Behavior of imperfect band-limited coronagraphic masks in a high-contrast imaging system

Erkin Sidick and Daniel W. Wilson

We investigate the behavior of imperfect band-limited occulting masks in a high-contrast imaging system through modeling and simulations. Grayscale masks having 1D Sinc^2 (linear-Sinc^2) amplitude transmission coefficient (Sinc^4 intensity transmittance) profiles as well as optical density and wavelength-dependent parasitic phases are considered occulters. We compare the behaviors of several, slightly different occulter transmittance profiles by evaluating the contrast performance of the high-contrast imaging testbed (HCIT) at the Jet Propulsion Laboratory (JPL). These occulters include a measured occulter, a standard Sinc^2 occulter, and several of its variations. We show that when an occulting mask has a parasitic phase, a modified Sinc^2 transmittance profile works much better than the standard Sinc^2 mask. We examine the impact of some fabrication errors of the occulter on the HCIT's contrast performance. We find through modeling and simulations that starlight suppression by a factor of more than 10^{10} is achievable at least monochromatically on the HCIT with the occulting mask and the optics currently being used on the testbed. To the best of our knowledge, this is the first time that we investigate the behavior of a real (or fabricated) focal plane occulting mask in a high-contrast imaging system. We also briefly describe the approach used at JPL in fabricating a grayscale occulting mask and characterizing its transmittance and phase profiles.

OCIS codes: 110.6770, 220.4830.

1. Introduction

The high-contrast imaging testbed (HCIT) at the Jet Propulsion Laboratory, California Institute of Technology, is the Terrestrial Planet Finder (TPF) coronagraph's primary platform for experimentation.[1,2] It is used to provide laboratory validation of key technologies as well as the demonstration of a flight-traceable approach to implementation. Its optical design corresponds to a classical Lyot coronagraph: attenuating the starlight by an occulting mask located at the focal plane without affecting the light of a planet if present, and blocking the ringlike residual light at the subsequent pupil plane by a Lyot stop. Wavefront control is performed by a high-density deformable mirror (DM) located at the pupil before the occulting mask to achieve monochromatic planet–star contrast ratio better than 10^{-10}.

High-contrast coronagraphic systems designed for space telescopes usually employ an occulting mask that is band limited and phase free. Kuchner and Taub[3] have shown that a conventional coronagraph with an ideal band-limited graded image-plane mask can, in principle, provide an arbitrarily large dynamic range without need for phase control or a severe Lyot stop. However, as far as we know, no one has made or described a process to make a focal plane occulting mask that is phase free. One type of occulting mask being used on the HCIT is a grayscale pattern written in high-energy beam sensitive (HEBS) glass using a high-voltage electron-beam lithography facility at JPL.[4] It is a linear (or 1D) mask, designed with a Sinc^2 function of a single variable, x. It has been found that this type of occulting mask not only displays an optical-density (OD) and wavelength-dependent parasitic phase, but its OD profile also differs from the original target—a perfect linear-Sinc^2 profile.[4,5] Nonetheless, these masks have been used to demonstrate near-TPF level performance in HCIT.[1,2] The OD profile discrepancy is caused mostly by fabrication errors and also to some degree by measurement errors. These errors in the transmission amplitude and phase of the occulting mask degrade the contrast performance. Especially when an occulting mask has parasitic phase, the mask no longer works only on the intensity of the incoming beam, instead, it modulates the phase of the incoming beam as well. To understand the contrast floor of the HCIT and similar coro-

The authors are with the Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109. E. Sidick's e-mail address is erkin.sidick@jpl.nasa.gov.

Received 28 September 2006; revised 7 November 2006; accepted 9 November 2006; posted 10 November 2006 (Doc. ID 75545); published 1 March 2007.

0003-6935/07/091397-11$15.00/0

nagraphic optical systems under various practical conditions, we have investigated the behavior of an ideal as well as several imperfect grayscale occulting masks by evaluating the narrowband contrast performance of the HCIT with an optical analysis tool. The details of the HCIT and this analysis tool used to perform our modeling and simulations are described in Refs. 1, 2, and 6. Recently our simulations have predicted that the contrast values obtainable on the HCIT with a narrowband (or monochromatic) illumination at $\lambda = 785$ nm wavelength are $C_m \sim 2 \times 10^{-11}$ (mean) and $C_4 \sim 5 \times 10^{-11}$ (at an angular separation of $4\lambda/D$), in contrast to the measured results of $C_m \sim 6 \times 10^{-10}$ and $C_4 \sim 8 \times 10^{-10}$, respectively.[6] We reported in that paper, based on our simulation results, that the measured, imperfect OD profile of the linear-Sinc^2 mask gives much better contrast performance as compared to a perfect linear-Sinc^2 design, but at the time we did not understand the causes of such performance enhancement. We have since then clarified that issue through detailed investigations. In this paper, we describe how various graded, imperfect band-limited coronagraphic masks perform, and compare to each other, in a high-contrast imaging system such as the HCIT. Our results are useful in choosing optimum occulting mask designs to maximize the contrast performance of the HCIT and other coronagraphs for finding terrestrial planets from space.

Both the amplitude and the phase errors of the various optical components and the DM corrective phase have different wavelength dependence, and the DM correction becomes ineffective as the optical bandwidth is increased. This imposes very challenging requirements on the reflectivity and the optical surface height of the various optical components,[7] as well as on the design and the fabrication of the occulting masks. Recently, we evaluated the broadband contrast performance of the current HCIT optical system and examined the behaviors of both dispersionless and dispersive occulting masks.[8] In this paper, we deal with only the monochromatic performance of the HCIT when equipped with both ideal and imperfect band-limited focal plane occulting masks.

2. HCIT Optical System and Components

In this section, we briefly describe the HCIT's optical system and components. The schematic of the HCIT layout in the x–z plane is shown in Fig. 1. Artificial starlight is created by a 5 μm pinhole illuminated by

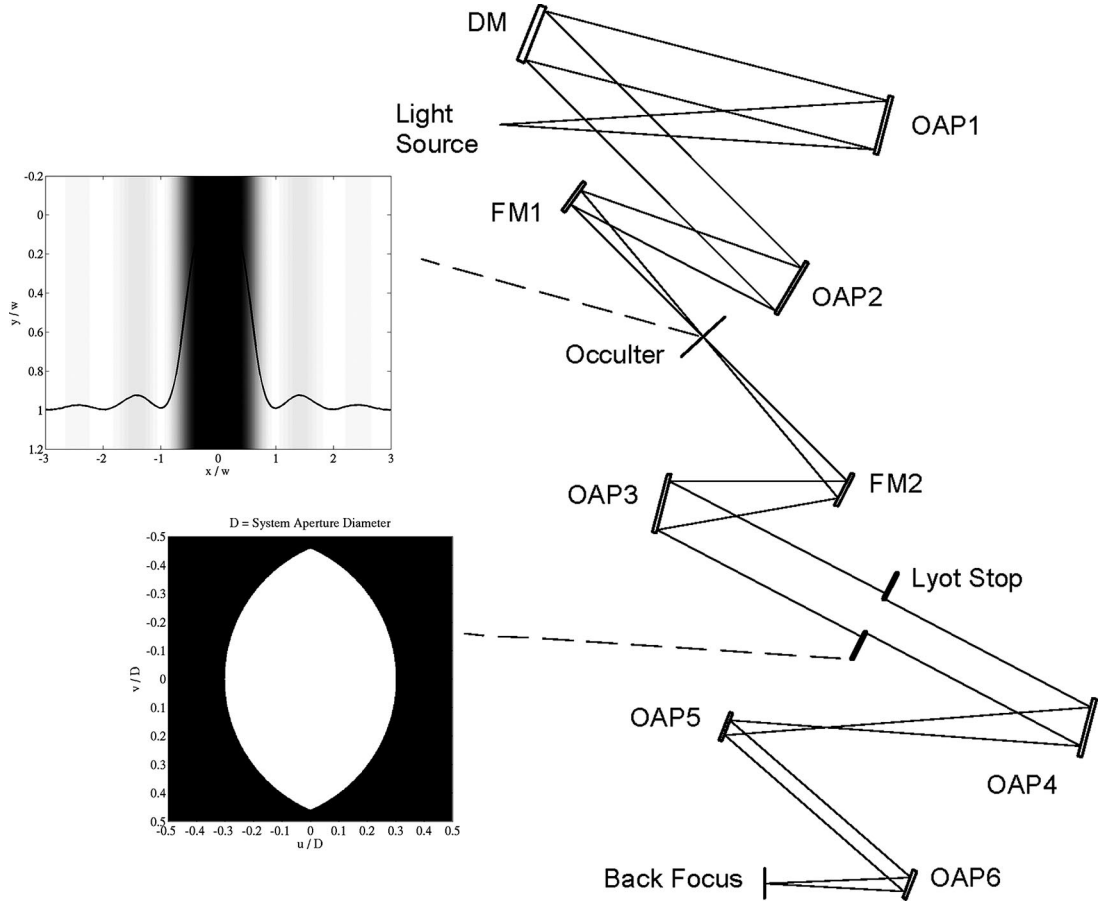

Fig. 1. Schematic of the top (x–z plane) view of the HCIT. The light source (starlight) is a 5 μm pinhole illuminated by an optical fiber, and a CCD science camera is located at the back focal plane for detecting the image of the starlight. DM = deformable mirror, OAP = off-axis parabolic mirror, and FM = flat mirror. The z axis is in the direction of the beam propagation, and the x axis lies on the paper plane perpendicular to the z axis.

an optical fiber. An off-axis parabolic mirror (OAP1) collimates the light from the pinhole and directs it to a high-density DM, which performs the wavefront control. A circular aperture mask on the DM defines the system pupil of the HCIT and has a diameter of $D = 30$ mm. After the DM, the collimated light is reimaged onto the focal plane of the occulting mask by OAP2 and a flat mirror (FM1). The occulting mask attenuates the starlight, and almost has no effect on the light of a planet if present. The back end of the system, from the occulting mask to the back focus plane, supports experimentation with diverse coronagraph configurations and apodizations. A flat mirror (FM2) and OAP3 recollimate the light passing through the occulter mask and form a same-size sharp image of the DM pupil at the Lyot plane. A Lyot stop blocks the ringlike residual light diffracted off the occulting mask while letting most of the planet light through. After OAP4 forms an image from the remaining stellar and planet lights, it is then magnified ($M \approx 3$) by the OPA5–OAP6 pair for proper sampling on the CCD science camera located at the back focal plane.

The optical prescription of our simulation tool and the various optical parameters are identical to those used on the testbed and shown in Fig. 1. We take into account the amplitude drooping of the illuminating beam exhibited at the system pupil, the independently measured phase errors of the various optics, and the OD and the λ-dependent parasitic phase of the occulting mask. Other errors and effects are not included in the analysis. They include, but are not limited to, the background scattered light on the testbed, uncertainty in the illumination beam amplitude nonuniformity, uncertainty on the phase errors of the actually used occulting mask, amplitude errors introduced by nonuniform reflectance of HCIT optics, polarization effects, detector noise, noise in the DM electronics, variations in the DM actuator gains and influence functions, and backreflections of light from the occulting mask glass and from the CCD camera screen. Other information about the HCIT setup and its optical components are given in Ref. 6 and will not be repeated here.

The original design of the occulting mask has a linear-Sinc^2 transmittance profile given by[3]

$$T(x) = \left[1 - \left(\frac{\sin(\pi x/w)}{(\pi x/w)}\right)^2\right]^2, \qquad (1)$$

where w is the occulter width. We use $w = 144$ μm in our numerical simulations unless specified otherwise (the occulter on the testbed also has $w = 144$ μm). The front-end f number of the HCIT is 28.5. Therefore when $\lambda = 785$ nm, this occulter width value corresponds to an occulter-plane focal ratio of $6.4\lambda/D$. The mask intensity transfer function associated with the above function has FWHM $= 1.1457w$, and it corresponds to a focal ratio of $7.4\lambda/D$. The Lyot stop corresponding to this type of occulters has an eye shape (see Fig. 1). The one currently being used on the testbed is a simple blackened sheet metal with a sharp edge. Its clear aperture is defined by the overlapping section of two circles with diameter D and is separated by a distance of ε. Presently $\varepsilon = 0.4D$ on the testbed, and we use the same value in our simulations.

3. Fabrication and Measurement of Grayscale Occulting Masks

A. Electron-Beam Fabrication

The occulting masks were fabricated by scanning-spot electron-beam (e-beam) exposure of HEBS glass (a proprietary product of Canyon Materials, Inc., San Diego, California, USA).[4] HEBS glass darkens with e-beam exposure, and the resulting absorption spectrum depends on the electron energy and the chemical formulation of the glass. For these experiments, a special HEBS glass formulation was developed by Canyon Materials to achieve high optical density when exposed at 100 kV using JPL's JEOL JBX-9300FS e-beam lithography system.

Prior to the occulting mask fabrication, two main physical effects must be calibrated and compensated: (1) the nonlinear optical density versus the dose sensitivity of the HEBS glass, and (2) the proximity effect—the exposure resulting from electron scattering. To measure the OD versus dose, we exposed 2 mm $\times$ 2 mm squares with constant dose and measured their transmission using a 785 nm diode laser and a picowatt optical power meter. The resulting 785 nm OD versus dose function is plotted in Fig. 2. The proximity effect is more complicated to calibrate. For this experiment we represented the proximity effect by a single Gaussian model, with the dose point-spread function (PSF) taking the form

$$\mathrm{PSF}(\mathbf{r}) = \delta(\mathbf{r}) + \frac{\eta}{\pi\alpha^2}\exp(-r^2/\alpha^2), \qquad (2)$$

where η is the strength, and α is the range. The incident beam can be represented as a delta function

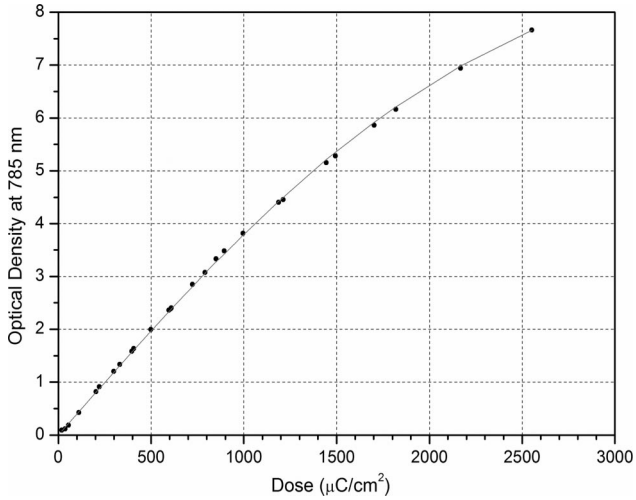

Fig. 2. Optical density of thick sensitive layer HEBS GI85 glass at 785 nm as a function of electron-beam dose. The data were fit with the ad hoc saturating function $y = Ax/(1 + Bx^2)$ for use in e-beam pattern preparation.

in Eq. (2) because the range of the proximity effect is large compared to the beam diameter, and we are not concerned with correcting feature sizes of the order of the beam size. The total dose delivered to the HEBS sensitive layer is thus the primary (incident) dose convolved with the PSF. To characterize the proximity effect, we captured transmittance images at 785 nm of the constant-dose calibration squares and extracted profiles across the edges. For the case of a broad area exposure, the total (incident + proximity) dose, ρ_{tot}, near the edge can be derived analytically as the convolution of Eq. (2) with a step function to give

$$\rho_{tot}(x) = \rho_{prim}\left\{\text{step}(x - x_{\text{step}}) + \frac{\eta}{2}\left[1 + \text{erf}\left(\frac{x - x_{\text{step}}}{\alpha}\right)\right]\right\}, \tag{3}$$

where ρ_{prim} is the value of the uniform primary dose. To fit the transmittance profile and determine the proximity effect range and strength parameters, Eq. (3) is converted to OD using the nonlinear OD function from Fig. 2, with the dose axis scaled by $(1 + \eta)$ to account for the fact that for broad area exposures, $\text{OD} = f\{\rho_{tot} = \rho_{prim}(1 + \eta)\}$. The results of the fitting indicated that $\eta \approx 1$ and $\alpha \approx 20$ μm. This means that the proximity effect contributes approximately the same dose as the incident beam and requires careful compensation to realize accurate transmittance profiles.

Once the nonlinear OD versus dose response and the proximity effect parameters are known, the desired OD pattern is converted to an e-beam dose pattern. To start, the desired OD profile is represented by an array of square or rectangular pixels (1.5 μm squares for the patterns described here). Each pixel is exposed by the e-beam at the same accelerating voltage, but with a dwell time that is proportional to the desired dose. First, the pixel ODs are converted to required total dose using the OD function, $\rho_{tot}(x, y) = f^{-1}\{\rho(x, y)\}(1 + \eta)$. Then the required primary dose for each pixel can be determined by deconvolving Eq. (2) from the desired total dose pattern. In practice, this is performed by using fast Fourier transform computations. The deconvolution of profiles that contain rapid changes in OD typically produces nonrealizable negative doses. The solution is to add a small uniform OD to the entire profile until there are no doses less than some small positive value that can exposed by the e-beam system.[4] The final pattern preparation step is to convert the pattern of primary pixel doses into the native JEOL e-beam format using JPL in-house software.

Prior to exposure, the HEBS glass is coated with 200 Å of chromium by thermal evaporation to serve as a discharge layer. We have found that the rapid exposure of HEBS glass can lead to pattern nonuniformity, likely because of substrate heating and/or bulk charging. Hence we expose the pattern at low current in multiple passes (typically five exposures at $\frac{1}{5}$ dose). After exposure, the chrome is stripped with Cyantek Corp. CR-7S etchant, and the sample is thoroughly cleaned.

B. Transmittance and Phase Measurement

The transmittance and/or OD profiles of the fabricated occulting masks were measured using the transmission imaging setup illustrated in Fig. 3. A monochromator with a tungsten–halogen source is used for illumination, and a pair of lenses image the output slit onto the occulting mask. The mask is oriented such that the residual wavelength variation (a few nanometers) across the slit is perpendicular to the 1D mask profile. The light transmitted through the mask is imaged by a microscope objective onto a cooled CCD camera (QImaging Retiga EXi). At wavelengths longer than 800 nm, a 650 nm long-pass filter was inserted just after the exit slit to block the second-order wavelengths from the monochromator (silicon CCD sensitivity starts at 400 nm). For each transmission image I_S, a dark image I_D and a reference image I_R (taken through unexposed glass) were captured using the same integration time and number of averages (typically 16). The transmittance image was then calculated from $T = (I_S - I_D)/(I_R - I_D)$ with the result shown in Fig. 4 for the Sinc^2 profile studied in this paper. The measured occulting mask profile used in the analysis of Section 4 was 50 row average cross section of the image as shown in Fig. 5. The magnification of the imaging system was calibrated by imaging a reticule of known feature spacing.

The measured phase of the exposed HEBS glass was reported in Ref. 5, and the phase versus OD data used in the current analysis are shown in Fig. 6 with different symbols representing the different wavelengths. We have developed an analytical model for this type of occulter phase in the form of $\phi = \phi(\text{OD}, \lambda)$, and have used it to obtain the occulter phase information at different wavelengths and at regions where

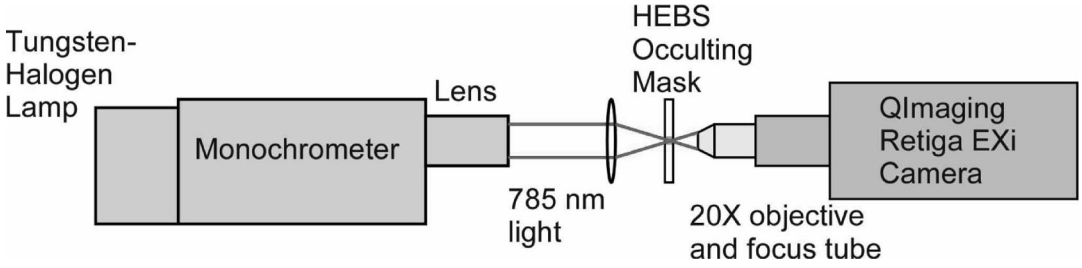

Fig. 3. Experimental setup for transmission imaging of HEBS occulting masks.

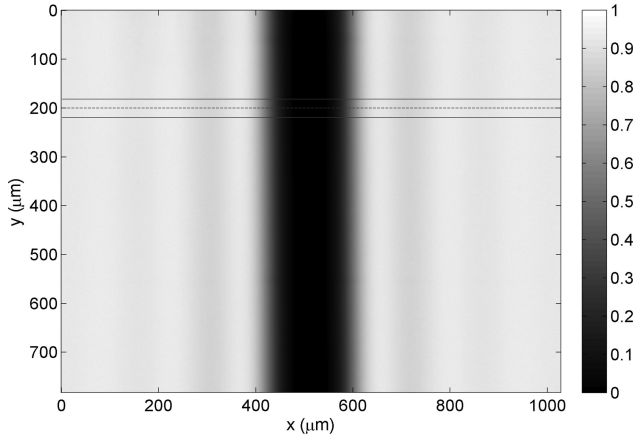

Fig. 4. Transmittance image of the Sinc^2 profile analyzed in this paper. Lines indicate the region of profile averaging for cross section.

OD $>$ 4.5. The solid curves in Fig. 6 were obtained using this analytical model. As we can see from Fig. 6, there are two distinctive regions in each ϕ(OD, λ) curve: a nonlinear region at OD $<$ 1.5 and a linear region at 1.5 $\leq$ OD $<$ 4.5. The occulter phase is not important when OD $\geq$ 8, and we assume ϕ(OD, λ) = ϕ(OD = 8, λ) for that region in our simulations. Figure 7 shows the OD and the phase profiles of an occulter, where the OD curve represents the data measured at λ = 785 nm, and the phase curve was obtained from the analytical model $\phi = \phi$(OD, λ) described above.

4. Results

A. Steps of Speckle-Nulling Simulations

A detailed and realistic optical model of the HCIT was set up by using its as-built prescription, and a speckle-nulling algorithm[1] was also implemented.

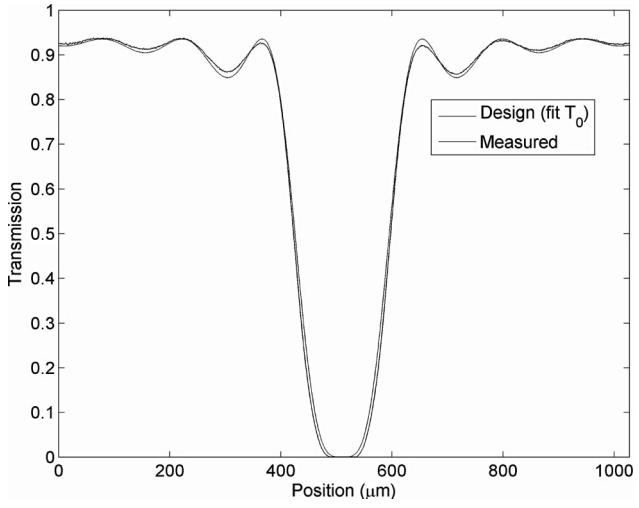

Fig. 5. Cross-section profile of Sinc^2 mask (50 row average between lines in Fig. 4). A fit of the design function $T(x) = T_0[1 - \mathrm{sinc}^2(x/w)]^2$ with maximum transmittance T_0 = 0.935 as a parameter is shown for comparison.

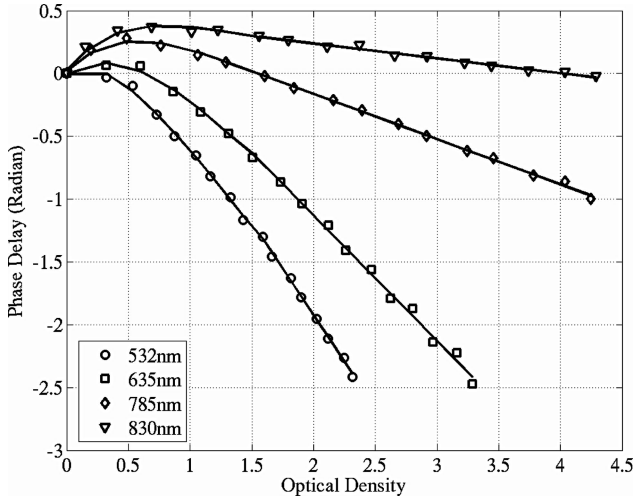

Fig. 6. Phase delay as a function of OD with wavelength as a parameter. The symbols represent the measured data of a HEBS glass (Ref. 5) and the corresponding curves obtained from an analytical model $\phi = \phi$(OD, λ).

This algorithm is identical in principle to what was implemented on the testbed, with some differences in the implementation detail. The fundamental principle of this algorithm was explained in Ref. 1 and will not be repeated here. In a typical speckle-nulling process on the testbed, the wavefront control software takes the starlight image captured by the CCD camera as an input, finds a new set of DM control commands from it, then updates the DM settings with the new set of control commands. In our simulations, on the other hand, the final image plane intensity map is obtained for a given set of parameters by running Modeling and Analysis for Controlled Optical Systems,[9] (MACOS) a versatile optical modeling tool developed at JPL and used in many flight projects. By carrying out end-to-end full diffractive propagation analysis with this optical model, we are not only able

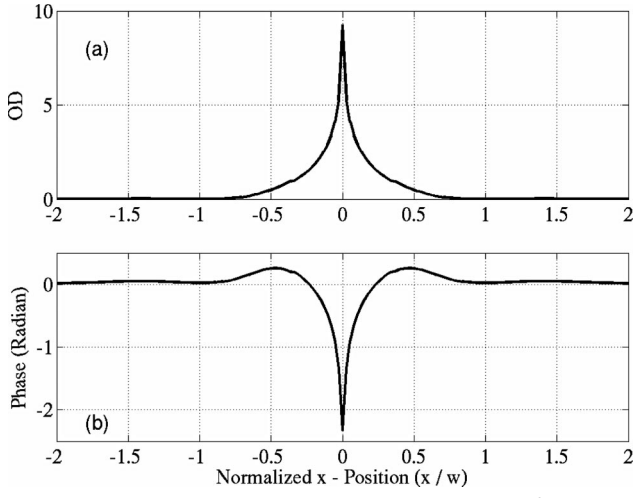

Fig. 7. OD and phase profiles of a linear (1D) Sinc^2 occulting mask. (a) OD profiles measured at λ = 785 nm; (b) corresponding phase profile obtained from the analytical model $\phi = \phi$(OD, λ).

to evaluate the HCIT's performance under various conditions, but also to simulate the speckle-nulling optimization process and make reasonably accurate predictions on its capability.

The DM used on the HCIT has 1024 (32×32) actuators arrayed on a 1 mm pitch. It can produce a high-contrast half-dark hole in an area defined by the controllable spatial frequency of the DM. The targeted high-contrast window in the HCIT coronagraph image is $[x_{\min} \quad x_{\max} \quad y_{\min} \quad y_{\max}] = [4 \quad 10 \quad -10 \quad 10] * \lambda/D$ in the final focal plane, where x and y are the position variables in the focal plane normalized by the corresponding focal distance. We use a mean contrast, C_m, and a contrast at $4\lambda/D$, C_4, as the ultimate metrics of the HCIT's performance. Here the C_m is defined as the contrast value averaged over the window of $[4 \quad 10 \quad -10 \quad 10] * \lambda/D$, and the C_4 as the one averaged over a smaller window of $[4 \quad 5 \quad -0.5 \quad 0.5] * \lambda/D$. The quantity contrast is calculated using Eq. (12) in Ref. 10 but without any averaging. That is, it is the star image with the nonocculted image and the occulter shape divided out.

In our simulations, the speckle nulling is carried out with the following several steps: (1) Adjust the de-space distance of the occulting mask from its nominal position so that the total energy at the final focal plane (or the scattered light) is minimized (usually 0–1 mm). (2) Carry out speckle nulling for five iterations in the full right half-plane, for five iterations in the full left half-plane, and repeat these two substeps one more time. This step reduces the amount of diffracted light leaking through the Lyot stop. (3) Continue the speckle nulling in the full right half-plane for 100 iterations. In this case, the size of the speckle-nulling window is determined by the controllable spatial-frequency bandwidth of the DM. (4) Continue the speckle nulling inside a reduced window of $[3 \quad 11 \quad -11 \quad 11] * \lambda/D$ for several hundred iterations. (5) Carry out the rest of the speckle-nulling iterations in an even smaller window of $[3 \quad 10 \quad -10 \quad 10] * \lambda/D$. There is an additional parameter aside from the speckle-nulling window that requires adjustment in the speckle-nulling algorithm, namely, the gain factor. It is used to scale the correcting phase map (or DM corrective surface pattern) obtained by processing the coronagraph image intensity. The right value for this gain factor depends on the implementation details of the algorithm and should be determined through numerical experimentation. The goal is to make the speckle-nulling process as efficient as possible. It should be pointed out that the MACOS-based speckle-nulling simulation process is fairly time consuming. This is because, although the spatial frequency and the orientation of the sinusoidal surface pattern of the DM used to null a specific speckle can be uniquely determined from the location of that speckle in the focal plane, the initial phase of the sinusoidal corrective phase is initially unknown and needs to be determined by trying several different values to see which one minimizes the speckle. In our case, we use an initial phase increment of 0.15 rad and calculate the star images six times to complete one speckle-nulling iteration. When doing so, for a grid size of 512×512 and with an Intel Xeon 3.2 GHz dual-processor workstation, it takes approximately 2 h for 100 speckle-nulling iterations.

B. Occulter Optical Density Profiles: Measured versus Design

In this section, we describe the difference of the measured occulter OD profile from the original target. The test setup used to measure the OD profiles of grayscale occulting masks studied here had a noise floor of OD ~ 3. It was found that the measured OD profile deviates from the target starting at OD ≈ 1, and such deviation becomes significant starting at OD > 3. Therefore the measured OD values in Fig. 7(a) corresponding to OD > 1 have been replaced by the data calculated from Eq. (1) with $w = 144.0$ μm. In this paper, we still call this measured and repaired OD profile a measured profile. However, it must be kept in mind that the central OD peak corresponding to OD > 1 of this measured occulter was actually obtained from the model. Here, the occulter width value of $w = 144.0$ μm was obtained by fitting the model in Eq. (1) to the measured data in the OD ≤ 1 data range. The original design had $w = 141.9$ μm. That is, the fabricated occulter turned out to be slightly wider than the original target. In addition, a close examination of the measured OD profile reveals that it differs from the design in the following two additional aspects: (i) Its sinusoidal ripples are damped (or smoothed) and do not hit the OD = 0 axis, as shown in Fig. 8(a). This might have been caused mostly by fabrication errors, and to a lesser extent by measurement errors. (ii) The measured OD versus x curve rises faster than the target at $0.35 < x/w < 0.7$, as shown in Fig. 8(b). The discontinuity on the measured curve in Fig. 8(b) near $x/w \sim 0.35$ is caused by the mismatch between the measured data and the

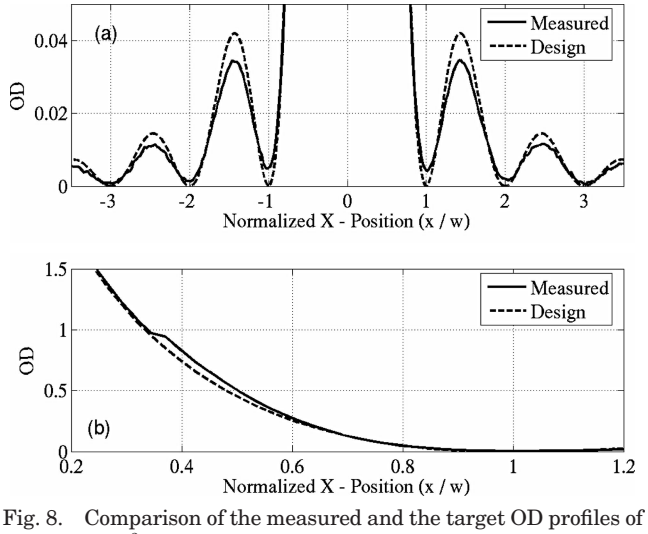

Fig. 8. Comparison of the measured and the target OD profiles of a linear-Sinc^2 occulting mask designed with $w = 141.9$ μm. (a) Ripples of the measured OD profile are smoothed (or damped) as compared to the design; (b) in regions where $0.3 < |x/w| < 1$, the measured OD profile deviates from the standard Sinc^2 design.

model. In the following, we compare the behavior of the measured occulter OD profile with that of the standard one by dividing the problem into several small categories.

C. Behavior of Phase-Free Occulter: Measured versus Design

In this section, we start comparing the behavior of the measured occulter with that of a standard mask. First, we examine the case where the occulters are phase free. That is, we take into account only the wavefront amplitude drooping of the artificial starlight and the phase errors of all optics, but ignore the parasitic phase of the occulter. If we carry out numerical speckle-nulling optimization on the HCIT optical model by utilizing the measured and the designed OD profiles of the occulter, respectively, we obtain two curves as shown in Fig. 9. They are the values of mean contrast, C_m, plotted as a function of the speckle-nulling iteration number. The figure legend indicates the corresponding occulter OD profile used in the simulation. Each iteration refers to one speckle-nulling cycle. The letter "a" indicates the point where the speckle-nulling window was changed from the full right half-plane to the $[3 \quad 11 \quad -11 \quad 11] * \lambda/D$ rectangle, and the letter "b" indicates the point where this window was further reduced to $[3 \quad 10 \quad -10 \quad 10] * \lambda/D$. As we can see, the contrast increases considerably faster in the first several speckle-nulling iterations following each size reduction of the speckle-nulling window. All of the mean contrast versus iteration number curves to be presented in the remainder of this paper display a similar characteristic, but we will no longer put the letters a and b in the figures. In the present case of the phase-free occulter, the standard Sinc^2 mask performs much better than the measured OD profile. At the end of 1000 iterations, the mean contrast obtained with the standard Sinc^2 mask is better by more than an order of magnitude as compared to that obtained with the measured OD profile.

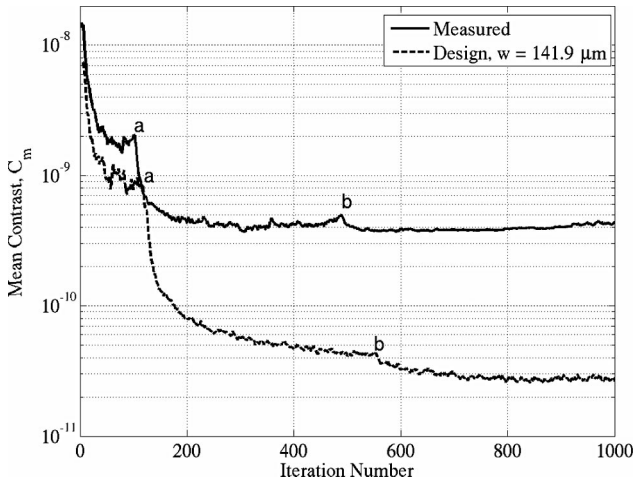

Fig. 9. Mean contrast, C_m, as a function of iteration number. The curves marked as measured and design, $w = 141.9$ μm correspond to the occulter OD profiles illustrated in Fig. 8. Letter a near Iteration = 100 indicates the point where the speckle-nulling window was reduced from the full right-half plane to $[3 \quad 11 \quad -11 \quad 11] * \lambda/D$, and the letter b near iteration = 500 indicates the point where it was further reduced to $[3 \quad 10 \quad -10 \quad 10] * \lambda/D$.

D. Behavior of Occulter with Parasitic Phase: Measured versus Design

We now examine the behaviors of realistic occulters, that is, the ones with parasitic phase errors. In this part of the simulations, both occulting masks, namely, the measured and design as explained in Fig. 8, are assumed to have parasitic phases given by the analytical model $\phi = \phi(\mathrm{OD}, \lambda)$, similar to the one shown in Fig. 7(b). For these two occulters, we now obtain the mean contrast versus iteration number curves as shown in Fig. 10. As we can see, the speckle-nulling process increases the mean contrast in the right half dark hole to $C_m = 1.6 \times 10^{-11}$ after 1000 iterations when the measured OD profile of the occulting mask is used. However, this obtainable contrast value decreases to $C_m = 8.8 \times 10^{-11}$ when the occulter is switched to the standard Sinc^2 mask. That is, the measured OD profile yields a factor of 6 improvement in mean contrast as compared to the standard Sinc^2 occulter. This result indicates that when the occulting mask has parasitic phase, the standard Sinc^2 occulter transmittance profile is no longer the best solution for a high-contrast imaging system like HCIT. This is one of the most significant results of this study.

At this point it is still not clear which of the following three factors about the occulter OD profile contributed to the above improved contrast performance of the HCIT: (i) the damped ripples in the occulter transmission amplitude; (ii) the occulter width w, and (iii) the nonstandard Sinc^2 shape of the central OD curve of the occulter. To clarify this issue, we repeated our numerical speckle-nulling optimization process for the following two new cases: (a) an occulter having a standard Sinc^2 OD profile with $w = 144$ μm at everywhere except in the central

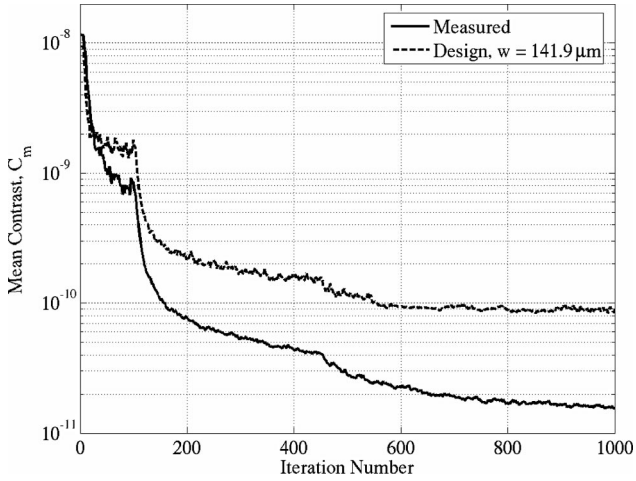

Fig. 10. Same as Fig. 9 except that the parasitic phase of the occulter is taken into account in each case.

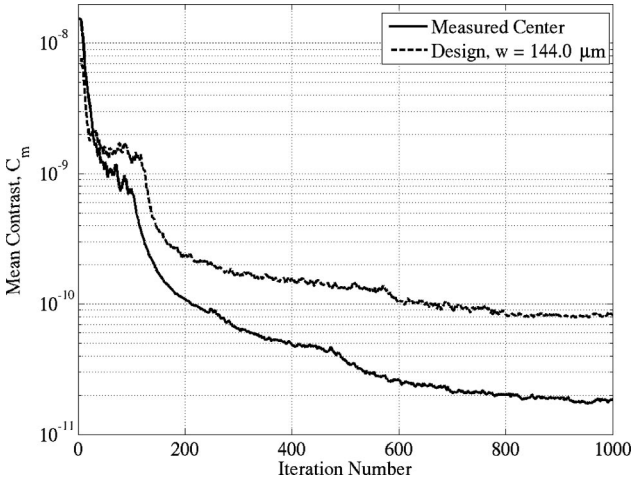

Fig. 11. Mean contrast, C_m, as a function of iteration number. The solid curve marked as measured center corresponds to an occulter having a standard Sinc^2 OD profile with $w = 144$ μm everywhere except in the central $|x/w| \leq 0.9$ region, where its OD profile is equal to the measured data as shown in Fig. 8(b). The dashed curve corresponds to a standard Sinc^2 OD profile with $w = 144$ μm.

$|x/w| \leq 0.9$ region, where the OD profile is equal to the measured data; (From here on, we call this occulter the "measured center" profile. The shape of the central peak of this profile is the same as the measured and "repaired" profile in Fig. 8(b), but its ripples are not damped or smoothed as in the case of the measured occulter.) and (b) A standard Sinc^2 occulter with $w = 144$ μm. Between these two cases, the former is chosen to explore what impact the damped ripples have on contrast performance, and the latter to find out the effects of occulter width. The results of the mean contrast versus the iteration number of these two cases are shown in Fig. 11. These results are very close to those shown in Fig. 10. That is, in the case of standard Sinc^2 occulter, the present change in the mask width from $w = 141.9$ to 144 μm does not make much difference in the contrast performance of the HCIT. On the other hand, the damped ripples of the occulter OD profile make some difference in contrast, but the difference is small, for example, at the end of 1000 iterations, $C_m = 1.8 \times 10^{-11}$ in the present case, and $C_m = 1.6 \times 10^{-11}$ in Fig. 10. That is, in the case of the measured occulter OD profile, the smoothed ripples yielded a contrast result that is slightly better than the case of the unsmoothed ripples. A similar phenomenon has been seen in another numerical test case that is not shown here. In that case, we generated a standard Sinc^2 OD profile with $w = 144$ μm and smoothed its ripples at $|x/w| > 0.9$ by averaging the data inside a moving window, thus creating damped ripples similar to those of the measured data. After 446 iterations of speckle nulling, we obtained $C_m = 2.5 \times 10^{-10}$ and $C_4 = 1.0 \times 10^{-9}$, respectively, with this smoothed OD profile, in contrast to $C_m = 1.5 \times 10^{-10}$ and $C_4 = 7.5 \times 10^{-10}$ obtained with the perfect Sinc^2 profile and with the same number of iterations. That is, for the standard Sinc^2 occulter, the smoothed ripples of the occulter transmittance reduced the contrast values as compared with the perfect occulter transmittance profile, but the decrease is very small. These results lead us to the following general conclusion: Damped ripples in the occulter's transmittance do not have much impact on the monochromatic contrast performance of a high-contrast imaging system. This is another important finding of this study. It eliminates the fears that existed before about the possible detrimental effect of the smoothed transmission amplitude ripples found in fabricated grayscale occulters and allows one to relax some of the requirements on the mask fabrication tolerance.

The comparison of the results in Figs. 10 and 11 show that the measured occulter is the clear winner among the four occulter profiles that have been examined. In other words, the third factor, that is, the non-Sinc^2 shape of the OD curve in the $0.3 < |x/w| \leq 1$ regions, is responsible for the great improvement in the HCIT's contrast performance. This result is very important. It indicates that when the occulting mask has a parasitic phase, the standard Sinc^2 OD profile is not the optimum choice for a high-contrast coronagraph.

E. More Imperfect Band-Limited Occulters

After determining what in the occulting mask OD profile contributes to the contrast performance improvement of the HCIT, we constructed a new occulter OD profile. This new profile has a Sinc^2 profile with $w = 144$ μm in all regions except in two intervals at $0.2 < |x/w| \leq 1$, where it is represented by a fifth-order polynomial. For the left-hand side of the central peak, it is given by

$$\mathrm{OD}(X) = -3.77X^5 - 14.74X^4 - 19.11X^3 - 6.36X^2 + 4.55X + 2.79, \quad X = x/w. \tag{4}$$

The polynomial curve portion of this profile starts from $x/w = -1$ and ends at $x/w \approx -0.21$ where it touches the Sinc^2 curve. The right-hand side of the central peak is equal to the mirror image of the OD(X) given by Eq. (4). In this paper we call this mask the "Modified Sinc^2" occulter. Part of its OD profile is shown by the solid curve in Fig. 12(a). The standard Sinc^2 occulter is also included in this figure for comparison. To confirm that this new occulter works the same as the measured occulter with parasitic phase described in Subsection 4.D and to gain a better understanding about how much the contrast performance of the HCIT is sensitive to the shape of the central peak of an occulter's OD profile, we chose two additional OD profiles as shown in Fig. 12(a). These two new occulters have the same OD profile as a standard Sinc^2 at $|x/w| > 1$, but a different profile inside the $|x/w| \leq 1$ region. Among them, the $|\mathrm{Sinc}|^{1.75}$ profile is given by

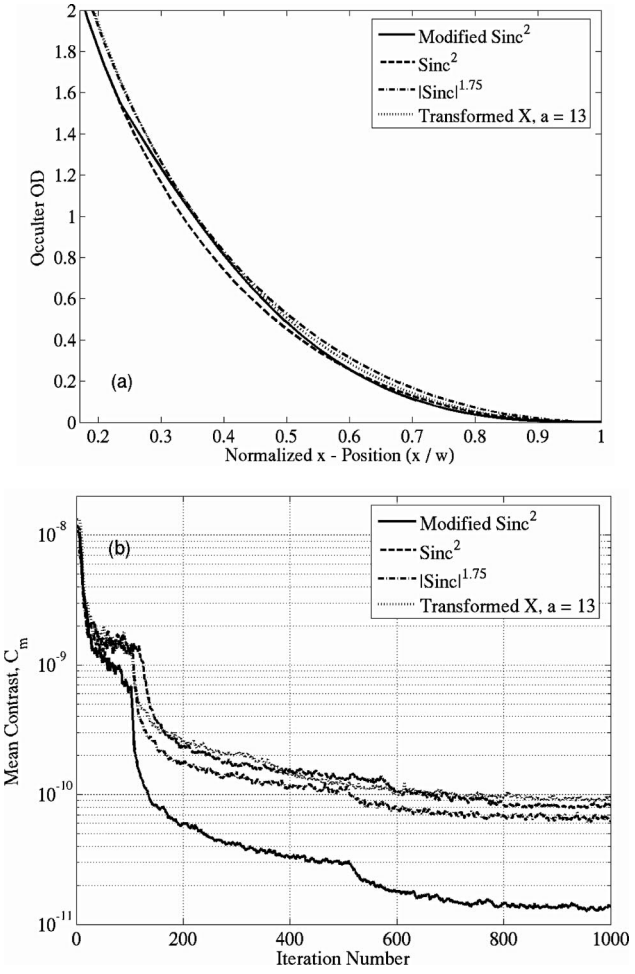

Fig. 12. (a) OD profiles of several occulting masks with w = 144 μm. The modified Sinc^2 profile is identical to the Sinc^2 everywhere except in the $\sim 0.25 < |x/w| \leq 1$ regions, where it is defined by a fifth-order polynomial whose coefficients were obtained from a fit to the measured OD profile in Fig. 8(b) in the same regions. (b) Predicted contrast performance of the HCIT versus iteration number corresponding to the four different types of occulting masks shown in (a).

$$T(x) = \begin{cases} \left[1 - \left|\dfrac{\sin(\pi x/w)}{(\pi x/w)}\right|^{1.75}\right]^2, & |x/w| \leq 1 \\ \left[1 - \left(\dfrac{\sin(\pi x/w)}{(\pi x/w)}\right)^2\right]^2, & |x/w| > 1 \end{cases}, \quad (5)$$

and the transformed X profile is given by Eq. (1) with x replaced by a transformed x, or x_t[11]:

$$x_t(x) = \begin{cases} x - \dfrac{1}{a}(X^5 - 2X^3 + X), & |X| \leq 1 \\ x, & |X| > 1 \end{cases}, \quad (6)$$

where $X = x/w$ is the same as in Eq. (4). When the parameter a is properly chosen, this transformed x results in a wider central transmittance peak compared with the standard linear-Sinc^2 profile. In Fig. 12(a), the transformed X profile was obtained with a = 13. Again, the occulter phases corresponding to all of these occulter OD profiles are obtained from our analytical model $\phi = \phi(\mathrm{OD}, \lambda)$ with λ = 785 nm.

The predicted contrast performance of the HCIT obtained with the above four different occulting masks are shown in Fig. 12(b). Here the result of the Sinc^2 occulter (dashed curve) is the same as the result of design, w = 144.0 μm occulter (dashed curve) in Fig. 10, and is again included here for comparison. As can be clearly seen from this figure, the modified Sinc^2 occulter again performs the best, confirming the earlier result of the measured occulter shown in Fig. 10. The results of the other three occulting masks are comparable, with the $|\mathrm{Sinc}|^{1.75}$ occulter being slightly better than the other two.

Figures 13(a)–13(c) show some examples of the predicted contrast and the corrective phase results obtained for the two best performing cases of the occulting mask: (1) the standard Sinc^2 without parasitic phase and (2) the modified Sinc^2 with parasitic phase. The predicted mean contrast values of these two cases have already been shown as functions of iteration number in Figs. 9 and 12(b), respectively. Figure 13(a) compares the logarithm-scale contrast maps, and Fig. 13(b) shows their slices along the $y = 0$ axis ($x > 0$ portion only). As we can see, after 1000 iterations of speckle nulling, the modified Sinc^2 occulter gives a mean contrast value that is more than twice as good as compared the phase-free Sinc^2 case. The corresponding corrective phase maps, or DM corrective surface patterns, are shown in Fig. 13(c). In this figure, the circular areas having nonzero data values correspond to the system clear aperture, and the faintly visible eye-shaped areas correspond to the Lyot stop clear aperture.

5. Discussion

In general, the speckles at the location of a planet in a coronagraph are caused by two sources: (1) quantum noise of the diffracted starlight and (2) speckle noise attributable to the scattering of the starlight by optical defects such as amplitude and phase errors of various optical components.[12] The optical setup of the HCIT, or any other coronagraph, minimizes the quantum noise of the diffracted starlight, while the speckle-nulling process implemented via a single DM or multiple DMs minimizes the speckle noise. The optical defects that scatter the starlight include, but are not limited to, the reflectivity variations and the phase errors of various mirrors, as well as the parasitic phase of the occulting mask. Our analysis does not take into account the amplitude errors caused by the reflectivity variations of various optics, but it includes the amplitude-drooping effect exhibited by the artificial starlight at the pupil plane. It also includes the phase errors of all mirrors and the parasitic phase of the occulting mask. When some coronagraph optics are positioned at nonpupil locations, as in the case of the HCIT, the propagation of light from the nonpupil optics to the pupil transforms the phase modulation caused by surface deformation into amplitude modulation, and likewise transforms amplitude modula-

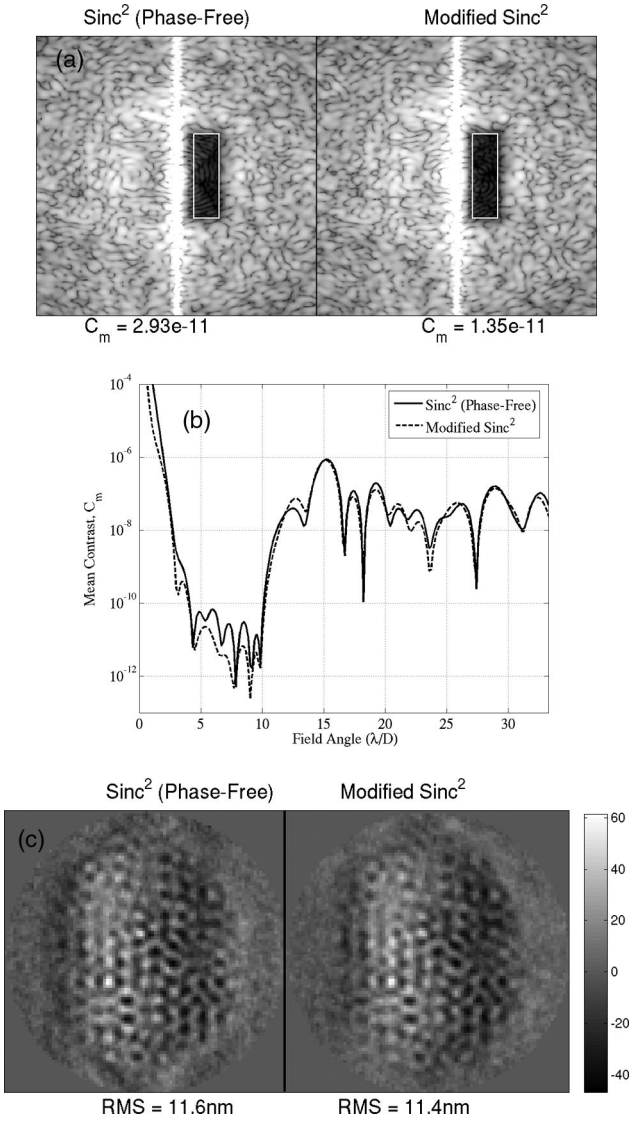

Fig. 13. Contrast and corrective phase examples corresponding to the band-limited (phase-free) Sinc^2 and imperfect band-limited, modified Sinc^2 occulters. (a) Predicted contrast maps. The white rectangular box indicates the boundary of $[3 \quad 11 \quad -11 \quad 11] * \lambda/D$ window. Shown as x label is the mean contrast C_m of the corresponding contrast map. (b) Slices of the contrast maps along the $y = 0$ axis ($x > 0$ half only). (c) Corrective phase maps (or DM corrective surface patterns) that yield the contrast results shown in (a) and (b). The unit of the color bar is nanometers. Shown as x label are the rms values of the corresponding phase map.

tion into phase modulation.[7] The wavefront phase solution of a single DM can correct both of these phase and amplitude errors simultaneously, but only over half of the image field, as shown in Fig. 13(a).

In a Lyot coronagraph such as the HCIT, a perfectly band-limited (or phase-free) occulting mask works only on the intensity of the incoming beam. However, if the occulter has parasitic phase, it not only blocks most of the incoming starlight but also modulates the phase of the unblocked beam leaving the occulter. That is, in this case the occulter works on both the intensity and the phase of the incoming beam. Therefore one can expect that in a coronagraph the effect of the occulter parasitic phase is to increase the speckle noise in the image plane, making the creation of a half dark-hole harder. The contrast results of the standard Sinc^2 (i.e., design) occulter in Figs. 9 and 10 are consistent with such a line of understanding. However, the results of the measured and the modified Sinc^2 occulters in Figs. 9, 10, and 12(b) do not support this reasoning. As we can see from Figs. 9 and 10, the measured occulter with parasitic phase does a much better job than its phase-free counterpart does. That is, the results of the design and the measured occulters contradict each other. In addition, if we compare the result of the modified Sinc^2 occulter with parasitic phase, Fig. 12(b), to the result of the phase-free standard Sinc^2 occulter in Fig. 9, we find that the former performs slightly better than the latter. What is the explanation for this?

If we look at the results of the measured occulter only, the following mechanism comes into one's mind as a possible answer: Owing to the parasitic phase of the occulter, some of the residual beam diffracted off the occulter gets pushed out of the Lyot stop clear aperture, giving rise to the reduced speckle noise in the image plane. In that case, the same thing should be true for the design occulter as well. In reality that is not the case. To understand the effect of occulter phase on image-plane speckle noise prior to speckle nulling, we calculated the total energy inside a central area of $[-10 \quad 10 \quad -10 \quad 10] * \lambda/D$ at the final focal plane, and plotted it as a function of occulter de-space for four different cases (This procedure is followed before each speckle-nulling process is started. That is, the occulter position is optimized before speckle nulling in this way). The results are shown in Fig. 14. In this figure, $\phi = 0$ or $\phi = \phi(\mathrm{OD}, \lambda)$ indicates an occulter with or without parasitic phase, respec-

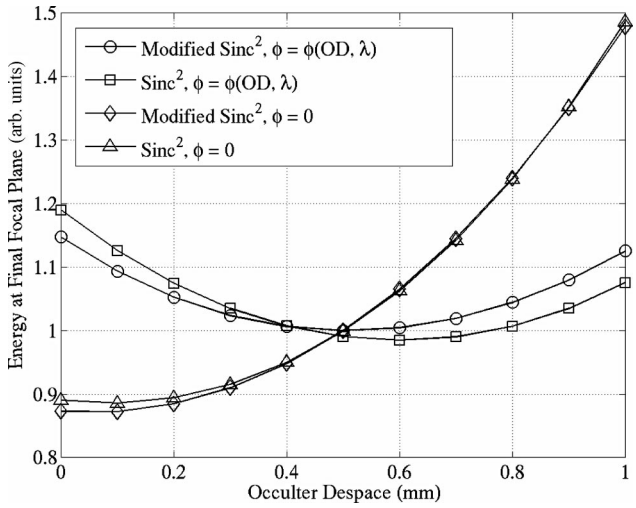

Fig. 14. Normalized total energy inside the $[-10 \quad 10 \quad -10 \quad 10] * \lambda/D$ region of the final focal plane as a function of occulter de-space. The case where $\phi = 0$ corresponds to an occulter without parasitic phase, and the case where $\phi = \phi(\mathrm{OD}, \lambda)$ to an occulter with parasitic phase.

tively. We find that the occulter phase increases the speckle noise in the final image plane in both cases, and the energy at its minimum point is the worst in the case of imperfect the modified $Sinc^2$ occulter as compared with the remaining three cases. That is, the best performing occulter actually introduces the greatest amount of speckle noise in the focal plane before speckle nulling. This leaves us with only one possible explanation: The combination of the modified $Sinc^2$ occulter's OD and phase profiles makes speckle nulling more effective as compared with the other three cases. To the best of our knowledge, all previous theoretical studies on space-based high-contrast Lyot coronagraphs have assumed phase-free, or constant-phase occulting masks. However, the fabrication of such an occulter is extremely challenging and has not yet been demonstrated. The findings of this paper indicate that by optimizing the OD profile of a HEBS glass occulter with parasitic phase, one can achieve contrast performance on the HCIT better than that obtainable by using a perfect (or phase-free) $Sinc^2$ occulter.

6. Conclusion

In this paper we examined in detail the behavior of several grayscale band-limited focal plane occulting masks. Our goal was to understand how an imperfect occulting mask with parasitic phase error performs in a high-contrast imaging system such as JPL's HCIT. We presented modeling and simulation results on the monochromatic contrast performance of the HCIT when utilizing occulting masks with standard and modified $Sinc^2$ OD profiles as well as several of their variations. Our analysis showed that the parasitic phase error of an occulting mask tends to increase the speckle noise in the final image plane. When the OD profile of a grayscale occulter having parasitic phase error is optimized it can perform better than its phase-free counterpart, and even better than a perfect (phase-free) $Sinc^2$ occulting mask. In another words, we showed that it is possible to eliminate the detrimental effect of an occulting mask's parasitic phase by optimizing its OD profile. The combination of the OD and the phase profiles of a modified $Sinc^2$ occulter proposed in this paper makes a speckle-nulling optimization process more effective and yields better contrast performance on the HCIT compared with its phase-free counterpart. We have also found that the damped sinusoidal ripples in a fabricated occulter's amplitude transmission coefficient do not have a significant detrimental effect on the HCIT's performance, and the HCIT is not sensitive to small fabrication errors (several micrometers) in occulter width. These findings are important when choosing occulting mask technologies and for adequately determining the fabrication tolerance of grayscale occulting masks used in a high-contrast imaging system such as NASA's Terrestrial Planet Finder coronagraph.

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. The authors thank Stuart Shaklan for his useful suggestions.

References

1. J. T. Trauger, C. Burrows, B. Gordon, J. J. Green, A. E. Lowman, D. Moody, A. F. Niessner, F. Shi, and D. Wilson, "Coronagraph contrast demonstrations with the high-contrast imaging testbed," in *Optical, Infrared, and Millimeter Space Telescopes,* J. C. Mather, ed., Proc. SPIE **5487,** 1330–1336 (2004).
2. A. E. Lowman, J. T. Trauger, B. Gordon, J. J. Green, D. Moody, A. F. Niessner, and E. Shi, "High-contrast imaging testbed for the Terrestrial Planet Finder coronagraph," in *Optical, Infrared, and Millimeter Space Telescopes,* J. C. Mather ed., Proc. SPIE **5487,** 1246–1254 (2004).
3. M. Kuchner and W. Traub, "A coronagraph with a band-limited mask for finding terrestrial planets," Astrophys. J. **570,** 900–908 (2002).
4. D. W. Wilson, P. D. Maker, J. T. Trauger, and T. B. Hull, "Eclipse apodization: realization of occulting spots and Lyot masks," in *High-Contrast Imaging for Exo-Planet Detection,* A. B. Schultz and R. G. Lyon, eds., Proc. SPIE **4860,** 361–370 (2003).
5. P. G. Halverson, M. Z. Ftaclas, K. Balasubramanian, D. J. Hoppe, and D. W. Wilson, "Measurement of wavefront phase delay and optical density in apodized coronagraphic mask materials," in *Techniques and Instrument for Detection of Exoplanets II,* D. R. Coulter, ed., Proc. SPIE **5905,** 590511 (2005).
6. E. F. Shi, S. Basinger, D. Moody, A. E. Lowman, A. C. Kuhnert, and J. T. Trauger, "Performance of TPF's High-Contrast Imaging Testbed: Modeling and simulations," in *Space Telescopes and Instrumentation I: Optical, Infrared, and Millimeter,* J. C. Marther, H. A. MacEwen, and M. W. M. de Graauw, eds., Proc. SPIE **6265,** 62653L (2006).
7. S. B. Shaklan and J. J. Green, "Reflectivity and optical surface height requirements in a broadband coronagraph. 1. Contrast floor due to controllable spatial frequencies," *Appl. Opt.* **45,** 5143–5153 (2006).
8. E. Sidick, A. C. Kuhnert, and J. T. Trauger, "Broadband performance of TPF's High-Contrast Imaging Testbed: Modeling and simulations," in *Advanced Wavefront Control: Methods, Devices, and Applications IV,* Proc. SPIE **6306,** 6306–6332 (2006).
9. *Modeling and Analysis for Controlled Optical Systems User's Manual,* Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California.
10. J. J. Green and S. B. Shaklan, "Optimizing coronagraph designs to minimize their contrast sensitivity to low-order optical aberrations," in *Techniques and Instrument for Detection of Exoplanets,* D. R. Coulter, ed., Proc. SPIE **5170,** 25–37, 2003.
11. B. D. Kern, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109 (personal communication, 2006).
12. P. J. Borde and W. A. Traub, "High-contrast imaging from space: Speckle nulling in a low-aberration regime," Astrophys. J. **638,** 488–498 (2006).