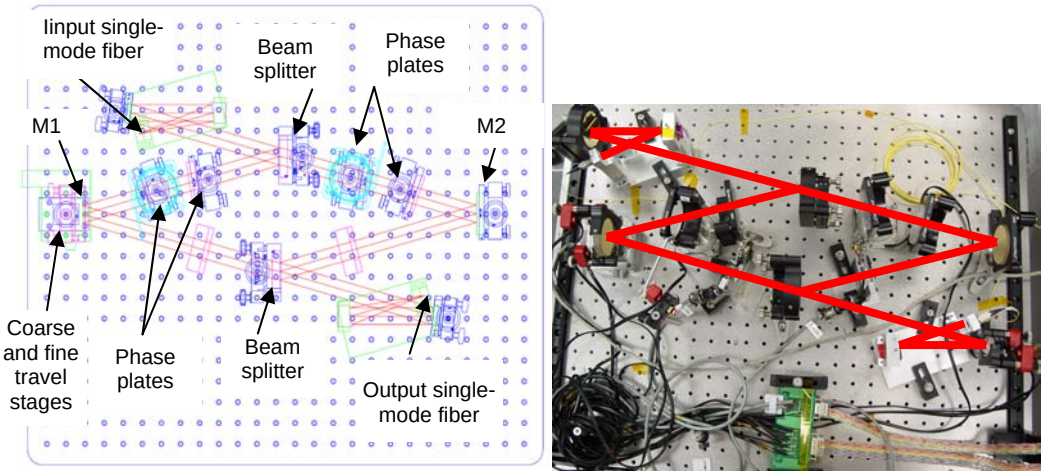


Figure 7: *Nulling Test bed set up.*

The nulling interferometery experiment at JPL employs a modified Mach-Zehnder design. Maintaining a high degree of symmetry is necessary for deep nulling, so a meticulous approach is needed for both construction and alignment. The nuller core is composed of two bare-gold coated mirrors, one in each arm, and two highly symmetric, dielectrically-coated, beampslitters. The core mirrors and beamsplitters are aligned such that their coated surfaces are parallel to an arcminute—to mitigate pupil rotation (to be discussed later). The optics of the nuller are 50mm- in diameter, with each arm extending for roughly 650mm of path length.

To minimize path length fluctuations due to vibration, the nuller is positioned on three tiers of isolation. The testbed breadboard is directly mounted on visco-elastic, sorbothane bumpers internal to the vacuum chamber. Sorbothane is a material specifically designed for vibration dampening. Furthermore, the vacuum chamber itself is mounted on larger, weight-bearing sorbothane pucks. The vacuum chamber is then positioned on a 24-inch-thick Newport optical table, which is supported by pneumatic vibration-isolating legs. The whole lot is then positioned on a vibration-isolating pad. The interferometer is placed inside of a vacuum chamber—to mitigate the affects of atmospheric turbulence. However, we do not evacuate the chamber since we are able to achieve the required stability at atmosphere. We coarsely tune each glass type using interference fringes as a thickness metric. We also balance the deconstructive fringes directly adjacent to the central null, in order to achieve a symmetric white light interferogram. We finely tune our dispersion by using null depth as a direct metric. The upshot is that there is a panoply of solutions for differential glass thickness, and properly tuning the differential thickness with our rotation stage resolution is in fact feasible. The particular solution we use for our differential glass thicknesses are 481μm of Fused Silica and 393μm of BK7

NULLING ERROR ANALYSIS

Each error term can be treated as an independent contributor to the overall null degradation. We use empirically determined values to estimate each error term's contribution, which linearly sum to yield our expected experimental null depth. We then check our experimental null result against that of the total null. This allows us to characterize the state of our interferometer's performance, and to evaluate which errors are contributing the lion's share. Table 1 shows a summary of the individual error terms.

An error assessment of our nulling interferometer includes the following measurable quantities: pupil rotation, birefringence, intensity mismatch, optical path fluctuations, and dispersion. Where pupil rotation, birefringence, and dispersion can be considered static errors, and intensity mismatch and optical path fluctuations dynamic errors Pupil rotation is purely geometrical and time-independent. It is caused by slight, out-of-plane folds that deviate from the parallelism of the mirror's surfaces. The end result is polarization leakage from one perpendicular state to another at the point of recombination. The first concern is getting the mirror surfaces parallel to within an arcminute. We accomplish this using a Fizeau interferometer to reflect off of each mirror's surface. We also place a mask at the interferometer input, at measure the relative offset induced by each interferometric arm.

Table 1. *Nulling Interferometer Error Budget*

Error Term	Equation	Variable	Empirical Value	Null Limit	% Contribution (broadband)
Pupil rotation	$N_{Pr} = \Theta^2/4$	Θ	0.01 deg	7.6×10^{-9}	5.8
Birefringence	$N_{Opd} = (2\pi\ \Delta OPD_{pol}\ /\lambda)^2/4$	ΔOPD_{pol}	0.04 nm	9.7×10^{-9}	7.4
Intensity mismatch	$N_{Im} = \Delta I^2/4$	ΔI	0.009% (laser) 0.03% (broadband*)	2.25×10^{-10} 2.3×10^{-8}	17.0
Optical path fluctuations	$N_{Opd} = (2\pi\ \Delta OPD/\lambda)^2/4$	ΔOPD	0.06 nm rms	8.7×10^{-8}	66.1
Dispersion	$N_D = (2\pi\ \Delta OPD_\lambda\ /\lambda)^2/4$	ΔOPD_λ	0.014 nm	4.9×10^{-9}	3.7
Laser null (638 nm)	$N = N_{Pr} + N_B + N_{Im} + N_{Opd}$	-	-	1.05×10^{-7}	-
Broadband null limit (15% WL)	$N = N_{Pr} + N_B + N_{Im} + N_{Opd} + N_D$	-	-	1.3×10^{-7}	-

We can adjust the beamsplitters to ensure a zero shear constraint. These two approaches allow us to tune and measure the pupil rotation to better than 0.01 degrees, and using the equation from Table 1, we estimate a null limit of 7.6×10^{-9}.

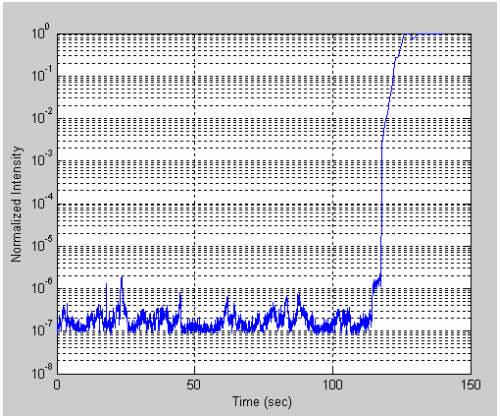


Figure 8: *This data set shows our laser nulling result (638 nm) of 1.25 $\times 10^{-7}$ averaged over 10 seconds, or 1.11 $\times 10^{-7}$ averaged over 3 seconds. Null depths of roughly 8 and 9 million to one respectively. The first 120 seconds of data are at the deconstructive null, while the constructive peak follows.*

Birefringence is another time-independent error, by which polarization phase delay occurs when overly stressed regions of a transmissive optic, such as a beamsplitter, cause anisotropic regions in the substrate. Improper beamsplitters mounting, namely those mounts that use a setscrew, can significantly limit our null depth. We use a Malus' Law setup (i.e., two cross polarizers with the sample between them and a large area photodetector) to test for stress-induced biregrince. It is a strikingly clear affect that can be witnessed visually. To mitigate birefringence in our beamsplitters and dispersion plates we either use a spring loaded mounting mechanism or simply glue the optic to the mount. In this manner, we have managed to reduce the polarization phase delay to 0.04 nm or a null limit of 9.7×10^{-9}. We use null depth as a direct metric to evaluate polarization effects.

We used two edge filters, one with throughput longward of 590 nm and one with throughput shortward of 710 nm, to achieve a 15% white light bandpass profile, as shown in Figure 9. Broadband nulling with our 15% white light source, has yielded nulls as deep as 1.06×10^{-6} averaged over a 3-second span, of which, the experimental results can be seen in Figure 6. The error analysis for each term is shown in Table 2.

Dispersion is still the dominant contributor to the lack of null depth. Experimentally this was seen as a lack of precision in the adjustment of the dispersion plates. As shown in Figure 10, the deepest null depth can be achieved over a range of differential thickness between the two dispersion plates, and when both plates can be tuned with infinite precision. The same curve is plotted on the right hand side of the figure given the step size resolution in our experiment. By replacing one of the glasses with a finer resolution stepping device, we should be able to achieve the required result.

ACKNOWLEDGEMENTS

We wish to acknowledge the contributions of a number of engineers and scientists who made contributions to this effort prior to the start of this current TPF instrument concept study. R Bartos for mechanical engineering support and Dr. J. Green for simulations of wavelength dependent effects.

Table 2. *Error analysis for the white light nulling experiment.*

Source Null	Pupil Rotation	Intensity Mismatch	OPD Fluctuations	Birefringence	Dispersion
Value achieved	0.01 Deg	0.1%	0.1 nm, rms	0.04 nm	0.15 nm
Contribution to Null	7.62E-9	2.5E-7	2.42E-7	9.70E-9	4.89E-9
Net Null:	1.06E-6 (950K:1)				
% Contribution	1%	24%	23%	1%	51%

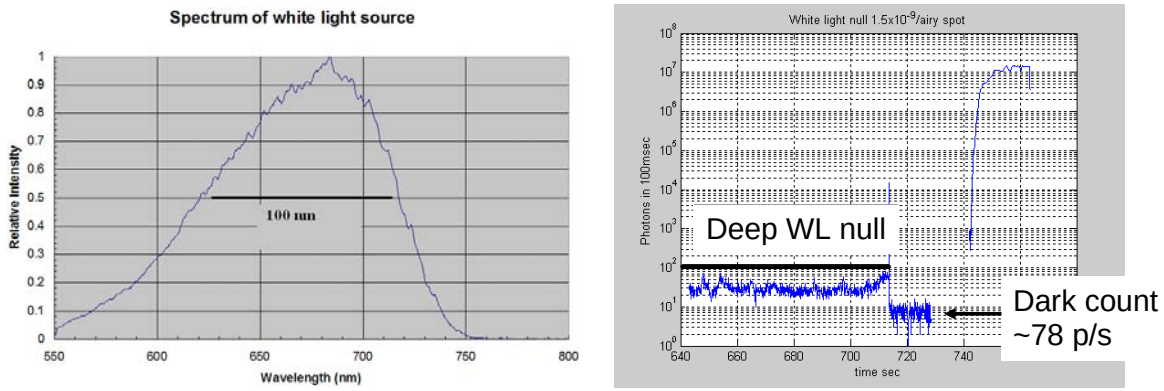


Figure 9: *(Left) The 15% white light spectrum after propagation through one arm of the interferometer. The full width at half maximum is 100 nm. (Right) The resulting null plot showing nearly 1M:1 contrast (1.06 × 10^{-6})*

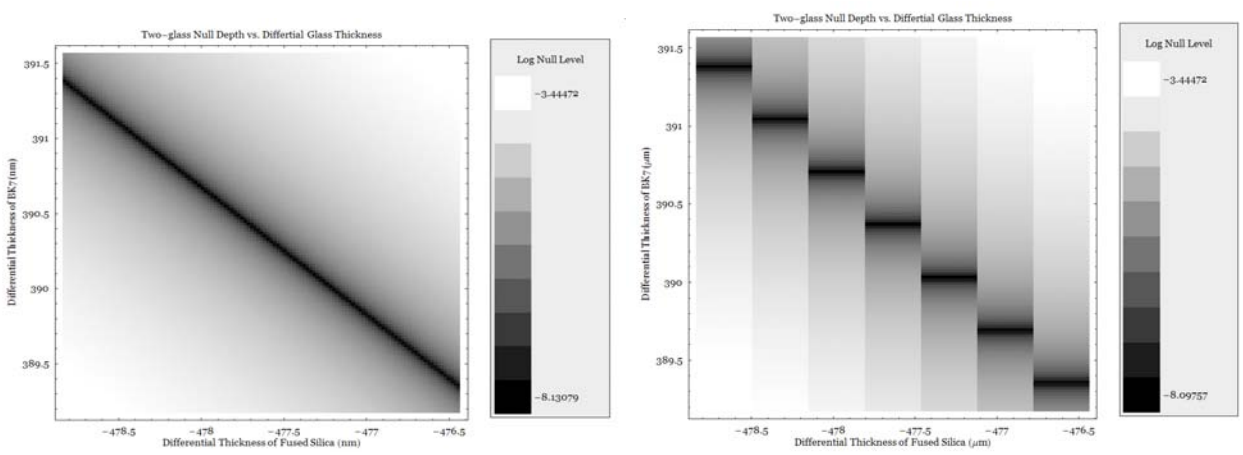


Figure 10: *Finding the optimum differential spacing in a two glass null. (Left) null depth surface for the 2 glass null. (Right), Same null depth but using the step resolution limits in our experimental equipment. At any large "step" in one glass, the corresponding optimum null can be found with same device with sufficient resolution.*

REFERENCES

Morgan, R. Ph.D., (2001), Thesis, University of Arizona. Optical Sciences.

Levine, B.M., et al., (2006), "The Visible Nulling Coronagraph—Architecture Definition and Technology Development," 207th Meeting of the AAS, Washington, DC.

Levine, B. M., et al, (2006), Proceedings of the SPIE, Volume 6265, pp. 62651A.

Shao, M., PI, (2006), "TPF-C Instrument Concept Study, Visible and Infrared Nulling Coronagraph Spectrometer," A Report Submitted to the Terrestrial Planet Finder Coronagraph Science and Technology Development Team, 7 March 2006, p4.

Samuele, R., et al., (2007), "Experimental Progress and Results of a Visible Nulling Coronagraph," IEEE Aerospace Conf. Paper #1366.

Lyot Coronagraphs with Band-Limited Masks

Brian Kern
Jet Propulsion Laboratory, California Institute of Technology

INTRODUCTION

Band-limited masks in Lyot coronagraphs have received more theoretical and experimental attention than any other coronagraph design under consideration for TPF-C, and 8th-order band-limited masks were the assumed architecture for Flight Baseline 1. The best contrast achieved to date has been with band-limited masks, on JPL's High Contrast Imaging Testbed (HCIT) (see Trauger, these proceedings). This paper will attempt to illustrate the advantages and disadvantages of band-limited masks, compared to the other potential TPF-C coronagraph architectures; the primary disadvantage is modest throughput and point-spread function (PSF) width, the primary advantages are robustness to aberrations, mechanical simplicity, and maturity.

INNER WORKING ANGLE AND THROUGHPUT

Band-limited masks for coronagraphs have been proposed in Kuchner & Traub (2002) and Kuchner, Crepp & Ge (2005). The fundamental feature of these coronagraphs is a mask that absorbs much of the on-axis light, diffracting the rest to an opaque Lyot stop. This gives zero on-axis transmission, with modest transmission off-axis. Most of the band-limited attention has gone to linear occulters, whose transmission is a function of only one spatial variable; coronagraphs with linear band-limited masks have typical transmissions of ~ 30–45% at 4 λ/D, shown in Figure 1. Because the Lyot stop must be smaller than the system pupil, the resulting off-axis PSF is larger than the PSF incident on the occulting mask. It is this transmission and diffractive efficiency (a relative measure of PSF size) that determine the total exposure time needed for an observation.

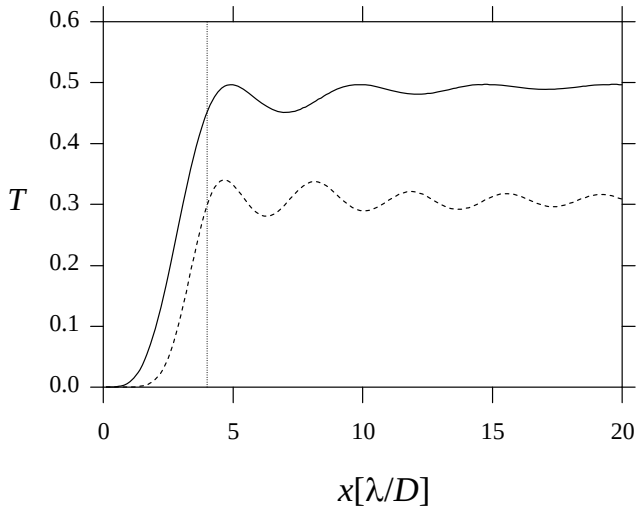


Figure 1: *Transmission (T) versus image plane position (x) for sample linear 4th-order (1-sinc^2) mask (solid line) and 8th-order m=1, l=3 mask (dashed line). Each mask is chosen to maximize transmission at 4 λ/D (shown as vertical line).*

The combination of mask width and Lyot stop size can be varied to optimize transmission at a given image plane location. Smaller inner working angles (IWAs) can be reached using narrower occulting masks, but Lyot stops become narrower, decreasing transmission and increasing PSF width. A 4th-order (1-sinc^2) mask can be designed to give 0.45 transmission at 4 λ/D, or 0.16 at 2 λ/D. An 8th-order $m = 1$, $l = 3$ mask can give 0.30 at 4 λ/D, or 0.03 at 2 λ/D.

Since the release of the TPF-C FB1 specifics, there has been a drive to reconsider smaller telescopes, and coronagraph architectures that accommodate smaller IWAs. As mentioned above, a 4th-order mask operating at 2 λ/D can give 0.16 transmission. However, other concerns arise when decreasing the IWA. One that has received little attention is the contrast degradation due to finite stellar size, shown graphically in Figure 2. Because stars appear as incoherently illuminated discs, the stellar limb passes through the coronagraph off-axis, and leaves residual light in the image plane. This effect is more troubling at smaller IWAs or with narrower masks (neither shown in Fig. 2). Already at 4 λ/D, a 4th-order mask can observe only 66 of the TPF-C top 100 stars to a contrast of 10^{-11}; going to smaller IWAs decreases this number further. By contrast, the 8th-order mask can observe all 100 stars to 10^{-11}, because it is less sensitive to light at very small off-axis angles. Other 4th- and 8th-order coronagraph architectures, such as the optical vortex or visible nuller, should show similar sensitivities to stellar size.

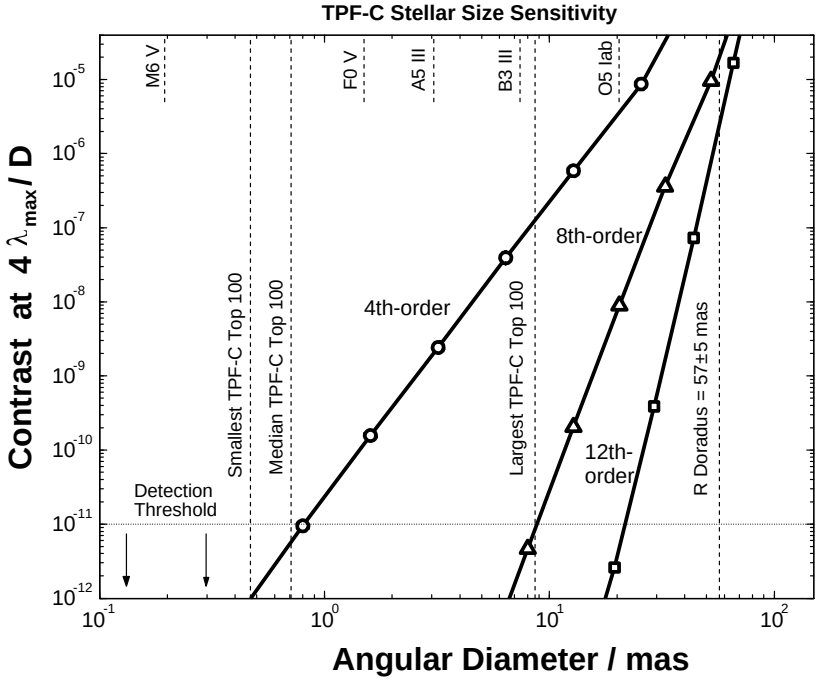


Figure 2: *Limiting contrast versus stellar angular diameter, for 4th-order (circles), 8th-order (triangles), and 12th-order (squares) masks, from Crepp (2006). Also shown are the 10^{-11} assumed detection threshold (horizontal line), and the smallest, median, and largest diameters of the TPF-C top 100 stars (vertical lines). Representative spectral types are shown along the top, at their angular diameters at 10 pc. An 8 m telescope is assumed.*

REQUIREMENTS / SENSITIVITY TO ABERRATIONS

The TPF-C dynamics error budget (Shaklan et al. 2005) and optical surface requirements (Shaklan, Green & Palacios 2006) establish the baseline requirements for 10^{-10} contrast measurements. The dynamics error budget, in particular, assumes an 8th-order linear occulter, which was chosen because its insensitivity to low-order aberrations (see Shaklan and Green 2005) allows a significant relaxation in pointing and rigid-body requirements. A sampling of the relative requirements of 4th- and 8th-order masks is shown in Table 1, taken from the TPF-C STDT. To first order, this table also represents the comparison between band-limited masks and other 4th- and 8th-order architectures, such as the optical vortex and visible nuller.

Table 1: *Maximum allowed dynamic RMS wave front errors to maintain 10^{-12} contrast per error term, for 4th- and 8th-order masks, in units of 550 nm waves. The relaxation ratio is the ratio of the requirements for 4th and 8th-order masks, showing that 8th order masks allow a significant (i.e., orders of magnitude) relaxation in pointing / stability requirements.*

Aberration	8th order	4th order	relaxation ratio
Tilt	4.3e-2	2.7e-3	16
Focus	3.8e-3	2.9e-5	132
Astigmatism	5.3e-3	4.0e-5	132
Coma	1.0e-3	5.6e-5	185
Trefoil	2.3e-3	1.7e-5	132
Spherical	3.5e-5	4.9e-6	7.2

Separate from the dynamic wavefront errors, elliptical polarization induced by reflection of flat wavefronts off curved mirror surfaces (or by reflection of a curved wavefront off flat mirror surfaces) results in low-order spatially varying phase differences between orthogonal polarizations. The resulting wavefront errors in orthogonal polarizations differ, and DM corrections cannot separately correct the two polarizations; corrections applied to one polarization will worsen the other. This has been characterized for TPF-C in Balasubramanian et al. (2005). Downstream from the primary mirror, two separate polarization channels can be constructed, each with its own coronagraph, to separately correct the two polarizations. This construction allows for independent correction of orthogonal polarizations, and so allows the two co-polarized wavefronts to be independently corrected; however, when observing an un- or partially-polarized thermal source, the cross-polarized wavefronts cannot be separately corrected (without installing separate polarization channels upstream of the primary mirror). In all cases, the aberrations introduced between polarizations are low-order, and 8th-order masks are therefore much less sensitive to these terms than are 4th-order masks. The net result is that 4th-order coronagraphs require either separate polarization channels or specialized mirror coatings (that minimize the polarization signature) to achieve 10^{-10} contrast, while 8th-order coronagraphs require neither. As with the dynamic requirements, the selection of an 8th-order architecture (whether band-limited mask or optical vortex) significantly relaxes the polarization requirements.

The bandwidth of band-limited mask coronagraphs is limited by either occulting mask errors or the broadband behavior of wavefront correction. Any wavefront correction using DMs has a limited bandwidth, as considered in Shaklan, Green & Palacios (2006). In the absence of wavefront errors, there is no inherent bandwidth limit to band-limited masks. The masks themselves will be chromatic at some level, with both the modulus and phase of the complex amplitude transmission profile varying with wavelength. This will be discussed further in Sections 4 and 5. But aside from the mask transmission, there is no aspect of these coronagraphs that imposes any additional bandwidth constraints beyond the need to make wavefront corrections using DMs.

MODELING / MODEL VALIDATION

Detailed optical broadband modeling, simulating the contrast performance of band-limited masks, has been done in four different contexts at JPL:

- PROPER, an IDL-based library (Krist)
- MACOS + proprietary Matlab code (Sidick)
- Python proprietary code (Moody)
- Fortran proprietary code (Hoppe)

The Fortran models used by Dan Hoppe were generated specifically to explore the unusual waveguiding effects of metallic binary occulting masks (Kuchner and Spergel 2003), and incorporated a more complicated EM propagation analysis to address those issues. The other three models vary in the specifics of their optical propagation techniques (although all include full Fresnel propagation to all intermediate optics), specification of surface figure errors, and computational optimization. The full diffraction calculations performed by the models avoid the limitations present in semi-analytic approximations; for example, the derivation of optical surface

requirements for TPF-C (Shaklan, Green & Palacios 2006; Shaklan & Green 2006) uses 2nd-order Taylor expansions to approximate wavefront behavior, an approximation that is avoided in these models.

The most meaningful application of these models is in conjunction with the High Contrast Imaging Testbed (HCIT) at JPL. HCIT is a Lyot coronagraph with a DM for wavefront correction, enclosed in a vacuum chamber, currently using (but not restricted to) band-limited masks. The purpose of HCIT is to validate model predictions and demonstrate the key technologies required for TPF-C. The level of maturity of this model-to-experiment connection, and the depth of experience with band-limited masks, make band-limited masks stand apart from other coronagraph architectures.

An excellent example of the interplay between modeling and experiment is the analysis of polarization properties of binary band-limited masks. These binary masks were constructed of aluminum on glass, with a variable transmission profile formed by varying the duty-cycle of a high-frequency (smaller than Fλ) on-off pattern (aluminized to bare glass), as shown in Figure 3. The anisotropic nature of the high-frequency structure of the mask causes different phase shifts for orthogonal linear polarizations, with the relative phase shift varying with transmission. This is tested by correcting one polarization, and measuring the contrast in the orthogonal polarization, as shown in Figure 4. The experiment clearly validates the prediction that the orthogonal polarization will have significantly worse contrast, above 10^{-7}.

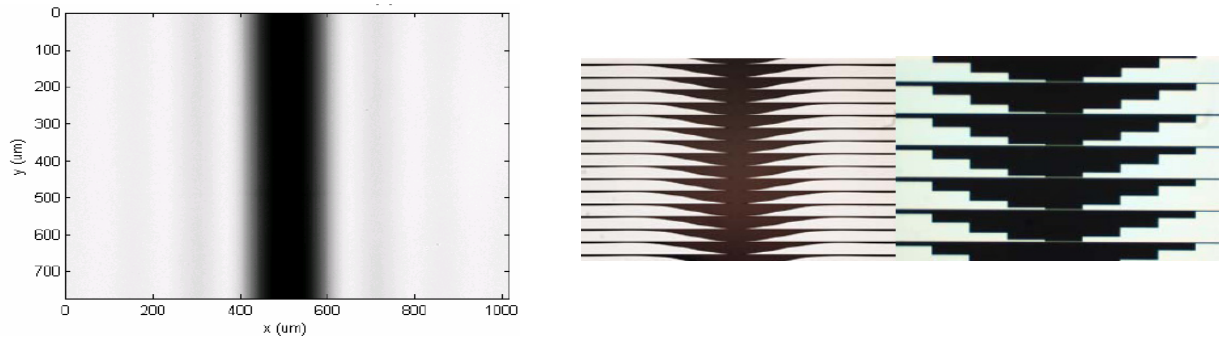


Figure 3: *Continuous and binary representation of 1-sinc^2 4th-order band-limited mask. Left panel is measured transmission of continuous, analog band-limited mask, right panels are SEM images of continuous and stepwise varying binary masks. Black areas of the right panels are aluminized.*

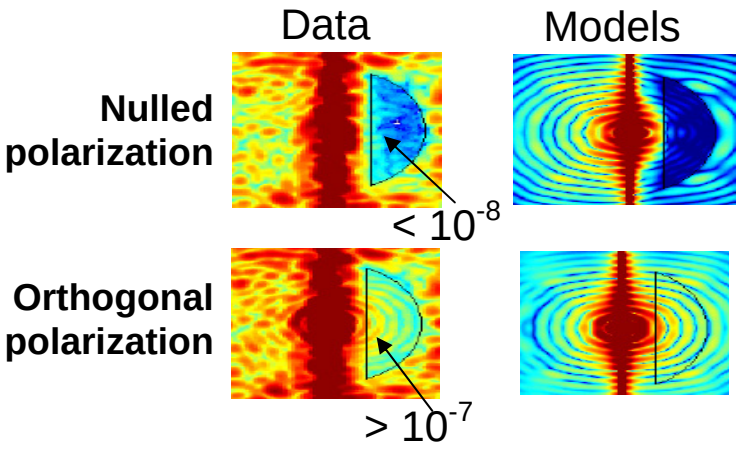


Figure 4: *Coronagraph image plane in binary mask polarization experiment and model. Each panel is the coronagraph image plane, with occulted source at the center, after creating a dark hole using a DM. Only one linear polarization was used for correction, and the contrast in the orthogonal polarization was measured. The character of the agreement with the model, that the contrast differed between polarizations by >10×, gives confidence that the binary mask modeling captured the pertinent physical effects.*

A significant amount of modeling goes toward understanding the effects of imperfect mask materials, and identifying mitigation strategies. For example, the analog masks in use on HCIT are written on High Energy Beam Sensitive (HEBS) glass. An occulting mask is written on HEBS glass by exposing it 100 keV electrons, with the total electron dose varying with position to give the desired absorption pattern. The transmitted phase varies with the optical density (OD) written (i.e., phase as well as OD varies with electron exposure), and because the OD varies spatially, the transmitted phase has spatial structure. This is reported in Halverson et al. (2005), and shown in Figure 5. The complex-valued nature of the occulter transmission means that the patterns intended to be band-limited in the absence of varying phase, are not band-limited when written on HEBS. The optical consequence of this is that in the Lyot plane, where a band-limited mask in a perfect coronagraph would produce diffracted light only near the edge of the pupil image and no light inside the Lyot stop, some amount of light is distributed throughout the Lyot plane when the occulter is not band-limited. This extra light is not distributed uniformly, but is brightest near the edges of the pupil and fainter near the center of the pupil. A Lyot stop with a smaller opening will reject more of this light. Wavefront correction can also redistribute this light, so that a dark hole can be created in the image plane. Predicting the ultimate contrast behavior using different Lyot stop sizes and wavefront correction requires detailed modeling.

A further complication arises in that the relationships of both OD and phase to electron exposure vary with wavelength. An occulter written to provide the correct OD at one wavelength will have a different OD profile at another wavelength, as well as a different phase profile. The models employed here predict the ability of wavefront control to maintain good broadband contrast in the presence of these variations.

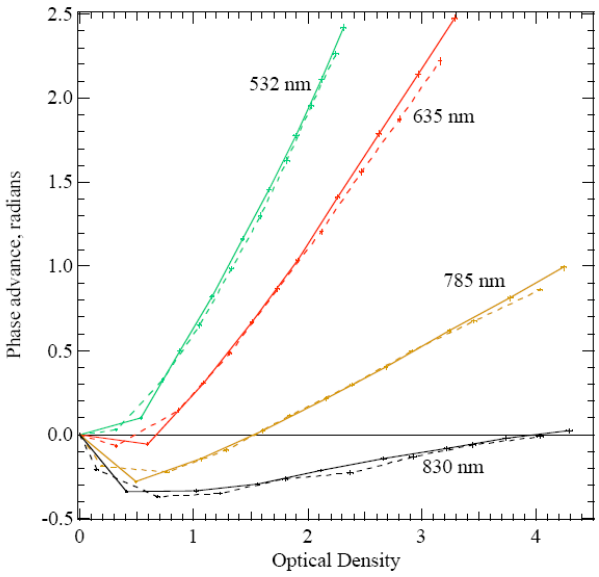


Figure 5: *Phase versus optical density of HEBS glass, at 4 wavelengths, from Halverson et al. (2005). When an occulting mask is written to have a band-limited absorption profile, the resulting variations in transmitted phase cause the mask not to be band-limited.*

MASK FABRICATION

As described in Section 4 above, one of the challenges of mask fabrication is finding a material (or a combination of materials) that allows sufficient control of OD and phase as a function of wavelength, so that a mask may be close enough to band-limited to achieve a target contrast. Two technologies have been tested at a significant level in the context of TPF-C and HCIT. The first is binary masks (see Fig. 3), with alternating areas of opaque and transparent areas, having a duty cycle that determines effective transmission on diffraction-sized spatial scales. The second is HEBS glass exposed by electron beams, to give a continuous analog absorption profile. Neither of these technologies are expected to deliver broadband ($\delta\lambda/\lambda \sim 10\%$ or wider) contrast at the 10^{-10} level, based on detailed modeling (and supported by experiments on HCIT).

Current mask investigations are exploring continuous analog metallic occulters as well as combined metallic and dielectric occulters. The modeling of these occulters, whose materials should be readily manufacturable, currently predict contrast values between 10^{-10} and 10^{-9}, for $\delta\lambda/\lambda = 10\%$. Further improvements, to the 10^{-10} level, are anticipated with additional investigations into metallic properties, currently underway.

Tolerances on random occulting mask errors (or any other image-plane errors) have been calculated by Lay et al. (2005) for contrast levels appropriate to TPF-C. These requirements are consistent with the current limits of superpolished surfaces.

DISCUSSION

Relative to other coronagraph architectures, band-limited masks offer modest throughput, excellent robustness to dynamic aberrations, low optical complexity, and unrivaled experimental verification. The experimental verification, and the maturity of the models that accompany band-limited masks, make this approach relatively low-risk compared to other architectures, at their current state of development.

The theoretical band-limited mask throughput is better than that of shaped pupils, similar to that of the visible nuller, and worse than those of optical vortices, phase-induced amplitude apodization, or external occulters. The diffractive efficiency (PSF width) compares in the same ways with other architectures.

The construction of occulting masks accurate enough in phase and OD for broadband performance at TPF-C levels has not been demonstrated, but is not expected to lie outside of current state-of-the-art. Current band-limited masks have achieved the best experimental contrast to date, below 10^{-9} monochromatically. The requirements on other optics are the same as for other coronagraph architectures, with the exception of the external occulter (which has very different optical requirements).

To summarize, band-limited masks have modest theoretical performance, but offer the most evidence of any architecture to date that they will ultimately be capable of delivering TPF-C performance, a challenging task.

REFERENCES

Balasubramanian, K., Hoppe, D.J., Mouroulis, P.Z., Marchen, L.F., Shaklan, S.B. 2005, "Polarization compensating protective coatings for TPF-Coronagraph optics to control contrast degrading cross polarization leakage," Proc. SPIE 5905.

Crepp, J. 2006 in prep.

Halverson, P.G., Ftaclas, M.Z., Balasubramanian, K., Hoppe, D.J., Wilson, D.W. 2005, "Measurement of wavefront phase delay and optical density in apodized coronagraphic mask materials," Proc. SPIE 5905.

Kuchner, M.J., Traub, W.A. 2002, "A coronagraph with a band-limited mask for finding terrestrial planets," ApJ 570, 900-908.

Kuchner, M.J., Spergel, D.N. 2003, "Notch-filter masks: practical image masks for planet-finding coronagraphs," ApJ 594, 617-626.

Kuchner, M.J., Crepp, J., Ge, J. 2005, "Eighth-order image masks for terrestrial planet finding," ApJ 628, 466-473.

Lay, O.P., Green, J.J., Hoppe, D.J., Shaklan, S.B. 2005, "Coronagraphic mask tolerances for exo-Earth detection," Proc. SPIE 5905.

Shaklan, S.B., Green, J.J. 2005, "Low-order aberration sensitivity of eighth-order coronagraph masks," ApJ 628, 474-477.

Shaklan, S.B., Marchen, L., Green, J.J., Lay, O.P. 2005, "The Terrestrial Planet Finder Coronagraph dynamics error budget," Proc. SPIE 5905.

Shaklan, S.B., Green, J.J. 2006, "Reflectivity and optical surface height requirements in a broadband coronagraph. 1. Contrast floor due to controllable spatial frequencies," Appl. Opt. 45, 5143-5153.

Shaklan, S.B., Green, J.J., Palacios, D.M. 2006, "The Terrestrial Planet Finder Coronagraph optical surface requirements," Proc. SPIE 6265.

Sivaramakrishnan, A., Koresko, C.D., Makidon, R.B., Berkefeld, T., Kuchner, M.J. 2001, "Ground-based coronagraphy with high-order adaptive optics," ApJ 552, 397-408.

Optical Vortex Coronagraphy

David M. Palacios[1], G.A. Swartzlander Jr.[2], and D. Mawet[3]
[1]*Jet Propulsion Laboratory, California Institute of Technology*
[2]*The School of Optical Sciences, The University of Arizona*
[3]*Institut d'Astrophysique et de Geophysique, Universite de Liege*

INTRODUCTION

Optical Vortex Coronagraphy is a new type of high contrast imaging technique under investigation for the Terrestrial Planet Finder Coronagraph (TPF-C). An optical vortex coronagraph (OVC) replaces the amplitude mask typically found in a Lyot Coronagraph with a spiral etched phase plate called an optical vortex mask (OVM). Ideally an OVC cancels light from a star over the entire pupil.[1] This allows for higher planet light throughput and a tighter inner working angle. In addition, an OVC is relatively insensitive to low-order wave front aberrations of the entrance pupil.[2] Recent experiments[3] have demonstrated the low-contrast application of an OVC. An achromatic OVM for low-contrast applications has also been proposed.[4] Using a similar technique it may be possible to produce an achromatic OVM suitable for high contrast applications. In this proceeding, we will address the high contrast performance of an OVC limited by the manufacturability of the OVM. An ideal OVM contains a smoothly varying spiral. However, the spiral phase shape of a real OVM will be pixilated by the lithography technique used to etch it. We simulated this effect with a currently manufacturable non-ideal OVM composed of a grid of square, 0.2 μm × 0.2 μm pixels and present the predicted high contrast performance of an OVC.

WHAT IS AN OPTICAL VORTEX?

Vortices are a ubiquitous feature of waves in nature that are known to occur in a variety of physical systems[5–7]. In optical systems, vortices manifest themselves as dark cores of destructive interference that occur at phase singularities. We can use this dark core to attenuate an on-axis star so nearby planets may be detected.[8] An optical vortex embedded in the center of a monochromatic light beam propagating in the z direction may be described by the electric field:[7]

$$E(\rho, \phi, z, t) = A(\rho, z)\exp(im\phi)\exp(i\omega t - ikz) \quad (1)$$

where $(\rho, \phi\ z)$ are cylindrical coordinates, $A(\rho, z)$ is a circularly symmetric amplitude function, m is a signed integer known as the topological charge of the vortex, and $k = 2\pi/\lambda$ is the wave number of a monochromatic field of wavelength λ. At any fixed instant of time, the vortex field described by Equation (1) produces helical surfaces of constant phase given by $m\phi - kz = \text{const}$, where m is a signed integer known as the topological charge, which is given by:[7]

$$m = (1/2\pi)\oint \nabla\phi\, ds \quad (2)$$

where $\nabla\phi$ is the phase gradient and ds is a line enclosing the helical axis of the vortex. Along the axis of this helix ($\rho = 0$) the amplitude vanishes owing to destructive interference and a dark vortex core is created, i.e., $A(0, z) = 0$.

Optical vortices may be created by a variety of different methods.[7] The technique most germane to coronagraphy is the use of an OVM, a type of diffractive optical element we will describe in the next section. Optical vortex holograms and sub-wavelength gratings are also promising methods and will be discussed briefly in a later section.

AN OPTICAL VORTEX MASK

An optical vortex is created when a planar (m = 0) monochromatic beam is passed through a transparent diffractive phase mask having a thickness that varies azimuthally. The thickness of an OVM is given by,[9]

$$dz = m\lambda_0 d\phi / 2\pi(n_s - n_0) \tag{3}$$

where λ^0 is the wavelength for which the mask is intended, $d\phi$ is an azimuthal rotation about the center of the OVM (in units of radians), n_s is the refractive index of the substrate, and n_0 is the index of refraction of the surrounding medium. A simulated 3-D plot of an OVM described by Equation (3) is displayed in Figure 1.

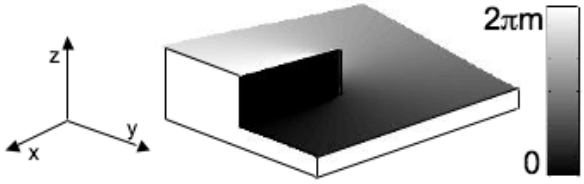


Figure 1: *A three dimensional plot of an optical vortex mask of topological charge m. The thickness of the mask (the z dimension) increases in the azimuthal direction forming a spiral shape on the upper surface of the mask. The thickness of the mask is represented by the phase shift through the mask in radians.*

The light transmitted through an OVM will gain an azimuthally varying phase as stated above, and in addition will also gain a vortex core of zero amplitude along the axis of the helical phase front. The vortex core created by an OVM has a physical extent, not unlike the eye of a hurricane. The vortex core extent may be derived using the simple ray model illustrated in Figure 2.

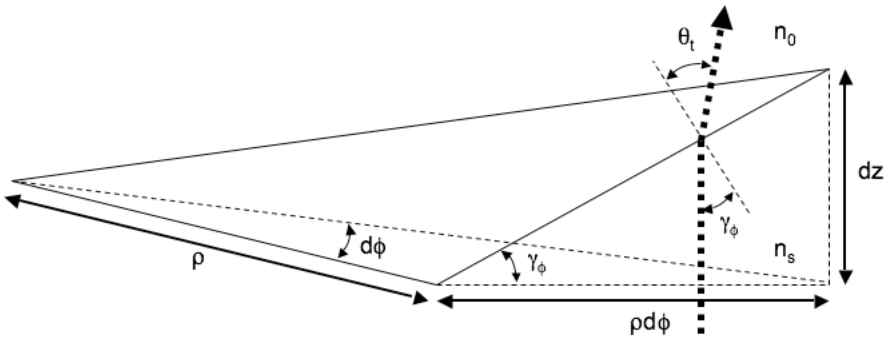


Figure 2: *A ray diagram of light transmission through an azimuthal wedge of an optical vortex mask. The thickness (dz) of the wedge increases as the angular width of the wedge ($d\phi$) increases in the azimuthal direction. A ray (dashed arrow) is perpendicularly incident on the mask at a radial position ρ. The ray is transmitted through the mask substrate of refractive index n_s and is incident on the spiral front surface of the mask at an angle γ_ϕ from the surface normal. The ray refracts at an angle θ_t from the surface normal into the surrounding medium of refractive index n_0.*

The figure depicts a ray perpendicularly incident on the flat backside of the OVM a distance ρ from the center of the mask and at an angle $d\phi$ from the edge of the deepest etch. The ray passes through the OVM glass substrate of refractive index n_s and refracts at the spiral shaped output side of the OVM into the surrounding medium of refractive index n_0. Using geometry it is possible to show that the angle of incidence (γ_ϕ) between the ray and the spiral surface is given by,

$$\gamma_\phi = \tan^{-1}\left[m\lambda_0 / 2\pi\rho(n_s - n_0)\right] \tag{4}$$

Substituting Equation (3) into Equation (4) yields:

$$\gamma_\phi = \tan^{-1}\left[dz/\rho d\phi\right] \tag{5}$$

By examining Equation (5) we can deduce that the closer the ray is to the center of the mask ($\rho \to 0$) the larger the value of γ_ϕ. The ray will totally internally reflect (TIR) when $\gamma_\phi = \gamma_{TIR} = \sin^{-1}[n_0/n_s]$.[10] It is possible to derive the vortex core size, ρ_c, for an optical vortex created by an OVM by solving for ρ in Equation (5) and setting $\gamma_\phi = \gamma_{TIR}$:

$$\rho_c = m\lambda_0/2\pi(n_s - n_0)\tan[\gamma_{TIR}] \tag{6}$$

All rays incident on the mask within ρ_c totally internally reflect and the transmitted amplitude is zero.

However, this does not completely represent the amplitude transmission of an OVM. When an electromagnetic field is incident on an interface between two dielectric mediums the amplitudes of the reflected and transmitted portions of the field will vary with the angle of incidence and the polarization state of the impinging light.[10] Referring to the geometry depicted in Figure 2, if the electric field is perpendicular to the plane of incidence (defined by the incident ray and the spiral surface) it is radially polarized (E_ρ) and if the electric field is parallel to the plane of incidence it is azimuthally polarized (E_ϕ). The relative amplitude of the transmitted radially polarized light for $\rho > \rho_c$ is given by:

$$\frac{E_\rho}{E_0} = 1 - \frac{n_s\cos(\gamma_\phi) - \frac{\mu_s}{\mu_0}\sqrt{{n_0}^2 - {n_s}^2\sin^2(\gamma_\phi)}}{n_s\cos(\gamma_\phi) + \frac{\mu_s}{\mu_0}\sqrt{{n_0}^2 - {n_s}^2\sin^2(\gamma_\phi)}} \tag{7}$$

where μ_0 is the magnetic permeability of free space and μ_s is the magnetic permeability of the OVM substrate material. The relative amplitude of the transmitted azimuthally polarized light for $\rho > \rho_c$ is given by:

$$\frac{E_\phi}{E_0} = 1 - \frac{\frac{\mu_s}{\mu_0}{n_0}^2\cos(\gamma_\phi) - n_s\sqrt{{n_0}^2 - {n_s}^2\sin^2(\gamma_\phi)}}{\frac{\mu_s}{\mu_0}{n_0}^2\cos(\gamma_\phi) + n_s\sqrt{{n_0}^2 - {n_s}^2\sin^2(\gamma_\phi)}} \tag{8}$$

Radial line plots of the amplitude transmission of OVMs of topological charge, m = 2, 4, and 6 are compared in Figure 3.

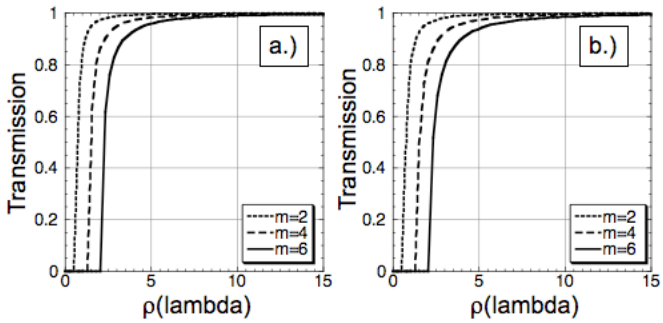


Figure 3: *Radial line plots of the amplitude transmission of optical vortex masks of topological charge, m = 2, 4, and 6 for a) radially polarized light and b) azimuthally polarized light.*

Although the line plots of the radial and azimuthal polarizations seen in Figure 3 look very similar they are slightly different. When linearly polarized light is considered, this leads to a small asymmetry in the vortex core shape. For linear polarized light, the vortex core will be slightly elongated along the axis perpendicular to the

polarization state. This will have a small effect on the performance of the coronagraph that should be noted, although it will be a much smaller effect than the discretization of the OVM.

Previous work on an OVC assumed an ideal OVM with a smooth varying spiral. However, a real OVM will possess a discretized spiral because the etching techniques used to carve the spiral have a minimum etchable feature size. The OVM is a mode converter, and converts the planer non-vortex mode of the telescope into a desired vortex mode of topological charge, m. When the OVM is not manufactured to the desired ideal smooth spiral, the mode conversion will not be complete and other modes of topological charge, $m' \neq m$, will also be created. This includes a non-vortex mode (m = 0), which will cause stellar light leakage in the OVC. Using present ion-beam lithography techniques it is possible to etch an m = 6 OVM for a wavelength of λ_0 = 600nm, with a minimum etchable feature size of approximately 50nm.* Due to simulation array size constraints (4096 × 4096 arrays) the numerical simulations presented in this proceeding contain an OVM with a mask feature size of 0.2 microns.

CORONAGRAPH ARCHITECTURE

A simple unfolded architecture of an OVC is depicted in Figure 4. Light from a distant star is imaged by a telescope represented by L_1. An OVM of topological charge m is placed near the focus of L_1 and a collimator represented by L_2 re-images the entrance pupil of the telescope system after the starlight is transmitted through the OVM. A circular Lyot stop is placed at the re-imaged pupil blocking unwanted starlight at the edge of the pupil. The remaining light is then re-imaged by L_3 with the on-axis starlight greatly attenuated compared to the off-axis light from a planet.

For this proceeding, we assume circular instead of elliptical symmetry since the present TPF-C telescope design contains beam-circularizing optics to increase the efficacy of the wave-front control system. Therefore, ignoring polarization effects, the optimal Lyot stop is a circular aperture.

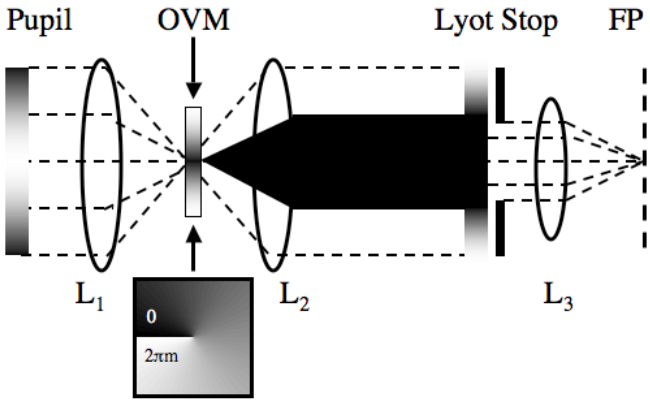


Figure 4: *A simple unfolded model of an optical vortex coronagraph. Lens (L_1) represents the telescope optics, which focus the light from the entrance pupil onto an optical vortex mask (OVM) of topological charge, m. Lens (L_2) collimates the light forming an exit pupil where a Lyot stop is placed. A third lens (L_3) re-images the light to the final image at (FP).*

PERFORMANCE SIMULATIONS

The high contrast performance of an OVC was simulated with an ideal imaging system possessing the following parameters:

- All simulations were performed on a 4096 × 4096 grid array.
- The OVM was designed to produce either an m = 2, 4, or 6 charged vortex at λ_0= 600nm.

* Using advanced techniques currently under investigation at JPL it is possible to smooth features without the use of smaller etchable feature sizes.[11-13]

- D=100 pixels.
- The focal spot incident on the OVM was assumed to be an ideal Airy disk with a first zero radius of 51 pixels.
- The star was imaged onto the OVM with an f# = 27 telescope, which matches the design of HCIT at JPL.
- The amplitude transmission function of the OVM was an incoherent superposition of Equation (7) and Equation (8).
- $\mu_s = \mu_0 = 1$.
- $n_s = 1.5$ and $n_0 = 1$.
- The OVM was discretized with a mask feature size of 0.2 microns.
- Contrast as defined by Green[14] et al. was used as a metric of system performance.

Simulated OVC Lyot plane images for an on-axis star are shown in Figure 5.

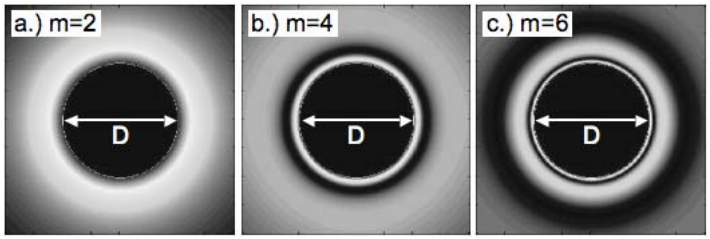


Figure 5: *Images of the Lyot plane of an optical vortex coronagraph with an optical vortex mask of topological charge a) m = 2, b) m = 4, and c) m = 6. The residual stellar light is redistributed outside the pupil of diameter, D.*

One remarkable feature of an ideal OVC is its ability to null over the entire pupil when m is an even integer.[1, 15] The discretization of the OVM causes leakage of stellar light reducing this effect. However, even with the discretized OVM the power inside the pupil was reduced by several orders of magnitude and we were able to verify the Lyot plane intensity profile predicted by Foo et al.[1] (for the m = 2 case) and by Mawet et al.[15]

TPF-C is presently designed to have an inner working angle of 4λ/D. This requires the use of an 8m diameter primary mirror, which is costly and difficult to manufacture. If TPF-C could have an inner working angle of 2λ/D, a 4m diameter primary mirror could be used, which would greatly reduce the cost and difficulty in manufacturing the mirror. Since an OVC is able to null such a large area of the pupil we optimized the performance of the OVC to obtain a contrast below 10^{-10} at 2λ/D. When the OVC contains a discretized OVM the Lyot stop size must be reduced to optimize the system for high contrast performance. The average contrast between 2–3λ/D vs. the Lyot stop diameter is plotted in Figure 6(a) for m = 2, 4, and 6 OVCs.

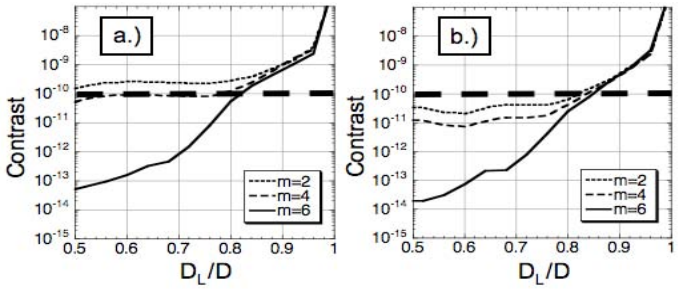


Figure 6: *Plots of the contrast averaged between a) 2–3λ/D and b) 4–5λ/D vs. Lyot stop diameter, D_L, for optical vortex coronagraphs of topological charge, m = 2, 4, and 6. The Lyot stop diameter is normalized by the pupil diameter, D. The thick dashed line at 10^{-10} represents the contrast requirement of the Terrestrial Planet Finder Mission.*

Using these results, we may conclude that a contrast below 10^{-10} at 2λ/D may be obtained by an m = 4 or m = 6 OVC by reducing the Lyot diameter to 0.8D. To compare these results to other coronagraph architectures, the average contrast between 4–5λ/D vs. the Lyot stop diameter is plotted in Figure 6(b). The results show a similar trend although better starlight suppression is obtained in general and now the m = 2 OVC performs better than the required 10^{-10} contrast required for terrestrial planet detection.

Another important quantity to consider is the amount of planet light that is transmitted through the system to the final image plane. A coronagraph may obtain high levels of contrast between a star and a planet but still fail to detect a planet because so few planet photons make it through the system. Besides this minimum detectable signal limitation, it is always desirable to increase the signal from the planet because this will decrease the image integration time. This increases the scientific gain that may be achieved within the mission's lifetime. Planet light throughput may be defined as the amount of planet light power transmitted through the Lyot stop. It's reduced primarily from the Lyot stop but may also be reduced by the on-axis vortex core as well. As a result, the planet light throughput will decrease for OVCs that employ higher m value OVMs, because the size of the vortex core increases with topological charge (see Fig. 3). The planet light throughput vs. Lyot stop size for m = 0, 2, 4, and 6 OVCs are plotted in Figure 7 for (a) a planet/star separation of $2\lambda/D$ and (b) a planet/star separation of $4\lambda/D$.

The m = 0 plot in Figure 7 is the no OVM case and represents the maximum power the Lyot stop alone will allow through the system. These results verify our expectation that the planet light throughput decreases as the value of m increases. The plots also show that the Lyot stop throughput itself is the dominant degrader of planet light throughput for smaller Lyot stop sizes. For an m = 4 OVC with $D_L = 0.8D$, a planet separated from its host star by $2\lambda/D$ had a throughput of $0.53P_T$, where P_T is the total planet light power. A planet separated from its host star by $4\lambda/D$ had a throughput of $0.62P_T$.

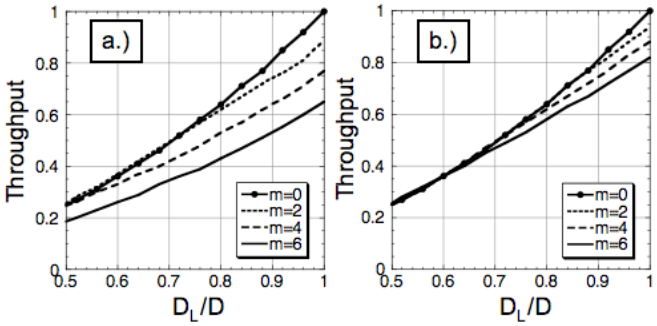


Figure 7: *Plots of the planet light throughput for a planet located at a) $2\lambda/D$ and b) $4\lambda/D$ vs. Lyot stop diameter, D_L, for optical vortex coronagraphs of topological charge, m = 0, 2, 4, and 6. The Lyot stop diameter is normalized by the pupil diameter, D. The m = 0 case represents the throughput limited by the Lyot stop without an optical vortex mask present.*

VORTEX MASK ACHROMATIZATION

In the search for terrestrial planets few photons may be wasted since the planet signal is so faint. Consequently, TPF-C is required to have a large operating spectral bandwidth to enhance the detected signal. In order for an OVC to be a feasible solution for TPF-C an achromatic design must be implemented. The topological charge imposed by an OVM on light having an arbitrary vacuum wavelength, λ, may be determined from the expression

$$m(\lambda) = m_0 \frac{\lambda_0}{\lambda}\left[\frac{n_s(\lambda) - n_0(\lambda)}{n_s(\lambda_0) - n_0(\lambda_0)}\right] \tag{9}$$

Equation (9) is a dispersive relation, showing that the topological charge varies continuously with the wavelength of the beam. If material dispersion can be ignored, $m(\omega) = m_0\omega/\omega_0$ and thus the vortex spectrum is given by $C_l(\omega) = sinc(m_0\pi\omega/\omega_0 - l\pi)$.[16] If $|\omega - \omega_0|/\omega_0 << 1$, or equivalently $|\lambda - \lambda_0|/\lambda_0 << 1$, the zero-order topological spectrum (corresponding to a non-circulating transmitted field) may be written $C_0^2(\omega) \approx (\delta\lambda/\lambda_0)^2$, where $\delta\lambda = \lambda - \lambda_0$. Thus, when the wavelength is mismatched from λ_0, a fraction of the transmitted beam is unaffected by the OVM.

Achromatic optical components may be created by combining elements that have different refractive properties. An achromatic OVM requires the condition $m(\lambda) = m_0$ to be satisfied across a wide bandwidth of wavelengths. Using Equation (9) the achromatic condition may be expressed $m(\lambda)/m_0 = 1$. The successful fabrication of an achromatic OVM for use in the vicinity of λ_0 therefore depends on finding suitably matched materials.

Toward this end the refractive index may be expressed as a truncated Taylor series expansion:

$$n(\lambda) = n(\lambda_0) + n'(\lambda_0)(\lambda - \lambda_0) + (1/2)n''(\lambda_0)(\lambda - \lambda_0)^2 \tag{10}$$

where $n'(\lambda_0) = \left(\partial n / \partial \lambda\right)\Big|_{\lambda_0}$ and $n''(\lambda_0) = \left(\partial^2 n / \partial \lambda^2\right)\Big|_{\lambda_0}$. For many optical glasses the first two terms in Equation (10) are sufficient to describe the refractive index in the vicinity λ_0 across a bandwidth of roughly 100 nm. Combining Equation (9) and Equation (10) the achromatic condition may be written as an error parameter

$$\varepsilon = \left(n_s - n_0\right)_{\lambda_0} - \lambda_0\left(n_s' - n_0'\right)_{\lambda_0} + (1/2)(\delta\lambda)^2\left(n_s'' - n_0''\right)_{\lambda_0} \tag{11}$$

where $\varepsilon = 0$ satisfies the condition $m(\lambda)/m_0 = 1$. The first term in Equation (10) cannot be made equal to zero; otherwise $\Delta z = \infty$ (see Equation (3)). In practice the value of Δz should be made smaller than the characteristic diffraction length of the beam. Setting $\delta\lambda = 0$, a pair of materials that satisfies $\varepsilon \approx 0$ at a single wavelength, λ_0, may be explored. For optical glasses the second term in Equation (10) is on the order of 10^{-2} and thus the index difference between the glasses must also be small. This difference will result in small Fresnel reflections at the interface, but will also require a value of Δz on the order of $100\lambda_0$. To obtain smaller values of Δz other materials must be explored. Dispersion equations such as the Sellmeier formula may be used to determine the refractive index and its derivatives for a given material. An analysis of more than 100 Schott glasses shows the trend, $|\varepsilon| \approx 1.2 \times |n_s(\lambda_0) - n_0(\lambda_0)|$. Thus large index differences (which are desirable to achieve small values of Δz) require large achromatic errors for these glasses.

The relative topological charge, $m(\lambda)/m_0$, and the topological transmission spectrum, C_l^2, may be numerically computed across the spectrum for a given pair of glasses. For example Figure 8 shows the values for Schott glasses N-LASF44 and N-SF14, and for the parameters $m_0 = 2$ and $\lambda_0 = 550$ nm.

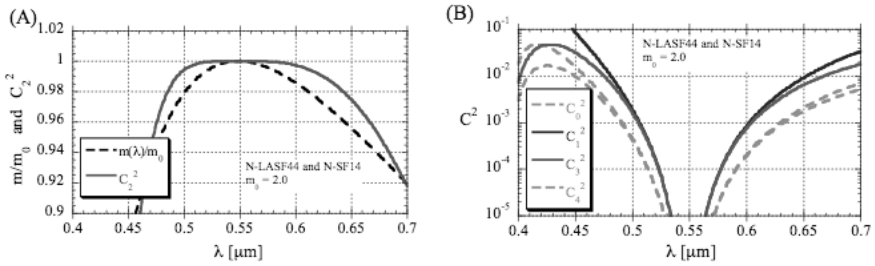


Figure 8: *Achromatic vortex lens having (m_0 =2) at a single wavelength. The matched materials are Schott glasses N-LASF44 and N-SF14. (A) Topological charge and transmission into the intended mode. The integration of $C_2^2(\lambda)$ across a bandwidth of 130 nm indicates that 99.8% of the beam is converted into the $l = m_0 = 2$ mode. (B) Transmission into neighboring modes, indicating no unintended modes at the design wavelength, $\lambda_0 = 550$ nm.*

The bandwidth of the element depends on the fidelity of the transmission into the desired mode, m_0. For example, the 99.8% bandwidth, $\lambda_{max} - \lambda_{min}$, may be established by use of the integral

$$\left(\lambda_{\max} - \lambda_{\min}\right)^{-1} \int_{\lambda_{\min}}^{\lambda_{\max}} C_{m_0}^2 \, d\lambda = 0.998 \tag{12}$$

The data in Figure 8 has a 99.8% bandwidth of 130 nm. In comparison, the 99.8% bandwidth of a chromatic fused silica vortex phase mask in air is 20 nm. As expected, Figure 8(B) shows no transmission into neighboring modes ($l \neq 2$) at λ_0.

OTHER VORTEX CORONAGRAPH ARCHITECTURES

The benefits of an OVC: high planet light throughput, low sensitivity to aberrations, tight inner working angle, are properties inherent to vortex fields and are not dependent on the vortex creation method. However, achromatization and mask manufacturability concerns will depend on the vortex creation method. Above we have explored vortex creation by a helical phase plate, but there are two other optical vortex creation techniques that have been proposed as OVM alternatives. They are the annular groove phase mask (AGPM) and the holographic vortex mask (HVM).

ANNULAR GROOVE PHASE MASK

It is possible to use a diffractive grating instead of a phase plate to induce the helical phase structure of an optical vortex. When the period of a grating is smaller than the wavelength of the incident light, it does not diffract as a classical spectroscopic grating. Instead, all the incident energy is forced to propagate only in the zeroth order, leaving the incident wave front free from any further aberration. This type of grating behaves like a homogeneous medium with unique characteristics that can be used to synthesize birefringent achromatic waveplates. The key point is that by carefully controlling the geometry of the grating structure, one can tune the form of birefringence and induce a precise phase shift between the orthogonal polarization components of the incident light. The phase shifting optical function can then be integrated on the surface of a substrate to implement the phase shifting spatial distribution of an AGPM[15] as shown in Figure 9(A).

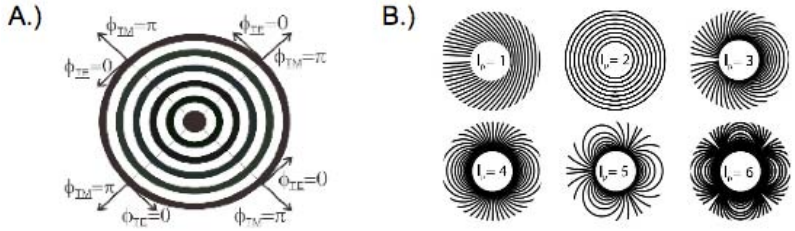


Figure 9: *A) m = 2 AGPM implementation. The AGPM consists of a concentric circular surface-relief grating with rectangular grooves with depth h and a periodicity Λ. B) grating geometry for different topological charges (1–6).*

This technique has been explored in depth for application in the infra-red requiring 10^{-6} starlight suppression.[15] However, to apply this technique to TPF-C, a high topological charge AGPM as shown in Figure 9(B) would have to be developed for use in visible light. High topological charged AGPM coronagraphs for the visible spectrum are still under development.

HOLOGRAPHIC VORTEX MASK

It is also possible to create an optical vortex with a hologram,[7] which may be created by interfering the desired vortex beam with a plane wave. The resulting intensity pattern (see Fig. 10(A)) is then etched onto a substrate to produce a holographic diffraction grating.

To further improve the design the holographic mask may be etched directly onto a prism creating a vortex grism. This device is inherently broadband because the same topological charged vortex is produced for all wavelengths and the prism compensates for the angular dispersion of the diffraction orders. A possible coronagraph concept making use of this technique is depicted in Figure 10(B). This technique holds several advantages over a helical OVM. First it is inherently broadband and second, it doesn't require the lithographic etches to penetrate as deep into the mask substrate. This technique is currently under investigation.

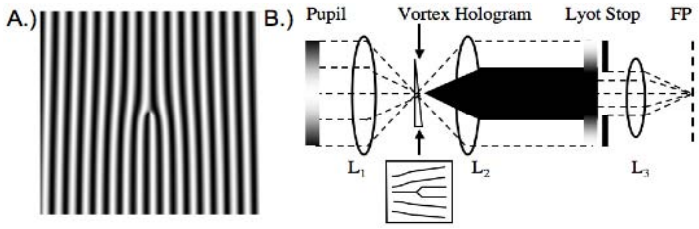


Figure 10: *A) An m = 1 vortex hologram. B) A holographic vortex coronagraph employing a vortex grism.*

CONCLUSIONS

An OVC employing an optical vortex mask etched by today's ion-beam technology is able to achieve the 10 order of magnitude starlight suppression required of TPF-C. Our simulations verify the Lyot plane behavior of an OVC predicted by Foo et al.[1] and Mawet et al.[15] An m = 4 OVC with $D_L = 0.8D$, reached the required 10^{-10} starlight suppression and obtained a planet light thoughput of $0.53P_T$ for a planet separated from its host star by $2\lambda/D$. An OVC may hold several key advantages for TPF-C. These advantages include high planet light throughput, manufacturability, lower aberration sensitivity,[2] and possible broadband operation.[4] Vortex masks with topological charge ranging from m = 1–3 have been manufactured.[12,13] Using present ion-beam lithography techniques, it may be possible to construct an m = 6 vortex mask designed for an optical wavelength of $\lambda_0 = 600$nm. In addition, alternative architectures may relax achromatic OVC design tolerances. Two alternative architectures have been suggested, the annular groove phase mask and the holographic vortex coronagraph, both of which are presently under investigation.

ACKNOWLEDGMENTS

David M. Palacios would like to thank S. Shaklan and D. Wilson for their suggestions and comments. This work was carried out at: the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration, the University of Arizona supported by a grant from the U.S. Army Research Office, and at the University of Liege.

REFERENCES

1. G. Foo, D.M. Palacios, G.A. Swartzlander Jr., Opt. Lett. **30**, 3308 (2005).
2. D.M. Palacios and S. Hunyadi, To appear in Optics Letters (2006).
3. Jae Hoon Lee, Gregory Foo, Eric G. Johnson, and Grover A. Swartzlander Jr., Phys. Rev. Lett. Article # 053901 (2006).
4. G.A. Swartzlander Jr., Opt. Lett. **31**, 2042 (2006).
5. L.M. Pismen, *Vortices in Nonlinear Field*s, Oxford Science Publications (Clarendon Press, Oxford, 1999).
6. Sheldon Green eds., *Fluid Vortices*, (Kluwer Academic Publishers, Dordrecht, Netherlands, 1995).
7. M. Vasnetsov and K. Staliunas, eds., *Optical Vortices*, vol. 228, Horizons in World Physics, (Nova Science, Huntington, NY, 1999).
8. G.A. Swartzlander Jr., Opt. Lett. **26**, 497 (2001).
9. F.B. Colstoun, G. Khitrova, A.V. Fedorov, T.R. Nelson, C. Lowery, T.M. Brennan, B.G. Hammons, and P.D. Maker, Chaos, Solitons, & Fractals **4**, 1575 (1995).
10. J.D. Jackson, *Classical Electrodynamics*, (John Wiley and Sons, NY, 1962).
11. D. W. Wilson, R. E. Muller, P. M. Echternach, and J. P. Backlund, Proceedings of SPIE, **5720**, 68 (2005).
12. Z. G. Chen, M. Segev, D. W. Wilson, R. E. Muller, and P. D. Maker, Opt. Lett. **22,** 1751 (1997).
13. Z. G. Chen, M. Segev, D. W. Wilson, R. E. Muller, and P. D. Maker, Phys. Rev. Lett. **78**, 2948 (1997).
14. Green, J.J., and Shaklan, S.B., Proc. SPIE **5170**, 25 (2003).
15. D. Mawet, P. Riaud, O. Absil, and J. Surdej, ApJ. **633**, 1191 (2005).
16. G.A. Swartzlander Jr., Opt. Lett. **30**, 2876 (2005).

The Phase-Induced Amplitude Apodization Coronagraph

Olivier Guyon[1], Eugene Pluzhnik[1], Shinishiro Tanaka[1], Basile Gallet[1], Stephen Colley[1], Robert Vanderbei[2], Stephen Ridgway[3], Roger Angel[4], Robert Woodruff[5], Mark Warren[6]

[1]Subaru Telescope, [2]Princeton University, [3]NOAO, [4]University of Arizona, [5]Lockheed Martin, [6]Axsys Technologies

ABSTRACT

The Phase-Induced Amplitude Apodization Coronagraph (PIAAC) uses lossless amplitude apodization (performed by reflection on aspheric mirrors) to produce a high contrast stellar PSF.

It combines nearly 100% throughput with small inner working angle (<2 λ/D), preserves the angular resolution of the telescope, is sufficiently robust to stellar angular size and can be designed to have very good achromaticity. The theoretical performance of the PIAAC would enable TPF-C science with a telescope half the size of what would be required if a coronagraph utilizing conventional apodization were used.

In this paper, we show how the PIAAC functions and what is its expected performance. We also discuss PIAA optics manufacturing challenges, design trade-offs and results from our ongoing laboratory demonstration.

PIAAC 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 is 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 Figure 1, 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).

EXPECTED PERFORMANCE

The PIAAC coronagraph performance is quantified and compared with other coronagraphs in the "Theoretical analysis of coronagraphs" paper in this volume. We summarize here the main characteristics of the coronagraph:

- **nearly 100% throughput** (see section "Design trade-offs: (2) Mild pupil apodization with conventional apodizer")
- 1λ/D angular resolution
- <2 λ/D IWA
- full 360 deg search area
- **good achromaticity** (see section "Design trade-offs: (2) Mild pupil apodization with conventional apodizer")

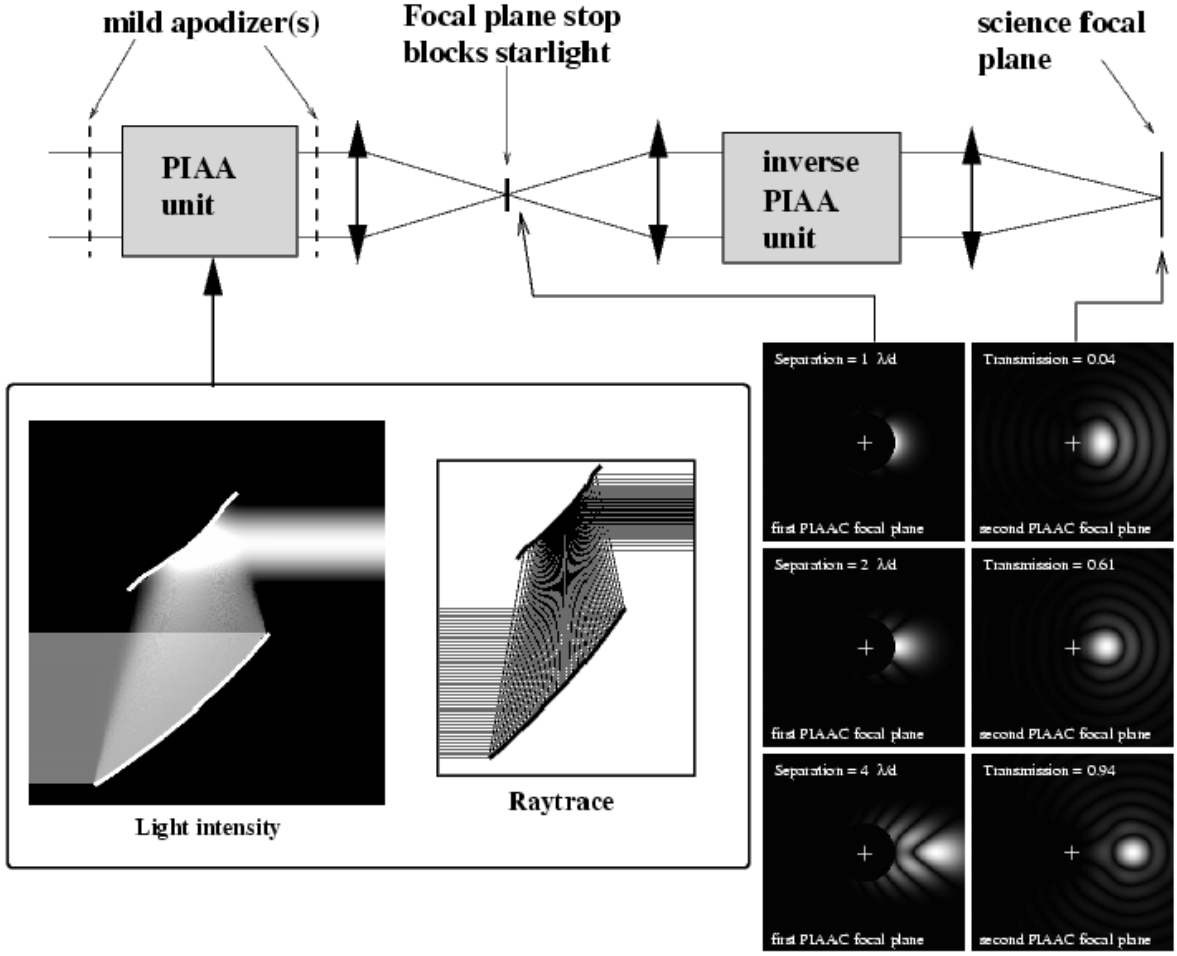


Figure 1: *Schematic representation of the PIAAC. The telescope light beam enters from the left and is first apodized by the PIAA unit. Mild apodizer(s) are used to perform a small part of the apodizations, and are essential to mitigate chromatic diffraction propagation effects and to allow for the design of "friendly" aspheric PIAA mirrors. A high contrast image is then formed, allowing starlight to be removed by a small occulter. An inverse PIAA unit is required to "sharpen" the image of off-axis sources.*

DESIGN TRADE-OFFS

(1) Focal plane occulter design and size

The PIAAC is a low-IWA coronagraph, and is therefore quite sensitive to low-order aberrations. Fortunately, this also means that low order aberrations can be efficiently and rapidly measured from the science focal plane.

For the PIAAC, as well as for any coronagraph with a focal plane occulter, the occulter should preferentially be reflective, so that the bright starlight can be used to measure low order aberrations before they start to impact science observations. We have designed such a mask, and further optimized the concept by only using the light falling within the ~0.5 λ/D to ~1.5λ/D radius interval to maximize its sensitivity (in this mask design, the very center of the mask is opaque out to 0.5 λ/D, the mask is reflective from 0.5 to 1.5λ/D and transmissive outward of 1.5 λ/D).

Thanks to this scheme, low-order aberrations can be measured to the required accuracy within a fraction of a second: the wavefront stability requirements for low order aberrations is therefore greatly relaxed.

(2) Mild pupil apodization with conventional apodizer

A PIAAC in which the apodization is performed entirely by the aspheric optics faces a very serious challenge: the outer edge of the first PIAA mirror (PIAA M1) has a small radius of curvature over a very small region. Since a large outer region of the output apodized beam contains a very small fraction of the total light in the pupil, the remapping needs to expand a very narrow annulus at the outer edge of PIAA M1 into a broad annulus on PIAA M2: this explains why the outer edge of PIAA M1 exhibits this sharp narrow "bend". This brings 2 problems:

– PIAA M1 becomes very difficult to manufacture
– The very sharp feature at the edge of M1 creates unwanted diffraction effects which are unfortunately chromatic, and therefore reduce the spectral bandwidth over which the PIAAC can be used.

A solution to both problems is to share the apodization between a PIAA system and a mild apodizer at the output of the PIAA. Since the problem arises from the fact that the apodized pupil is very dark at its edges, it is solved by designing the PIAA to maintain the outer edge of the apodized beam at typically 1% of the surface brightness of the center of the beam. A mild apodizer then further reduces the flux in this already fairly dark part of the apodized beam. The conventional apodizer removes a small fraction of the total light (~10%) and has therefore a limited impact on the system throughput, IWA and angular resolution. A detailed analysis of this trade-off [Pluzhnik et al. 2006] shows that PIAA systems can be designed to be achromatic at the 1e- 10 level over the whole visible spectrum if one is willing to lose approximately 20% of the incoming light.

OPTICS MANUFACTURING

The PIAA optics are highly aspheric. The outer edge of PIAA M1 is the most challenging feature: it is strongly curved over a relatively small radius interval. With the PIAAC "hybrid" design described above (combination of PIAA apodization with a conventional apodizer), the PIAA optics shapes are within current manufacturing capabilities. High precision polishing of aspheric optics has been developed for extreme-UV lithography, and is now available from several vendors.

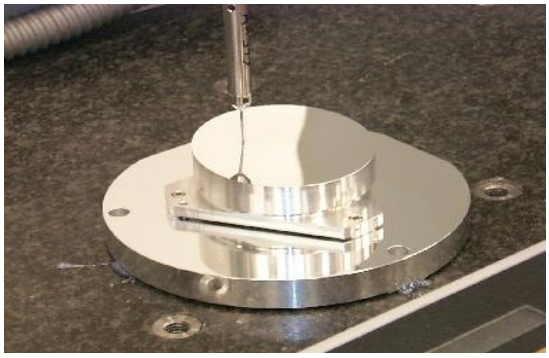


Figure 2: *PIAA M2 after diamond turning (prior to polishing)*

While each PIAA mirror can first be polished independently of the other (using a computer generated hologram—CGH—for null testing), the final step of fine polishing should aim to a flat wavefront in the 2-mirrors PIAA system.

The PIAA mirrors used in our laboratory demonstration (see next section) were polished by Axsys Technologies (MI, USA). The main fabrication steps are:

– diamond turning of an aluminum substrate
– Nickel coating of the aluminum parts
– Diamond turning of the nickel layer
– Hand polishing of each mirror against a CGH reference
– System assembly
– Fine hand polishing of one of the mirrors within the system to obtain a flat wavefront

The beam size was chosen to be large enough (75mm diameter) to accommodate the manufacturing process described above. The system wavefront obtained was 25nm RMS. Existing computer polishing and metrology processes would lead to significantly higher quality optics.

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. 3).

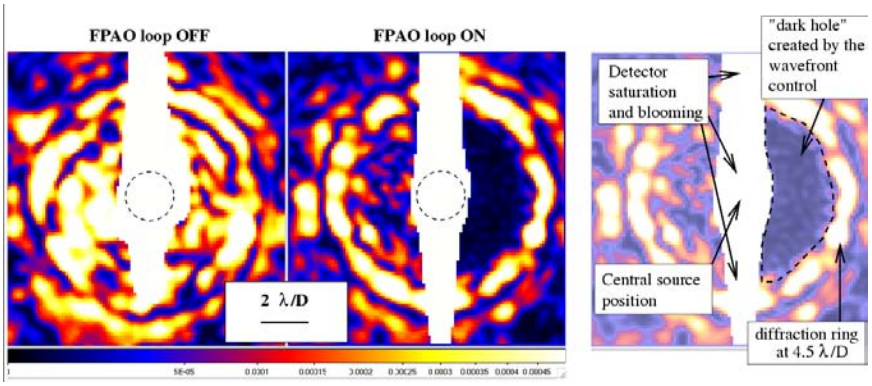


Figure 3: *Preliminary results from the PIAA coronagraph laboratory demonstration. A heavily stretched version of our laboratory PSF image (left) shows that almost all starlight is concentrated within 1.5 λ/D radius (the large vertical structure is due to charge bleeding on the detector). The light beyond 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.*

In the new optical layout (see Fig. 4), a pupil plane is made available for the conventional apodizer, and a focal plane is accessible for the focal plane occulter.

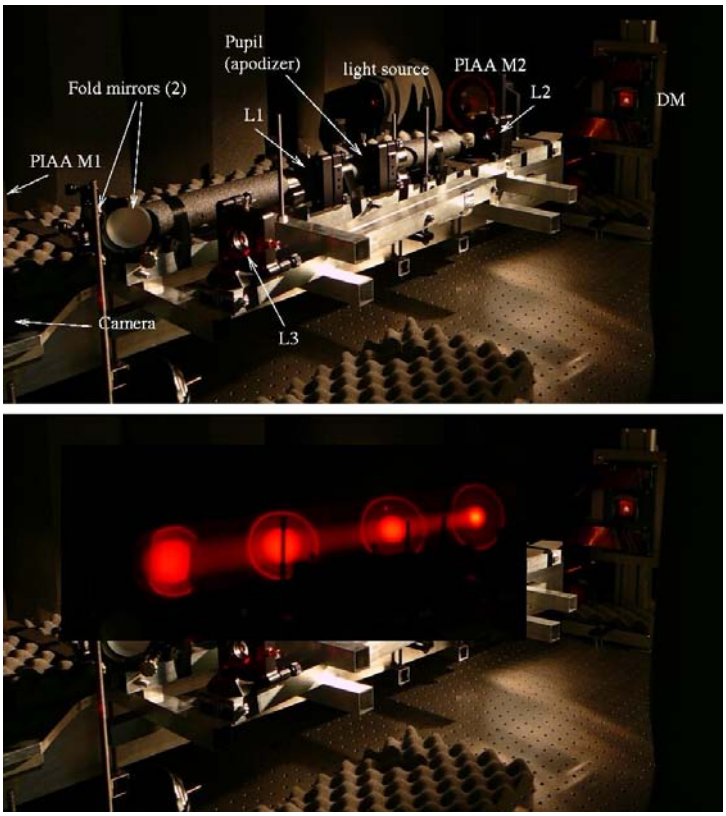


Figure 4: Layout of the PIAA laboratory experiment. In the bottom panel, a long exposure acquired while moving a white screen reveals the beam apodization betwen PIAA M1 and PIAA M2.

CONCLUSIONS / FUTURE WORK & STUDIES

The PIAA coronagraph is theoretically very efficient, and could allow TPF-like science goals with a moderate size visible telescope (about 3m). Preliminary laboratory results have confirmed the basic principle and demonstrated high throughput, low IWA coronagraphy at the ~1e 6 contrast level. We are currently improving our laboratory testbed to reach higher contrast levels.

PIAA differs from more conventional coronagraphs in 2 aspects, which each deserve further study:

- PIAA has a lower IWA: low order aberrations need to be kept very low. This may require a dedicated low-order WFS using light reflected by the central occulter.
- PIAA remaps the wavefront, which has strong implications for the overall design of the wavefront control optics. Placement of the DM(s) within the optical train (before and/or after the PIAA optics?) is critical.

More information about PIAAC in the following papers:

"Diffraction-based Sensitivity Analysis of Apodized Pupil-mapping Systems", Belikov, Ruslan; Kasdin, N. Jeremy; Vanderbei, Robert J., 2006 ApJ, 652, 833

"Exoplanet Imaging with a Phase-induced Amplitude Apodization Coronagraph. III. Diffraction Effects and Coronagraph Design", Pluzhnik, Eugene A.; Guyon, Olivier; Ridgway, Stephen T.; Martinache, Frantz; Woodruff, Robert A.; Blain, Celia; Galicher, Raphael, 2006 ApJ, 644, 1246

"Diffraction Analysis of Two-dimensional Pupil Mapping for High-Contrast Imaging", Vanderbei, Robert J., 2006 ApJ, 636, 528

"Exoplanet Imaging with a Phase-induced Amplitude Apodization Coronograph. II. Performance", Martinache, Frantz; Guyon, Olivier; Pluzhnik, Eugene A.; Galicher, Raphael; Ridgway, Stephen T., 2006 ApJ, 639, 1129

"Pupil Mapping in Two Dimensions for High-Contrast Imaging", Vanderbei, Robert J.; Traub, Wesley A., 2005 ApJ, 626, 1079

“Laboratory Demonstration and Numerical Simulations of the Phase-Induced Amplitude Apodization", Galicher, Raphael; Guyon, Olivier; Otsubo, Masashi; Suto, Hiroshi; Ridgway, Stephen, 2005 PASP, 117, 411

"Exoplanet Imaging with a Phase-induced Amplitude Apodization Coronagraph. I. Principle", Guyon, Olivier; Pluzhnik, Eugene A.; Galicher, Raphael; Martinache, Frantz; Ridgway, Stephen T.; Woodruff, Robert A., 2005 ApJ, 622, 744

"Two-Mirror Apodization for High-Contrast Imaging", Traub, Wesley A.; Vanderbei, Robert J., 2003 ApJ, 599, 695

"Phase-induced amplitude apodization of telescope pupils for extrasolar terrestrial planet imaging", Guyon, Olivier, 2003 A&A, 404, 379

ACKNOWLEDGMENTS:

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.

New Worlds Observer: An Occulter Based Concept for Terrestrial Planet Finding

Jonathan Arenberg[1], Amy Lo[1], and Charles Lillie[1], Webster Cash[2], Charley Noecker[3]

[1]Northrop Grumman Space Technology, [2]University of Colorado, [3]Ball Aerospce

New Worlds Observer (NWO) is a multi-spacecraft concept that is capable of detecting and characterizing extra-solar planetary systems, as well as obtaining general astrophysics observations. NWO is a rapidly emerging concept only recently introduced into the discussion as a viable concept[1, 2]. This architecture offers an affordable, modest risk option to achieving NASA's long standing goal of identifying and studying Earth-like planets around other stars.

The NWO concept consists of an external occulter and a generic space telescope. The occulter has specific requirements on its shape and size, while the telescope is an ordinary diffraction limited astronomical instrument comparable to the Hubble Space Telescope. This system requires a deep space orbit like that for JWST or Spitzer. The occulter is a specially shaped opaque screen (binary mask) that creates a deep shadow in the target star's light, in which the telescope is flown. This deep shadow created by the occulter is the essential element of the New Worlds concept that enables the detection and characterization of extra-solar planets with an ordinary astronomical telescope.

Mission concepts to detect extra-solar planets using an external occulter have a long history[3,4,5,6,7]. What is novel about New Worlds is the combination of deep, broad band starlight suppression (10^{-10} over an octave or more), a relatively small occulter, and a manufacturable binary-mask implementation[8]. This leap in starlight suppression, discussed below, enables the design of a practicable mission able to detect and characterize extra-solar terrestrial planets.

We briefly review the key factors in the performance of our mission enabling occulter, including our choice of a mask shape, which allows a closed form algebraic analysis[8] and a numerical evaluation of the diffraction problem. Following this overview of the occulter's performance and basic scaling laws, we introduce a concept for a mission design. The conclusion lists the ongoing and planned work on New Worlds.

The New Worlds Observer occulter provides high starlight suppression because its shape manages diffraction very well. It does so by using an edge shape corresponding to a carefully chosen radial profile of attenuation, averaging over azimuth angles at each radius. Our NWO attenuation profile is the offset hypergaussian appodization function, given by

$$\begin{aligned} A(\rho) &= 0 && \rho < a \\ A(\rho) &= 1 - e^{-\left(\frac{\rho - a}{b}\right)^n} && \rho \geq a \end{aligned} \tag{1}$$

where $A(\rho)$ is the appodization (starlight transmission) as a function of radial distance, ρ, from the occulter center; a is the radius of the inner solid disk; b is the 1/e length of the hypergaussian function; and n is the order of the hypergaussian. An example of a NWO occulter is illustrated in Figure 1.

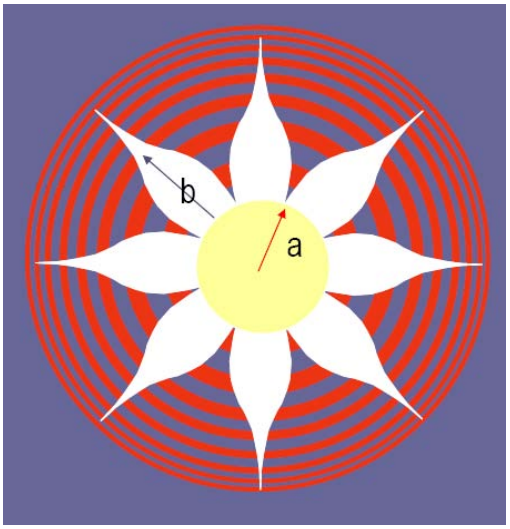


Figure 1: *The NWO Occulter consists of a central solid circle (yellow disk) and (white) petals following a hypergaussian contour. The red rings indicate the Fresnel zones (at which the field contributions in the Fresnel propagation integral reverse sign) due to diffraction of the starlight by the central disk;* a *is the radius of the inner disk; and* b *is the 1/e length of the petals' hyper Gaussian function.*

The occulter operates in the Fresnel optical regime, different from much of the previous works which were been developed for the Fraunhofer regime. The apodization is derived by imposing a complete cancellation of the electric field in the center of the occulter shadow some distance *F* behind the occulter. The binary partition function was also developed in the Fraunhofer regime[8] but was realized later to be also applicable for NWO in the Fresnel regime, given by the following formulae:

$$S = \left\{ 0 \le \rho \le a, \theta \in \Theta \right\}$$
$$\Theta = \bigcup_{q=0}^{P-1} \left[\frac{2\pi q}{P} + \frac{A(\rho)}{2}, \frac{2\pi(q+1)}{P} - \frac{A(\rho)}{2} \right] \qquad (2)$$

where *S* is the parameter set indicating the coordinates where the occulter is opaque (in polar coordinates (ρ, θ); *P* is the number of petals, and *A(ρ)* is the apodization function.

The on-axis cancellation achieved by the occulter, seen in Figure 1 and described by equation (2), is enabled by the sign reversals illustrated by the red Fresnel zones in Figure 1. In fact, the on-axis attenuation is lost in the Fraunhofer limit (infinite distance). Using the Fresnel integral to propagate the mask function of Equation (2) forward by a distance *F* to the telescope, Cash has shown[8] that the fraction of on axis starlight that is found behind a New Worlds occulter is:

$$R = \left[n! \left(\frac{\lambda F}{ab} \right)^n \right]^2 \qquad (3)$$

where F is the occulter to telescope separation. This simple and elegant result captures the basic performance and scaling of the occulter. If the occulter is designed for a given wavelength, it works better (with more suppression, *R*) at all shorter wavelengths. Equation (3) also shows that for a given sum of *a*+*b*, the best performance (smallest *R*) is obtained when *a* and *b* are equal. The hypergaussian index *n* plays a more subtle role, linking suppression performance above and below the design wavelength with the on-axis extinction of star light.

Finally, Equation (3) can be seen as relating the occulter Fresnel number (given by $(a+b)^2 / \lambda F$) to the extinction of starlight, and leads trivially to the understanding that the occulter operates in the optical near field, namely in the Fresnel regime.

If, for simplicity, we assume a particular occulter with designated *n*, *λ*, *a*, *b* and *F*, the scale of the tolerances can be derived from a perturbation analysis[8]. Table 1 summarizes the tolerances of the occulter. Specific values for the occulter tolerances will be the result of a rigorous derivation including all sources of performance degradation[9].

An external occulter designed to find terrestrial planets must:

- achieve starlight suppression adequate to control statistical and systematic noise in the planet pixel
- create a shadow of high suppression that is large enough to cover the telescope
- subtend a small enough angle to see close (< 100 mas) to the parent star
- meet these requirements over a wide optical bandwidth.

Table 1: *Tolerances for the NWO occulter.*		
Tolerance	**Category**	**Value**
Position	Lateral	Several Meters
	Axial	100's Km
Angle	Rotation	Any rotation
	Pitch/Yaw	Many Degrees
Shape	Truncation	1 mm
	Scale	10%
	Blob	3 cm^2
Hole	Single Hole	3 cm^2
	Pinholes	total area < 3 cm^2

In addition, the occulter must be practical to manufacture:

- the occulter should be binary for ease of manufacturing; graded transmission is difficult to achieve.
- the occulter must be small enough to be deployed by existing technology
- the occulter must be insensitive to microscopic errors to allow manufacturing of a large optical element.

The architecture of NWO is illustrated in Figure 2. A four-meter class telescope located on a separate spacecraft is positioned 10,000 to 100,000 km away from at least one occulter spacecraft carrying a starshade that is tens of meters in diameter.

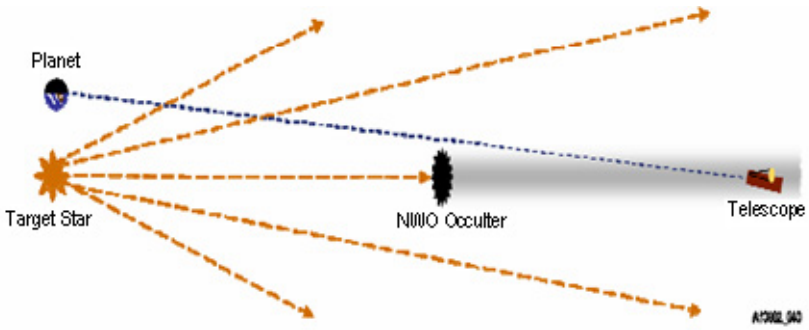


Figure 2: *New Worlds Observer Architecture*

The occulter's function is to create a shadow into which the telescope is flown, The petals of the occulter are designed to suppress the Spot of Arago (on-axis bright spot) that would occur if the occulter was a simple disk. These petals diffract light so that there is destructive interference for on an on-axis source and no effect for off-axis light.

When focused by the telescope, the residual starlight appears on-axis as an extended image whose size is on the order of the IWA convolved with the telescope PSF. This defocusing of on axis light produces an extra factor in suppressing the light from the star. Thus, it is more proper to discuss the suppression of starlight in the telescope's image plane, rather than at its entrance aperture. For telescopes and occulters sized for a TPF class mission, this additional factor is 10 to 100.

The effect of this angularly dependent focus is depicted in the rightmost image of Figure 3, where we show three simulated focal plane images of an Exo-planet detected with an NWO occulter + telescope system. The faint dot pattern may be thought of as the occulter + telescope system's "point spread function" for the central star. Incomplete destructive interference produces gradual roll off of the intensity of the PSF, and is the chief source of the "background" signal outside of the IWA. The off-axis planet light competes with this background signal, but because the residual starlight is spread out over many pixels, the intrinsic on-axis contrast ratio only needs to be at the 10^{-8} level.

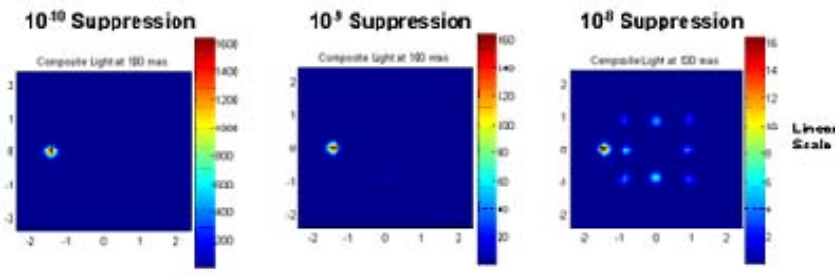


Figure 3: *An external occulter needs a smaller contrast ratio than an internal occulter because it spreads the star's residual light over many pixels. The leftmost image shows the planet signal with a 10^{-10} contrast ratio. The rightmost figure shows the same scene with a 10^{-8} contrast ratio. The "PSF" of the central star is visible in the rightmost image. Even with only a 10^{-8} contrast ratio, NWO can obtain a clear detection of the extra-solar planet.*

The occulter's diameter and its separation from the telescope, *F*, must be chosen such that the IWA is sufficiently small to meet the science requirements. This same occulter must also be large enough to provide sufficient starlight suppression. These conditions can generally be met, but such a solution implies a minimum occulter size[10].

During manufacture, deployment, and operation, the occulter is subject to various perturbations that cause it to lose some performance capabilities, whether in IWA or suppression. For example, manufacturing errors may cause imperfect interference and reduce the suppression ratio. We can recover these capabilities by building larger occulters. Figure 4 illustrates the minimum occulter size needed to achieve the indicated IWA and suppression ratios for an ideal or perfect occulter. The phase space of manufacturable occulters is continuous, such that a larger occulter situated farther away can always improve the performance in both IWA and suppression by the desired amount. This phase space is constrained first by manufacturability, where the size of occulter is limited by current deployment capabilities to around 60 m. It is also constrained by the occulter's propulsion system capability, where the total distance traveled by the occulter is limited by its fuel capacity, which restricts the maximum mass of the occulter and separation between the occulter and the telescope.

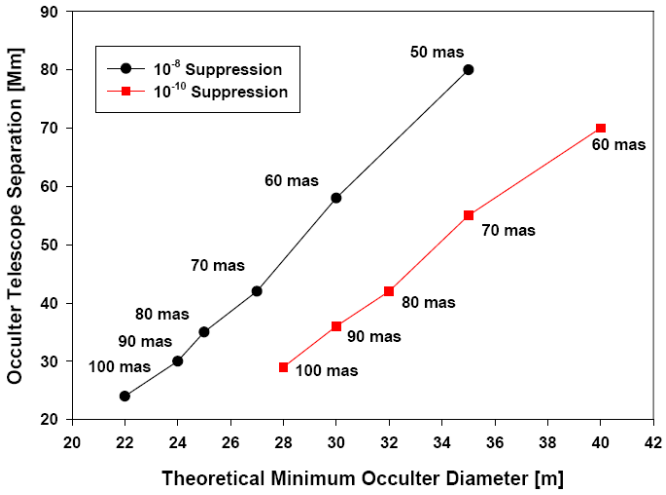


Figure 4: *The NWO design is flexible; there are many occulter sizes that can meet the science requirements. By making larger occulters, we can compensate for other performance degrading error sources.*

Multiple occulters increase the mission coverage and functionality. One of the main drawbacks of the NWO concept is the somewhat limited number of pointings due to the need to move the occulter over large distances. If, instead of one, there are two or more occulters, the observing efficiency and operational capability are greatly increased. We envision a two occulter NWO concept that is capable of detecting and characterizing terrestrial planets in the solar neighborhood. Figure 5 illustrates this dual occulter concept.

The first occulter is a small occulter, with an effective diameter of 25 m. This occulter is separated by ~30,000 km from the telescope, and is the survey occulter, capable of looking at one star every 6–7 days. This occulter obtains the first extra-solar system detections, and over the course of a 5 year mission, occults at least 130 stars two times, for ~300 pointings.

The second occulter is a large occulter, with an effective diameter of 50 m. This occulter observes a system once every 17 days and is capable of remaining aligned with the telescope and the star for more than 100 hours. This allows spectroscopy of Exo-planets to be performed for the detection of absorption or emission features of interest. In 100 hours, we estimate the occulter and telescope will be able to observe the molecular oxygen line with a signal to noise ratio of well over 10. This larger occulter would enable longer integrations, and let us see in closer to the star, due to its smaller IWA. The longer integration time also allows us to more effectively remove exo-zodi contributions. The velocity change (ΔV) requirements (which determine fuel requirements) of the two occulters are given in Table 1. Due to the large ΔV requirements, a solar electric propulsion system could be an expedient method to enable the necessary number of pointings.

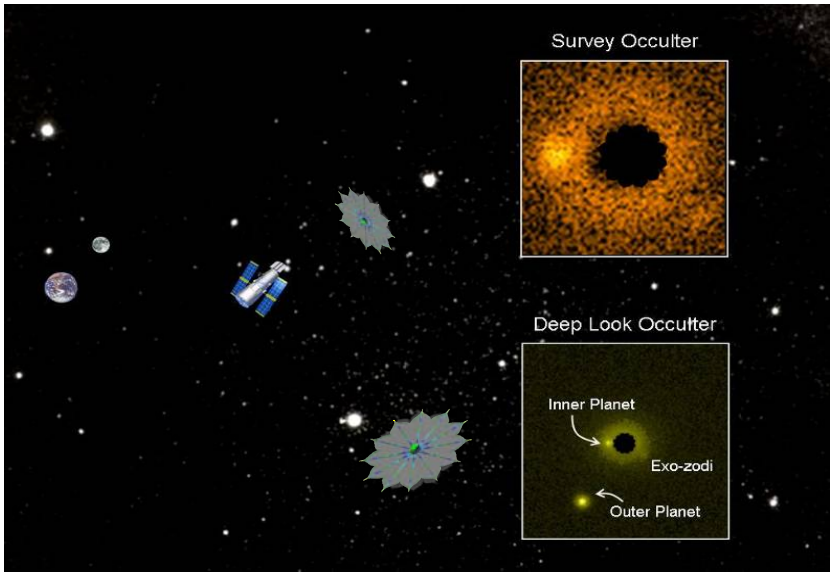


Figure 5: *In the full NWO terrestrial planet finding mission concept, two occulters work in tandem to survey nearby stars and detect and characterize Exo-solar planets.*

We have introduced the New Worlds Observer concept. The key enabling technology, the occulter with its hypergaussian shaped petals, has been shown to be capable of reaching the IWAs and broad band starlight suppression levels that are required for a mission to find extra-solar terrestrial planets. Such an occulter can be built and can be aligned, maintained and moved to another target with current technologies. A mission using an external occulter requires only an ordinary, modest-sized (diameter 3–4 m) telescope to detect and characterize extra-solar terrestrial planets. As currently envisioned, NWO will spend half of its time in the search for Exo-Earths and the other half devoted to general astrophysical observations. The fact that the NWO telescope is a generic astronomical instrument will make it easier to find partners for a highly capable dual-purpose mission.

Table 2: *ΔV Requirements for the Dual Occulter NWO Concept*

	Small Occulter	**Big Occulter**
Per Pointing ΔV	41 m/s	77 m/s
Total Pointings (1 yr)	61	21
Total Pointings (5 yr)	304	107
Total Required ΔV	12.4 km/s	8.2 km/s

The New Worlds concept is developing rapidly and much effort and time is being spent to advance the concept. Among the current and future steps planned are additional laboratory measurements with higher fidelity occulters, continued study of the orbital dynamics of the mission concept, building a subscale occulter model and continued study and simulation of the optical properties of the occulter.

The New Worlds team looks forward to sharing its progress.

REFERENCES:

1. Cash, W, et al, Proc. SPIE **5899**, pp. 274–284 (2005)
2. Arenberg, J, and Cash W, IAU Colloquium **200** (2006)
3. Spitzer, L., American Scientist, **50**, pp. 473–484 (1962)
4. Marchal, C., Acta Astronautica, **12**, pp. 195–201 (1985)
5. Cocks et al, Proc. SPIE **2807**, pp. 74–85 (1996)
6. Copi, C, J, and Starkman, G, Astrophysical Journal, **532**, pp. 581–592 (2000)
7. Jordan, I. et al, Proc. SPIE **5487**, pp. 1391–1401 (2004)
8. Cash, W, Nature, **442**, pp.51–53, (2006)
9. Arenberg, J et al, Proc. SPIE **6265** (2006)

"Discover the origin, structure, evolution, and destiny of the universe, and search for Earth-like planets"

… from the NASA Science Plan for 2007, Astrophysics Sub-goal

For more information on NASA's Navigator Program, visit:

http://planetquest.jpl.nasa.gov

National Aeronautics and Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

www.nasa.gov

JPL Publication 07-02 7/07