A Nulling Coronagraph for TPF-C

Michael Shao[1], Bruce Martin Levine[1], James Kent Wallace[1], Glenn S. Orton[1], Edouard Schmidtlin[1]; Benjamin F. Lane[2]; Sara Seager[3]; Volker Tolls[4] ; Richard G. Lyon[5]; Rocco Samuele[6]; Domenick J. Tenerelli[7], Robert Woodruff[7]; and Jian Ge[8]

[1]Jet Propulsion Laboratory/California Institute of Technology, [2]Massachusetts Institute of Technology, [3]Carnegie Institute of Washington, [4]Harvard Smithsonian Center for Astrophysics, [5]Goddard Space Flight Center, [6]Northrup-Grumman Space Technology Corporation, [7]Lockheed Martin Space Systems Company, [8]University of Florida.

ABSTRACT

The nulling coronagraph is one of 5 instrument concepts selected by NASA for study for potential use in the TPF-C mission. This concept for extreme starlight suppression has two major components, a nulling interferometer to suppress the starlight to ~10^{-10} per airy spot within 2 λ/D of the star, and a calibration interferometer to measure the residual scattered starlight. The ability to work at 2 λ/D dramatically improves the science throughput of a space based coronagraph like TPF-C. The calibration interferometer is an equally important part of the starlight suppression system. It measures the measures the wavefront of the scattered starlight with very high SNR, to 0.05nm in less than 5 minutes on a 5mag star. In addition, the post coronagraph wavefront sensor will be used to measure the residual scattered light after the coronagraph and subtract it in post processing to 1~2x10^{-11} to enable detection of an Earthlike planet with a SNR of 5~10.

Keywords: Exoplanets, terrestrial planets, nulling interferometry, coronography

1. INTRODUCTION

This paper outlines a concept study for a high contrast instrument for the Terrestrial Planet Finder Coronagraph (TPF-C) mission. 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.

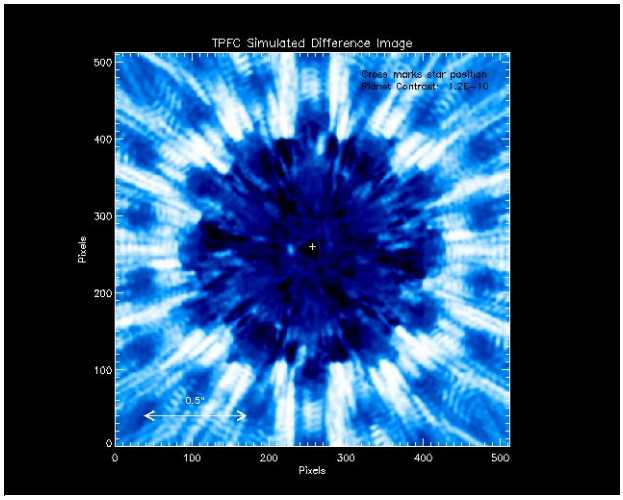

Figure 1: Simulated image from TPF-C of an Earth-like planet orbiting around a G2V-type star at 10pc.

The nulling coronagraph can, in theory, achieve the required 1x10^{-10} suppression of starlight within an IWD of 2 λ/D. Coverage of the 0.5 ~1.7μm wavelength range is done in intervals of 25% bandwidth. The search for planets will be conducted at short wavelengths, where the IWA is smaller. Figure 1 shows a simulation of an image from a nulling coronagraph during the search phase. The smaller IWA of a nuller lets us extend spectroscopy to 1.7μm for a larger number of targets.

This paper is divided into sections on Science Drivers, Concept Description, and Concept Analysis. The Science Driver section develops the case for a TPF-C mission based on a greater number of targets than is currently in the baseline TPF-C Science requirements. The following section then describes the basic principles of the nulling coronagraph instrument

and why it is capable of making detections at smaller IWA's. We conclude with a section on the Calibration wavefront sensor, or the estimation of post-SSS wavefront errors by amplitude and phase measurement of the stellar leakage.

2. SCIENCE DRIVERS

2.1 How many stars can be surveyed by the nuller?

The nulling coronagraph, aka nuller, has a couple of advantages over the baseline 8th order mask Lyot coronagraph. One is higher photon throughput and the second, more important, is the ability to work at 2 λ/D rather than 4λ/D. This section quantifies the science advantage of these two properties of the nuller. A quantitative description of throughput and IWA inner working angle is given in the instrument description section.

2.2 Higher throughput

We used the matlab script supplied by the TPF-C project to calculate the SNR of the nulling coronagraph. The only modification to the code was the equivalent coronagraphic transmission and lyot stop efficiency. All other parameters such as detector QE, collecting aperture etc., were unchanged. The graph below compares the nuller to an 8th order mask lyot coronagraph. The star list for the comparison was the default TPF-C starlist of stars where an Earth would be detectable in the habitable zone at or greater than 62mas separation (4 λ/D).

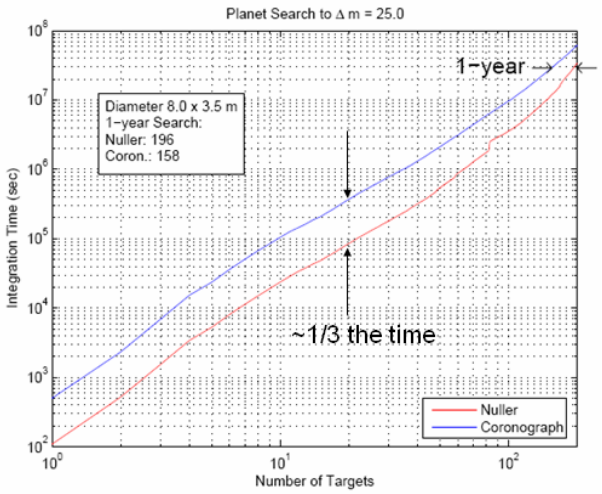

Figure 2-1

If TPF-C were to repeatedly visit a moderate number of stars, 30~50, the nuller would be able to visit those objects between 2~3 times as often for a fixed integration time, increasing the geometric completeness (and the number planets found) by roughly a factor of ~2.5.

The search for Earthlike planets will be conducted at the shortest wavelength band 0.5μm~0.6μm because that gives the smallest inner working angle (IWA).

The comparison of throughput was made using a starlist that is observable by the nuller working at 2 λ/D and the lyot coronagraph at 4 λ/D. In the next sections we describe the science advantage of being able to observe planets closer in at 2 λ/D.

2.3 Number of potential targets at 2 λ/D

In assessing the number of potential targets for the nulling coronagraph, we searched through the list of Hipparcos stars <30pc from the Sun brighter than 7 mag and made the following assumptions:

- Every star has an Earth in the middle of the habitable zone
- The orbit is circular
- We're looking for the planet at max elongation, 90deg phase angle
- The albedo of the exo-earth is the same as our Earth 30%
- The angular separation is orbit radius divided by the distance
- We assume (consistent with TPF-C project assumptions) the Sun-Earth has a contrast of 1.2×10^{-10} and the contrast varies as $1/r^2$ where r is the orbit radius in AU.
- A detectable planet has to have a contrast better than 1.0e-10. (this favors stars less luminous than the Sun)

We can calculate the number potential targets at 2 λ/D and 4 λ/D for several scenarios. The scenarios are 1) discovery at 0.5~0.6μm, 2) detection of Oxygen at 0.78μm, and 3) detection of near IR molecular signatures at 1.6μm. In this analysis we calculate "potential" targets, it is possible that the nominal TPF-C mission lifetime would not permit us to go through the whole target list.

Table 1 below lists the number of potential target stars around which an Earth can be detected at various wavelengths for a nuller (2 λ/D) and the baseline coronagraph (4 λ/D). A graphical representation of the number of detectable earths is shown in Figure 2-2 and the Oxygen and CO_2 feature detection is shown in Figure 2-3.below.

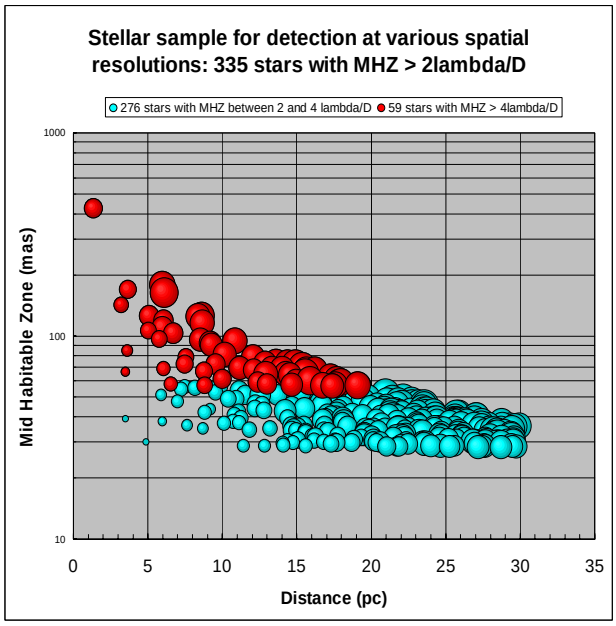

Figure 2-2: Number of potential earth-like planets as a function of field angle.

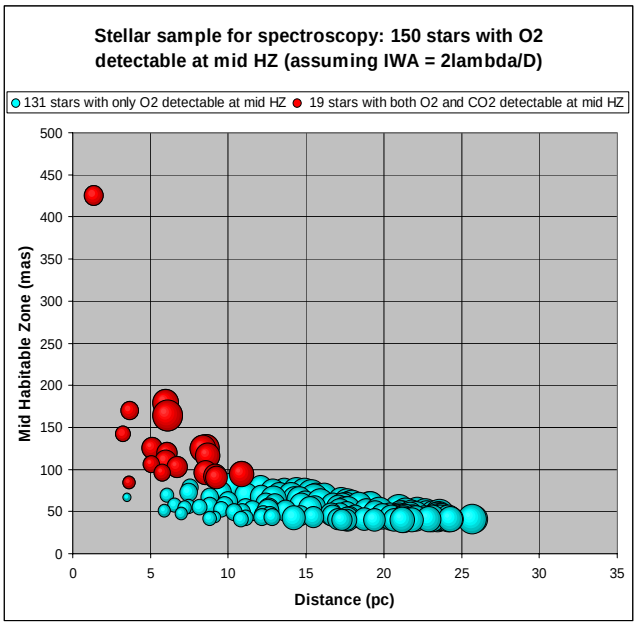

Figure 2-3: Number of stars with detectable Oxygen and CO_2 atmospheres.

2.4 Spectral features of exo-Earth at 0.8 and 1.0~1.6μm

The number of potential Earths where oxygen at 0.76μm can be detected is much larger at 131 stars for the nuller versus 19 stars for a coronagraph working at 4 λ/D. While the detection of atmospheric oxygen is perhaps the most important, extending the wavelength coverage beyond 800nm enables us to identify many other potential biomarkers.

# potential targets versus wavelength			
wavelength	Detect Earth (0.5~0.6um)	Det Oxygen 0.78um	NIR 1.6um
2 λ/D	276	131	20
4 λ/D	59	19	3

Table 1

The 0.9-1.7μm micron range is essential to characterize a range of different terrestrial planets (see Figure 2-4), and the TPF-C 0.5-0.8μm region was designed with only modern Earth in mind. The spectral region beyond 1μm is essential for detecting the terrestrial gases CO_2 and CH_4. CO_2 makes up more than 95% of Venus and Mars atmospheres, and is considered an identifier of terrestrial planets (since giant planets do not have CO_2). Yet there are no CO_2 features shortward of 1μm that are accessible to TPF-C, even in the case of Venus. The numerous strong CO_2 features between 1-1.5μm will allow CO_2 identification. CH_4 is an important biomarker gas because it was expected to be present on early Earth from methanogenic bacteria (e.g., 3Gyr ago CH_4 may have been 1000x more abundant in Earth's atmosphere than it is today).

Although CH_4 can also be produced abiotically, its presence on a terrestrial planet is still significant because it has an extremely short lifetime due to photolysis. There is only one very weak feature in the 0.5-0.8μm range, an additional CH_4 feature at 0.9 μm and a much stronger one at 1μm. Newly discovered terrestrial planets have the potential to be extremely different from Earth, Venus and Mars---planets that appear very similar over a short wavelength span in the visible may show more substantial differences over an extended wavelength region 0.5-1.6μm.

CH_4 and CO_2 are detectable in the spectral region when their abundances are large compared to the present Earth. Here "large" means comparable to conditions on the early Earth, or during periods of "hothouse Earth" or during periods of later methane bursts. Thus these species have a good chance of being detected at these wavelengths on an early Earth-like planet. Our instrument will measure the 1.6μm feature (NIR feature in Table 1) of CH_4, because this is the shortest-wavelength feature which we can identify at present as being strong enough to permit detection.

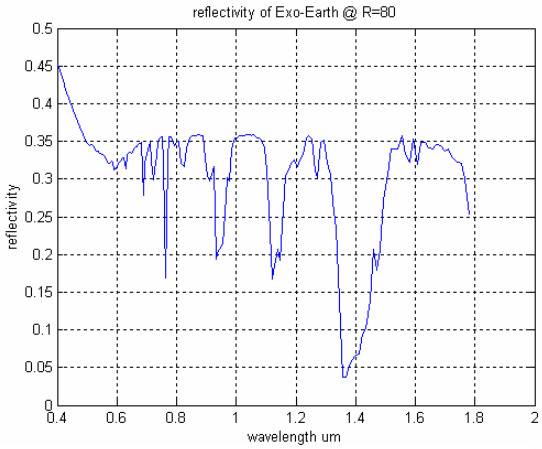

Figure 2-4: Theoretical spectrum of an exo-Earth at R=80.

2.5 Spectral features of exo-Jupiters

The spectral characterization capability of the instrument is, of course, not limited to terrestrial planets. Some of the first objects to be separated from their primaries will be giant planets. Determination of the properties of these bodies is strongly enabled by near-infrared capabilities, as can be illustrated by a substantial body of studies of gas giants of various sizes and ages which orbit solar-like stars of spectral class: G2V (Sudarsky, et al. 2000, and references therein). These models have been used to discern detectable spectral differences that might be used to discriminate fundamental properties, such as composition and age by Burrows, et al. (2006). In addition, the appropriate models of a Jovian-mass body at Jupiter's distance from the sun and a similar effective temperature and age have been vetted against the measured reflection spectrum of Jupiter (Karkoschka, 1994).

For the sake of illustration, we focus on the differences in the spectra between two bodies of Jovian size and age. These are shown in Figure 2-5 for a body 1 AU from a solar-like star and for a body which is 4 AU from a solar-like star. This figure emphasizes the differences found in the models of Burrows, et al. (2006) and Sudarsky, et al. (2000) for gas giants (i) closer than 1.6 AU to their primaries and (ii) further than 1.6 AU from their primaries.

The Jupiter proxy sitting closer to its primary star is brighter by about a full order of magnitude, not only because of its closer proximity, but because of the presence of H_2O condensation forming the predominant upper-level bright cloud layer. The spectral location of water vapor absorption bands are indicated in the figure in red, as appropriate for the 1 AU body; the most distinguishable of these in the spectrum is at 0.94μm. At 4 AU, most of the spectral radiation is light reflected from the primary by the clouds, but at 1 AU (and for bodies even closer to their primaries) most of the radiation longward of 0.8μm consists of thermal emission. For planets at 1 AU and greater, the presence of H_2O clouds slightly attenuates the variation with wavelength in the planetary spectra. For bodies closer than 1 AU from the primary (not shown in Figure 2-5), alkali-metal absorption is easily detected in the visible spectrum. In fact, the sodium-D line was detected by Charbonneau, et al. (2002) in the transit spectrum of EGP HD209458b. On the other hand, distinctions between characteristics of bodies between 5 and 1 AU, "close-in" Jupiters and more familiar Jupiters, are primarily given by the spectrum longward of 0.8μm, as evident in Figure 2-5.

In particular, the spectral variations between methane and water vapor absorption bands become prominent or subdued, depending on orbital distance. For Jupiter-sized bodies further out that 4 AU, the spectra will be very similar to the 4-AU body, so long as the temperature profile is dominated by equilibrium with heat from the interior which is a remnant of the heat of formation. This is the case with Jupiter, Saturn, and Neptune in our own solar system, even though the sizes and compositions of these planets make their effective

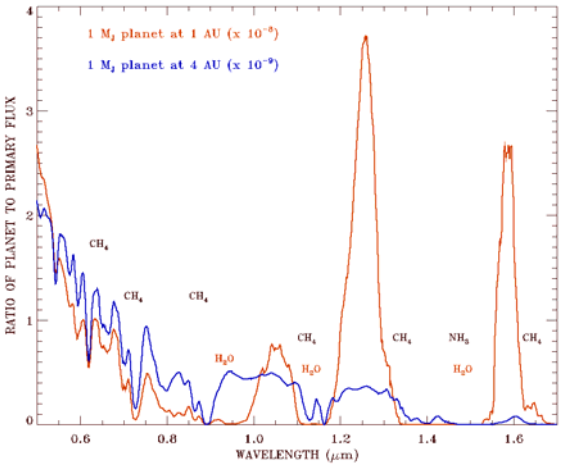

Figure 2-5

temperatures much lower than Jupiter's. For bodies which are colder, either from lower mass or age, the effective level of upper-cloud formation for water vapor and ammonia will be deeper in the atmosphere. Like Uranus and Neptune, a methane condensate cloud may form, as well. For Uranus and Neptune, this is not an optically thick cloud; in general, the spectra of these bodies will be dominated by saturated methane absorption. Uranus and Neptune both illustrate that this is so much the case that the visible color of the planet shifts toward the blue-green because of red absorption by CH4 bands in the visible. The near-infrared reflectivity is nearly zero everywhere.

For giant planets in elliptical orbits, Sudarsky, et al. (2005) note that the greatest time-dependent spectral variation will be in the "switching on" or "switching off" of cloud condensation and its attendant opacity increase or decrease, respectively. According to Sudarsky, et al. a Jupiter-size body - with a condensation/evaporation time scale much shorter than the orbital period - and an eccentricity of 0.3 and semi-major axis of 1.5 AU will get sufficiently cold at apoastron for an ammonia cloud to condense and form a cloud above the water vapor cloud. A planet closer in might switch from a cloud-free atmosphere at periastron to one with water vapor clouds at apoastron. However, their models assume a time scale for a temperature change (which governs condensation and evaporation) that is much shorter than the orbital period, and this is not necessarily true. In fact the radiative time scale for change in Jupiter's atmosphere at depths characterizing thermal emission or cloud reflectivity (0.1 - 2 bars) has been characterized as being on the order of half of its orbital period (Conrath, et al. 1990). For this reason, spectral differences differentiating periastron and apoastron may be significantly smoothed out by the long atmospheric response.

3. NULLING CONCEPT AND ARCHITECURE

3.1 Nulling Interferometer

3.1.1 Using a Nulling Interferometer Instrument with a Single Aperture Telescope

In the infrared domain, the nulling interferometer approach is more conventional, however, we apply the principles of nulling interferometry to the visible light domain with a conventional single aperture telescope as an alternate starlight suppression method. A nulling interferometer based instrument is used as a high-contrast imaging system to suppress both diffraction and scattering (Shao, 1991).

In principle, a nulling interferometer effectively projects an interference pattern on the sky over the star-planet system to be imaged (Figure 3-1). 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=500μm wavelength.

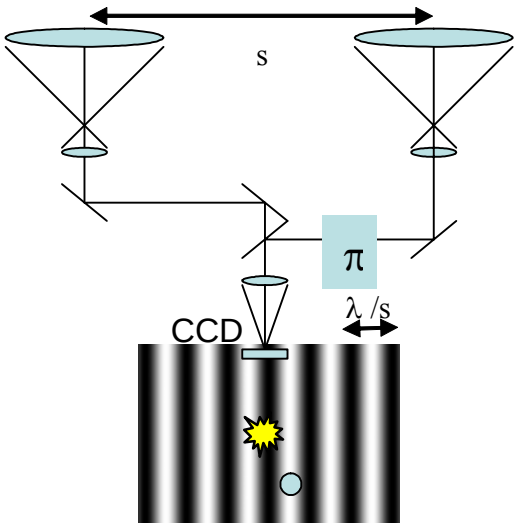

Figure 3-1: Imaging with nulling interferometry. A fringe pattern is superimposed over the star and planet to be imaged, the star is placed at the bottom of a deep, achromatic destructive (null) fringe, while planet on the constructive fringe is transmitted.

The implementation for TPF-I 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.

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 3-2 below). Thus, for this experiment, the final layout for the beam combiner consists of identical two-glass pairs of (rotatable) dielectric phase retarders (Morgan, et. al., 2000), 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. In theory, the nulling coronagraph suppresses starlight to 10^{-10} within 1.5~2 λ/D while a more traditional Lyot coronagraph or apodized aperture telescope which has an inner working distance of 3~4 λ/D (Levine et. al., 2003; Wallace, Shao, Levine and Lane, 2003).

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

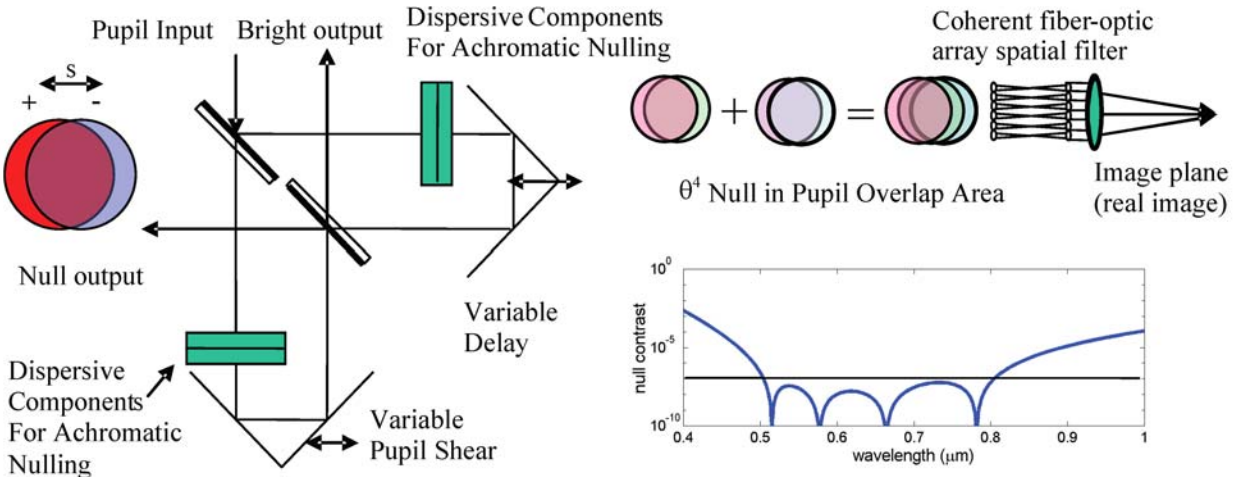

Figure 3-2: Visible nulling coronagraph instrument concept. Left) A concept for single-input symmetric nulling interferometer. Note that the two pupils emerge from the null output displaced (or sheared) by a distance, s, proportional to the baseline of the instrument. Lower right) Calculation of 'pseudo-achromatic' 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 θ^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.

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 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.

3.2 Spatial Filter Arrays

Spatial filter arrays (SFA) are a means to mitigate both amplitude and wavefront aberrations in an input beam. A SFA consists of a lenslet array followed by a bundle of single-mode fibers, followed again by a lenslet array matched to the input lenslet array as shown in Figure 3-3. Each fiber is mated to one pair of lenslets, one on the input and the other on the output side. SFAs are also realizable by replacing the fiber bundle by an array of pinholes. An input collimated beam, with amplitude and wavefront aberrations, is input from the left and is decimated into an array of sub-aperture beamlets via the lenslet array (Figure 3-4 left). Each of the beamlets is brought to focus onto the input end of a single-mode fiber (Figure 3-4 right). Light which is matched to a fiber fundamental eigenmode will be coupled into the fiber; unmatched light will be suppressed and/or coupled into the cladding modes and either is absorbed, reflected, or otherwise coupled out of the cladding. Some light will also reflect off the front face due to Fresnel losses. Light coupled in will ultimately evolve into and propagate only in the fundamental spatial mode of the fiber. At each fiber's exit an expanding single mode beam is incident on each lenslet, and the output lenslet array stitches the beamlets into a coherent wavefront. The average wavefront error (piston) over each input sub-aperture will vary, due to the input aberrations resulting in stepwise phase differences (rightside of Figure 3-3). In addition sub-aperture amplitude and wavefront errors higher order than piston will result in intensity losses and thus the output sub-aperture beamlets will also vary in intensity.

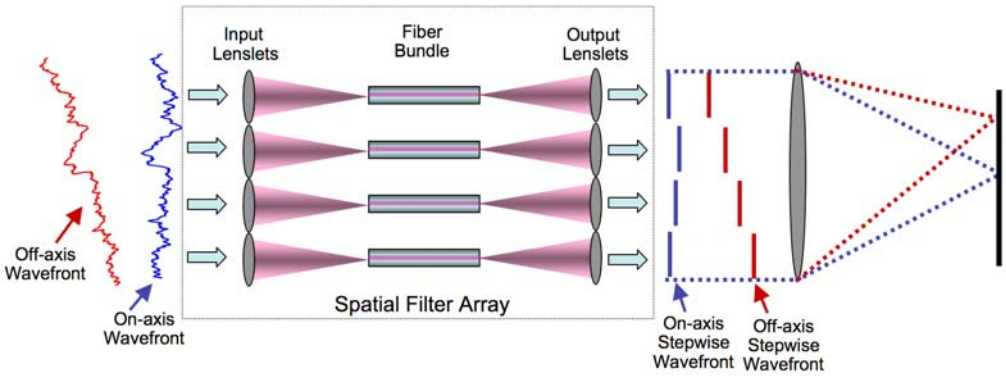

Figure 3-3: Principle of Spatial Filter Array. Input wavefront and amplitude aberrations are converted into intensity variations and stepwise phase on output.

The net effect is that amplitude and wavefront spatial frequencies will be filtered. Local tip/tilt variations, over each input sub-aperture, will result in mis-steering the beamlets into the fiber and will thus reduce intensity on output. Light from an on-axis source will have a stepwise phase (shown in blue in Figure 3-3) resulting in an on-axis image with sidelobes. The first image sidelobe will occur at $N\frac{\lambda}{D}$ (in angular sky coordinates), where N is the number of sub-apertures across the diameter of the beam and D is the diameter of the telescope aperture and in effect this defines the full useable field-of-view at $2N\frac{\lambda}{D}$. Reflecting off a deformable mirror (DM) prior to entering the input side of the SFA, with at least 3 actuators per sub-aperture with a continuous facesheet DM, or with discrete DM, and then sensing piston, tip and tilt from the output side and feeding back to the DM results in a high quality output beam.

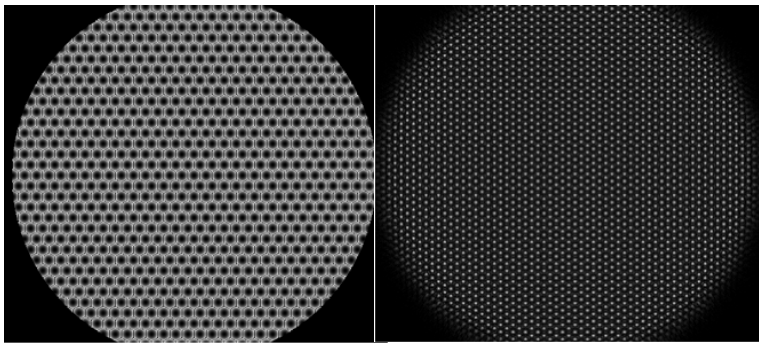

Figure 3-4: *Hexagonal Pack Lenslet Array and Intensity at Fiber Entrance.s* Left, Optical phase introduced by an array of hex pack lenslets. Right, Intensity pattern at fiber entrance, one Airy spot per fiber.

For a fiber of core index n_{core} and cladding index of n_{clad} the fiber's numerical collection aperture (NA) can be derived from the conditions for total internal reflection to be $NA = \sqrt{n_{core}^2 - n_{clad}^2}$ where $NA = \frac{1}{n_o SinU}$ and U is ½ angle of acceptance into the fiber. Matching the F/# of an individual lenslet to the fiber NA ($NA = 1/2F/\#$) gives an

approximation to the ideal lenslet F/# as $F/\# = \frac{1}{2\sqrt{n_{core}^2 - n_{clad}^2}}$. The actual relationship, derived in (Ruiler and Cassaing, 2001), is shown to be $F/\# = 0.70425\frac{d_{core}}{\lambda}$. For a 9 micron core diameter with core index of $n_{core} = 1.471$ and cladding index of $n_{clad} = 1.4702156$ the optimal lenslet F/# at $\lambda = 0.6$ microns, for a circular aperture lenslet, is $F/\# = 10.56$.

Coupling into a fiber varies with wavelength. Figure 3-5 shows the effect wavelength on the coupling into and out of a fiber for both circular lenslets and hexagonal lenslets. The lenslet F/# has been optimized for optimal coupling for a circular lenslet at $\lambda = 0.6\mu m$. Figure 3-5 is calculated by evaluating the normalized inner product of the fibers eigenmode with an input unaberrated beam and thus represents and upper bound. The inner product is given by:

$$\Omega = \frac{\langle E(r) | TEM_{00}(r) \rangle}{\sqrt{\langle E(r) | E(r) \rangle \langle TEM_{00}(r) | TEM_{00}(r) \rangle}} \quad (1)$$

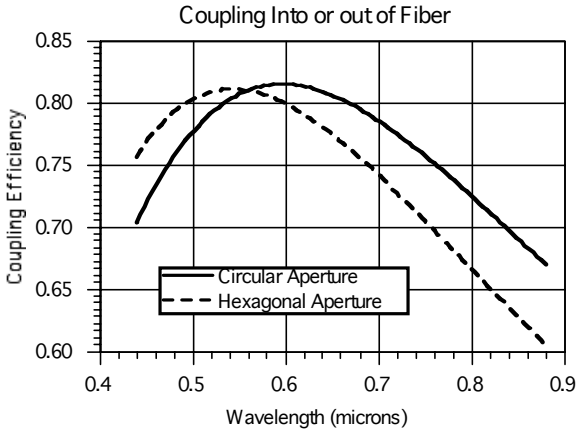

Figure 3-5: *Coupling losses in Fiber Input/Output*

where the intensity coupling is $\eta = |\Omega|^2$ and the bracket notation refers to a 2D integral from 0 to 2π and 0 to ∞ in the focal plane of the lenslet and $TEM_{00}(r)$ is the fundamental eigenmode, assumed a Gaussian beam herein, but is actually a modified Bessel function which closely approximates a Gaussian near its center but with arger tails (Shaklan and Roddier, 1988). In practice this inner product is numerically difficult due to sampling considerations and the range of the integral. The solution to this is to evaluate the inner product in the plane of the lenslet array by back propagating the eigenmode to this plane; thus the integral now only extends over the sub-aperture. The theoretical peak coupling efficiency is 81.4% on input, but since the beamlets are also apertured on output this coupling efficiency is applied twice and the resultant peak throughput, lenslet to fiber to lenslet, is 66.3%.

Coupling into the fiber also decreases monotonically with a tilted wavefront entering the fiber in Figure 3-6. Shown is the intensity coupling efficiency into the fiber versus local tilt angle at the fiber. Thus e.g. if we assume an 8 meter aperture telescope operating at a wavelength of 500 nm and that the beam is compressed to 2 cm beam at the fiber bundle and a planetary source at 4 λ/D then the local tilt at the fiber bundle is 20.63 arcseconds and thus the planetary light throughput loss is only ~1%.

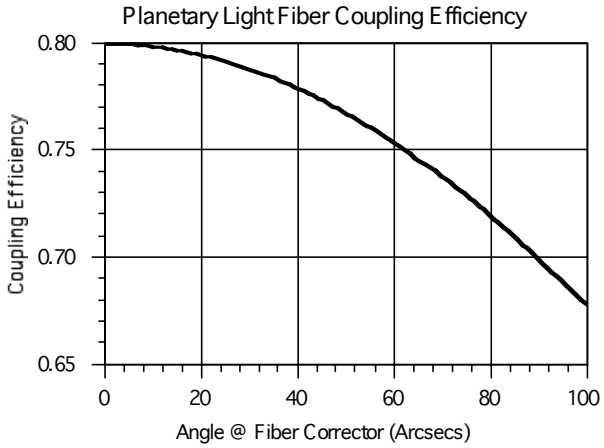

Figure 3-6: *Coupling losses vs. Angle @ Fiber*

Figure 3-7 shows plots through the fiber at varies distances during propagation. Within the first 20 – 30 microns of propagation the Airy disk pattern is still clearly visible, but it becomes successively more suppressed with increasing distance and ultimately after ~100 microns the Airy rings are rapidly damped and a stable mode propagates. There is still some power in the cladding giving long tails on the Gaussian. It was found in this simulation, that after ~100 microns of propagation the beam had reached to an eigenmode to better than 10^{-13} and that the distance is a strong function of the extinction in the cladding.

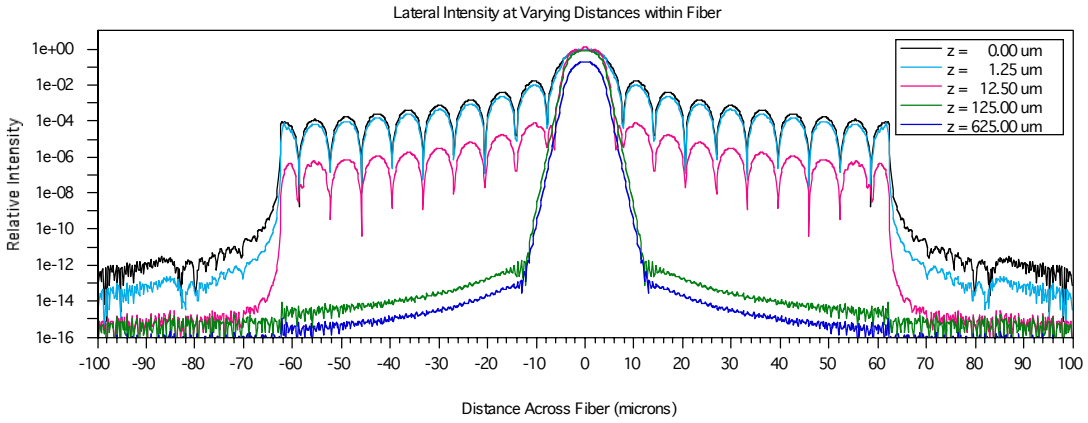

Figure 3-7: *Attenuation of higher order modes with propagation down fiber*

Analytical approaches also exist for describing the distance to a stable fiber mode and are based upon defining and calculating the losses into the cladding modes (Wallner, et. al., 2002). Wallner, et.al. shows that the minimum normalized filter length is given by:

$$\frac{z_0}{\rho} = \frac{1}{\alpha\rho}\ln\left\{\frac{A(1-\eta)}{\eta}\right\} \tag{2}$$

where η is the intensity coupling into the fiber, ρ is the core radius, and A is the ratio of the power in the fundamental mode to the power in the leaky modes with the worst case assumption that all the power not coupled to the exciting mode is coupled to the leaky mode with the lowest attenuation. α is determined from the imaginary part of the propagation constant, $\alpha = 2\beta_{leaky} = \frac{1}{\rho}\sqrt{\frac{V^2}{2\Delta} - U^2}$ where $V = k\rho n_{core}\sqrt{2\Delta}$ and Δ is the core-to- cladding index difference and U is the core parameter. For most fibers of interest $\frac{V^2}{2\Delta} >> U^2$ and thus $\alpha \approx \frac{V}{\rho\sqrt{2\Delta}}$. This yields propagation distances of 10 – 100 microns for most conventional single-mode fibers when operating in the visible.

3.3 Segmented vs. Continuous Facesheet Deformable Mirrors

Conventional deformable mirrors (DM) have developed along the lines of continuous facesheets where the mirrored surface is a continuous membrane with a periodic array of force actuators located behind the mirrored surface. An alternative approach are the recently developed MEMS based segmented DMs. These consist of MEMS based mirrorlets each of which is mounted on typically 3 actuators and each mirrorlet moves, as a rigid body, independently of each other, and herein are known as MEMS mirror arrays (MMA).

In a continuous facesheet DM the wavefront is typically decomposed into a linear superposition of the influence functions via least squares fitting. An influence function represents the response (deformation) of the surface to a unit of force. In MMAs each region of the DM is independent and thus fitting to local piston, tip and tilt over each mirrorlet is required. There are gaps between segments and reflection off the gap edges causes diffraction of the beam. If relay optics are not used then significant mixing of the beamlets over each mirrorlet can occur due to free-space Fresnel diffraction effects from the gaps. A Boston Micromachines DM with 1027 hexagonal mirrorlets (37 concentric rings) of 517μm side-to-side with 2 μm gaps would diffract according to 3 characteristic Fresnel numbers: for the full mirror, for

the segment widths and for the gap widths. Each of which has a differing distance to its far zone. Figure 3-7 shows this effect via numerical evaluation of a Fresnel diffraction integral for a collimated beam reflecting off the mirror for distances of z = 0.1 mm, 1 mm, 1 cm, 0.1 meter and 1 meter. The MMA diffraction can be characterized by the Fresnel numbers, $F_N = d^2/\lambda z$, where d is the characteristic scale size of the diffracting structure, λ is the wavelength and z is the distance of propagation. For the 3 characteristic scale sizes, namely the gap width (2 um), the segment width (517 um) and the width of the entire DM (~1.6 cm), the Fresnel numbers are given by:

$$F_N = \begin{cases} \text{Gaps}: & 8/micron \\ \text{Segmt}: & 5.35\times10^5/micron \\ \text{DM}: & 5.12\times10^8/micron \end{cases} \qquad (3)$$

The far-zone is when the Fresnel is < 1 thus the transition from Fresnel to far zone occurs for propagation distances of 8 microns, 535 mm and 512 meters for the gaps, segment width and full DM respectively. This can clearly been seen in Figure 3-8. After a distance of 0.1 mm the gaps have diffracted to near sinc functions, and for distances of 1 cm to 10 cm the segments are in the Fresnel zone but are subsequently lost within the Fresnel zone of the full MMA at 1 meter. After 1 meter the diffraction is dominated by Fresnel rippling of the full aperture of the MMA.

To mitigate these effects an oversized 2F imaging optic needs to be placed between the MMA to relay its coherent beam onto the lenslet array of the SFA. The propagation distance would then be z=4F and the diameter of the relay optic is driven by the angular diffraction of the gap by λ/d. To collect most of the power the ½ angular diameter of the relay optics needs to be $\frac{D/2}{2F} > 3\lambda/d$ over the propagation distance to the relay optic of 2F. Thus $D > 12\frac{\lambda F}{d}$ and since z=4F => F = z/4 then $D > 3\frac{\lambda z}{d}$ where d is the gap width, z is the distance from the DM to the SFA and l is the wavelength. For a distance of 0.5 meters and 2 um gaps in the visible the relay optics diameter should be a relatively large $D > 0.375$ meters with the aperture size scaling linearly with distance z. Using a relay closer to the MMA and/or multiple relays allows the aperture size to shrink, or alternatively collecting less of the diffracted light and allowing slightly more increased rippling would allow smaller apertures.

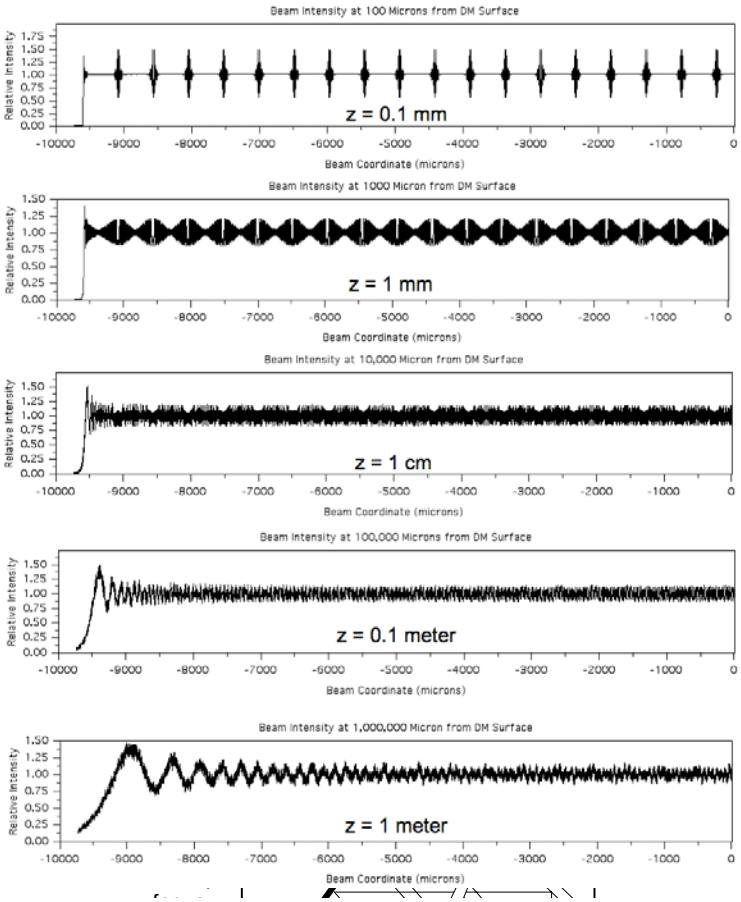

Figure 3-8: Beam Propagation of Segmented DM

3.4 CALIBRATION SYSTEM

3.4.1 Impact of the post-SSS (calibration) wavefront sensor

Although starlight is suppressed to 10^{-10}, the resulting speckle pattern needs to be subtracted to a much lower level, 2 x 10^{-11} for a 5 sigma detection of a planet. The post-starlight suppression system (SSS)

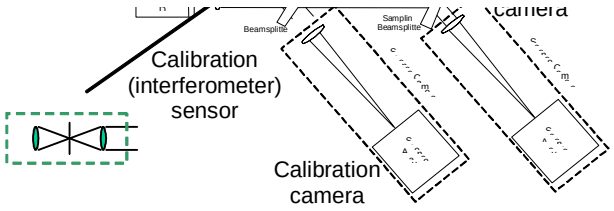

Figure 3-9: Concept for a calibration wavefront sensor with a nullling coronagraph.

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.

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.

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. (Variations of this approach have been described by Guyon, 2004.) A Mach-Zehnder Interferometer forms the heart of this concept (Figure 3-9).

It comprises two beams: a reference beam and an unknown. The reference beam is created from the light that passes

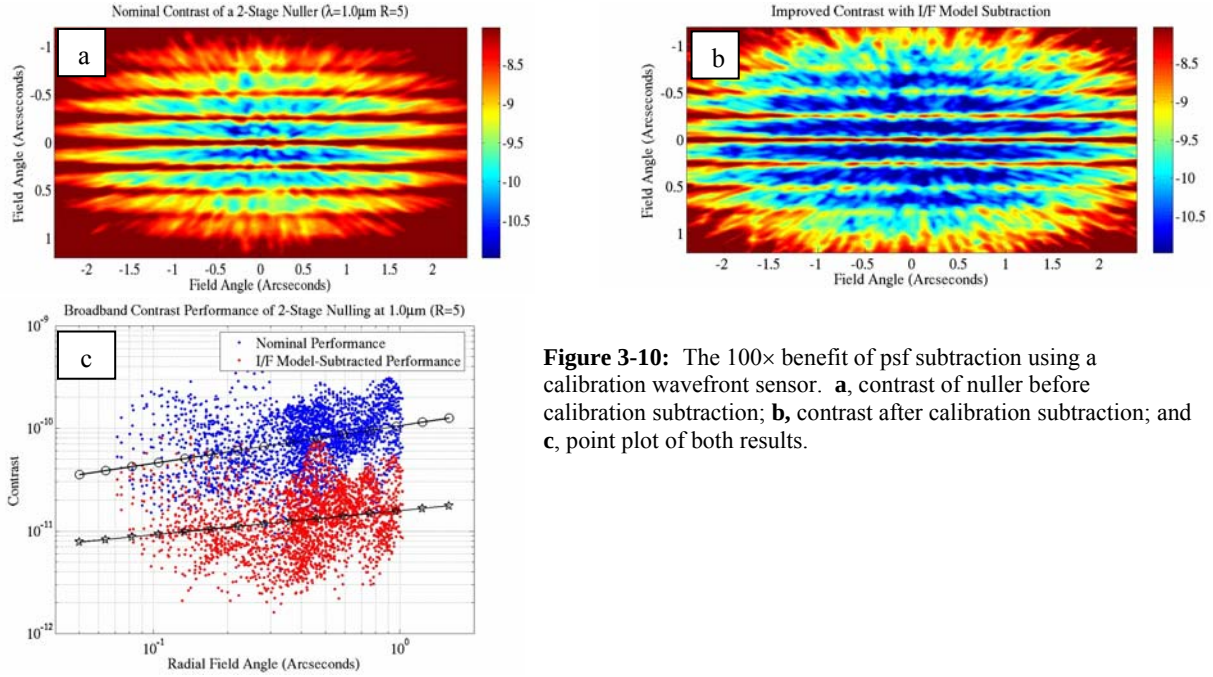

Figure 3-10: The 100× benefit of psf subtraction using a calibration wavefront sensor. **a**, contrast of nuller before calibration subtraction; **b,** contrast after calibration subtraction; and **c**, point plot of both results.

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 simulation of the 8 m × 3.5 m TPF telescope (Figure 3-10). 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.

3.5 Nulling Architecture Concept for TPF-C

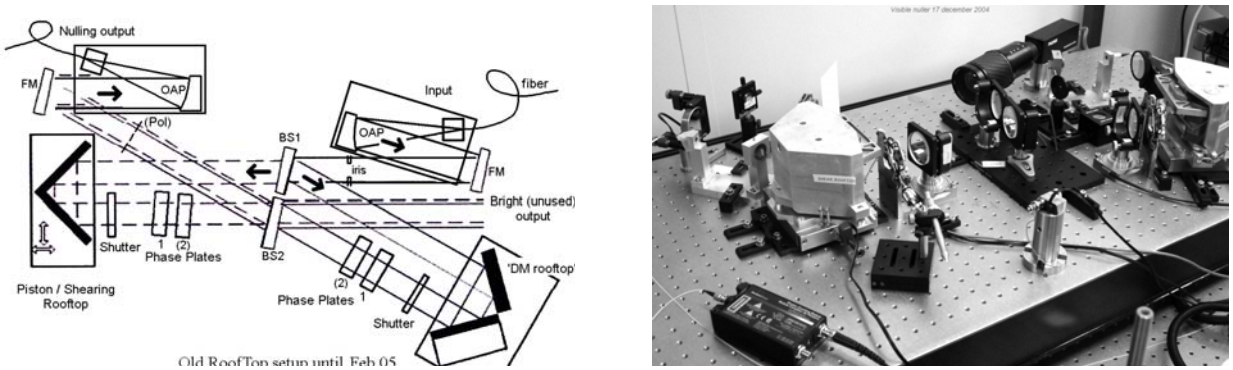

Figure 3-11: Initial nulling interferometer test bed using roof top mirrors. Left, diagram of major components. Right, setup breadboard inteterferometer.

In this section we provide a conceptual design for starlight suppression using a nulling interferometer-based starlight suppression system. The nuller is based on the principles shown in Figure 3-3. The visible nulling testbed at JPL is primarily a modified Mach Zender type interferometer, with single pupil input. Initial results were obtained with configurations based on rooftop where arms fold on themselves with 2 mirrors. The first nuller design was a planar rooftop (aka 2 mirror nuller) whose shear and path length could be independently varied (see Figure 3-11 above).

In order to accommodate a pupil plan at a future deformable mirror, the rooftop setup was modified into 'Boomerang' configuration shown in Figure 3-12 below**.**

The benefit of using a rooftop is that OPD and shear can be adjusted independently, for OPD the rooftop must be shifted longitudinally, whereas for shear the rooftop must be shifted sideways.

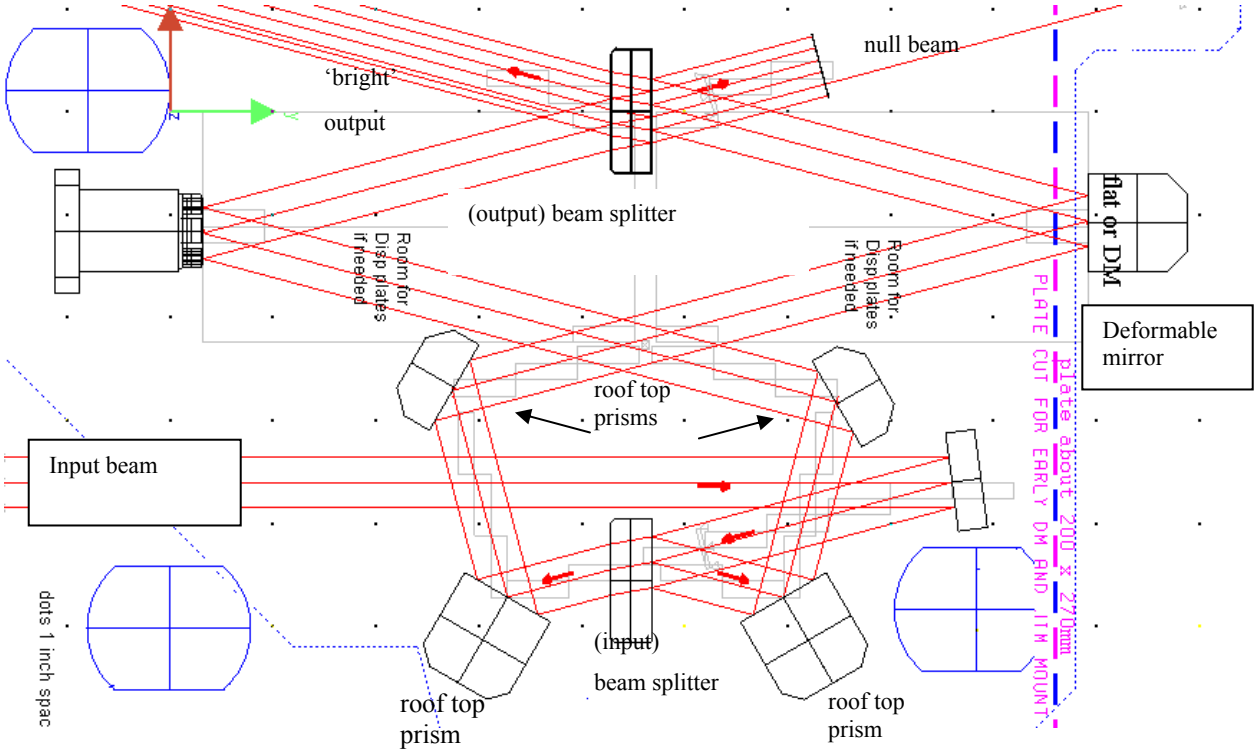

Figure 3-12: A prototype'Boomerang' nulling interferometer showing the major features in a TPF-like configuration.

The candidate TPF configuration uses 2 of these 'Boomerangs' stacked atop the other. The entire instrument fits within a 50cm circular envelope. The nulled output of the first interferometer is directed to a second identical nuller via a vertical rooftop mirror. The second nuller works in reverse, consequently the output of the second interferometer is located near the input position at the first beamsplitter. The top down view shown in Figure 3-13 shows the optics from the same perspective but showing a view highlighted to the calibration camera, light from the science leg is split and directed to the lower nuller by a turning mirror and a vertical roof mirror (needed to maintain the same spatial orientation as the bright output) to a pair of off axis parabolic (OAP) mirrors. This OAP mirror pair is identical to the one (just below) that focuses the bright output of the interferometer, spatially filters the light through a pinhole, and then recollimates the reference light. Both beams are then recombined at a beam splitter which lies below the fiber array. Both outputs are redirected through a triangular mirror to a calibration camera which simultaneously images both interference patterns. By modulating the phase delay in these beams, one needs two images to capture 4 phases (0 $\pi/2$, π, 3 π /2) in the calibration camera.

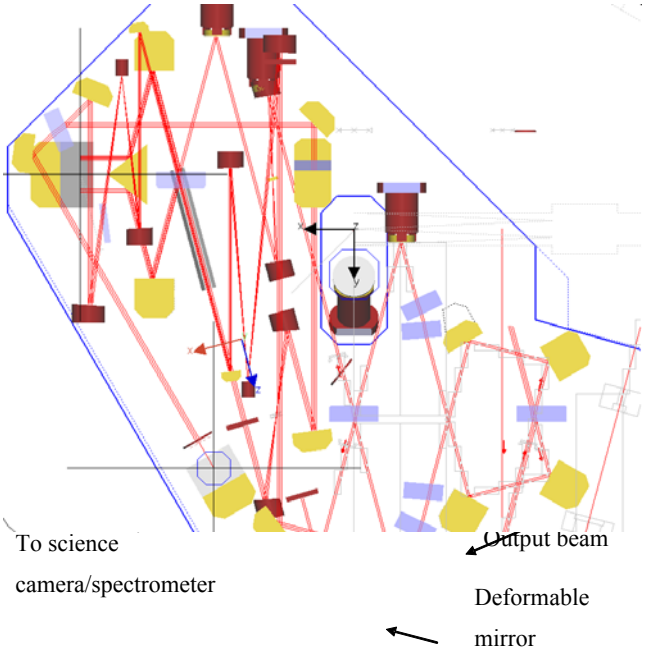

Figure 3-13: (Left) A top down view of a Boomerang configuration double nulling interferometer with a view of lower Calibration leg revealed.

SUMMARY

The nulling coronagraph, due to its ability to work close to the star, 2 l/D, offers very significant science advantages over the baseline coronagraph for the TPF-C mission. This paper describes a conceptual design for such an instrument for the TPF-C mission. Other papers at this conference will describe the status of the technology development work underway. Last of all, this concept studied for the TPF-C project is being applied to a sounding rocket project, PICTURE, that is planned for a launch early in 2007. The calibrator interferometer concept is part of the Gemini Planet Imager, an extreme AO coronagraph being built for the Gemini Telescope. It is also forms the basis for the design of an extreme AO coronagraph for the Thirty Meter Telescope Project. These other projects are described in other papers at this conference.

ACKNOWLEDGEMENT

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract to the National Aeronautics and Space Administration.