Nulling interferometers: the importance of systematic errors and the X-Array configuration

Oliver P. Lay* and Serge Dubovitsky
Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109

ABSTRACT

The nulling interferometers proposed for planet detection are arrays of collector telescopes whose amplitudes and phases are carefully controlled to generate a null response at the star. Perturbations in the amplitude and phase response of the instrument lead to time-dependent fluctuations in the stellar leakage that can mimic a planet signal. Understanding these non-linear systematic errors is important, since they drive most of the instrument requirements for missions such as the Terrestrial Planet Finder and Darwin.

We show that 'amplitude-phase' errors are the dominant source of instrument noise. They are unaffected by the technique of phase chopping, increase rapidly at short wavelengths, are largely independent of the size and transmission efficiency of the collector optics, and depend only weakly on the nulling configuration and distance to the target system. Detection of an Earth around a G-type star like the sun requires ~1.5 nm of path control and ~0.1% control of the amplitude, integrated over all frequencies, including DC.

This paper also introduces the X-Array – a new nulling configuration with 4 collectors and a central combiner arranged in an X pattern. This has a number of advantages over the standard dual Bracewell layout, and over other configurations that have been proposed.

Keywords: Optical interferometry, nulling, planet detection

1. INTRODUCTION

There is currently a large effort to find and study Earth-like planets around other stars. A number of indirect detection schemes for the detection of planets are either in use, or planned in the near future. These include radial velocity measurements, astrometric wobble, and transit photometry.[1,2] None of these approaches detects light originating from the planet itself. The challenge for such a direct detection is the combination of large contrast ratio between the star and planet (~ 10^7 in the mid-IR; ~ 10^{10} in the visible) and the small angular separation (0.1 arcsec or 0.5 μrad for the Earth-Sun at a distance of 10 pc).

The technique of nulling interferometry[3,4] has been proposed to meet this challenge, and is being pursued by both National Aeronautics and Space Agency (the Terrestrial Planet Finder mission[5] and the European Space Agency (the Darwin mission[6]). In its simplest form – the single Bracewell configuration[3]– the light from two collectors, separated by a baseline b, is combined coherently. The response of this simple interferometer on the plane of the sky is a series of sinusoidal fringes, with angular spacing of λ / b. By adding a phase delay to one arm, the starlight can be made to interfere destructively, shifting the response such that the star falls in a null of the fringe pattern. At the same time, a planet, offset from the star by an angle $\lambda / 2b$, falls on the peak of the response. As the baseline is rotated in the plane perpendicular to the star direction, the star remains nulled and the planet moves through peaks and nulls of the response, giving an output photon rate that is modulated in time. Appropriate demodulation of this output gives the flux of the planet and the position relative to the star. The high angular resolution of the interferometer is used to resolve the small angular separation, and the technique of nulling greatly reduces the photon noise from the star. The primary disadvantage of the single Bracewell configuration is that the response on the sky is symmetric on either side of the star. As a result, the planet position angle has a 180-degree ambiguity, it is difficult to separate the planet signal from other sources of emission, and it is not possible to implement the technique of phase chopping (discussed below).

* oliver.p.lay@jpl.nasa.gov; phone 1 818 354-2521; fax 1 818 393-4950

A number of interferometer configurations with more than two collectors have been proposed to overcome these limitations.[7-11] The instrument is the optical equivalent of a phased array, with an amplitude and phase shift applied to the light from each collector element before combination. The nominal amplitudes and phases are chosen to give a null response on-axis. The use of a single-mode spatial filter (e.g. single-mode fiber) before detection relaxes a number of requirements[12], and a dispersing element splits the light into a number of spectral channels on the detector. In addition to the star and planet(s), there will also be emission from warm dust in the target system, a product of the break-up of comets and asteroids, known as the Exo-Zodiacal (EZ) disk. The analog in our own solar system is the Local Zodiacal (LZ) dust, which provides a blanket of foreground emission.

The sensitivity of the nulling interferometer determines how much integration time is required to achieve a given signal-to-noise ratio (SNR) for a given planet at a given distance. There is both random noise and systematic noise. Previous analyses[2,7,8,13] have concentrated on the random noise component, consisting of photon shot noise and the electronic noise of the detection process. The key quantity is the mean null depth, which is relatively straightforward to analyze.

The emphasis of this paper is the systematic noise, arising from fluctuations in the null depth. These effects are shown to drive most of the requirements on the instrument, and understanding them is central to the design of a nulling interferometer. Section 2 describes the error analysis for nulling interferometers, and is based in large part on a more extensive analysis described in [14], in which many of the derivations can be found for the results that are stated here. Section 3 introduces a new nulling configuration, called the X-Array. This 4-element, 2-dimensional array has a number of advantages over the standard dual Bracewell configuration. The key results are summarized in Section 4.

2. ERROR ANALYSIS

2.1. Baseline decomposition

Consider an array of collector telescopes and associated beam trains that deliver light to a beam combiner prior to passage through a single-mode spatial filter (SMSF) and detection. The fully phased condition occurs when all path lengths from an on-axis target source to the detector are exactly matched, and the electric fields from each collector sum coherently to produce a high photon rate at the detector. The null condition can be achieved by the addition of phase shifts ϕ_j to each collector beam, such that the electric fields for an on-axis source now sum to zero. The rate at which photons are detected is given by

$$N = \iint B_{sky}(\mathbf{s}) R(\mathbf{s}) P(\mathbf{s}) \, d\Omega, \tag{1}$$

where $B_{sky}(\mathbf{s})$ is the brightness distribution on the sky for a spectral channel centered on wavelength λ and bandwidth $\Delta\lambda \ll 1$ (units: photons / s / m^2 / sr); s is a unit vector whose direction represents position on the sky (no units); $P(\mathbf{s})$ is a field-of-view taper function resulting from the size of a collecting aperture and the input response of the single-mode spatial filter (no units; value of unity on axis); $R(\mathbf{s})$ is the response of the interferometer on the sky, excluding the taper (units: m^2); and $d\Omega$ is a differential element of solid angle on the sky (units: steradian). The response of the phased array can be written as a sum over all pairs of collectors [ref]:

$$R(\theta_x, \theta_y) = \sum_j \sum_k A_j A_k \cos\left(\phi_j - \phi_k + \frac{2\pi}{\lambda} \left[x_{jk}\theta_x + y_{jk}\theta_y \right] \right), \tag{2}$$

where θ_x and θ_y are small angular offsets from the nominal line-of-sight (units: rad), (x_{jk}, y_{jk}) give the spacing between collectors j and k (units: m), and the x- and y-axis are chosen in convenient directions within the array, in the plane perpendicular to the line-of-sight to the target. The amplitude A_j is the electric field amplitude response for collector j, scaled such that A_j^2 represents the detected photon rate from collector j alone, for a source with flux of 1 photon / m^2. The ability to express the response as a sum over the constituent baselines of the array is a powerful concept; we need only solve the signal and noise analysis for the case of a single baseline, and then sum these solutions together to obtain the answer for a configuration of arbitrary complexity. Figure 1 shows how the response of a standard Dual Bracewell configuration (Fig. 7) can be synthesized from the responses of its individual baselines.

The integral over the sky in Eq. (1), combined with the phase term in square brackets in Eq. (2), is equivalent to a Fourier transform, and it can be shown that the detected photon rate for the phased array is given by

$$N = \sum_j \sum_k A_j A_k \left\{ \cos\left(\phi_j - \phi_k\right) \overline{B}_{sym,jk} - \sin\left(\phi_j - \phi_k\right) \overline{B}_{asym,jk} \right\}, \tag{3}$$

where the horizontal bar denotes a 2-dimension Fourier Transform of the sky brightness distribution, evaluated for the baseline jk. The sky brightness distribution has been split into two components – one that is center-symmetric about the line of sight, and one that is anti-symmetric. The target star, local zodiacal (LZ) dust (the foreground of dust in our solar system) and exo-zodiacal (EZ) dust (the dust present in the target system) are all nominally center-symmetric, and therefore couple to those baselines for which $\cos(\phi_j - \phi_k) \neq 0$; a point-like planet makes equal contributions to the symmetric and anti-symmetric brightness distributions. It is the only contributor to baselines that have $\sin(\phi_j - \phi_k) \neq 0$. In the language of astronomical interferometry, the $\overline{B}_{jk}$ terms of Eq. (3) are the source visibility fluxes for baseline jk.

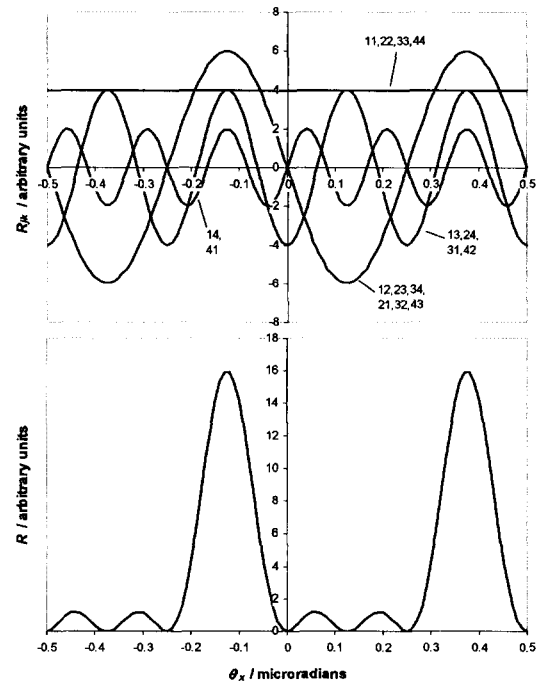


Figure 1: Response of a Dual Bracewell nulling interferometer (see Fig. 7) on the sky. (a) Contributions from each collector pair jk. (b) Resulting total response. The collector amplitudes $A_j = 1$.

2.2. Array rotation

Without array rotation, the photons detected from the planet, star, LZ and EZ dust would all appear at zero frequency, and it would not be possible to separate the contributions. As the array is rotated, the response pattern on the sky rotates with respect to the targets. Since the star is circularly-symmetric, and the LZ has uniform brightness, neither signal is modulated. The EZ appears as an inclined disk, and is therefore elongated. Because of its center-symmetry the EZ photon rate is modulated at even harmonics of the array rotation frequency. The symmetric component of the planet brightness also appears at even harmonics, but the anti-symmetric component appears at odd harmonics, as long as there area baselines in the array for which $\sin(\phi_j - \phi_k) \neq 0$, i.e. ones with phase differences that are not integer multiples of π.

Figure 2 schematically illustrates the signals and noise in the frequency domain. The planet signal is shown at the first and second harmonics of the rotation frequency (in general the planet signal will also appear at higher harmonics of the rotation frequency). The EZ appears at even harmonics only, and the star and LZ appear at DC. Two distinct types of noise are also shown. The random photon noise is the Poisson or shot noise resulting from the random arrival of photons, with a power spectral density equal to the expected photon rate $\langle N \rangle$. The systematic noise results from instrument instabilities: small variations in the instrument response, particularly at the on-axis null, modulate the number of stellar, LZ and EZ photons that get detected. This noise is likely to fall off with frequency. Only noise that competes with planet signal is important. As the observation time T_{obs} is increased, the planet signal becomes narrower and taller (preserving the area) and the signal-to-noise ratio (SNR) increases as $T_{obs}^{1/2}$. Increasing the rotation frequency shifts the planet signal to higher frequencies; the random noise contribution is unchanged, but the systematic noise is decreased. We will show that it is this systematic noise that drives most of the requirements on instrument performance.

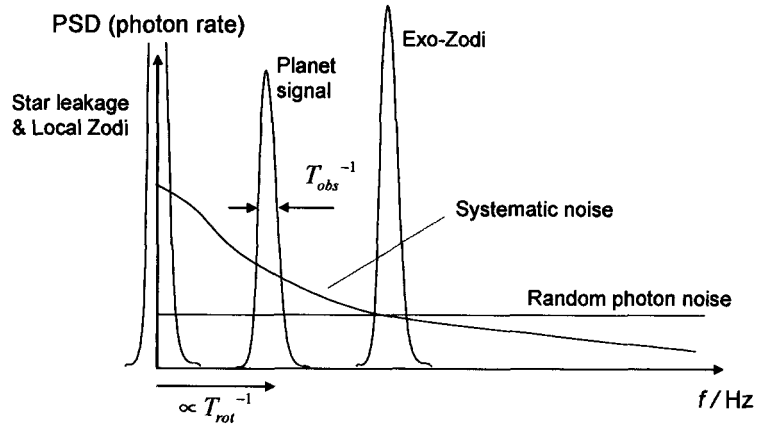


Figure 2: Schematic representation of the power spectrum for the planet signal, stellar leakage, Local Zodi, Exo-Zodi and the associated random and systematic noise.

2.3. Perturbation analysis

The photon rate at the detector can be written as $N = N_0 + \delta N$, where N_0, which we term the 'geometric leakage', represents the photon rate for an ideal instrument without perturbations, and the 'perturbation leakage' δN is the additional photon rate due to perturbations. The geometric leakage can be pictured as the part of the sky brightness distribution that 'leaks' through the ideal instrument response pattern R (Fig. 3a shows the stellar contribution), and is given by Eq. (3) with the nominal values for A_j, ϕ_j, x_j and y_j.

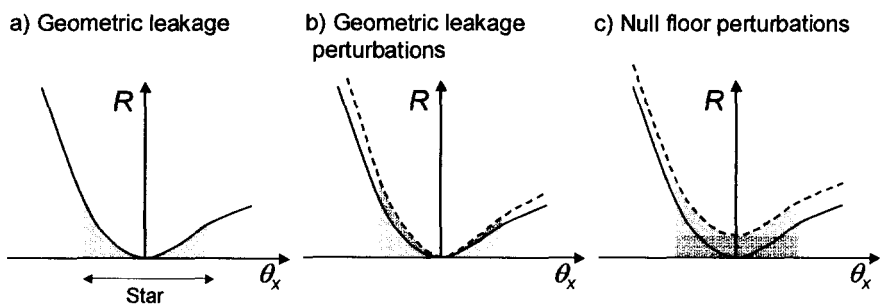


Figure 3: Schematic representation of different contributions to the photon rate. Light gray shading corresponds to the diameter of the star. (a) Hatched area shows 'geometric leakage' of star through the interferometer response R. (b) First order fluctuations in phase, amplitude and collector separation preserve the on-axis null, but lead to perturbations in the geometric leakage (dark shaded area). (c) Second order fluctuations in phase and amplitude give rise to perturbations in the 'null floor' (dark shaded area).

We now derive the perturbation leakage δN. Equation (3) gives the detected photon rate as a function of the array amplitudes A_j, phases ϕ_j, and the collector positions (x_j, y_j) which are implicit in the terms. Perturbations in these quantities lead to perturbations in the detected photon rate. A first-order analysis is not sufficient, however. This is because the amplitudes and phases have been chosen deliberately to minimize the photon rate. Since we are operating at the null condition for an on-axis source, we must expand to second-order in amplitude and phase:

$$\delta N \approx \sum_j C_{Aj}\delta a_j + C_{\phi j}\delta\phi_j + C_{xj}\delta x_j + C_{yj}\delta y_j + \sum_j\sum_k \left[C_{AAjk}\delta a_j \delta a_k + C_{A\phi jk}\delta a_j \delta\phi_k + C_{\phi\phi jk}\delta\phi_j\delta\phi_k \right], \qquad (4)$$

where $\delta a_j = \delta A_j / A_j$ is the fractional change in amplitude from collector j. The C terms are sensitivity coefficients, calculated by taking the appropriate partial derivative. For example,

$$C_{Aj} = A_j \frac{\partial N}{\partial A_j} \approx 2A_j \sum_k A_k \cos\left(\phi_j - \phi_k\right) \overline{B}_{sym,jk} \tag{5}$$

$$C_{A\phi jk} = A_j \frac{\partial^2 N}{\partial A_j \partial \phi_k} \approx \begin{cases} 2A_j A_k \sin\left(\phi_j - \phi_k\right) \overline{B}_{sym,jk} & j \neq k \\ -2A_j \sum_l A_l \sin\left(\phi_j - \phi_l\right) \overline{B}_{sym,kl} & j = k \end{cases} \tag{6}$$

The contribution from the asymmetric component of the brightness distribution has been neglected, since the planet flux is orders of magnitude lower than the star, LZ and EZ. We distinguish between two classes of perturbation leakage: 'geometric leakage perturbations' and 'null floor leakage'.

Geometric leakage perturbations arise from the linear terms in Eq. (4). Figure 3b shows an example for the stellar contribution. The central null is preserved, and the error depends on the size of the source. Contributions from the star, exo-zodi and local zodi are all important. While the star has the highest flux, it is also the most compact of the three, and therefore most effectively nulled. The photon rate in Eq. (4) is proportional to the perturbation, making it straightforward to calibrate out the slow planet-like component of the error in post-processing, given knowledge of the fluctuation, i.e. it is sufficient to measure these errors in real-time; it is not necessary to control them in real-time. These errors are also 'collector-based', with an independent contribution from each collector.

Null floor perturbations arise from the non-linear, second-order terms in Eq. (4). Figure 3c shows an example for the stellar contribution. There is no longer an ideal on-axis null, and errors are obtained even for point-like sources. The stellar photons dominate this class of error. In contrast to the collector-based geometric leakage perturbations, these null floor perturbations are 'baseline-based', with a unique contribution from each baseline pair. Furthermore, the non-linear nature of the errors renders calibration of the errors in post-processing ineffective; instead, we must try to minimize the perturbations in real-time.

The relative importance of geometric leakage perturbations versus null floor perturbations depends on the size of the fluctuations, and the types of source present. Collector configurations that produce a broad central null (e.g the OASES configuration[8]) have lower stellar leakage and are less susceptible to geometric leakage errors, compared to a Dual Bracewell configuration. All configurations, however, are vulnerable to null floor perturbations.

Perturbation leakage contributes to both random and systematic noise. The mean photon rate that determines the level of random photon noise in Fig. 2 is given by

$$\langle N \rangle = N_0 + \langle \delta N \rangle \approx N_0 + \sum_j \left[C_{AAjj} \langle \delta a_j^{\,2} \rangle + C_{\phi\phi jj} \langle \delta \phi_j^{\,2} \rangle \right], \tag{7}$$

where we have assumed that the amplitude and phase perturbations are uncorrelated between the collectors. We use angular braces to denote an ensemble average over all realizations of the noise at a given point in time. Note that there are no linear or cross-terms, since they average to zero.

The power spectral density (PSD) of the systematic noise shown in Fig. 2 is obtained from Eq. (4):

$$\begin{aligned} \widetilde{\delta N} \approx & \sum_j C_{Aj}^2 \widetilde{\delta a}_j + C_{\phi j}^2 \widetilde{\delta \phi}_j + C_{xj}^2 \widetilde{\delta x}_j + C_{yj}^2 \widetilde{\delta y}_j \\ & + \sum_j \sum_k \left[C_{AAjk}^2 \widetilde{\delta a}_j * \widetilde{\delta a}_k + C_{A\phi jk}^2 \widetilde{\delta a}_j * \widetilde{\delta \phi}_k + C_{\phi\phi jk}^2 \widetilde{\delta \phi}_j * \widetilde{\delta \phi}_k \right] \end{aligned}, \tag{8}$$

where the tilde denotes a PSD, and * indicates a convolution operation. The non-linearity of the second-order terms complicates the analysis, leading to the convolution of the disturbance power spectra. Perturbations of different frequency mix together to produce photon fluctuations at the sum and difference frequencies. An amplitude fluctuation with frequency 100 Hz from collector 1, δa_1, can mix with a phase fluctuation with frequency 101 Hz from collector 2, $\delta\phi_2$, to produce a fluctuation in the detected photon rate at frequency 1 Hz. Perturbations of all frequencies contribute to the systematic noise at the planet signal frequency in Fig. 2. These mixing effects cannot be calibrated out in post-processing; instead the perturbations must be minimized in real-time.

An example calculation shows the relative magnitudes of the different contributions. The following parameters were used. Target system: Earth-Sun system at 15 pc, with angular separation of 50 mas, solar-system level of Exo-Zodiacal dust, observed through the Local Zodiacal dust at ecliptic latitude of 30 degrees. Instrument: dual Bracewell configuration (Fig. 7) with 20 m spacing, collector diameter of 4 m, throughput of 10% and a single rotation of 50000 s. Values are for an optical bandwidth of 0.5 μm, centered on a wavelength of 10 μm. Perturbations: amplitude 0.1% rms, phase 0.001 rad rms, both with '1/f' power spectra cut-off below 0.02 mHz and above 10 kHz; collector position, 1 cm rms, white noise up to a cut-off at 0.64 mHz. The results are shown in Table 1, under 'Without phase chop'.

Table 1: Comparison of performance with and without phase chopping

	Without phase chop			With ideal phase chop		
Planet Signal / s^{-1}	**0.097** *			**0.097**		
Photon noise / s^{-1}	**0.070**			**0.070**		
Stellar null floor leakage		0.006			0.006	
Stellar geometric leakage		0.058			0.058	
Local Zodi		0.040			0.040	
Exo Zodi		0.022			0.022	
Systematic noise / s^{-1}	**0.071**			**0.049**		
First order		0.037			0.025	
δA			0.026			0.000
$\delta\phi$			0.024			0.025
δx			0.010			0.000
δy			0.000			0.000
Second order		0.061			0.042	
$\delta A\,\delta A$			0.025			0.000
$\delta A\,\delta\phi$			0.042			0.042
$\delta\phi\,\delta\phi$			0.025			0.000
SNR, one rotation (50,000 s)	**0.971**			**1.132**		

* after even harmonics removed

The geometric leakage from the star is comparable to that from the Local Zodi and Exo-zodi, and all are substantially larger than the stellar null floor leakage, which represents the contribution of the imperfect instrument to the random photon noise. The systematic noise, which is all due to the instrument, is dominated by second-order effects, in particular the amplitude-phase cross-terms. Although the break-out of this error is not shown, it is dominated by the stellar photons, as expected from Eq. (6). The 1 cm collector position error is not a major contribution.

It is the systematic noise that drives the requirements on amplitude and phase control, and in turn most of the requirements on the instrument. Control of the amplitude at the 0.1% level, and phase to 0.001 rad (1.5 nm at λ = 10 μm), for each beam train, integrated over all frequencies, is a very challenging task. Because of the frequency mixing, both static and dynamic perturbations are important. The amplitude of the electric field that is coupled into the single mode spatial filter depends on a number of factors, including reflectivity of the mirrors, absorption of transmissive optics, beam shear, and wavefront errors such as tilt, focus, astigmatism and the higher order aberrations. There is additional non-linearity in the process to convert tilt to amplitude, for example, further complicating the analysis, and a detailed discussion is beyond the scope of this paper. The phase of the electric field obviously depends on path length changes introduced by thermal drifts and structural vibrations, but also on mismatches between beam trains of the birefringence and dispersion introduced by the large number of optical elements. Both amplitude and phase must be actively controlled with auxiliary sensors and actuators, which introduce their own noise.

The analysis so far has neglected several other important sources of error. Rotation of one beam train relative to another (along the direction of propagation) results in a misalignment of the polarization states at the detector, which degrades the null depth. This effect is discussed in [ref??] and turns out to be directly analogous to an amplitude perturbation. Thermal emission from the instrument and stray light from the sun or target star both contribute to the random photon noise and systematic noise, variations in the detector gain can also mimic the planet signal. These three are all linear errors, and can, in principle, be measured in real-time and calibrated out in post-processing. But the planet signal is ~ 1 photon / s, and we require knowledge of the fluctuations at least an order of magnitude below this. Measurement of the fluctuations in thermal emission and stray light to this accuracy, in the presence of photon noise from the Local Zodi

dust, would take longer integration times than for the detection of the planet signal itself. The solution to this problem is the technique of phase chopping.

2.4. Phase chopping

Phase chopping introduces an additional modulation to the planet signal that allows it to be isolated from variations in thermal emission, stray light and detector gain. The instrument is switched between two states, that differ in the way the collectors are phased. The useful planet signal is contained in the asymmetric component of the brightness distribution (the symmetric component is contaminated by the Exo-Zodi emission). The two chop states are chosen such that the response for one chop state is the same as the response for the other, rotated through 180 degrees about the line of sight. This is accomplished by choosing phases for one state that are the negative of the other. For example, for the dual Bracewell configuration the phases would be $\{0, \pi/2, \pi, 3\pi/2\}$ and $\{0, -\pi/2, -\pi, -3\pi/2\}$. The latter is equivalent to $\{0, 3\pi/2, \pi, \pi/2\}$. Both states have the same symmetric response, but 'opposite' asymmetric responses. Any thermal emission and stray light should be incoherent between the beam trains, and therefore identical for the two chop states. Taking the difference in photon rates for the two states will therefore remove variations in these quantities, while preserving the asymmetric component of the planet brightness. The 'chopped' photon rate is given by

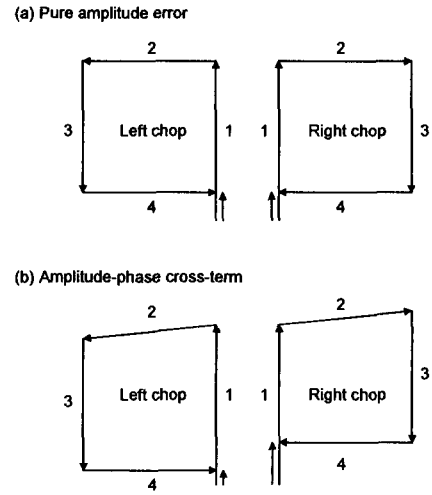


Figure 4: (a) Electric field phasors due to an on-axis point source, at the output of the single-mode spatial filter, from the 4 collectors of a Dual Bracewell configuration, for the left and right chop states. An amplitude error on collector 1 adds equally to the photon rate (proportional to the square of the dashed residual electric field) obtained in the left and right chop states, and is removed when the difference is taken. (b) Example of an amplitude-phase cross-term. The combination of an amplitude error on collector 1 and a phase error on collector 2 adds more photons to the right chop state than the left. This asymmetry leads to an error when the chop states are differenced.

$$N^- = \sum_j \left\{ C^-_{Aj}\delta a^+_j + C^-_{\phi j}\delta\phi^+_j + C^-_{xj}\delta x^+_j + C^-_{yj}\delta y^+_j \right\} + \sum_j \sum_k \left\{ C^-_{AAjk}\delta a^+_j \delta a^+_k + C^-_{A\phi jk}\delta a^+_j \delta\phi^+_k + C^-_{\phi\phi jk}\delta\phi^+_j \delta\phi^+_k \right\}. \quad (9)$$

The '-' superscript denotes half the difference between chop states; the '+' superscript denotes half the sum between chop states. For example, if the chop states are denoted by Left and Right, then $\delta a^+_j = \left(\delta a^L_j + \delta a^R_j\right)/2$ represents the amplitude perturbations that are common-mode between the two states (which ideally would be all of them). The 'chopped response' on the sky has no symmetric component, and couples entirely to the desirable asymmetric component of the sky brightness. As a result, the power spectrum of the chopped output contains only odd harmonics of the rotation frequency. While they have been removed from the chopped output, note that the star, EZ and LZ still contribute to the photon noise exactly as for the case without chopping.

The chopped sensitivity coefficients can be evaluated using the phases for the two states. The property of inverted phases between the states, substituted into Eqs. (5), (6), and their equivalents for the other variables, lead to the following results: $C^-_{Aj} = C^-_{xj} = C^-_{yj} = C^-_{AAjk} = C^-_{\phi\phi jk} = 0$, $C^-_{\phi j} = C^L_{\phi j} = -C^R_{\phi j}$, $C^-_{A\phi jk} = C^L_{A\phi jk} = -C^R_{A\phi jk}$. This implies that the chopped output is not sensitive to the first-order systematic errors in amplitude, and position, or to second-order errors in amplitude or phase. It is, however, just as sensitive to first-order phase errors and second-order amplitude-phase cross terms as before.

$$N^- = \sum_j C^-_{\phi j}\delta\phi^+_j + \sum_j \sum_k C^-_{A\phi jk}\delta a^+_j \delta\phi^+_k. \quad (10)$$

This is a key result for this paper. The consequences are illustrated by the example in Table 1, under the column 'with phase chopping'. The planet signal and random noise are unchanged. The overall systematic noise has been reduced, but the dominant amplitude-phase cross terms remain, in addition to the first-order co-phasing error. Figure 4 shows graphically how phase chopping removes pure amplitude errors, but does not remove the amplitude-phase errors.

Phase chopping is vulnerable to additional errors. If the instrument is switched alternately between the two chop states at frequency f_{chop}, then any photon fluctuations at f_{chop} are 'synchronously detected' and lead to systematic noise in the chopped output. It is therefore desirable to have a high chop frequency, at least to the point where detector read noise becomes an issue. The chop mechanism itself may also introduce asymmetries; for example, if delay lines are used to introduce the switch in phase, and one setting introduces a slightly different tilt from the other, then the stellar leakage will differ between the two states and couple into the chopped output. There are many combinations of these error terms, and a more complete analysis can be found in [14].

The analysis in this section has been for a simple binary chop – switching between two states. It is possible to show that there is no linear combination of arbitrary chop states that can remove the systematic errors without also removing the planet signal.

2.5. Combining spectral channels

All the previous analysis has applied to a single narrow spectral channel. While this is appropriate for the spectroscopic characterization of planets, the initial detection of planets requires broadband operation to maximize sensitivity. For the Terrestrial Planet Finder mission the bandwidth is currently specified to be 6.5 – 17 μm. If, however, a single detector were used to cover this entire range, the interferometer response on the sky, which scales with wavelength, would become smeared out. Instead, the band is divided into a number of spectral channels for both spectroscopy and detection, and the channels are then combined in post-processing to obtain broadband sensitivity. This also allows the channels to be weighted to maximize the overall signal-to-noise ratio.

It can be shown[14] that when the channels are weighted in the optimal way to maximize the overall sensitivity, then the combined SNR over all channels is given by

$$SNR_{tot} = \left\{ \sum_{\text{indep } n} SNR_n^{\,2} \right\}^{1/2} , \tag{11}$$

where SNR_n is the SNR for channel n. This equation assumes that the noise in the spectral channels is independent, i.e. that there are no correlations between channels. While this is always true for random photon noise, the systematic noise is likely to be correlated across adjacent channels that have similar wavelength. At some point, the systematic errors do become uncorrelated as the difference in wavelength increases. There are two effects to take into account: (1) as the wavelength is decreased, the planet signal shifts to higher harmonics, and therefore couples to higher harmonics of the systematic noise, which are uncorrelated with the lower harmonics; (2) different contributions to the systematic error have different dependences on the wavelength. For example, an amplitude-phase error arising from the combination of a wavefront error (λ^{-2} in amplitude) and a path length error (λ^{-1} in phase) has an overall dependence of λ^{-3}, and will be dominant at short wavelengths. An amplitude-phase error arising from a shear error (λ^{0} in amplitude) and a coating dispersion error (λ^{0} in phase, for this example) is achromatic, and will be more dominant at the longer wavelengths compared to the λ^{-3} contribution. A full analysis of this problem requires detailed simulation, and is beyond the scope of this paper. In order to estimate the detection sensitivity, we can approximate the behavior by assuming a value for the number of noise-independent spectral channels across the band, which in general will be different from the physical number of channels in the spectrometer.

2.6. Dependence on wavelength and distance

In this section we explore the dependence of the signal and noise terms on wavelength and distance to the target system. The values are calculated for the dual Bracewell configuration with phase chopping, and the target system specified in Tables 1 and 3.

Figure 5a shows the dependence on wavelength, for a series of 5 noise-independent spectral channels, each 2 μm wide, spanning the band 7 – 17 μm. The values are for the example parameters specified in Section 2.3. The planet signal rises towards the long wavelengths as a result of the Planck function for a 265 K black body. The systematic noise contributions from the amplitude-phase cross-term and the co-phasing error rise rapidly at short wavelengths. The random photon noise is dominated by stellar geometric leakage at short wavelengths and by the Local Zodi at long wavelengths. The resulting SNR is a maximum at about 14 μm, and falls off rapidly at short wavelengths. This has two

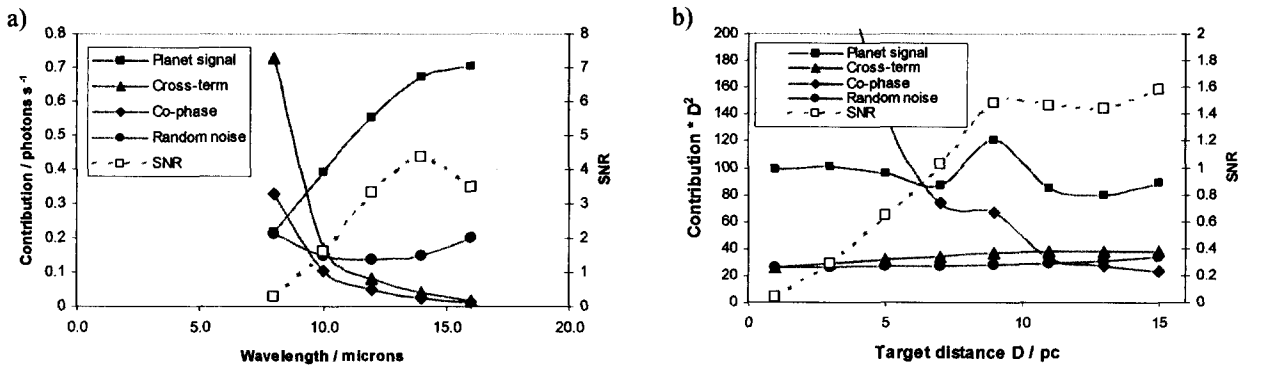


Figure 5: (a) Wavelength dependence for planet signal, two systematic noise terms, random photon noise, and the resulting signal-to-noise ratio, after a single array rotation of 50,000 s for an Earth-Sun system at 15 pc. Spectral channels are 2 μm wide. Further assumptions are listed in Tables 1 and 3. (b) Distance dependence for planet signal, two systematic noise terms, random photon noise, and the resulting signal-to-noise ratio, after a single array rotation of 50,000 s with a spectral channel spanning 9 – 11 μm for an Earth-Sun system. The effect of the inverse square law has been removed by multiplying the contributions by the square of the distance.

implications. First, spectroscopy becomes much more challenging at $\lambda < 10$ μm, and will require long integration times. Second, the angular resolution of the interferometer should be assessed using a wavelength of ~ 14 μm, since this is where the maximum sensitivity is obtained.

Figure 11 illustrates the dependence on distance to the target system, for a spectral channel spanning 9 to 11 μm. The effect of the inverse square law with distance has been removed for clarity (the ratios of the quantities are preserved). As a result, the planet signal appears flat with distance. As the distance increases, the angular separation of the planet from the star decreases. This reduces the frequency of the planet signal harmonics, which then couple to lower harmonics of the systematic noise, accounting for the slight rise in the amplitude-phase cross-term with distance. The stellar geometric leakage dominates the photon rate in this spectral channel, scaling with distance as D^{-4}. The associated photon noise scales as D^{-2} (appearing flat in the plot), and the co-phasing error scales as D^{-4}, dominating at distances shortward of 10 pc. It is the co-phasing error that drives the SNR down for nearby stars, despite the increased number of signal photons. Performance for nearby stars can be improved by reducing the array size, re-phasing the array[11], or using a configuration with a broader null[8,9].

3. THE X-ARRAY CONFIGURATION

In this section we introduce a new nulling configuration, which we call the X-Array. It offers a number of advantages over other configurations that have been proposed.

The ideal nulling configuration has 3 desirable properties: (1) a broad null to minimize stellar leakage; (2) high angular resolution to clearly separate the planet signal from the star, lumps in the Exo-Zodi emission, and any other planets that may be present; and (3) high modulation efficiency. The configurations proposed so far can, at best, meet 2 of these 3 criteria. The evenly-spaced dual Bracewell design has high modulation efficiency, but is forced to compromise between stellar leakage and angular resolution – a large array size gives good resolution but high leakage, and vice versa. Configurations with higher order θ^4 and θ^6 nulls, such as ESA's Bow-tie configuration and the OASES family proposed by [8] are capable of broad nulls and high angular resolution with a sufficiently large array size, but suffer from low efficiency.

The X-Array meets all 3 criteria. It is a simple modification of the standard Dual Bracewell design, as shown in Fig. 7. The amplitudes and phases of the 4 collectors follow the same scheme. The array is simply opened out to form a rectangle or 'X'. The key feature is that the 'nulling' baselines (those with a phase difference of π) are now independent of the 'resolution' or 'imaging' baselines (those with a fractional- π phase difference), as depicted in Fig. 8a. The nulling baselines determine the leakage, and the resolution baselines determine the angular resolution in the chopped output map. The response of this configuration on the sky is shown in Fig. 8b and 8c. The short nulling baselines, length b, lead

to wide fringes with angular spacing λ / b. These are crossed in the orthogonal direction by a series of narrow fringes, spaced by λ / B, where B is the length of the resolution baseline. The nulling and imaging properties of the array have been completely decoupled: stretching the X along the horizontal direction increases the angular resolution without impacting the stellar leakage. Phase chopping for the X-Array is implemented in the same way as for the Dual Bracewell; the phases of the second chop state are indicated in parentheses in Fig. 8a, resulting in a sky response that is a reflected-through-the-center version of Fig. 8c. The average modulation efficiency – a measure of the efficiency with which planet photons are converted into output signal – of the X-Array is 0.22, compared to 0.28 for the Dual Bracewell and 0.10 for the Bow-tie.

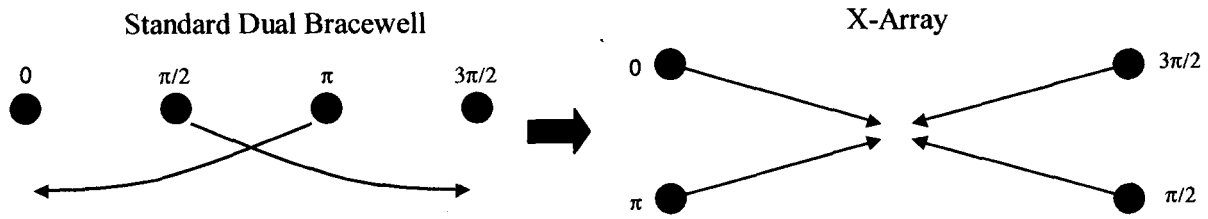


Figure 7: The X-Array can be considered to be an 'opened out' version of the standard Dual Bracewell configuration. The amplitudes and phasing of the collectors are preserved.

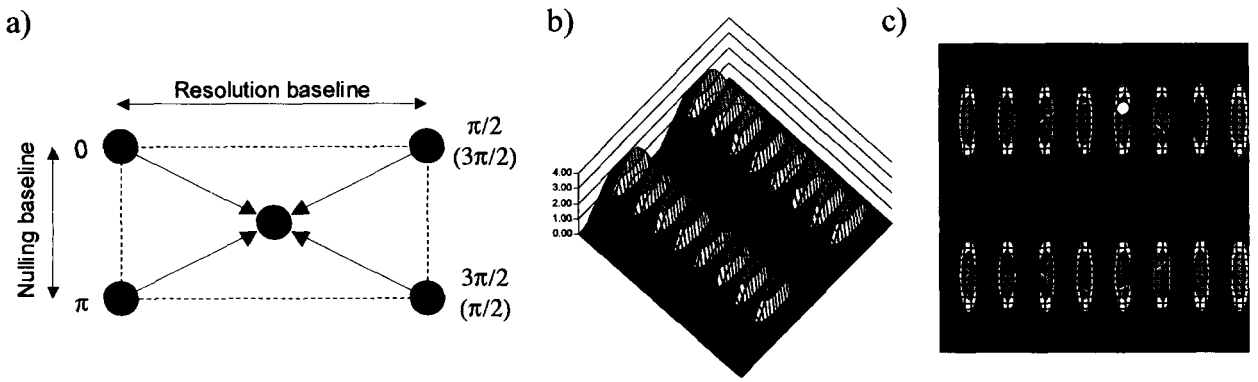


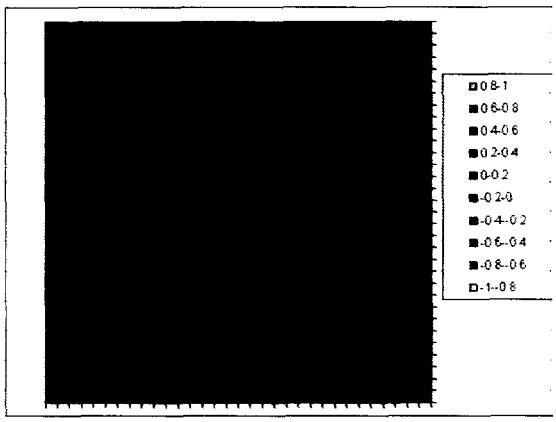
Figure 8: (a) The X-Array decouples the nulling baselines (π phase difference) from the resolution baselines (fractional-π phase difference). The different phases for the second chop state are indicated in parentheses. (b) Response of the X-Array on the sky showing wide and narrow orthogonally intersecting fringes. (c) X-Array response on the sky showing the locus of a planet as the array is rotated about the line of sight to the star.

Figure 9 shows simulated correlation maps for a 2-planet system, for a standard Dual Bracewell design, 95 m long, and an X-array with dimensions of 100 m by 22 m. The X-Array has a much smaller point spread function than a dual Bracewell of equivalent size. The Full Width to Half Maximum is 12 mas for the X-Array, compared to 30 mas for the Dual Bracewell. This is because the 4 resolution baselines have lengths 100 m, 100 m, 102 m and 102 m for this X-Array, compared to 31 m, 31 m, 31 m and 93 m for the Dual Bracewell. The two nulling baselines are 62 m long for the Dual Bracewell, compared to 22 m for the X-Array. The additional stellar leakage incurred by the Dual Bracewell leads to a doubling of integration times for targets at 15 pc. These and other comparisons are made in Table 1.

Table 1: Comparison of X-Array and standard Dual Bracewell nulling configurations

X-Array	Standard Dual Bracewell
+ High resolution AND short integration times (low leakage)	- High resolution OR short integration times (low leakage)
+ Simple, direct collector-to-combiner beam relay	- Two-step beam relay needed to balance path lengths
+ Identical, interchangeable collectors	- All collectors different (left vs right, inner vs outer)
+ One configuration meets all needs	- Must switch to low-resolution config for nearby stars
+ Uniform collector propellant usage	- Outer collectors require more propellant
- Modulation efficiency = 0.22	+ Modulation efficiency = 0.28
- Array angles will vary if aspect ratio changed	+ Array angles fixed

a) Standard Dual Bracewell



b) X-Array

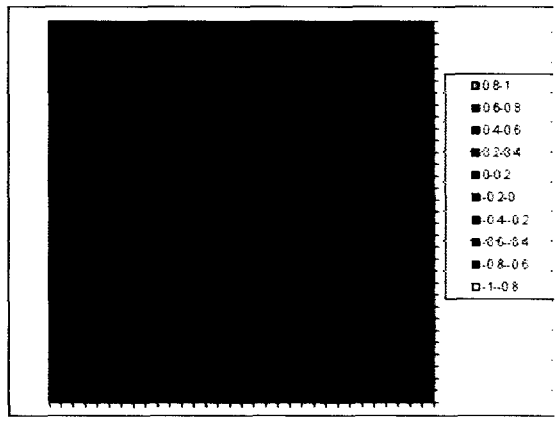


Figure 9: Comparison of raw correlation maps for the linear Dual Chopped Bracewell configuration and the X-Array. Both arrays span approximately 100 m, and the grid spacing is 10 milliarcseconds. Two planets are present, with offsets of (+70 mas, 0) and (=15 mas, +40 mas) relative to the center. The Dual Bracewell has a synthesized beam diameter of approximately 30 mas, compared to 12 mas for the X-Array.

It is worth noting the connection between the X-Array and the configuration proposed by Velusamy and Beichman[7]. The X-Array consists of two 2-element Bracewell nullers, separated by the resolution baseline, and combined at a central beam combiner spacecraft. The Velusamy design uses two 3-element Degenerate Angel Cross nullers instead of the Bracewell nullers, for a total of 6 collectors and a central beam combiner spacecraft.

4. SUMMARY

This paper presents a generalized error analysis for nulling interferometers. The key points are listed below.

1. The response of nulling interferometer configurations can be decomposed into a sum over all baselines in the array. Results for a single baseline can then be extended to configurations of arbitrary complexity.

2. It is the systematic error (null stability), not the random error (null depth) that drives most of the requirements on the instrument. Detection of an earth-like planet at mid-infrared wavelengths, for the example case described, requires that the phase of the signal from each collector is controlled to ~ 1 millirad (~ 1.5 nm at 10 μm), and the amplitude to ~ 0.1%.
    a. Systematic errors are caused by time-varying leakage through the null, due to perturbations in the amplitudes, phases and polarization angles of the electric fields from each collector telescope. Stellar photons are the dominant source of systematic noise.
    b. The perturbation analysis is dominated by non-linear, second-order terms, and in particular by amplitude-phase cross terms. The disturbance power spectra mix with each other, and perturbations at all frequencies, including DC, make an impact.

3. The technique of phase chopping is needed to remove slow drifts in stray light, thermal emission and detector gain that can mimic the planet signal. While removing some systematic errors, it has no impact on the dominant amplitude-phase cross-terms, and the co-phasing error. There is no phase chop scheme that can remove the systematic errors without also removing the planet signal
    a. The systematic errors rise rapidly at short wavelengths. Since the planet signal is also falling at short wavelengths, the instrument has most of its sensitivity at longer wavelengths, and spectroscopy at $\lambda <$ 10 μm will be very challenging.

b. The impact of the amplitude-phase error on SNR is approximately independent of the distance to the target system. The co-phasing error increases with the angular size of the star and becomes dominant for nearby systems.

The work described in this paper was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

1. N. Woolf and J. R. Angel, "Astronomical searches for earth-like planets and signs of life", Annu. Rev. Astron. Astrophys **36**, 507-537 (1998).
2. G. W. Marcy and R. P. Butler, "Detection of extrasolar giant planets", Annu. Rev. Astron. Astrophys **36**, 57-97 (1998).
3. R. N. Bracewell, "Detecting nonsolar planets by spinning infrared interferometer", Nature **274**, 780-781, (1978).
4. E. Serabyn, "Nulling interferometry and planet detection", in *Principles of long baseline interferometry*, P. Lawson, ed, (Jet Propulsion Laboratory, 2000), pp. 273-292.
5. D. Coulter, "NASA's Terrestrial Planet Finder mission: the search for habitable planets," in *Towards Other Earths: DARWIN/TPF and the search for extrasolar planets*, M. Fridlund and T. Henning, eds., (ESA Publications Division, Noordwijk, Netherlands, 2003), pp. 47-54.
6. A. Karlsson and L. Kaltenegger, "The technology of DARWIN", in *Towards Other Earths: DARWIN/TPF and the search for extrasolar planets*, M. Fridlund and T. Henning, eds., (ESA Publications Division, Noordwijk, Netherlands, 2003), pp. 41-46.
7. C. A. Beichman, N. J. Woolf, and C. A. Lindensmith, *Terrestrial Planet Finder: a NASA Origins program to search for habitable planets*, (Jet Propulsion Laboratory, 1999).
8. J. R. P. Angel and N. J. Woolf, "An imaging nulling interferometer to study extrasoloar planets", ApJ **475**, 373-379, (1997).
9. B. Mennesson, and J. M. Mariotti, "Array configurations for a space infrared nulling interferometer dedicated to the search for earthlike extrasolar planets", Icarus **128**, 202-212, (1997).
10. O. Absil, *Nulling interferometry with IRSI-Darwin: detection and characterization of Earth-like exoplanets*, Masters thesis, (University of Liege, Belgium, 2001)
11. M. C. Noecker and R. L. Linfield, "Error analysis for a compound nulling interferometer",in *Interferometry in space*, M. Shao, ed., Proc. SPIE **4852**, 412-422, (2003).
12. B. Mennesson, M. Ollivier, C. Ruilier, "Use of single-mode waveguides to correct the optical defects of a nulling interferometer", J. Opt. Soc. Am. A **19**, 596-602 (2002).
13. K.-C. Liu, D. W. Miller, and G. J. W. Mallory, "Derivation of stability requirements and control effort analysis for space nulling interferometers," in *Proceedings of IEEE Aerospace Conference* (Institute of Electrical and Electronics Engineers, New York, 2002), pp. 797-807.
14. O. P. Lay, "The impact of systematic errors on nulling interferometers", Appl. Opt. (submitted), (2004).