2012 TDEM: DEMONSTRATION OF STARSHADE STARLIGHT-SUPPRESSION PERFORMANCE IN THE FIELD. FINAL REPORT

9/10/2015

TIFFANY GLASSMAN (PI), MEGAN NOVICKI, MICHAEL RICHARDS, DANNY SMITH, STEVE WARWICK (NGAS)
KEITH PATTERSON (JPL)
ANTHONY HARNESS (CU)

Document #: 1469885

2012 TDEM: DEMONSTRATION OF STARSHADE STARLIGHT-SUPPRESSION PERFORMANCE IN THE FIELD. FINAL REPORT

9/10/2015
TIFFANY GLASSMAN (PI), MEGAN NOVICKI, MICHAEL RICHARDS, DANNY SMITH, STEVE WARWICK (NGAS)
KEITH PATTERSON (JPL)
ANTHONY HARNESS (CU)

E-SIGNED by Nick Siegler
on 2015-10-28 20:20:46 GMT

Approved _____   _____

Nick   Siegler   Date
Exoplanet Exploration Program Chief Technologist, JPL

E-SIGNED by Douglas Hudgins
on 2015-10-29 11:31:22 GMT

Approved _____   _____

Douglas   Hudgins   Date
Exoplanet Exploration Program Scientist, NASA HQ

CONTENTS

1 EXECUTIVE SUMMARY

The 2012 Starshade Field Testing TDEM Tests the optical properties of Starshades over a long baseline outdoor range to address two scientific milestones:

Milestone #1: Demonstrate, using a Starshade, contrast better than 10^{-9}, at all radii past the Starshade tips, in 50% bandwidth light.

Milestone #2: Demonstrate agreement between the measured and predicted contrast resulting from a range of Starshade shapes.

The outdoor range consists of a simulated star light source, a Starshade mounted on an automated center stand, and a commercial telescope mounted with an astronomical camera to observe the Starshade and source.
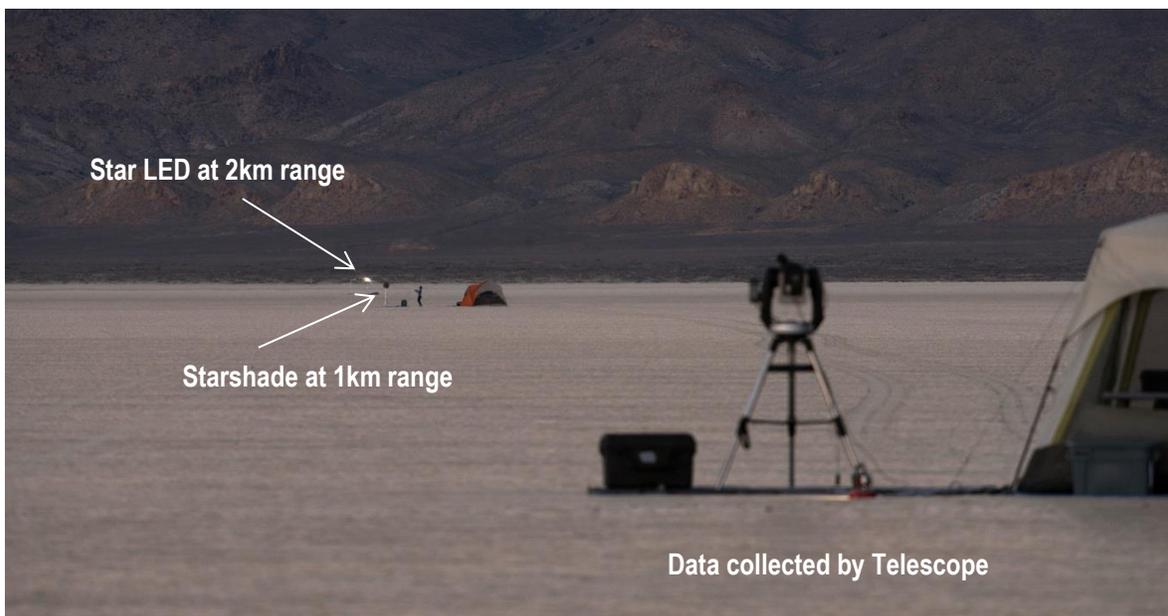
Figure 1-1 Field Test Setup. An outdoor range is used to allow extremely long baselines and give optical arrangements closer to in flight optics. The light source is 2km from the telescope; the Starshade is 1km from the telescope.

For most tests 58cm tip to tip Starshades were used with two basic shapes, one with a Hypergaussian edge shape and the other with a Numerically Determined (IZ5) edge shape (Section 2.3). Variations on these shapes were also used, particularly to address milestone #2 (Section 5).

Tests are conducted over five nights, closest to a new moon for minimum stray light. Measurements are taken with the simulated star blocked by the Starshade and with the Starshade out of alignment (unblocked) to compare and assess the level of blockage achieved by the Starshade.

The optics tested here are not flight like. The use of the scaled down Starshade at these distances gives a Fresnel number of ~250, compared to a flight like number of ~16, in addition, the resolution of the telescope used results in 13 resolution elements across the Starshade, compared with 2-4 for a flight like

3

number. These effects combine to give contrast ratios significantly better than would be achieved with flight like values. Optical models used to predict the contrast performance of the Starshade in flight like conditions can be validated against actual performances at this different Fresnel number and resolution element, which is the essence of Milestone #2.

Contrast ratio was selected as the figure of merit for a Starshade as that correctly reflects the ability to detect light from a dim source at an angular separation from a bright source. Contrast ratio is defined as the ratio of signal brightness at a location offset from the star to the peak brightness of the unblocked star.

1.1 MILESTONE COMPLIANCE

During the course of the TDEM three field tests were performed, each scheduled for 5 nights of observation. These tests are referred to at the May 2014 test, the September 2014 test and the April 2015 test in this document.

During the April 2015 test, contrast ratios of better than 10^{-9} were measured for both Hypergaussian and Numerically Determined Starshade shapes (Section 4.2).
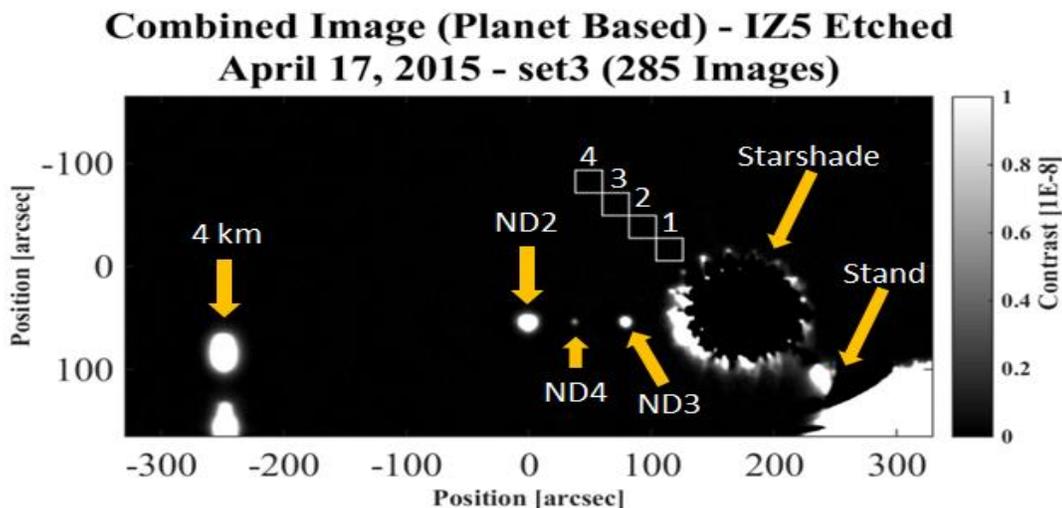
Figure 1-2 Milestone #1 Best Contrast. By combining over 47 minutes of images we are able to demonstrate a contrast ratio of 7.3×10^{-10} at the position of the box labeled 1. The simulated planet labelled ND4 is 6×10^{-9} as bright as the main simulated Star. The 4 km light sources are used to track atmospheric conditions as well as indicate the feasibility of a 4 km baseline test for possible future work (see section 7).

During the September 2014 test, flawed Starshades were measured and compared to predicted results generated using Northrop Grumman, JPL and University of Colorado models. In general, model predictions were within a factor of 5 of measured values. This level of disagreement between model predictions and measurements was expected, given the uncertainties in the measured results (see section 4.1). However significant differences between the different model predictions were seen, likely due to the modifications to the models to handle the spherical wavefront due to the light source at 1km rather than infinity. The differences between models are higher than expected based on previous model comparisons (see section 5.1 and reference [5]). Work is ongoing to resolve these differences (see section 7).
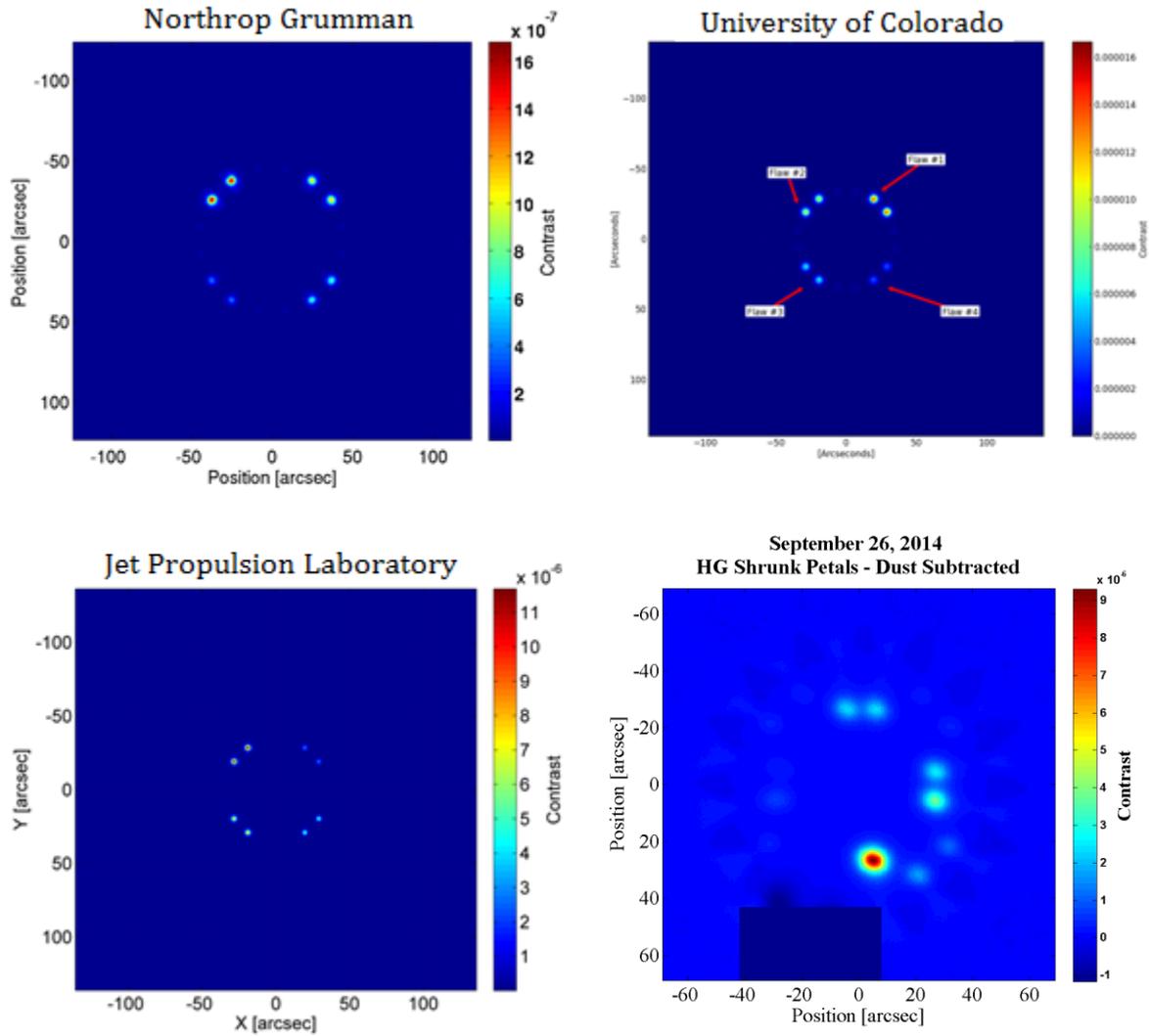
Figure 1-3 Milestone #2 Model Verification. Three independent models predicted the impact of intentional flaws on the optical performance of the Starshade. Shrunk Petal model results (top, bottom-left) and actual measurement (bottom-right) are shown here. Qualitative similarities between the contrast features in the model and the measurement are clear. Quantitatively they agree to within an order of magnitude (see Section 5.4.2). The bright peak and dark mask at the bottom of the measured image are due to the Starshade stand and should be ignored.

2 KEY TDEM DETAILS

2.1 TEST DESCRIPTION

Our field tests were conducted during new moon periods at Smith Creek dry lake bed in northern Nevada (Figure 1-1). Details of the tests include:

- Telescope-to-Starshade and Starshade-to-light source distances of 1 km each.
- Telescope: 20 cm Celestron.
- A cover for the Celestron was used with an aperture cut into the "upper" half of the cover. This allowed a fully unobscured aperture to be used, avoiding the standard secondary and support structure. The off-axis optical system should have minimal impact on the PSF as it is not near the edge of the aperture and PSF fitting is used to estimate the brightness of all sources. The unobscured aperture acts to minimize internal scatter in the telescope in the absence of a Lyot mask in the camera and greatly simplifies the modeling of scattered light in the system, as well as provides for the required small aperture to approach the required Fresnel number for the system.
- SBIG CCD camera with 3326x2504 pixels, each covering 0.55"x0.55".
- Filter wheel fitted with broadband color filters B, V, and R (Section 3.2), as well as a series of neutral density (ND) filters which reduced the signal across all wavelengths by factors of approximately 100 (ND2), 1000 (ND3), and 10000 (ND4).
- Motorized stand at the Starshade station (1 km from the telescope) to allow the Starshade to be easily and precisely moved in and out of the line of sight. The stand also allowed for small, accurate adjustments to get the best possible source-Starshade-telescope alignment.
- Main simulated 'star' light source: 1W LED (100Lumen) with focus lens located 2 km from the telescope (1 km from the Starshade station). The primary light source was a 1W LED focused using an 85mm lens into a beam ~2.5° wide, giving an apparent far-field brightness approximately 3600 times brighter than the 1W Without the lens. We use an expanding beam instead of a collimated source (like starlight) to reduce wavefront errors and simplify the optics. (The collimated source would need to be many times larger than the Starshade).
- At the 2.5° beam width, the virtual position of
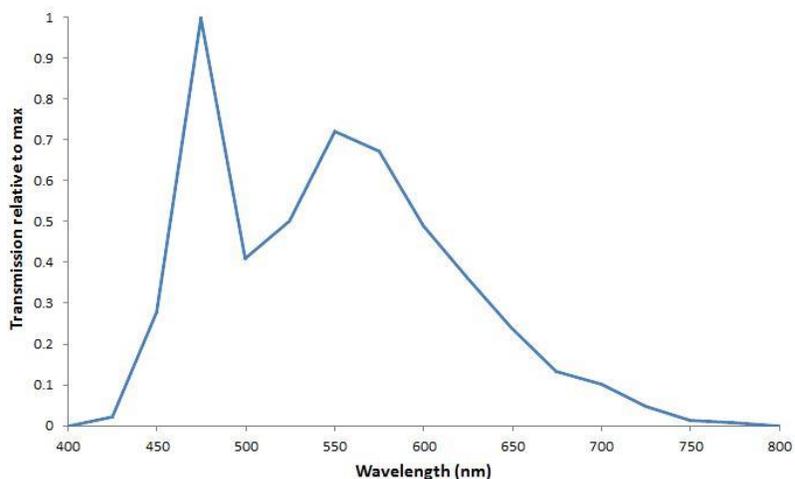
Figure 2-1: Spectral Weighting of Starshade Images from the Field Tests. Starshades operate over a wide bandwidth. For this test, the bandwidth is limited by the light source and optics used. The curve above shows the relative signal strength of the main light source across the visible band. The telescope transmission and detector quantum efficiency of our setup are included in this calculation.

the expanding beam source is approximately 1m further distant than the actual position of the source, a difference that is negligible over 1km.

- In addition to the LED transmission, the transmission curve in Figure 2-1 takes into account the transmission of our telescope and the quantum efficiency of our detector.
- The transmission curve for the system equates to a system bandwidth of over 55%
- Simulated 'planet' light sources: four 0.07W (4.5Lumen) LEDs (no focus lens) with neutral density filters (ND1-ND4) offset from the main light source in the image plane. The visibility of these sources when using the Starshade provides benchmarks for contrast performance and reference points for image stacking. The lower power of the simulated planet LEDs and the signal attenuation of the neutral density filters, along with increase in the power of the main light source due to the focus lens, places the simulated planets in the 10^{-6} to 10^{-9} contrast range.
- Two 0.07W (4.5Lumen) atmospheric monitoring LEDs placed 4 km from the telescope, well separated from the test baseline in the image plane were added to the setup for tests 2 and 3. These sources were monitored throughout our data-taking process in order to evaluate the atmospheric stability at a longer range for potential future tests.
- Our field test setup operates at a Fresnel number approximately 16x the optimal flight scale Fresnel number.

2.2 ENVIRONMENTAL EFFECTS SPECIFIC TO THE GROUND TESTS AND MITIGATION STRATEGIES

Our current site at Smith Creek dry lake bed in Northern Nevada is an improvement over the previous sites (El Mirage near Antelope Valley, CA, and Silurian Lake bed Near Death Valley, CA)[1]. It permits a range of >4km and is at an altitude of ~6000ft. The dust level and atmospheric conditions vary widely from good to poor depending on weather conditions, particularly daytime wind. We use the Full Width at Half- Maximum (FWHM) of the point spread function (PSF) of the brightest planet or unblocked source in the image to characterize our seeing. The point spread function is determined by fitting a 2-dimensional Gaussian distribution to the source. In Figure 2-2, we show the seeing conditions during one night of our test in September of 2014.

There are other factors that affect the observed contrast besides the shape of the Starshade. A critical aspect of our tests was to mitigate the factors that affect our test data.
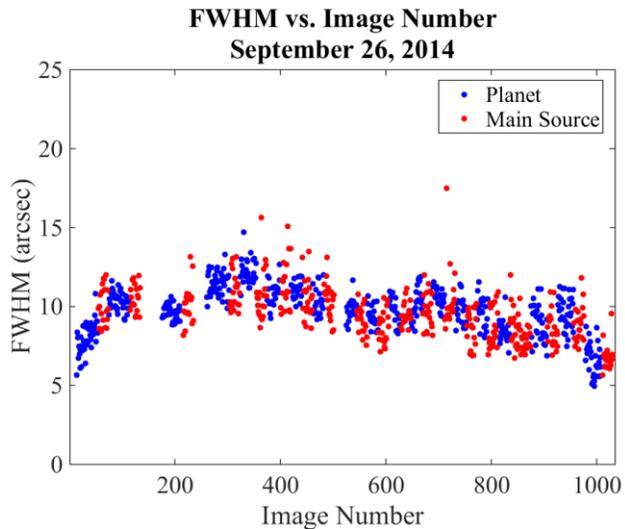
Figure 2-2. Typical Seeing Conditions at Smith Creek Dry Lake Bed. The seeing, as measured by the FWHM of the sources in the images, was typically between 6 and 18 arcseconds over our three tests. Shown here is the seeing for one of the nights during our second test.

The primary effects of the atmosphere on the observations come from forward scatter from dust, which produced a visible halo of light surrounding the Starshade; blurred images/variable flux of light sources from seeing; and refraction of the simulated star light source by variable temperature atmospheric cells, which led to a move in apparent position of the main source and planet LEDs. In addition, the brightness of the unblocked source varied by up to a factor of four during a single night with no change in the test setup. In our tests we worked to understand and control these factors enough that we are getting a useful validation of the numerical simulations and a prediction of the on-orbit performance.

Seeing: In our tests, images that are blurred such that source light leaks out from behind the Starshade are not used in analysis. Thus, blurring of the images due to seeing has little effect on the performance of the Starshade. Provided that the angular diameter of the blurred source remains smaller than the effective diameter of the Starshade, the amount of residual light entering the telescope is not changed. This drives the size of the Starshade to be tested at a known range (See section 4.2.3). The atmospheric distortion also causes variability in the flux of the simulated star and planets. This is controlled for by taking many blocked and unblocked images, allowing us to get good statistical uncertainty determinations on all measurements. In our results section, blocked and unblocked sets of images are paired so that the comparison is over a short period of time, usually less than 30 minutes. This best controls for long timeline variations.

Refraction: In our testing we have observed nights where the apparent vertical position of the simulated star has moved by up to 1 meter over 30 seconds to 1 minute. We believe that this apparent motion is caused by layers of warmer or colder air moving over the lake bed causing refraction. When these layers are passing through rapidly, the adjustments of the height of the Starshade needed to keep the system in alignment cannot be made in time to take useful images. This concern is mitigated by co-adding short exposures and throwing out misaligned images.

Dust: As part of these tests, we developed dust subtraction image processing techniques detailed in the Appendix. This subtraction technique has been refined to be effective on data with significant dust levels. Lowering the dust in the environment would minimize or eliminate the need for this step in image processing and would improve the accuracy of our measurements. While the amount of dust for ground tests can be reduced or even eliminated with the choice of test environment, the current reliance on dust subtraction is parallel to what will be needed to remove exo-zodiacal light in the flight system.

We examined the dust levels in the background of our images for each of our tests and found that the dust
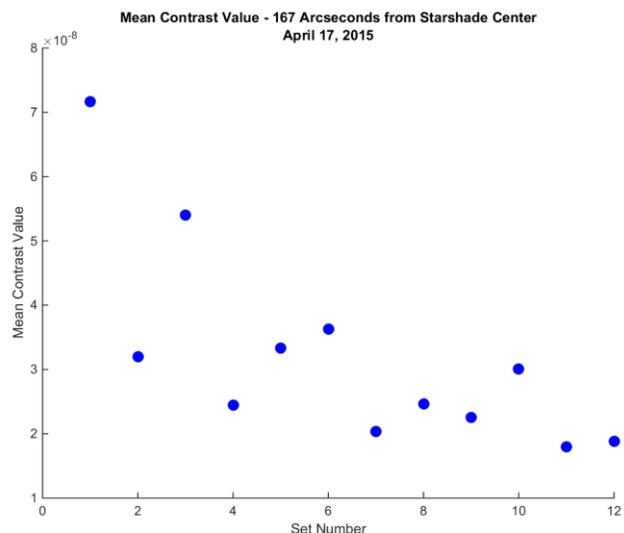
Figure 2-3. Dust Conditions at Smith Creek Dry Lake Bed. Shown here are the mean background levels 167 arcseconds away from the Starshade for observations taken on April 17, 2015. Each set number represents a group of images taken together in time; increasing set number corresponds to increasing time. The dust levels progressively decreased throughout this particular night.

levels can vary significantly during a run and even sometimes during a single night of observations. Figure 2-3 shows the dust levels for each 'set' of data (a data 'set' represents a group of images taken close together in time) taken on April 17, 2015. The dust for this night became progressively less as the night continued (larger set number indicates later in time), but this was not necessarily the case for all of the nights during this observing run. As the variation of the dust could be unpredictable, removing it from the images was necessary, particularly during nights when we wanted to combine data taken over an extended period of time.

2.3 STARSHADE SHAPES MODELED AND BUILT

The Starshades that we designed, built, simulated, and tested are based on two base designs. Both of these are designed to be close scale models to the full-scale Starshade designs. The particular shapes were chosen to get the best possible contrast performance in the field test conditions. These are the "base" Starshade designs that all other Starshade shapes will be based on.

The first base design is a Hypergaussian (HG) shape, where the petal width is given by a parametric equation:

$$W = \frac{2\pi}{P} \exp\left(-\left(\frac{r-a}{b}\right)^n\right)$$

Here, W is the width of the petals; P is the number of petals; and a, b, and n are parameters which govern the size and shape of the Starshade. For the design used here, P=16, a=12 cm, b=12 cm, and n=6. Since the tips and valleys for the HG design become very narrow in the ideal case, and we know the manufacturing technique limits them to about 100-200 μm width, we truncate the tips and valleys for all the HG Starshade simulations so that the petals extend between 16.91 cm and 28.11 cm from the center.

The second base design is a Numerically Determined design where the width of the petals is derived numerically and is given by a set of points along the edge[2]. The shape of these Starshades is derived from a numerical optimization technique based on specified input criteria. Figure 2-4 shows the petal width function for the particular Numerically Determined design used here, called the IZ5 design, developed at JPL. We performed very limited testing of another Numerically Determined Starshade developed at JPL, called the HS25. We used this design when testing Starshade performance dependence on wavelength (see Section 3.2).

For the tests conducted during this TDEM, we tested the two base designs fabricated using two different materials and methods. Carbon fiber Starshades made of 9 pieces (one central circle with half the petals attached, and then the other 8 petals
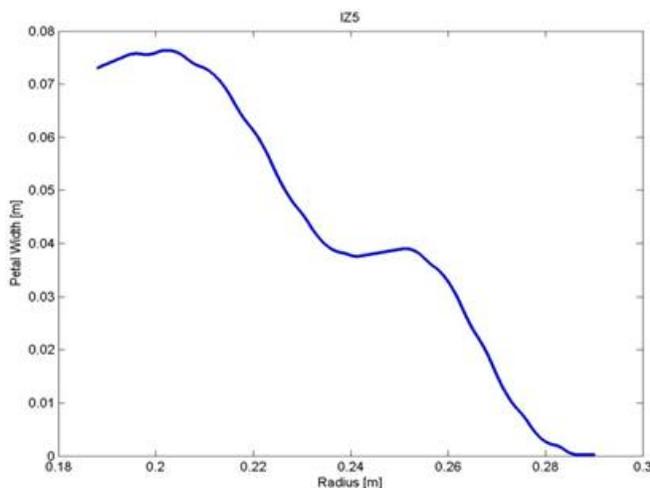
Figure 2-4: The petal width versus radius for the IZ5 design used in our tests. Numerically Determined petal shapes tailor the edge and allow for shorter petals.

inserted in between) allow for extremely narrow valleys[1], which is important for the Hypergaussian shape. This construction is expensive and any fiber wisps on the edge of the Starshade scatter light and impact the effectiveness of the shape, so for mass manufacturing we used a chemical etching process to cut the Starshades out of 10mil steel sheets. The etching process precision is limited by the thickness of the sheets, and therefore leads to valleys limited to 100 μm and tips limited to 200 μm, not quite as sharp as needed to properly approximate the flight-like design.

2.4 CONTRAST RATIO

Contrast ratio was selected as the figure of merit for a Starshade as that correctly reflects the ability to detect light from a dim source at an angular separation from a bright source. Contrast ratio is defined as the ratio of signal brightness at a location offset from the star to the peak brightness of the unblocked star. Suppression is defined as the ratio of the total starlight that enters the telescope with the Starshade in place to that without the Starshade. Suppression is primarily affected by the Starshade performance. Contrast ratio, on the other hand, gets better (decreases) with telescope resolution – as the Point Spread Function (PSF) of the star gets narrower and the peak higher, the contrast relative to the peak of the star gets better. In the flight configuration, where the telescope resolution element is about half the radius of the Starshade in the image plane, the telescope resolution contributes a factor of 100 to the contrast ratio (relative to the total suppression) at locations just beyond the edge of the Starshade. For the field test configuration, we are currently over-resolving the Starshade by a factor of about 8. The telescope resolution element is about 1/16 the radius of the Starshade in the image plane. In this case, the telescope resolution contributes a factor of $\sim 2 \times 10^6$ to the contrast ratio at the radius of the Starshade tips compared with a flight like telescope contribution of $\sim 1 \times 10^2$.

For the field tests, uncontrollable effects such as dust and background scattered light introduce errors that would make it very difficult to accurately calculate the total suppression of the measurement. While the effects of these error sources are not particularly large, they swamp the $\sim 10^{-8}$ effect that we are trying to measure. In other words, these effects would need to be modeled to extremely high accuracy to allow accurate subtraction of background light and therefore allow accurate full suppression measurements. Therefore we made the decision to use contrast as the best figure of merit for these tests.

Starshades with higher Fresnel number, such as in our test setup, will also produce better contrast. For the Starshade flaws, operating at a different Fresnel number will change the predicted brightness caused by those flaws. However, since the test and model can be matched at the same Fresnel number, this has no effect on the validity of the test. The wording of Milestone #1 was selected knowing that the field test could not accurately represent the flight configuration in this respect and is intended to apply to the configuration as tested (2km total range length, 58cm tip to tip Starshade, and 4cm telescope aperture). Therefore the 10^{-9} contrast goal in Milestone #1 is explicitly intended to be achieved in the configuration tested, including the extra benefit from the inflated Fresnel number and telescope resolution. In the future we plan to test and model the system at a range of Fresnel numbers (see Section 7) to ensure that the scaling of performance vs Fresnel number matches the predictions. Future tests that operate at the same Fresnel number as a flight system can prove out the specific performance predictions for flight.

2.5 TDEM TIMELINE

A site survey of possible lakebeds was conducted in April 2014. Possible locations had been identified in advance from satellite images. Smith Creek Dry lake bed in central Nevada was chosen as the best combination of altitude, darkness, least dust, and accessibility for test support vehicles.

The first test run was conducted in May 2014 over 5 nights shortly after the start of the TDEM contract, and was used primarily as an engineering run to confirm that the equipment designed for the test would operate as intended, and that the lakebed was a suitable location over 5 nights.

As part of this test run, we carried out initial contrast ratio measurements, and looked at the effects of tilt, spin and filter color on the performance of HG and IZ5 Starshades. These measurements are covered in more detail in Section 3.

The second test run was conducted in September 2014 over 5 nights, though two nights were lost to wind and rain at the test site. This test was used primarily to look at the impact of deliberate flaws on the Starshade shape. These images are compared to the predictions of performance made by JPL, University of Colorado (CU), and Northrop Grumman (NG) optical models in Section 5.4.

The third test run was conducted in April 2015 over 5 nights, and was dedicated to getting the best (smallest) contrast ratio possible to address Milestone #1.

3 ENGINEERING TEST RUN (MAY 2014 TEST)

Our May 2014 Desert Test (Table 3-1), the first test under contract, was conducted to verify the functionality of the test setup and environment, to develop a post-processing methodology, and to test the sensitivity of the Starshade's performance to variations such as tilt, spin, and wavelength.

3.1 OBSERVATION AND POST-PROCESSING STRATEGIES

In the May 2014 test we were able to develop effective data gathering and processing strategies to apply to subsequent tests.

The general observation strategy was to obtain a series of images of each Starshade along with accompanying unblocked images used to compute contrast. For each set of images the exposure time was adjusted to be the longest exposure allowed without saturating the brightest point in the image. A series of 20-60 images were taken in sequence with the Starshade in the line of sight and then the Starshade was moved out and the unblocked main light source

Test Duration	May 28th – June 1st 2014
Nights of Observation	5
Characteristic Seeing	6-35 arcseconds
Refraction	Worst at the end of night
Dust Levels	Moderate but highly variable
Other Factors	Significant insect activity visible in many of the images. Shutter malfunctions on approximately 5-15% of the frames, rendering them useless.

Table 3-1: May 2014 desert test characteristics

was imaged in a series of approximately twenty 0.1 second exposures through an ND4 filter. Typical exposure times for the blocked images (those with the Starshade in the line of sight) were 1-10 seconds. Our experimental data processing and contrast calculations were done as follows:

- Calibration images are taken and combined. Bias frames and flat-field images are averaged while the median pixel values of each dark set were taken. Each averaged flat-field image is normalized by its mean.
- Each image is bias-subtracted, dark-subtracted, flat-scaled, and converted to counts (filter transmittance and exposure time divided out).
- A series of images taken for a given Starshade shape and position is then co-aligned based on planet peak locations or the Starshade (depending on the final use of the data) and combined via a trimmed mean after rejecting obviously misaligned images (Refraction Error).
- Images with a FWHM of the reference planet more than 2σ from the mean FWHM of the series are also rejected (Seeing Error).
- Each individual image of the unblocked source is fit with a 2-d Gaussian distribution to estimate the peak of the source.
- The co-aligned Starshade image is divided by the median value of the peak of the unblocked source to convert to contrast units. Unblocked images are taken close in time to the corresponding blocked images to ensure that the measurements are taken in similar atmospheric conditions.

Unless stated otherwise this is the general methodology used in all of the tests mentioned in this report.

3.2 STARSHADE SENSITIVITY TO WAVELENGTH

We were able to image each Starshade using V, B, and R band filters as an investigation into the feasibility of a future field test to characterize the dependence of Starshade performance on source color. The results presented here should be considered preliminary. They are meant to illustrate the feasibility of a future test and to begin the development of a strategy to measure color dependence of different Starshade shapes effectively. In a future test more time could be allotted toward this goal, resulting in a much larger data set with more uniform conditions.

Figure 3-1. The total throughput (including telescope transmittance and detector QE) of our test setup for the R, V, and B color filters.

Figure 3-1 shows the throughput of our optical setup for each filter. The curves account for telescope transmittance and detector quantum efficiency as well as filter transmittance. Since the main light LED is not perfectly white, these throughputs must be weighted by the relative LED signal strength at different wavelengths in order to apply to our data. Taking this into consideration, it

was found that the R filter transmits 30% of the total signal, The V filter transmits 37% of the total signal, and the B filter transmits 19% of the total signal.

We chose the brightness of the tips and valleys of the Starshades (after filter transmittance has been divided out) as the relevant quantity to measure the color dependence of Starshade performance. For each combined color image (Figures 3-2 through 3-4), three annuli – one contains the valley features, one contains the tip features, and one just beyond the Starshade tips – are defined. The mean of each annulus is reported relative to the brightness of an open filter image in order to account for potential atmospheric differences between data sets. A mean background level is subtracted before the ratios are computed.

Since we use the brightness of the tips and valleys as part of our Starshade performance metric, the quality of our color data is extremely sensitive to misalignment and dust levels. This sensitivity can be mitigated in future tests via our post-processing dust subtraction technique (Section 8.1), by developing a better alignment method in the field, and by creating a post-processor that searches for equal-brightness valleys in images, thus guaranteeing alignment.

B Filter

Starshade	Valleys	Tips	Just Outside
IZ5	0.57	1.18	1.21
HS25	0.77	1.03	1.05
HG	1.63	1.97	1.99

V Filter

Starshade	Valleys	Tips	Just Outside
IZ5	1.07	1.38	1.41
HS25	0.80	0.80	0.83
HG	1.32	1.70	1.72

R Filter

Starshade	Valleys	Tips	Just Outside
IZ5	0.72	0.82	0.83
HS25	0.87	0.72	0.70
HG	0.73	0.72	0.71

Table 3-2. Shown in the tables above are comparisons of contrast calculated from images with B, V, and R filters (Figure 2-1) to contrast calculated with an Open filter. Each image was divided by the transmittance of its respective filter. This work was carried out as part of the engineering test run (April 2014) and as such was not generated or analyzed with the same rigor as the main test goals. Color response is recommended as one area for future follow on work.
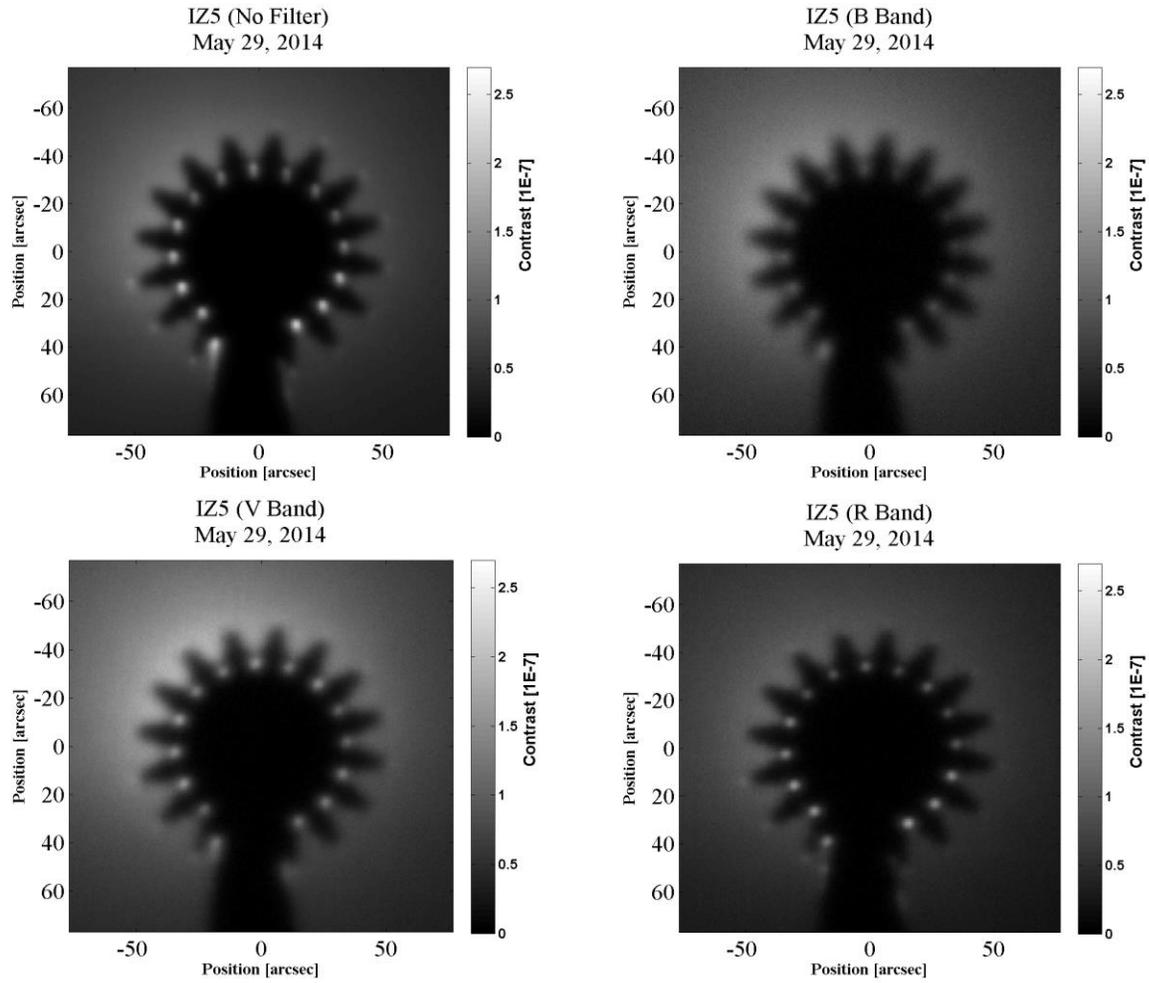
Figure 3-2. IZ5 Starshade in Different Filters. Shown in the figure are images taken using the IZ5 Starshade using different filters. Images are: top left, open filter; top right, B-band filter; bottom left, V-band filter; bottom right, R-band filter. The dust and atmospheric conditions were not constant throughout the series. The valleys appear to be brighter relative to the tips as the wavelength of the band increases.
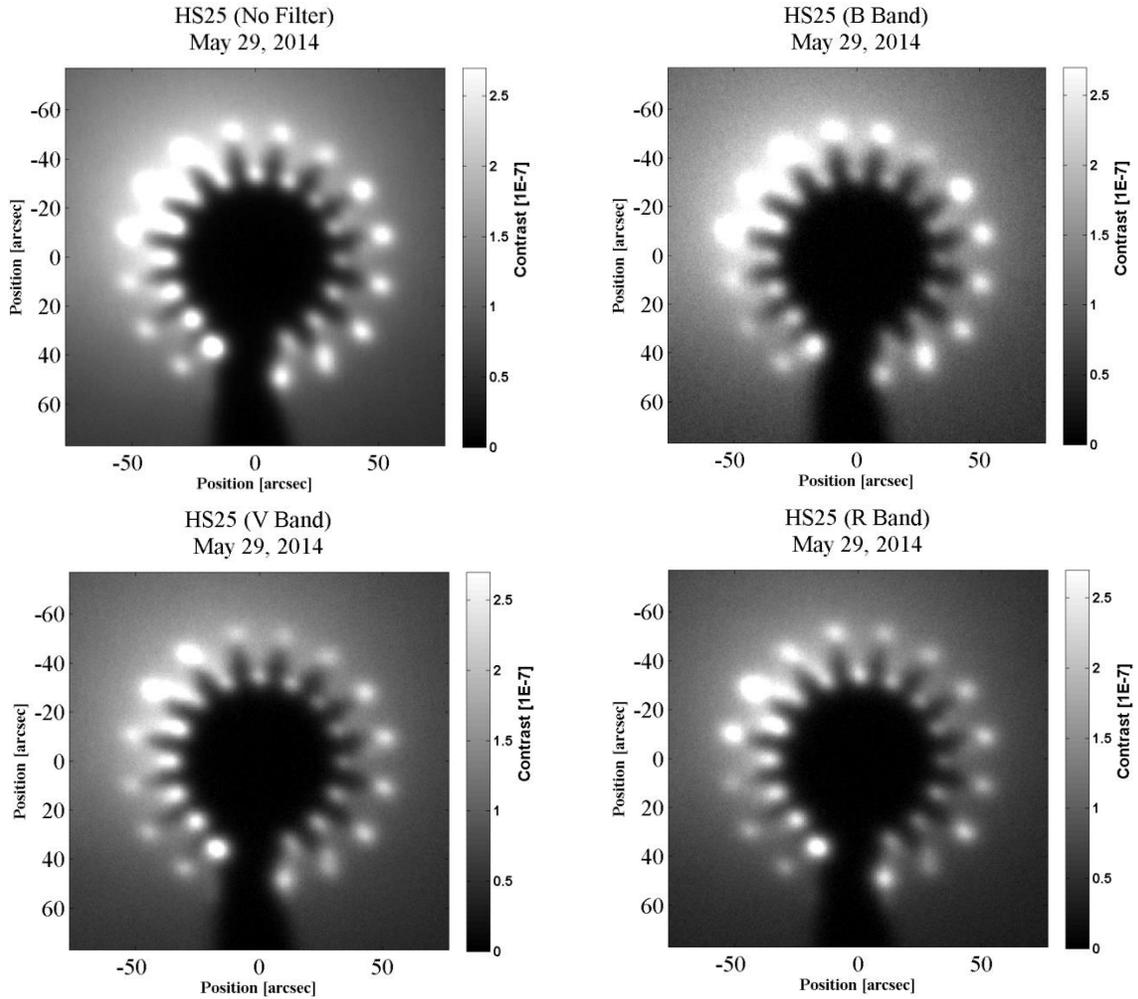
Figure 3-3. HS25 Starshade in Different Filters. Shown in the figure are images taken using the HS25 Starshade using different filters. Images are: top left, open filter; top right, B-band filter; bottom left, V-band filter; bottom right, R-band filter. The dust and atmospheric conditions were not constant throughout the series.
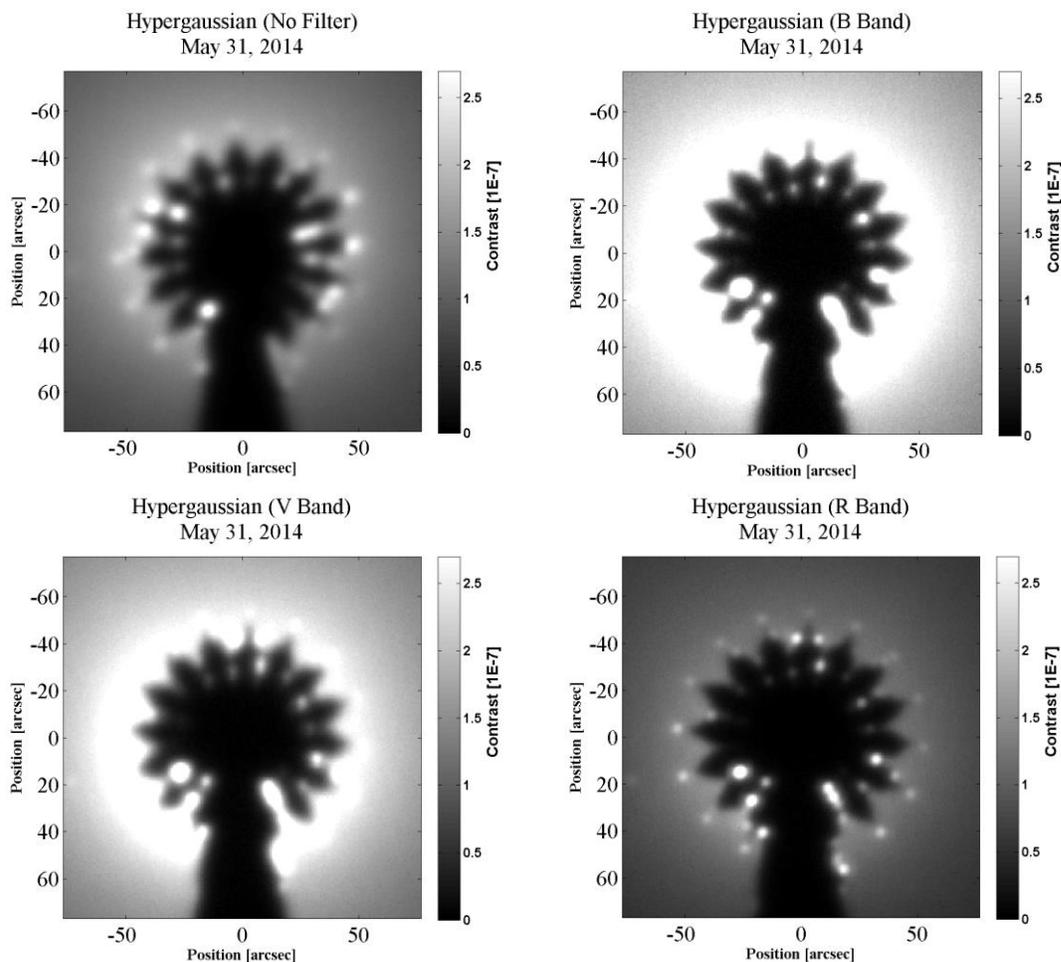
Figure 3-4: HG Carbon-Fiber Starshade in Different Filters. Shown in the figure are images taken using the HG Carbon-Fiber Starshade using different filters. Images are: top left, open filter; top right, B-band filter; bottom left, V-band filter; bottom right, R-band filter. The dust and atmospheric conditions were not constant throughout the series, but seem to be the dominant contributor to the differences between the images.

3.3 STARSHADE SENSITIVITY TO TILT AND SPIN

As the ability to tilt the Starshade with respect to the light path adds significant flexibility to the Starshade concept of operations, we chose to investigate the sensitivity of Starshades to that tilt using our field setup.

We measured the contrast performance of our Starshades at four different tilt angles (0, 5, 10, and 28 degrees with respect to the baseline) (Figure 3-5) in order to quantify the degree to which Starshade performance depends on tilt.

Our measurements indicate that tilting the Starshade with respect to the light path has no significant effect on Starshade performance. After analysis of co-added images at various tilt angles (0, 10, and 28 degrees, Figure 3-6), we could not detect any dependence of contrast performance on tilt angle (Table 3-3), and thus conclude that Starshade tilt should not be of particular concern when considering the feasibility of a flight-scale mission. Our 3σ detection limit (Section 4) instead is correlated with the value

of the mean background, indicating that dust levels are more impactful to Starshade performance than tilt. Whatever influence tilt has on Starshade performance, it is negligible compared to the dust levels in our images.
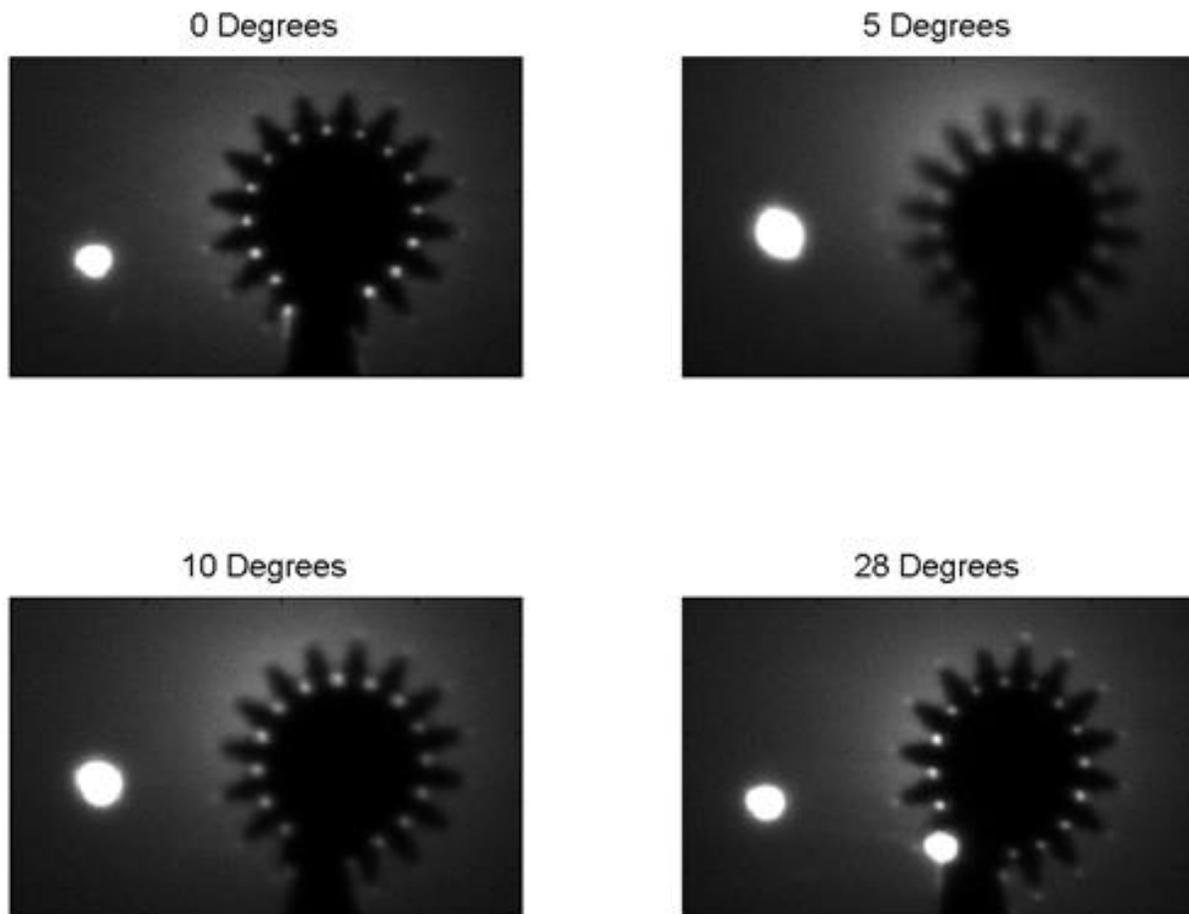
Figure 3-5: Basic Starshade suppression performs well even when the Starshade is angled significantly far away from face on. Single contrast images of the IZ5 Starshade blocking the main source, tilted 0, 5, 10, and 28 degrees from the baseline. Images were taken on May 29th, 2014. The planet seen to the left of the Starshade is the ND1 planet.
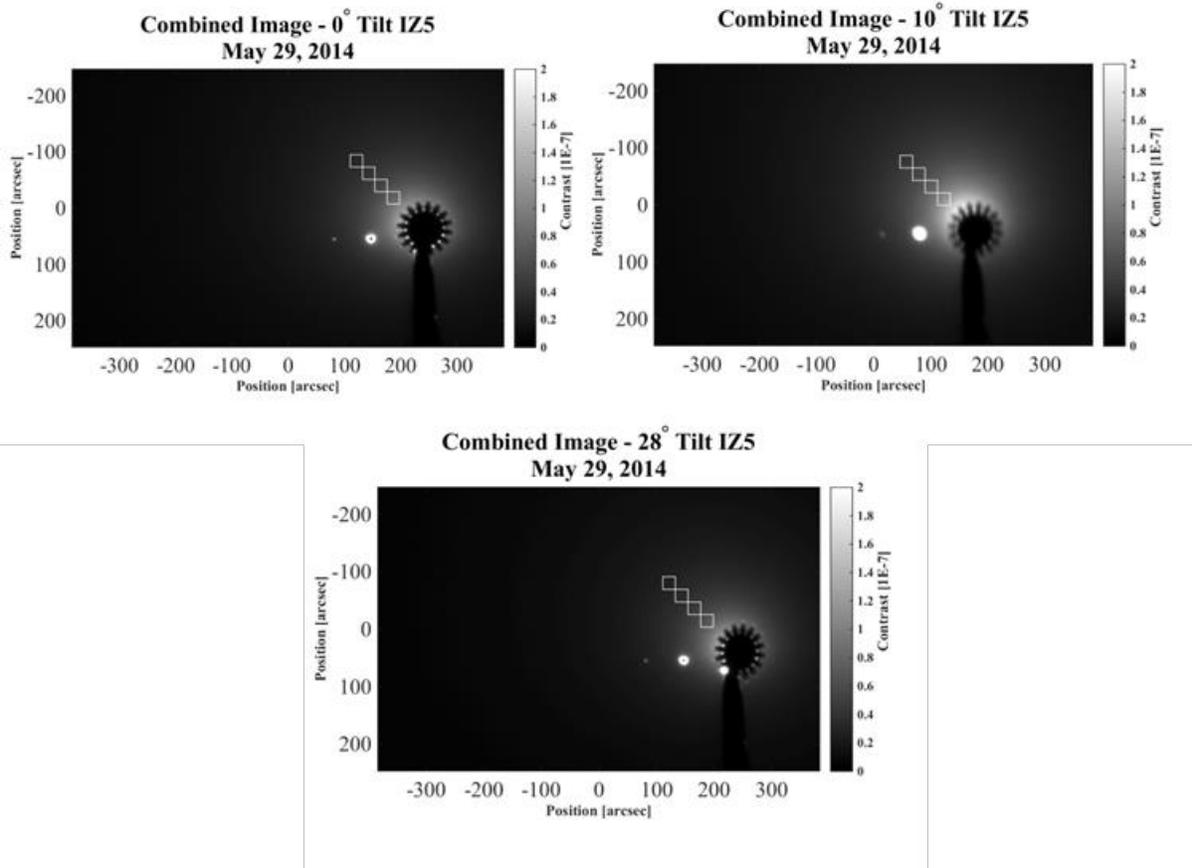
Figure 3-6: Co-added contrast maps for the IZ5 Starshade tilted at 0 (top left), 10 (top right), and 28 (bottom) degrees from the baseline. Any dependence that contrast performance has on tilt is negligible compared to its dependence on dust levels.

Tilt Angle (Degrees)	Mean Background (Box 1)	3σ Level (Box 1)
0	7.82E-08	2.18E-08
10	1.13E-07	3.48E-08
28	6.60E-08	1.80E-08

Table 3-3: Mean background and 3σ levels from box 1 (box closest to the Starshade in Figure 3-6) for the IZ5 etched Starshade at various tilt angles with respect to the light path. We find no correlation between Starshade performance and tilt angle. Dust levels are high in the data and may have swamped out any measureable dependence.

We also experimented with spinning the Starshade while integrating as a method to decrease the impact of Starshade shape errors on performance. These tests are meant to probe the optical effects of spinning Starshades. Potential spinning techniques of a flight system were not considered when taking data. Figure 3-7 shows two contrast maps of the IZ5 Starshade, one with the Starshade stationary and the other with the Starshade rotating at 60 RPM (1-10 Revolutions per exposure). The bright spots in the Starshade valleys seen in the stationary image arise from the manufacturing process, and are effectively shape flaws. In the spinning image these bright spots are blurred into a less-bright ring. Spinning the Starshade during integration spreads the signal from the flaws across a larger area on the CCD, reducing their detrimental effects on suppression without modifying the diffraction behavior of the Starshade.

As stated in Section 3.2, these tests were conducted from an engineering standpoint and are not representative of our highest quality data. Instead, they should both be viewed qualitatively as glimpses into avenues for future tests.
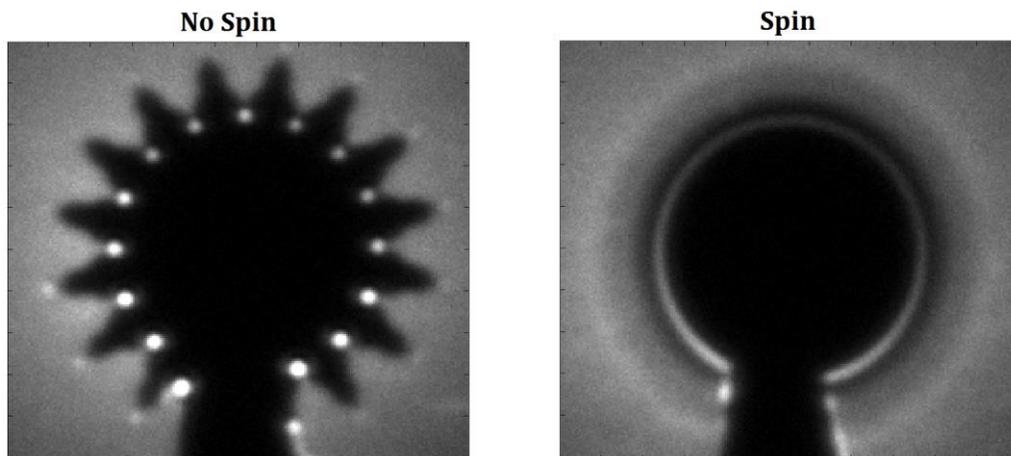
Figure 3-7: Contrast map for the IZ5 Etched Starshade stationary (left) and spinning at 60 RPM (right). The flaws in the valleys show up as bright points in the stationary image, and are spread out into a dimmer ring in the spinning image.

4 MILESTONE #1 (APRIL 2015 TEST)

Milestone #1: Demonstrate, using a Starshade, contrast better than 10^{-9}, at all radii past the Starshade tips, in 50% bandwidth light.

We addressed Milestone #1 in our April 2015 test (Table 4-1), where we focused solely on measuring the performance of our two "base" Starshade shapes, the Hypergaussian (HG) and IZ5 (Section 2.3). The atmospheric conditions of our April 2015 test were easily the best and most stable of the three tests, with characteristic seeing of 6-12 arcseconds for three of the five nights. In Figure 4-1 we show the FWHM of the Gaussian fits of one of the planets and the main source throughout a characteristic night of observation. Since the seeing approached our diffraction limit in the second half of most of our observation nights, we had the potential for extremely high data quality.

Test Duration	16th – 21st April 2015
Nights of Observation	5
Characteristic Seeing	6-12 arcseconds
Refraction	Worst at the beginning of night
Dust Levels	Moderate to low, typical dust brightness on par with the ND4 planet
Other Factors	Early evening had worse seeing and refraction, after about midnight each evening conditions improved significantly. Much colder than previous tests.

Table 4-1: April 2015 desert test characteristics

With five nights of observation to measure the performance of just two Starshade shapes, we were able to gather large data sets for each shape. The abundance of data and excellent seeing conditions throughout the night allowed us to stack many images from throughout a given night in order to create contrast maps with very large effective integration times, boosting our signal to noise ratio (SNR) and mitigating the effects of shot noise. In addition, the amount of data allowed us to be very selective with images to be included in our measurements, leading to final products less affected by atmospheric turbulence and dust than in previous tests.

Figure 4-1: The FWHM of the ND2 planet and the main source as a function of image number from data taken on April 17th, 2015. The seeing improved significantly as the night went on, coming close to the diffraction limit of our optical setup. Images toward the end of the night were easy to combine due to the consistency in the observations.

The methodology used to record and process data from the April 2015 test follows that described in Section 3.1. Some additional refinements were made given the goal of the test – pushing to get the best possible contrast estimate for our current setup and to meet Milestone #1 - therefore many images were taken with a single Starshade during a single night, during which the dust contribution and seeing were variable.

In addition to the deep observations of the two primary Starshade shapes, we also took images using a circular occulter at the distance of the Starshade stand, and images where the planets and 4km sources were on, but the main source was off and the Starshade was out of the line of sight.

In order to measure the contrast performance of our Starshades we defined a '3σ background level' to act as a performance metric. This detection limit is simply three times the standard deviation of the area of the contrast map under scrutiny, and was chosen to be close to the contrast level at which the detection of a source could be made with confidence. For each contrast map presented here we report the 3σ level for four 40 x 40 pixel boxes, starting at the Starshade edge and extending radially outward (Figure 4-3).

4.1 UNCERTAINTY ANALYSIS

Each contrast measurement presented in this paper also includes an uncertainty estimate. To estimate the uncertainty of our 3σ levels we considered the individual components of our data-taking process (Section 3.1). For each set of measurements of a given Starshade we took a corresponding set of unblocked measurements to find the peak of the main light source, which we then used to convert the Starshade measurements from counts to contrast units (Section 3.1). To calculate the uncertainties of our 3σ levels we found the uncertainties on our unblocked peak measurements and then performed the contrast conversions using the upper and lower bounds of the peak measurements, calculating how the different peak values affected the 3σ levels (Figure 4-2). This process is discussed in detail below.

The peak of the unblocked source was calculated by fitting a 2-d Gaussian distribution to the source in each image in the series. The peak of the series was taken to be the median value of the peaks of the individual images. The uncertainty on the peak of the unblocked source is:

$$\sigma_{UNB} = \sqrt{\sigma_{SHOT}^2 + \sigma_{PEAK}^2 + \sigma_{RMS}^2} \ \text{�}$$

Where σ_{SHOT} is the uncertainty from shot noise in the unblocked images, σ_{PEAK} is 1.25 times the standard deviation of the peaks within the unblocked series, and σ_{RMS} is the uncertainty from the 2-d Gaussian fit of the median peak. Our analysis indicates that uncertainty due to read noise and detector non-linearity is negligible compared to the other dominating terms. The uncertainties introduced by bias and dark subtraction were also found to be negligible (less than 0.1% of the peak value, Table 4-2). The typical values of our uncertainty contributors are shown in Table 4-2.

Uncertainty Source	Typical Contribution (% of Peak Value)
Shot noise (σ_{SHOT})	0.7%
Gaussian fitting (σ_{RMS})	0.2%
Atmospheric turbulence (σ_{PEAK})	20.1%
Bias Frames	Negligible
Dark Frames	Negligible
Flat fields	0.3%
Neutral density filter	0.8%
Dust subtraction	Not Calculated
Source-Starshade-Telescope Misalignment	50 – 100%

Table 4-2: Sources of Uncertainty on the Unblocked Peak Values. The rows above the triple line are uncertainty contributions computed directly for each data set and are estimated numerically. The other contributors are not included in our numerical estimate; they are either negligible or must be inferred. A detailed discussion of the misalignment uncertainty is included in Section 5.3.1.

By far the largest uncertainty contribution accounted for numerically comes from the variation of the peak of the unblocked source within a given set of observations. The peak of the unblocked source typically has a standard deviation of about 20% of the peak value, whereas the uncertainty contributions from Gaussian fitting and data reduction are typically below 1%.

Our best contrast results were achieved by averaging many images into one 'master image' with a very long effective exposure time. The master image was converted from counts to contrast units by dividing by the median unblocked peak measurement from the corresponding set. The unblocked uncertainty is propagated to the 3σ measurement by performing this final division with the upper-bound (measured value + uncertainty) and lower-bound (measured value – uncertainty) unblocked peaks rather than the true unblocked peak value. 3σ values are calculated for each of these extreme cases and are used to constrain the actual measurement. Figure 4-2 charts this process, using values taken from our best contrast result for the HG Carbon Fiber Starshade as an example. The 3σ level used is from the 40 x 40 pixel box just off of the Starshade edge (Box 1, See Figure 4-3). It is apparent from Figure 4-2 that the offsets of the upper- and lower-bound 3σ levels from the actual value are not equal. To remain concise in our presentation, we present our uncertainties in ± notation, using the larger of the two offsets.
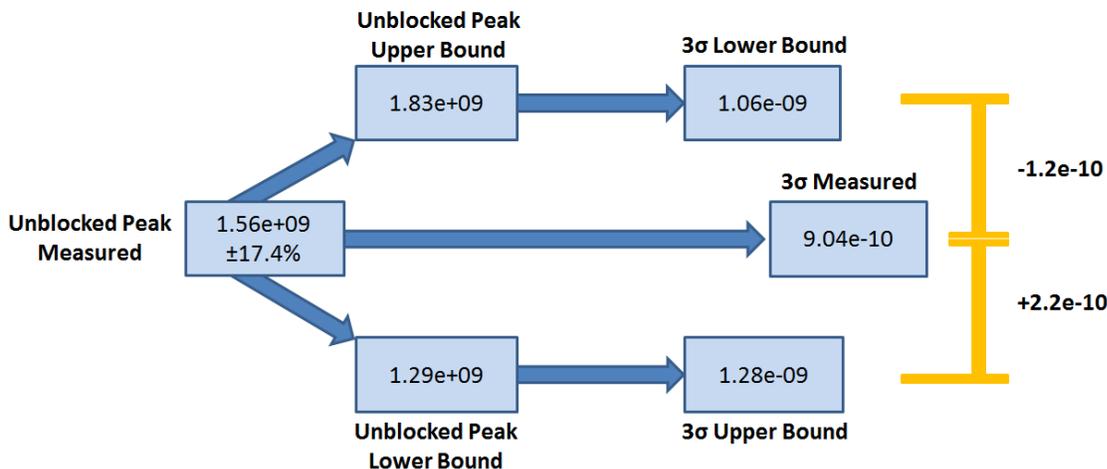
Figure 4-2. Process for propagating the uncertainty of the unblocked peak value to the 3σ level measurement. Upper-bound and lower-bound refer to the measured unblocked peak plus its uncertainty and minus its uncertainty, respectively.

The uncertainty contributions of our dust subtraction routine and of source-Starshade-telescope misalignment are not included in our numeric uncertainty estimate due to their difficulty to quantify for individual images. We considered the effects of misalignment on data gathered during our September 2014 test and ran misalignment simulations with our diffraction model in order to better understand the general uncertainties associated with misalignment. We discuss these considerations in detail in Section 5.3. After our efforts to understand misalignment, we believe that it can vary our contrast levels by 50 - 100%, making it an extremely large contributor to uncertainty and something to focus on mitigating in the future (Section 7).This potential 50 - 100% variance is considered generally rather than folded into our numerical uncertainty estimates.

4.2 RESULTS

We present our best contrast results for each Starshade shape here. To assure that only the best and most consistent images were included in the combination, we applied the following selection criteria to the image sets:

- The FWHM of a visible planet LED (ND3 planet for April 16th, ND2 planet for all other nights) was recorded for each image to estimate seeing quality (Figure 4-1). Images with planet FWHMs greater than 2σ from the mean of the set were discarded.
- Any images with significant saturation on the Starshade were discarded as this is generally indicates a misalignment between the source and the Starshade due to atmospheric refraction.

Sometimes the Starshade would appear to move relative to the position of the planets because the planets suffer greater refraction as the light travels through twice as much air. Since we aligned images using the planet peaks this results in the apparent blurring of the Starshade in some of our final contrast maps. The center of the Starshade was approximated in the final image and used to define 40 x 40 pixel background-sampling boxes a standard distance away from the Starshade along a radial line; the box centers were placed approximately 75, 105, 136, and 167 arcseconds away from the center. The mean and 3σ level within each box were then calculated and are used as the basis for our contrast estimates.

We used small boxes to calculate contrast levels because it aligns with post-processing strategies used in previous tests. In future tests it may be possible to extract better 3σ levels by replacing the boxes with larger-area rings at various radii (Section 7).

As the standard deviation of the background in a stacked image should decrease as a function of the square root of the number of images stacked, it was expected that as we combined more images together that our noise (detection limit) would decrease. This was not the case as there was an additional factor contributing to the variability of the background in our images: the dust in the atmosphere, which was not constant throughout the night. Because of the significance of the dust in our images, we developed a dust subtraction method that removes the contribution of the dust to our contrast sensitivity. Excluding the Starshade, planets, 4km sources, and the Starshade stand, a 40th degree polynomial was fit to the dust in each individual image, assuming a smooth distribution. This was then subtracted off the image and then the same selection process and co-aligning was applied as in the original set. A detailed description of our dust fitting and subtraction method, including a justification of our choice of using a 40th degree polynomial to fit the dust, is included in the Appendix (Section 8.1). Generally, removing the dust improved our standard deviations by about a factor of 10.

For each of the Starshades used in this test (HG Carbon Fiber [HGCF], IZ5 Etched [IZ5], and HG Etched [HGE]) (Section 2.3), we constructed the best set of images for a given night based on the dust-subtracted data. We present the results for each Starshade with and without the dust subtraction. The mean value of the dust is given for reference as appropriate. For those images in which the dust has been subtracted, the background mean no longer has a meaning and is therefore not reported.

4.2.1 BEST OVERALL CONTRAST

Our best contrast is shown in Figure 4-3, which is a combination of 285 10-second dust-subtracted images taken with the IZ5 Starshade. A cross-section of that image is shown in Figure 4-4. The rough edges of the Starshade are due to refraction 'moving' the Starshade in the image. The artifacts in the bottom right of the image are due to the Starshade stand. The bright spots on the left-hand side of the image are the two sources we had placed at 4km to monitor atmospheric conditions to gather site data for a potential longer baseline setup in a future test. The three planets clearly visible, from the edge of the Starshade out, are the ND3, ND4, and ND2 planet LEDs. The 3σ standard deviations of the boxes in the image are also shown in Figure 4-3. Box 1 is the box closest to the Starshade. A point source with a brightness of 10^{-9} would be detectable near the edge of the Starshade at a greater than 3σ level. With these data we have met our milestone.

In Section 8.2 we have included work we have carried out using a convolutional filter to further highlight objects in the image that have a known point spread function. Using this method, the ND4 planet LED, with a brightness $\sim6\times10^{-9}$ of the unblocked main source has a Signal to noise Ratio of approximately 60 standard deviations above mean.

Combined Image (Planet Based) – IZ5 Etched
April 17, 2015 – set3 (285 Images)
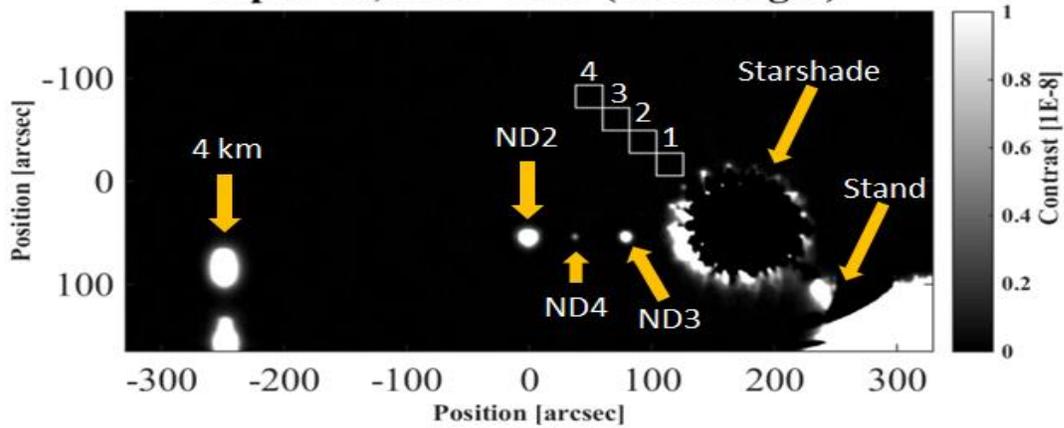
Box	3σ Level
1	7.31E-10±2.40E-10
2	7.17E-10±2.36E-10
3	7.36E-10±2.42E-10
4	7.44E-10±2.45E-10
Center	7.35E-08±2.42E-08
Far Away	7.51E-10±2.47E-10

Planet	Peak
ND2	8.65E-07
ND3	9,74E-08
ND4	6.09E-09

Figure 4-3: Best Overall Contrast. 285 images with a total exposure time of over 47 minutes, taken on April 17, 2015 with the IZ5 Starshade blocking the main source were dust-subtracted and then median combined. We have defined four boxes at increasing radius from the center of the Starshade, as well as a box on the center of the Starshade and a box at the edge of the image to measure the background levels. Visible in the image, from left to right, are two sources placed at 4km to monitor long-range atmospheric conditions, the ND2, ND4, and ND3 planets, the Starshade, and some artifacts due to the Starshade stand. The edges of the Starshade appear blurry due to the fact that the images were co-aligned based on the position of the planets, (Planet based image) and not the Starshade. The peak brightness of each of the planets in contrast is also shown. This image and statistics represent our best contrast results to date, where a source with a brightness of less than 10^{-9} would be detected at better than 3σ.
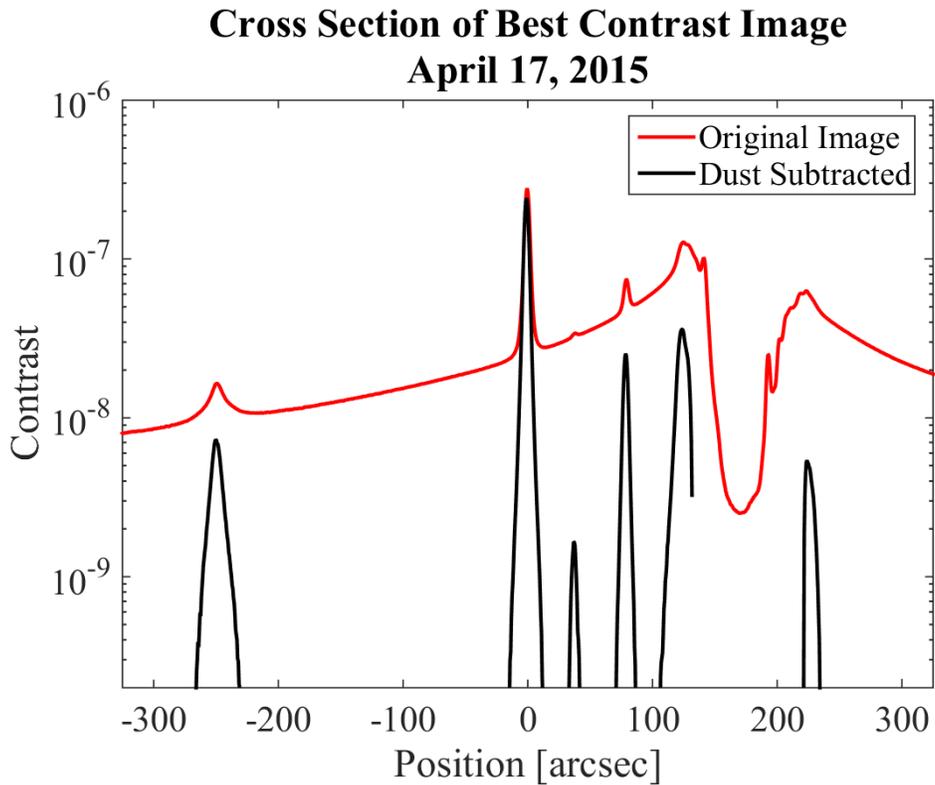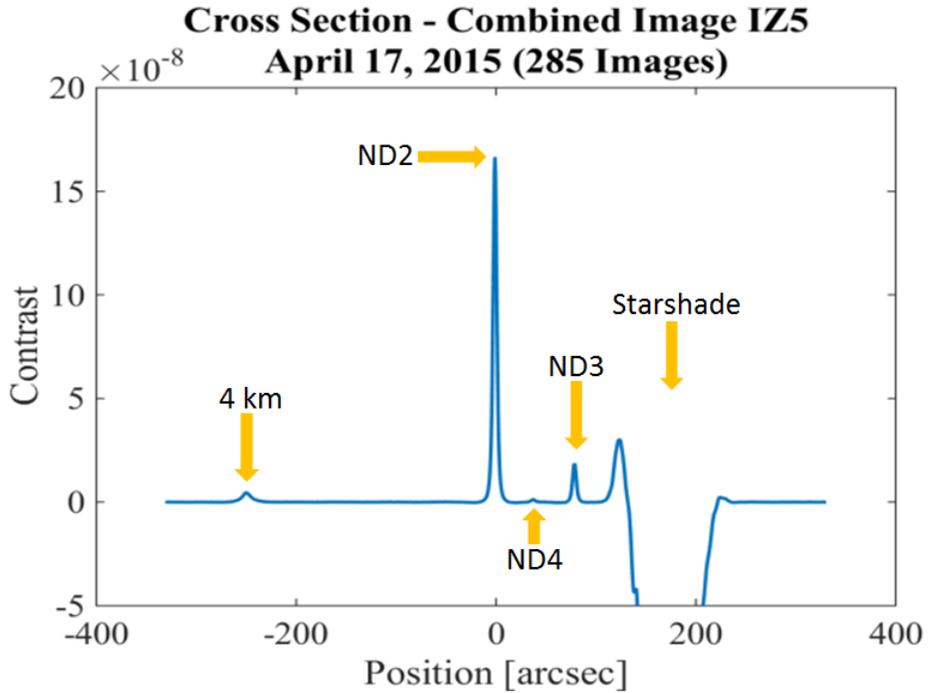
Figure 4-4. The mean of a 50 pixel slice through our best contrast image on a linear scale with dust subtracted (Top) and a log scale with and without dust (Bottom). The dust has been subtracted, so the background is approximately zero. The slice contains the edge of the 4km source, which is visible on the left.

4.2.2 BEST CONTRAST RESULT – OTHER STARSHADES

Shown in the following figures are the best contrast achieved using each of the different Starshades used in the April 2015 Test. The data were analyzed with and without dust subtraction and the results are shown for both methods.

Combined Image (Planet Based) - HG Etched April 20, 2015 - set1 (116 Images)

Figure 4-5: Best result, HG Etched with Dust Subtraction. 116 images (total exposure time of 19.3 minutes) taken on April 20, 2015 with the HG Etched Starshade blocking the main source were dust-subtracted and then median combined. The 4 km sources can be seen to the far left, and to the bottom right, artifacts from the dust subtraction can be seen where the Starshade stand is. The 3σ values for each box and the planet peak contrast values are shown in the figure. In this image, a source with a brightness of less than 10^{-9} would be detected at better than 3σ.

Box	3σ Level
1	9.6E-10 $\pm$ 2.12E-10
2	8.47E-10 $\pm$ 1.87E-10
3	7.88E-10 $\pm$ 1.74E-10
4	7.42E-10 $\pm$ 1.64E-10
Center	9.95E-09 $\pm$ 2.20E-09
Far Away	6.12E-10 $\pm$ 1.35E-10

Planet	Peak
ND2	4.20E-07
ND3	3.28E-08
ND4	3.11E-09

Combined Image (Planet Based) - HG Carbon Fiber April 16, 2015 - set1 (118 Images)
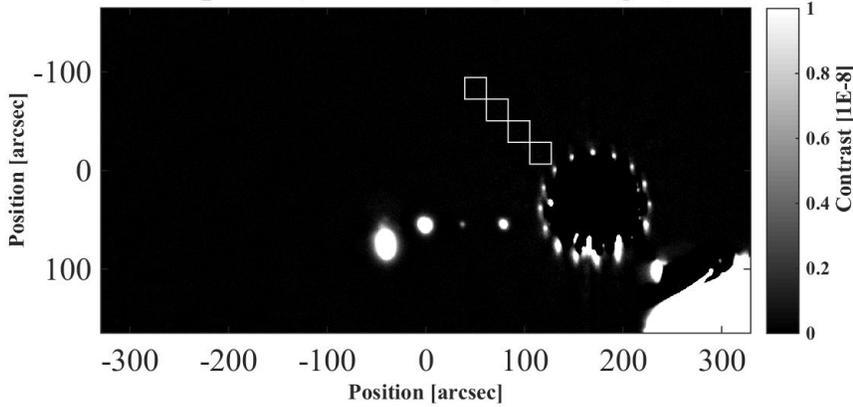
Box	3σ Level
1	1.06E-09 ± 2.23E-10
2	9.20E-10 ± 1.94E-10
3	8.51E-10 ± 1.79E-10
4	7.93E-10 ± 1.67E-10
Center	9.75E-09 ± 2.05E-09
Far Away	6.19E-10 ± 1.30E-10

Figure 4-6: Best result, HG Carbon-Fiber with Dust Subtraction. 118 images (total exposure time of 19.7 minutes) taken on April 16, 2015 with the HG Carbon-Fiber Starshade blocking the main source were dust-subtracted and then median combined. A single 4 km source can be seen just to the left of the ND2 planet, as 4 km source placement changed night to night. The artifacts from the dust subtraction on the stand to the bottom right are present, but varied in shape depending on dust levels. The 3σ values for each box and the planet peak contrast values are shown in the figure. In this image, a source with a brightness of less than 10^{-9} would be detected at close to 3σ.

Planet	Peak
ND2	7.02E-07
ND3	7.82E-08
ND4	5.91E-09

4.2.3 BEST CONTRAST RESULT– WITHOUT DUST SUBTRACTION

**Combined Image (Planet Based) - IZ5 Etched
April 17, 2015 - set3 (285 Images)**
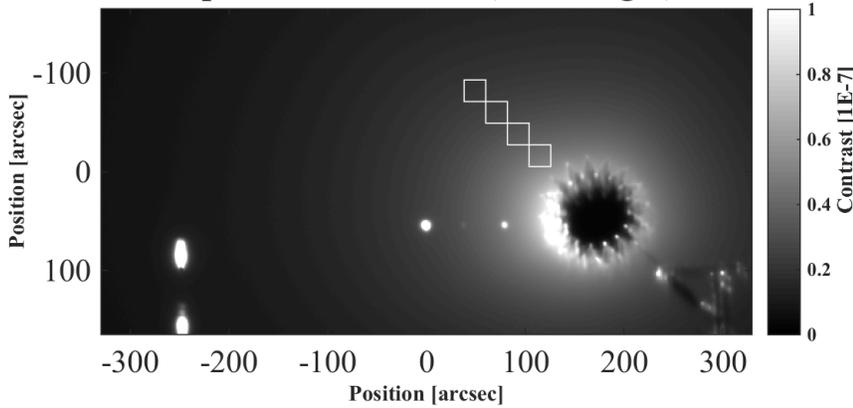
Box	Mean	3σ Level
1	4.58E-08	1.09E-08±3.58E-09
2	3.40E-08	6.12E-09±2.01E-09
3	2.69E-08	3.95E-09±1.30E-09
4	2.21E-08	2.66E-09±8.75E-10
Center	3.04E-09	5.15E-09±1.69E-09
Far Away	7.59E-09	5.22E-10±1.72E-10

Figure 4-7. Best result, IZ5 without Dust Subtraction. 285 images (total exposure time of 47.5 minutes) taken on April 17, 2015 with the IZ5 Starshade blocking the main source were median combined. This is the same group of images shown in the figure with dust subtraction. The 3σ values for each box and the planet peak contrast values are shown in the figure. In this image, a source with a brightness of less than 10^{-8} would be detected at close to 3σ.

Planet	Peak
ND2	8.65E-07
ND3	9.76E-08
ND4	6.10E-09

**Combined Image (Planet Based) - HG Etched
April 20, 2015 - set1 (116 Images)**
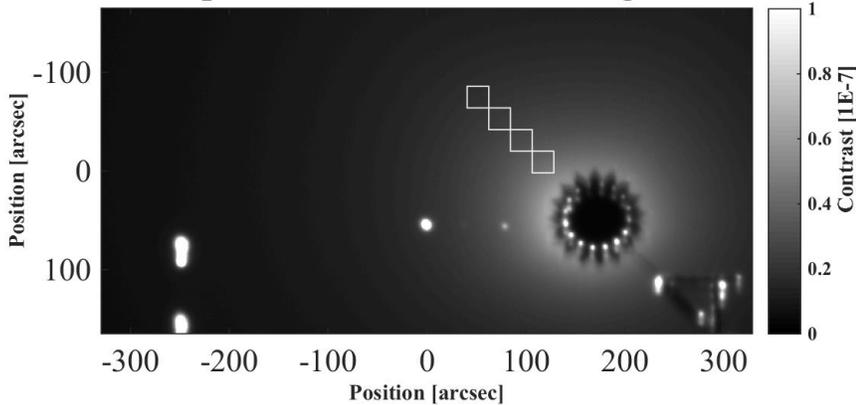
Box	Mean	3σ Level
1	4.05E-08	9.97E-09±2.20E-09
2	2.98E-08	5.18E-09±1.15E-09
3	2.34E-08	3.42E-09±7.56E-10
4	1.92E-08	2.37E-09±5.24E-10
Center	1.92E-09	6.50E-10±1.44E-10
Far Away	6.27E-09	6.51E-10±1.44E-10

Figure 4-8. Best result, HG Etched without Dust Subtraction. 116 images (total exposure time of 19.3 minutes) taken on April 20, 2015 with the HG Etched Starshade blocking the main source were median combined. This is the same group of images shown in the figure with dust subtraction. The 3σ values for each box and the planet peak contrast values are shown in the figure. In this image, a source with a brightness of less than 10^{-8} would be detected at close to 3σ.

Planet	Peak
ND2	4.20E-07
ND3	3.30E-08
ND4	3.06E-09

Combined Image (Planet Based) – HG Carbon Fiber April 16, 2015 – set1 (118 Images)

Figure 4-9. Best result, HG Carbon-Fiber without Dust Subtraction. 118 images (total exposure time of 19.7 minutes) taken on April 16, 2015 with the HG Carbon-Fiber Starshade blocking the main source were median combined. This is the same group of images shown in the figure with dust subtraction. The 3σ values for each box and the planet peak contrast values are shown in the figure. In this image, a source with a brightness of less than 10^{-8} would be detected at close to 3σ.

Box	Mean	3σ Level
1	5.27E-08	1.24E-08 ± 2.61E-09
2	3.89E-08	6.89E-09 ± 1.45E-09
3	3.07E-08	4.33E-09 ± 9.11E-10
4	2.51E-08	3.13E-09 ± 6.59E-10
Center	3.11E-09	2.05E-09 ± 4.31E-10
Far Away	7.00E-09	6.97E-10 ± 1.47E-10

Planet	Peak
ND2	7.02E-07
ND3	7.91E-08
ND4	6.01E-09

4.2.4 CIRCULAR OCCULTER

We took a short series of 22 images with a circular occulter (diameter of 58cm) placed into the line of sight at the Starshade station. The ring of light resulting around the occulter (Figure 4-10) was extremely bright, and therefore we needed to use an ND3 filter for our observations to avoid saturation. For comparison, the images taken with the Starshades had no filter and did not saturate. Shown in Figure 4-10 is our mean combined image (total exposure time of 3.7 minutes). The light of the main source was not sufficiently suppressed to see any of the planets in our configuration. Compared with the observations using the Starshades, the main source was suppressed a factor of 100-1000 times less.

Combined Image - Circle
April 18, 2015 (22 Images)
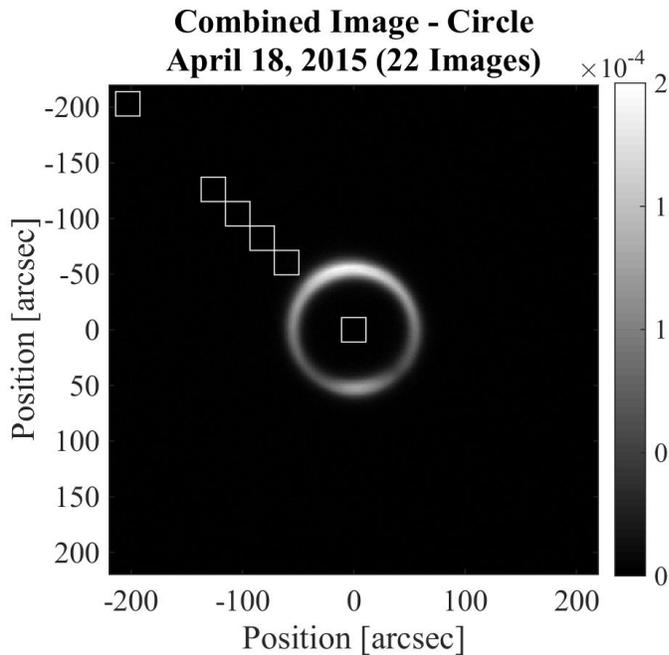
Figure 4-10. Circular Occulter. 22 images (total exposure time of 3.7 minutes) taken on April 18, 2015 with the circular occulter blocking the main source were median combined. The 3σ values for each box and the contrast values are shown in the figure. In this image, a source with a brightness of less than 10^{-6} would be detected at less than 3σ. This means we were not able to detect any of the planets in this image.

Box	Mean	3σ Level
1	2.98e-06	6.24e-06
2	1.85e-06	5.45e-06
3	1.72e-06	4.48e-06
4	1.62e-06	4.45e-06
Center	1.96e-06	4.54e-06
Far away	1.59e-06	4.52e-06

4.2.5 BACKGROUND CONTRIBUTION

The prominence of dust scattering in all of our tests prompted us to investigate the background to see if there were any additional contributions from the environment that were placing a lower limit on our achievable contrast level. Our method for testing this was to turn off the main source and move the Starshade out of the line of sight while leaving on the planet and 4km LEDs for reference. As seen in Figure 4-11, the dust halo around where the main source would be is completely gone, but there is still some dust scattering around the 4km sources. We did no dust subtraction on these images. The peaks of the planets were comparable to the peaks measured with the Starshades in place. We measured the standard deviation in boxes placed in approximately the same relative locations to the planets as in the blocked images. The background is quite uniform in the area with the Starshade and planets, and increases slightly far away from our observations, as evidenced by the mean values in the boxes, and the standard deviations are at or below the standard deviations of the mean combined dust-subtracted blocked images with the same number of contributing images. From this we conclude that a significant component of our detection limit is the dust in the field, and that the dust subtraction methodology is not providing an artificial improvement to our sensitivity.

**Combined Image - Planets Only
April 19, 2015 (61 Images)**

Box	Mean	3σ Level
1	3.39e-10	7.31e-10
2	3.43e-10	7.17e-10
3	3.43e-10	7.36e-10
4	3.43e-10	7.44e-10
"Center"	3.08e-10	7.35e-10
Far away	5.27e-10	7.51e-10

Planet	Peak
ND2	4.25E-07
ND3	4.00E-08
ND4	2.96E-09
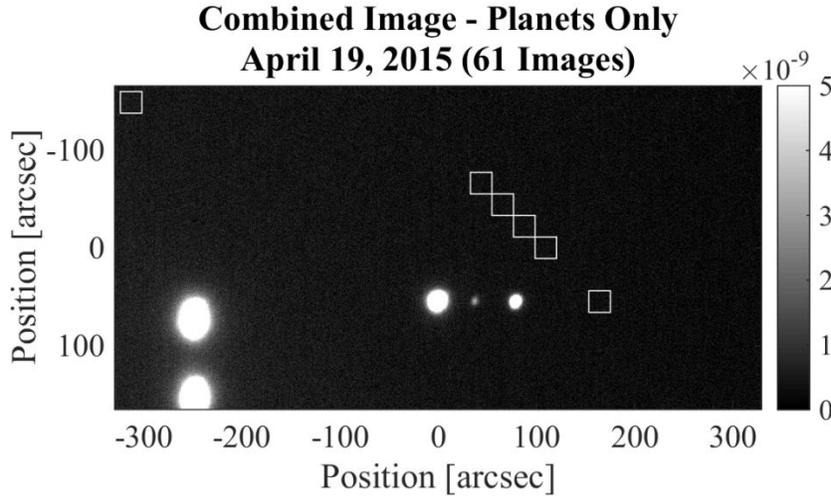
Figure 4-11. Main Source Off. 61 images (total exposure time of 10.2 minutes) taken on April 19, 2015 with the main source off and no Starshade were median combined. This was to determine the background levels arising from the scene itself. The 'center' of the image is approximately where a Starshade would be relative to the planets. The 3σ values for each box and the planet peak values are shown in the figure. In this image, a source with a brightness of less than 10^{-9} would be detected at close to 3σ. This data establishes a lower limit to our possible detection and is compared with the dust subtracted data to verify that we are not artificially improving our statistics beyond what is possible as a result of the dust subtraction.

5 MILESTONE #2 (SEPTEMBER 2014 TEST)

Milestone #2: Demonstrate agreement between the measured and predicted contrast resulting from a range of Starshade shapes.

We addressed Milestone #2 in our September 2014 test (Table 5-1), where we measured the performance of a number of Starshades with intentionally flawed shapes. These "error shapes" are deviations from both the HG and IZ5 Starshade shapes, and were chosen to be similar to shape errors that we might see in a flight Starshade structure. As part of a flight system, these types of errors might be caused by manufacturing, dynamics, deployment, thermal distortion, etc. For our purposes the flaws provide specific points of comparison between models and observation in order to better compare the two. For these test articles, we specify a particular error as a deviation from the base shape and then build that into the Starshade.

Test Duration	23rd – 26th September 2014
Nights of Observation	3
Characteristic Seeing	10-30 arcseconds
Refraction	Extreme on the first night, significant improvement over the next two nights.
Dust Levels	Very low on the first night, increasing steadily as the test progressed.
Other Factors	Two of the planned five nights were rained out.

Table 5-1: September 2014 desert test characteristics

The errors we chose to build have amplitudes much larger than we expect to see on orbit. We chose error amplitudes that we predicted to cause contrast peaks in the range of 10^{-5} to 10^{-8}. These are easy to see in the field test images and therefore can be well correlated with the simulated results.

We chose six families of errors for this particular test. These were chosen for relevance to the flight structure as well as ease of modeling and manufacturing. These errors are: tip truncation, valley truncation, petal clocking, petal width variation, displaced edges, and sinusoidal deviations along the edge. Additional shapes may be considered for future tests. See Section 5.4 for a breakdown of each error shape.

Comparisons of our test results to three separate simulations are included here. All of the simulations include the same basic observation parameters:

- Simulations include only expanding beam light from the main source, at the correct distance, diffracting from the Starshade edge. No background, scattered light, or other noise sources are included.
- Simulations were done at wavelengths from 400 to 800 nm in steps of 25 nm and averaged together, weighted according to the transmission curve of the optics in the system (light source and lens, telescope, and CCD QE) (See Figure 2-1)
- The simulations solve the Fresnel integral for the particular Starshade shape and calculate the resulting light distribution in the image (or focal) plane.

- Contrast is calculated by dividing the light distribution in the image plane by the peak of the PSF of the unblocked light source.

Table 5-2 lists all of the Starshades we built and the error shapes that were included in each one. We tried to apply the same error shapes, with the same amplitudes, to the HG and IZ5 designs whenever possible. The exception is the valley truncation error – the IZ5 valleys start past the points at which the HG valleys were truncated. So the valleys were truncated at larger radii for the IZ5 design. For the Starshades with added errors, deviations from the base shape were added to four petals, evenly spaced around the Starshade (at the 12:00, 3:00, 6:00, and 9:00 positions). For each Starshade, the errors are the same family, with a range of amplitudes. Error amplitudes are numbered such that #1 corresponds to the largest flaw and #4 the smallest, except for the sines error. For the sines error we varied two parameters (frequency and amplitude), so each of the 4 flaws have no inherent increasing order. Both the etched and Carbon fiber manufacturing designs for the Starshade were measured using a touch probe (one instance of each). Measurement errors as deviations from the CAD files were recorded at 2400 points around the Starshade. Typical excursions in the Starshade plane were ~0.005 inches and worse case was 0.012 inches for less than 20 consecutive points (~5cm along 1 petal edge). These measured Starshades were run through the NG model, which showed almost no difference in performance between theoretical and measured values. This assessment was confirmed by comparison of the size of the intentional flaws that needed to be used to generate measurable responses. See Table 5-2.

Base Shape	Error Type	Error Amp #1	Error Amp #2	Error Amp #3	Error Amp #4
HG	None				
HG	Tip Trunc.	0.2831 m	0.2781 m	0.2731 m	0.2681 m
HG	Valley Trunc.	0.175 m	0.180 m	0.185 m	0.190 m
HG	Petal Clocking	0.08 radians	0.06 radians	0.04 radians	0.02 radians
HG	Petal Width	-5%	-4%	-3%	-2%
HG	Edge Displacement	1.2 mm	0.9 mm	0.6 mm	0.3 mm
HG	Sines (freq, amp)	N=25, 0.15 cm	N=60, 0.1 cm	N=60, 0.15 cm	N=25, 0.1 cm
IZ5	None				
IZ5	Tip Trunc.	0.2831 m	0.2781 m	0.2731 m	0.2681 m
IZ5	Valley Trunc.	0.189 m	0.191 m	0.193 m	0.195 m
IZ5	Petal Clocking	0.08 radians	0.06 radians	0.04 radians	0.02 radians
IZ5	Petal Width	-5%	-4%	-3%	-2%
IZ5	Edge Displacement	1.2 mm	0.9 mm	0.6 mm	0.3 mm
IZ5	Sines (freq, amp)	N=25, 0.15 cm	N=60, 0.1 cm	N=60, 0.15 cm	N=25, 0.1 cm

Table 5-2: Starshade Shapes Built and Tested. Error amplitude #1 is generally the error that we expect to give the largest error signal, error amplitude #4 the smallest.

5.1 THEORETICAL MODELS

Independent high-fidelity numerical codes for modeling the diffraction performance of Starshades have been developed by Northrop Grumman (NG), NASA's Jet Propulsion Laboratory (JPL)[3], and Colorado

University (CU)[4], using a variety of numerical methods. We have performed detailed comparisons between these models and found that the results agree to within ~5% at flight-like scales [5]. All shape imperfections that arise from the manufacturing process are accounted for in the Starshade shapes input into the three optical models.

For comparison of model predictions to observations in the field, the models have been augmented to allow for a curved wave incidence (CWI) light source [6]. The NG CWI model and the original point source model converge to the same solution at a Starshade-source distance of 10^{16} meters, indicating the new model scales properly.

The NG model solves the Fresnel integral and calculates the resulting field in the image plane for light diffracting off of pieces of the Starshade outline. 32 fields in total are calculated (each half-petal of the 16 petal Starshade is considered on its own) and are then added together to produce the overall light distribution.

For the CWI beam used in our tests, the effective telescope to Starshade range (R_{eff}) is maximized by placing the Starshade at the halfway point between the telescope and the light source. In that case, the effective range [7] is $R_{eff}=\frac{1}{4}R_{tot}$. Validating the model at flight-like F# supports using that model for on-orbit configurations.

5.2 OBSERVATIONS AND ANALYSIS

The methodology used to record and process data from the September 2014 test is the same as described for the May 2014 test (Section 3.1). In addition to finding the unblocked light source peak, co-adding images, and converting images to units of contrast, we also had to find the contrast peaks created by the various Starshade flaws. A methodology for this additional step follows.

The brightest flaw expected for a given Starshade was positioned to be at the 12:00 position for one series of exposures and then the Starshade was rotated approximately 90° or 180° counter-clockwise (a 'clocked' measurement) to permit imaging of the flaw previously blocked by the Starshade stand. The same set of unblocked images was used in the analysis of each pair of images using the same Starshade. Specifics of the flaw fitting method are:

- Each flaw and planet is fit with a 2-d Gaussian distribution to estimate the peak of the flaw. This method works well for most shapes, but has disadvantages for flaws that are an unresolved composite of several contributions. For example, the simulation of the Starshade with sine wave flaws (Section 5.4.4) predicts several bright spots close together, which when blurred in observation merge together in a non-circular shape.
- Flaws for which we have multiple measurements available (about 40% of the time) are combined in a weighted average (see below).
- Our dust subtraction technique (see Appendix) is applied to each image and each flaw peak is then re-estimated using the same Gaussian-fitting technique. Two contrast estimates are presented for each flaw: one with the dust subtracted and one without.

As each Starshade has two sets of measurements, sometimes we have multiple measurements of the same flaws in different positions. In order to account for variation in the atmospheric effects between the two sets of measurements, we measured the peak of the brightest (ND1) planet in the two images. The flaw peak values were then scaled by the relative difference in the brightness of the ND1 planet and

any multiple measurements for the same flaws were averaged. For some of the images, it appeared that the flaw in the 6 o'clock position might be due to the intentional flaw on the Starshade, but often these measurements were corrupted by glint from the Starshade stand, so none of the 6 o'clock points were used in our analysis There will still be residual error present in the averaged flaw value due to the misalignment of the Starshade with the source (either from absolute misalignment or variable light refraction) as these effects are not circularly symmetric.

5.3   UNCERTAINTY ANALYSIS

There were two main measurements that we made for each contrast map in order to compare our data to model predictions: finding the peak of the unblocked main light source and finding the peaks created by the intentional flaws to the Starshade shape. Each set of measurements for a given Starshade had a corresponding set of unblocked measurements taken close in time in an effort to have both sets of measurements taken with the same atmospheric conditions. One set of unblocked images was used for both orientations of the Starshade to facilitate inter-comparison of overlapping flaw brightness as previously described. Both the unblocked peaks and the flaw peaks were calculated with a 2-d Gaussian fitting routine. We calculated the uncertainty in our measurements as described in Section 3.1, considering the contributions from calibration images, shot noise, standard deviations of the image sets (seeing uncertainties), and the RMS errors of the Gaussian fit itself. As in our April 2015 test, our analysis indicates that the uncertainty attached to our measurements is largely attributed to variations in peak brightness during measurement due to seeing, and that contributors such as the variability of calibration images, read noise, and detector non-linearity are negligible in comparison (Table 5-3).

While the main procedure for determining the uncertainties of peak measurements was the same for both, the blocked images (images taken with the Starshade blocking the main light source) were handled slightly differently from the unblocked images. The peaks in the blocked images are much fainter than those in the unblocked images, so to reduce noise the images were combined with a 25% trimmed mean as the first image processing step. Therefore, σ_{PEAK} is simply the standard deviation of the peaks in the individual images within the set (see Section 3.1). The same σ_{PEAK} is assumed for the dust-subtracted images, although the 2-d Gaussian fits were not available for all of the individual dust-subtracted images due to the computationally intensive nature of dust subtraction and the large number of images, however we did complete dust subtraction on the individual frames for two sets of blocked images. These two sets were chosen to be those with the largest and smallest standard deviations of the ND1 planet to probe the uncertainty range. For these sets, σ_{PEAK} was consistent between the original and dust subtracted images so we are confident that our σ_{PEAK} values are reasonable for the dust subtracted case.

The typical values of our uncertainty contributors are shown in Table 5-3. The uncertainty due to the dust subtraction is assumed to be the median value of the difference between the original and dust-subtracted peaks.

Uncertainty Source	Typical Contribution (percent of flaw peak)
Shot noise	1%
Gaussian fitting	1%
Atmospheric turbulence (seeing)	15 – 20%
Atmospheric refraction variation within series	2 – 3%
Detector non-linearity	Negligible
Bias	Negligible
Source-Starshade-Telescope Misalignment	50 – 100%
Dust Subtraction	3%
Neutral-density filter	0.8%

Table 5-3: Sources of Uncertainty. The rows above the triple line are uncertainties computed directly, while the others are either negligible or have to be inferred. As misalignment error is the most significant, it is addressed in detail below.

Contrast values with uncertainty estimates are included for each of the flawed Starshades in their respective tables in the results section. Uncertainties that arise from detector nonlinearity and from bias correction were found to be negligible and are thus not included in our estimate. The uncertainties arising from misalignment, dust subtraction, use of the ND filter, and atmospheric refraction were difficult to quantify on an image-to-image basis, so are not folded into the numeric estimate for each individual measurement. They are instead considered as an average over the image sets.

As the uncertainties of each individual flaw are critical to our comparisons with the model results, we show how including all of the uncertainty contributions in Table 5-3, rather than just the first three, could change our uncertainty calculation. Table 5-4 gives a numeric uncertainty estimate for every possible contributor using the #1 flaw on the HG tip truncation Starshade as an example. The value reported for this flaw peak in Section 5.4.1 is $3.09 \times 10^{-6} \pm 8.43 \times 10^{-7}$. Table 5-4 shows that variations introduced by misalignment, refraction, and dust subtraction could drive the uncertainty up to $\pm 3.1 \times 10^{-6}$, an inflation of almost 400%. This factor is not included in the data presented as part of this paper.

Source	Maximum Uncertainty Contribution (contrast units)
Shot noise	1.13E-7
Gaussian fitting	9.3E-8
Atmospheric turbulence (seeing)	2.81E-7
Atmospheric refraction variation within series	9.2E-8
Detector non-linearity	1E-9
Bias	6.25E-10
Source-Starshade-Telescope Misalignment	3.09E-6
Dust Subtraction	9.2E-8
Total Uncertainty	**3.1E-6**

Table 5-4: Uncertainty Estimation for One Flaw. The individual uncertainty contributions from each source are calculated for the #1 flaw of the HG tip truncation Starshade. Variations due to misalignment, refraction, dust subtraction, and the ND filter could drive our uncertainties up to 400% the reported values.

5.3.1 MISALIGNMENT

Misalignment of the source and the Starshade is difficult to estimate for each individual point as there are two contributions to the misalignment in each image: the misalignment of the telescope-Starshade-light source setup which is approximately constant over the entire data set, and variable light refraction due to atmospheric effects within each image set, which manifests as a misalignment. The precise error contribution for each measurement from misalignment is not available but can be estimated using a combination of factors.

Modeling of Starshade misalignment can roughly estimate the impact that misalignment has on the measured peak brightnesses of the flaws. For the HG Petal Width variation shape (Section 5.4.2) we ran six models with the main light source misaligned by 0 to 15 arcseconds. For each model we compared the peaks of the flaws to those of the aligned case, and recorded the percent difference. Figure 5-2 (bottom) shows our results for the flaw that the light source was moving toward. Models predict that misalignment has a significant effect on peak brightness.

Predicted

Measured

Figure 5-1: Model/experimental peak pairs for the HG Petal Width variation Starshade. Models predict equal-brightness peaks around a single petal, while experiment shows peaks of different brightness. This discrepancy can be used to estimate misalignment.

Simulations of the petal width variation Starshade in perfect alignment with the light source predict pairs of equal-brightness peaks around each flawed petal. The relative brightnesses of the peak pairs measured in the September 2014 test are not of equal brightness however (Figure 5-1), and thus indicate a constant offset in the y-direction of our setup. To quantify the degree of this constant misalignment, we analyzed the difference in brightness between the two peaks around a single flawed petal in the misalignment models discussed above (Figure 5-2 top), and compared those deltas to the deltas in our data. This analysis revealed that our absolute misalignment could have been as much as 9 arcseconds (about a 30% difference in brightness between the peaks around a single flawed petal, taken from Figure 5-2 top) and could affect our peak measurements by +/- 50 – 100 % over the entire data set (number range extrapolated from the 9 arcsecond misalignment estimate and Figure 5-2 bottom).

The degree of misalignment due to light refraction can also be estimated. For two of the sets of observations we determined the x- and y-coordinates of the ND1 planet based on a Gaussian fit of each individual frame of the composite image. The two sets were chosen to probe the boundaries of the additional misalignment expected within a given set of measurements. For the observations in question, the planet moved between 0.5 and 2.5 arcseconds in the x-direction and between 0.5 and 1.5 arcseconds in the y-direction. This misalignment varies across our data set, but contributes to the +/- 50 – 100 % effect of misalignment on our measurements described above.

Misalignment Deltas Between Peaks of a Single Petal
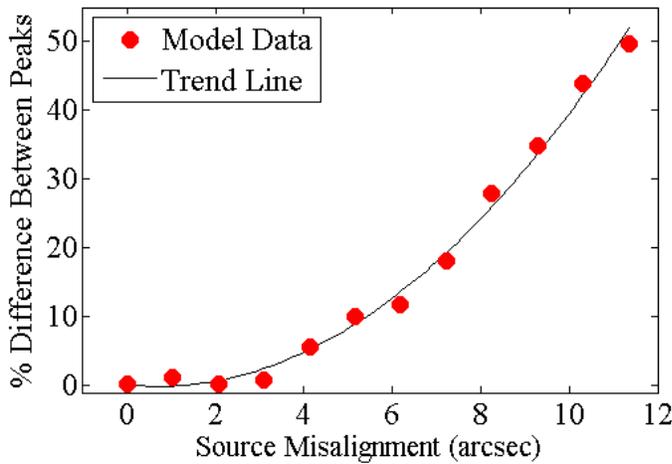
Percent difference of peak brightnesses about a single petal at various misalignments

Delta: 29.1%

9 arcsec Misalignment
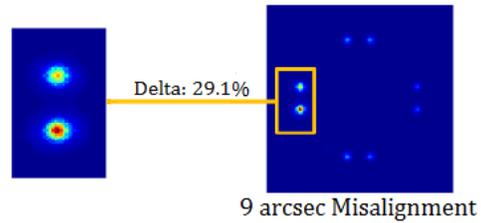
Misalignment Deltas of Total Flaw Brightness
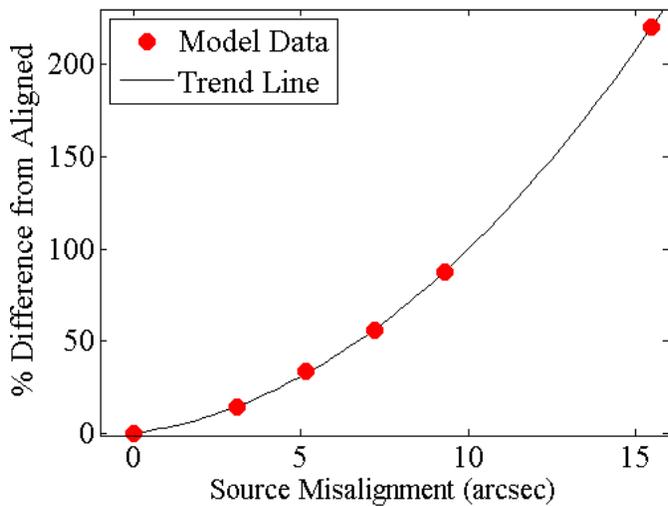
Percent difference of total flaw brightness between the aligned case and various misaligned cases

Aligned   15 arcsec Misalignment
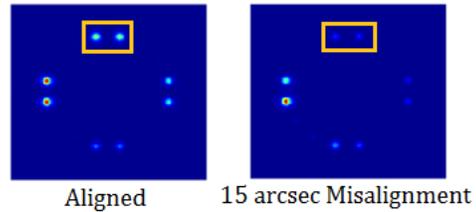
Figure 5-2: Misalignment model data. There are two steps to the misalignment modeling process. We modeled the change in brightness of peaks around a single petal (top graph) and also the change in total flaw brightness compared to the aligned case (bottom graph), both for various misalignments between 0 and 15 arcseconds. The top graph and the differences between peaks around a single flaw apparent in our experimental data (Figure 5-1) can be used to estimate the degree of misalignment in our test setup. The bottom figure illustrates how this misalignment could affect our contrast numbers. The trend lines fit are both quadratic functions with R^2 values greater than 0.99.

5.4   RESULTS

For each of the Starshade designs in our test, we calculated the expected optical performance over the wavelength range of our test using three independent simulation codes – one from the NASA Jet Propulsion Laboratory, one from Colorado University, and one from Northrop Grumman Aerospace Systems. In our field test we gathered data for comparison to the models.

Errors on the Starshade shape tend to create distinct point sources in the contrast images. We used the peaks of these point sources, measured for all flaws in both the modeled and experimental data, as a comparison metric. For each of the six flaw types we show a package of results below. We show the flawed Starshade design for the HG base shape, contrast maps from model and experiment, and point-to-point comparisons between model and experiment for both the original and dust-subtracted measurements. Since each Starshade has four petals that are flawed compared to the base design, this translates to four points of model/experiment comparison for each image. The noise and background levels in our images limit our ability to measure a peak for some flaws, so for each image we have *up to* four points of comparison. In the case of the IZ5 Petal Clocking Starshade all of the flaws were too faint to measure, so we have no comparison to the simulation for this Starshade. Note that the orientation of the Starshades varies between the various simulations and the experimental data, either rotated or flipped in the plane of the Starshade. Therefore the images don't directly line up as displayed. Also note that the non-dust-subtracted observations contain significant amounts of background light (due to scattering from dust in the air). This scattering is not included in the current model results. Our results are followed by a discussion of the data set as a whole and its relation to Milestone #2.

Each of the subsections below presents our results for a single flaw type on a single Starshade shape (HG or IZ5). We first show a parameterization of the Starshade shape with the flawed shape overlaid on top (see Figure 5-3 for an example). We then show theoretical contrast maps and flaw peak values computed by each of the three numerical models (Figure 5-4, Table 5-5). Finally, we show the corresponding contrast maps imaged in the desert and report our measured flaw peak values and their ratios to model predictions (Figure 5-5,Figure 5-6, Table 5-6).

5.4.1 Tip Truncation*:*

The tips of the petals are cut off to a specified radius. The amplitude given is the maximum radius of the petal for that tip. With no error the maximum petal radius is 28.11 cm for the base HG and 29.0 cm for the base IZ5.
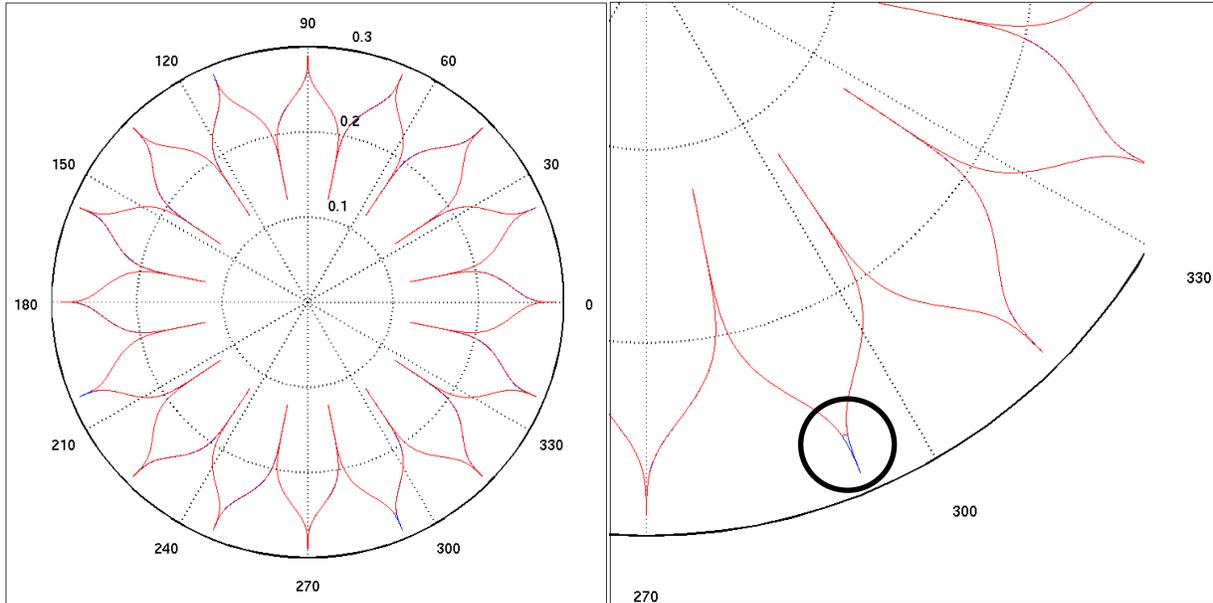
Figure 5-3: Tip Truncation Shape input for Optical Models and Starshade manufacturing. Truncation of the tips of four petals relative to the base HG design (blue). The shape with errors is in red, but is only different from the base shape at the four error locations. The right-hand figure is zoomed in to one petal to show details of the deviation. The circle indicates the location of the truncated tip.

HG Tip Truncation: Model Predictions
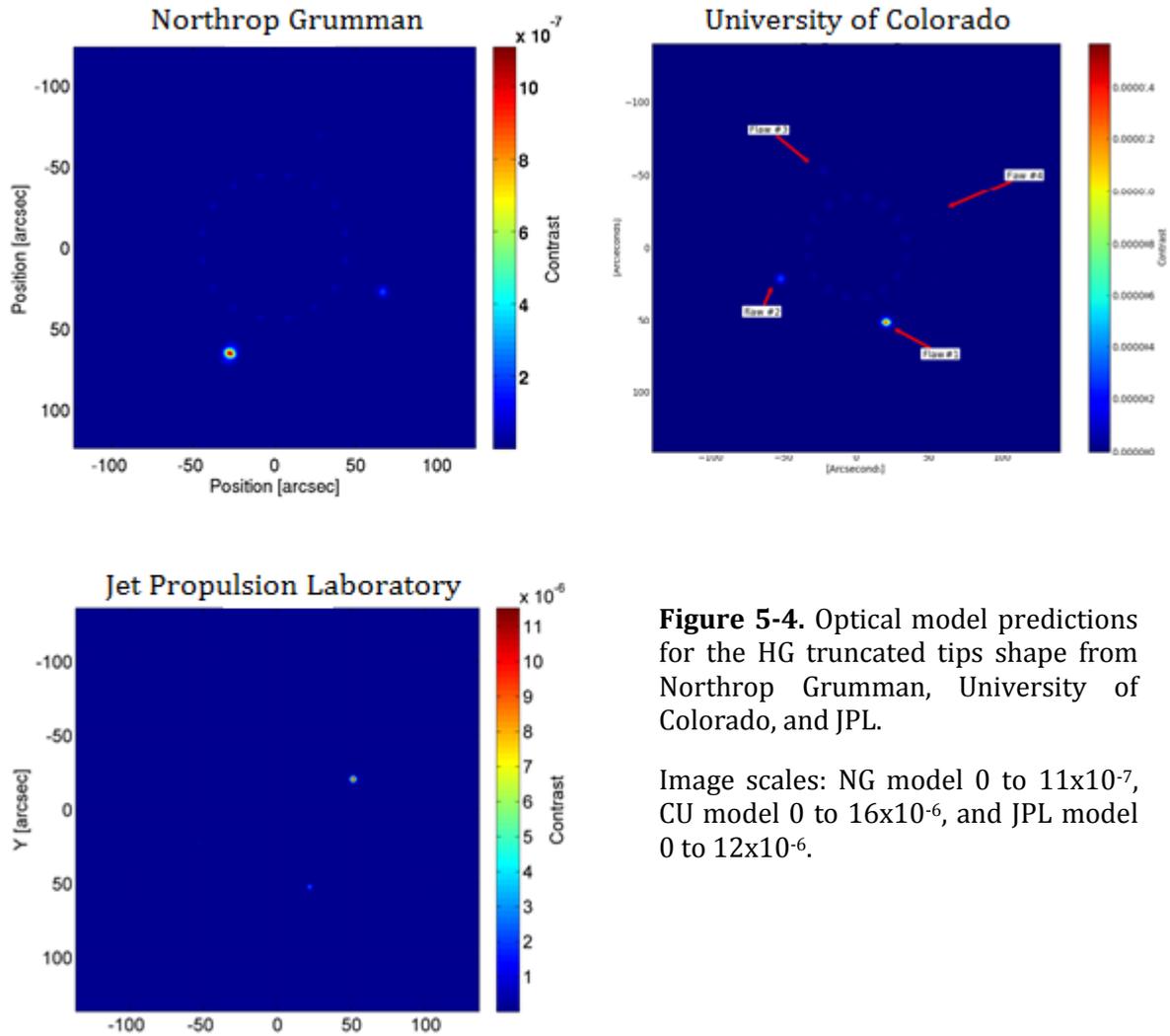
Figure 5-4. Optical model predictions for the HG truncated tips shape from Northrop Grumman, University of Colorado, and JPL.

Image scales: NG model 0 to 11×10^{-7}, CU model 0 to 16×10^{-6}, and JPL model 0 to 12×10^{-6}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	1.11E-06	1.57E-05	1.20E-05
Flaw 2	2.09E-07	3.61E-06	2.50E-05
Flaw 3	3.35E-08	6.32E-07	4.20E-07
Flaw 4	4.73E-09	1.62E-07	1.10E-07

Table 5-5. HG tip truncation flaw brightnesses computed by each of the models.

HG Tip Truncation: Measured Contrast Figures

With Dust
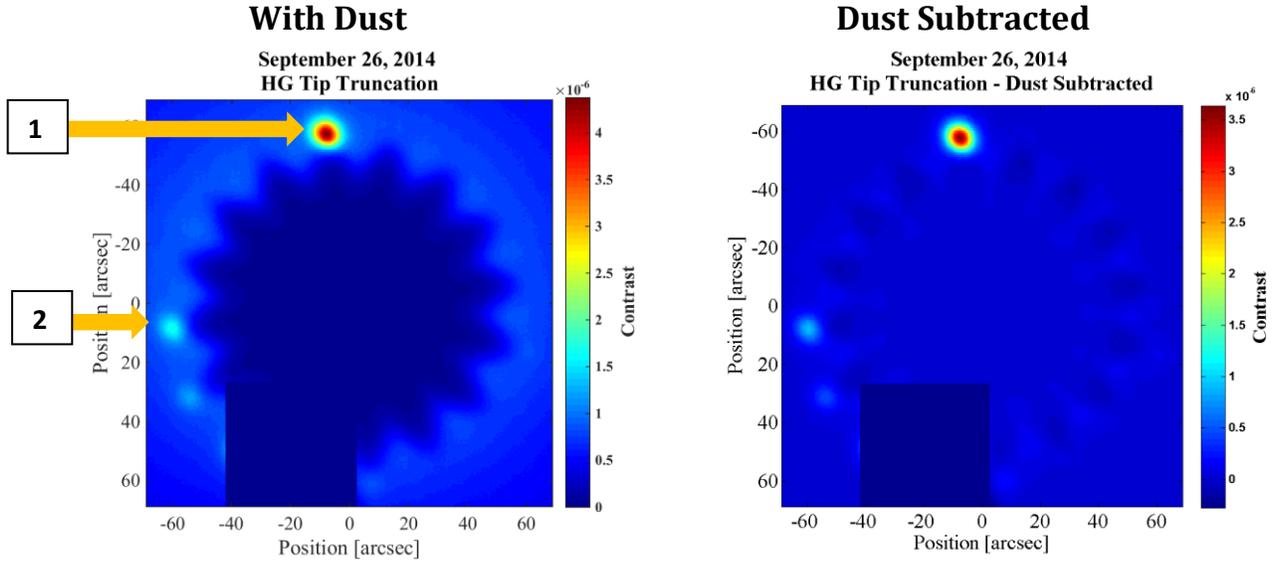
Dust Subtracted

Figure 5-5. Measured contrast maps for the HG edge shape with the tip truncation error, with dust (left) and dust-subtracted (right).

HG Tip Truncation: Measured Contrast Numbers

The HG Starshade with truncated tips had one additional flawed tip (visible in the 8:00 position in the Figure 4-6). Flaw brightness could only be calculated for the two brightest flaws. -

	Flaw #	Measured Peak	Uncertainty
With Dust	1	3.09E-06	± 8.43E-07
	2	9.01E-07	± 3.50E-07
	3		
	4		
Dust Subtracted	1	3.09E-06	± 8.42E-07
	2	1.09E-06	± 3.50E-07
	3		
	4		

Table 5-6. Measured contrast peaks and comparison to the models for the HG edge shape with the tip truncation error.

HG Tip Trunc: Model/Experiment Ratios
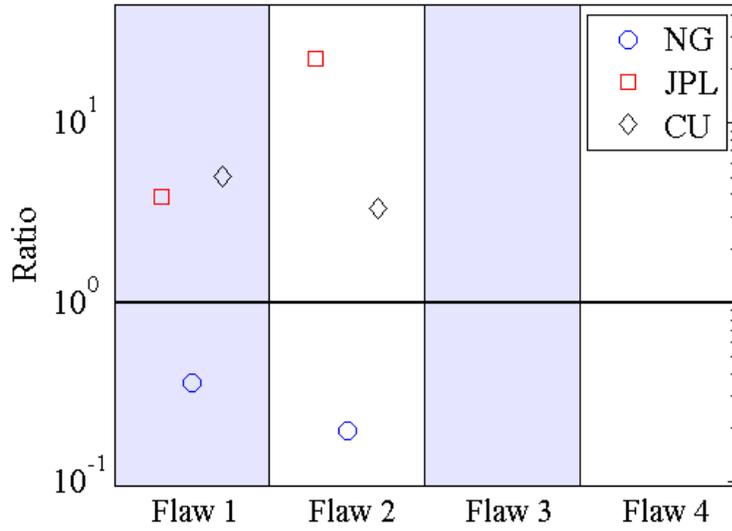
Figure 5-6. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the HG tip truncation Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

IZ5 Tip Truncation: Model Predictions
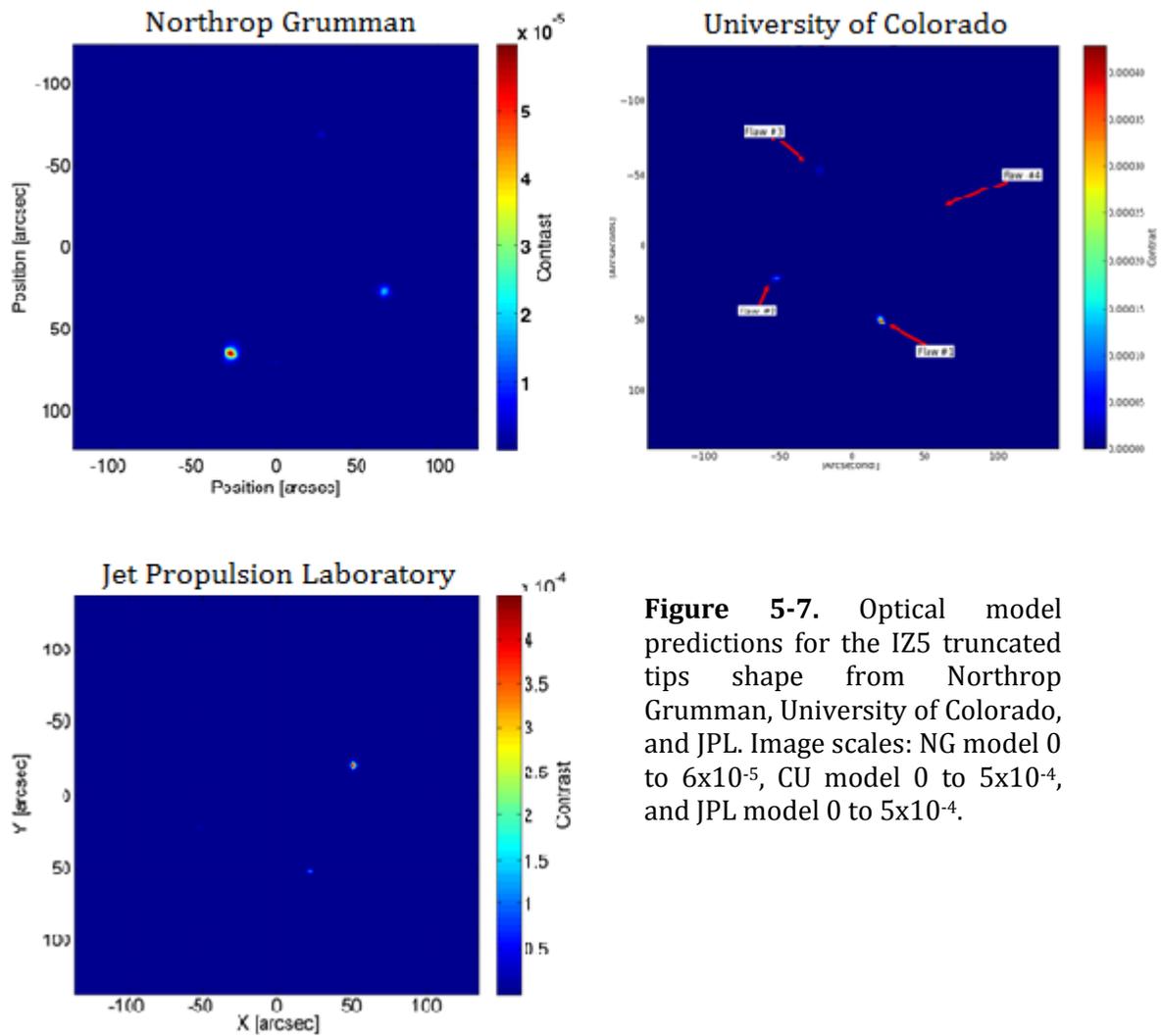
Figure 5-7. Optical model predictions for the IZ5 truncated tips shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 6×10^{-5}, CU model 0 to 5×10^{-4}, and JPL model 0 to 5×10^{-4}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	5.98E-05	4.30E-04	4.50E-04
Flaw 2	1.90E-05	1.32E-04	1.20E-04
Flaw 3	3.20E-06	2.93E-05	2.10E-05
Flaw 4	5.46E-07	4.54E-06	3.10E-06

Table 5-7. IZ5 tip truncation flaw brightnesses computed by each of the models.

IZ5 Tip Truncation: Measured Contrast Figures
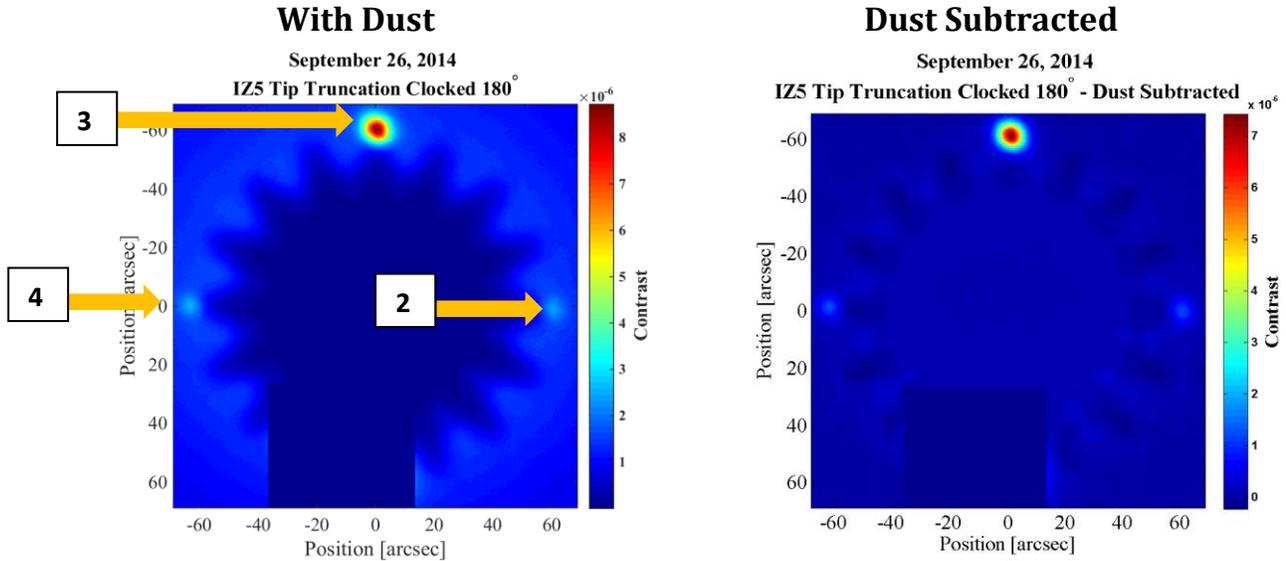
Figure 5-8. Measured contrast maps for the IZ5 edge shape with the tip truncation error, with dust (left) and dust-subtracted (right).

IZ5 Tip Truncation: Measured Contrast Numbers

The largest flaw on the IZ5 tip truncated Starshade produced a peak that overwhelmed the images – it was brighter even than the ND1 planet. Therefore, for the second set of these observations we clocked the Starshade 180° to place that flaw behind the stand. Clearly, the measured brightness of flaw 2 appears very out of family with the other flaws. There were no noted indications during the observations of why that would be.

	Flaw #	Measured Peak	Uncertainty
With Dust	1	6.48E-05	± 9.88E-06
	2	8.58E-07*	± 1.47E-07
	3	7.23E-06	± 1.19E-06
	4	1.12E-06	± 1.87E-07
Dust Subtracted	1	6.48E-05	± 9.89E-06
	2	8.58E-07*	± 1.49E-07
	3	7.23E-06	± 1.19E-06
	4	1.12E-06	± 1.88E-07

Table 5-8. Measured contrast peaks and comparison to the models for the IZ5 edge shape with the tip truncation error. * Flaw 2 measurement appears to be very out of family with the other flaws.
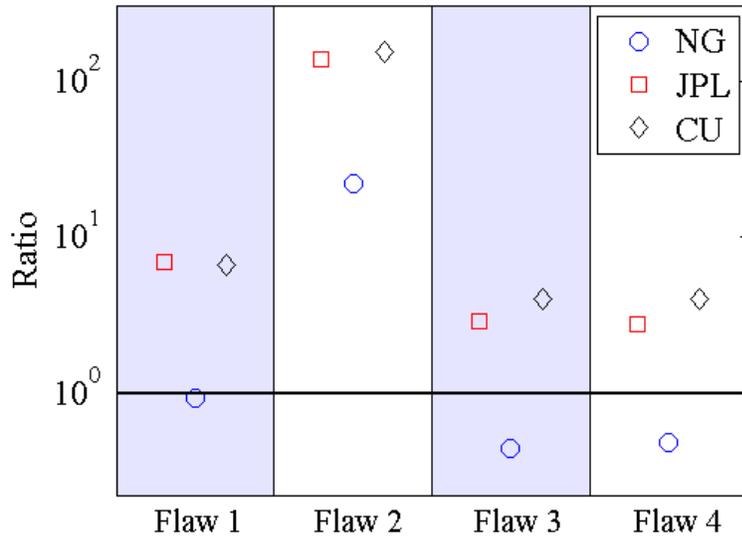
Figure 5-9. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the IZ5 tip truncation Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

5.4.2 PETAL WIDTH VARIATION (SHRUNK PETALS):

The width of the petal is changed by a given percentage. A negative amplitude means the petal width is smaller by that percentage at all radii.
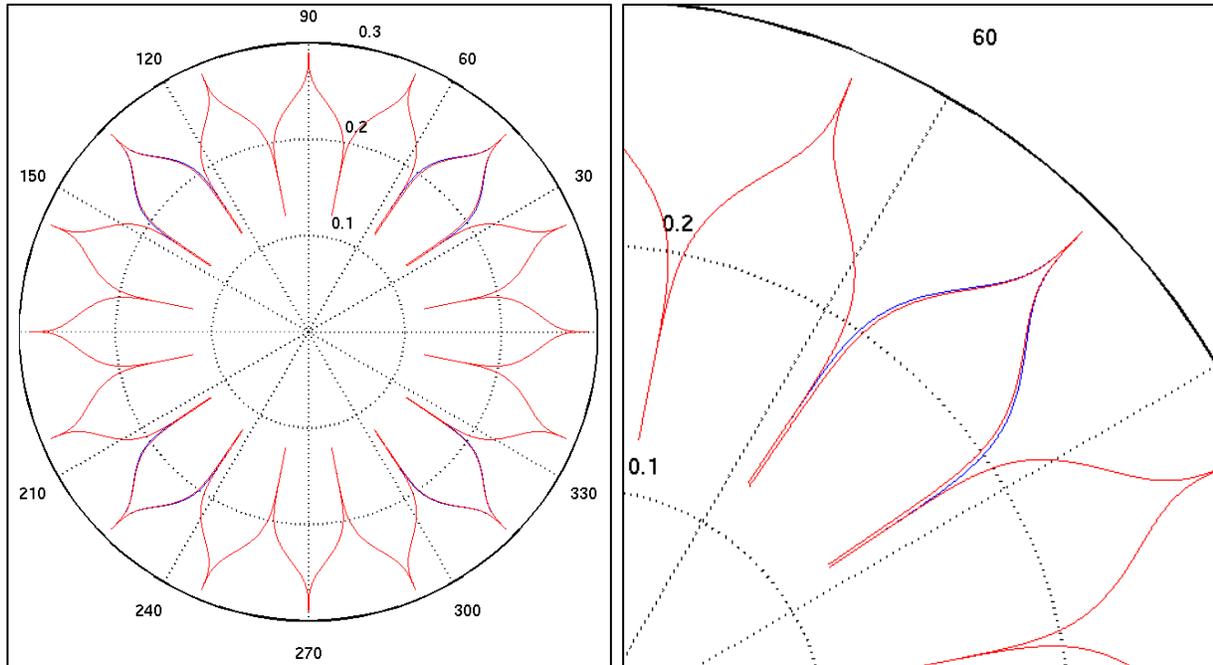
Figure 5-10: Petal Width Variation Shape input for Optical Models and Starshade manufacturing. Variations in the width of four petals from the base HG design (blue). The shape with errors is in red, but is only different from the base shape at the four error locations. The right-hand figure is zoomed in to one petal to show details of the deviation.

HG Shrunk Petals: Model Predictions
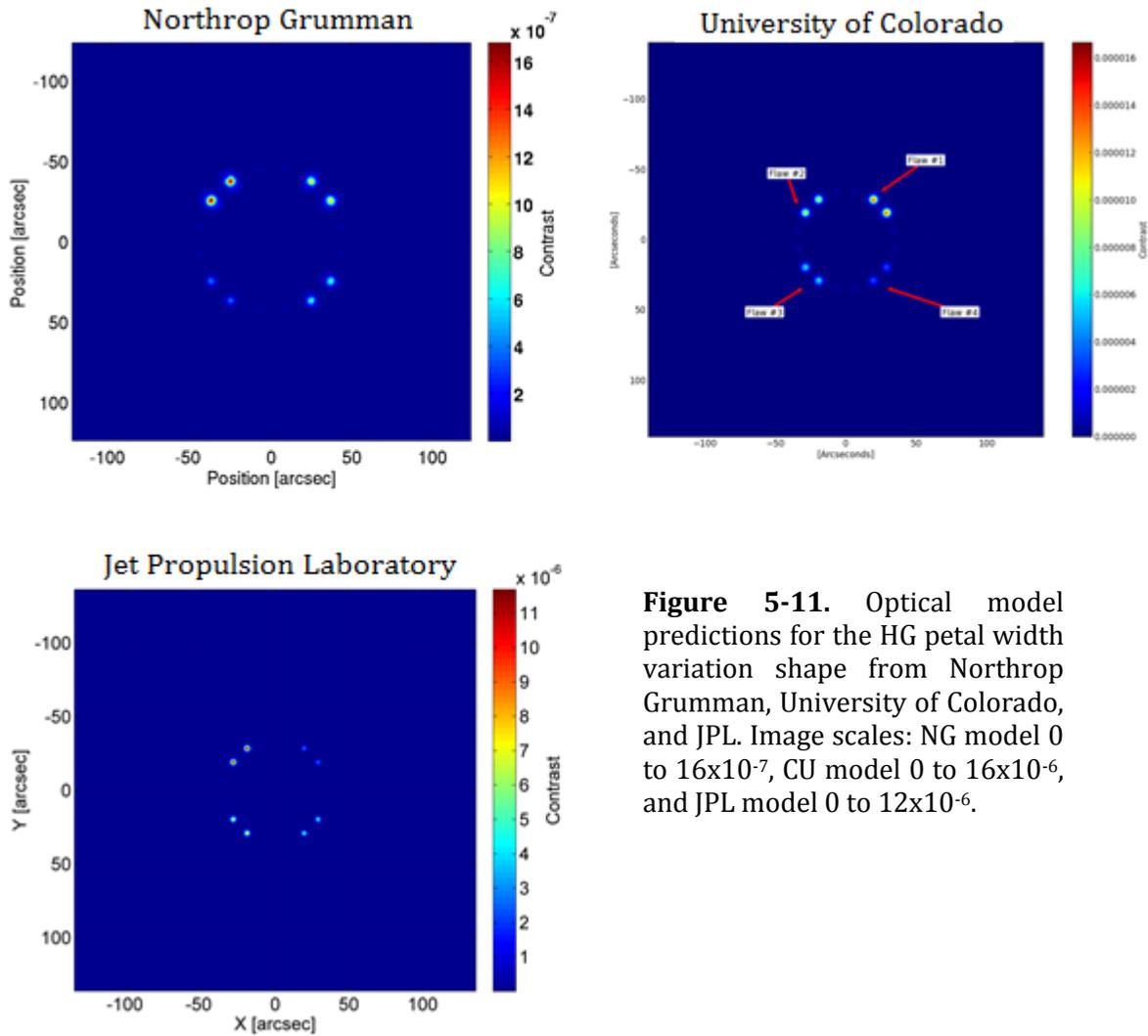
Figure 5-11. Optical model predictions for the HG petal width variation shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 16×10^{-7}, CU model 0 to 16×10^{-6}, and JPL model 0 to 12×10^{-6}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	1.68E-06	1.66E-05	1.20E-05
Flaw 2	1.15E-06	1.14E-05	8.10E-06
Flaw 3	7.02E-07	2.07E-05	5.10E-06
Flaw 4	4.04E-07	3.97E-06	2.90E-06

Table 5-9. HG petal width variation flaw brightnesses computed by each of the models.

HG Shrunk Petals: Measured Contrast Figures

With Dust
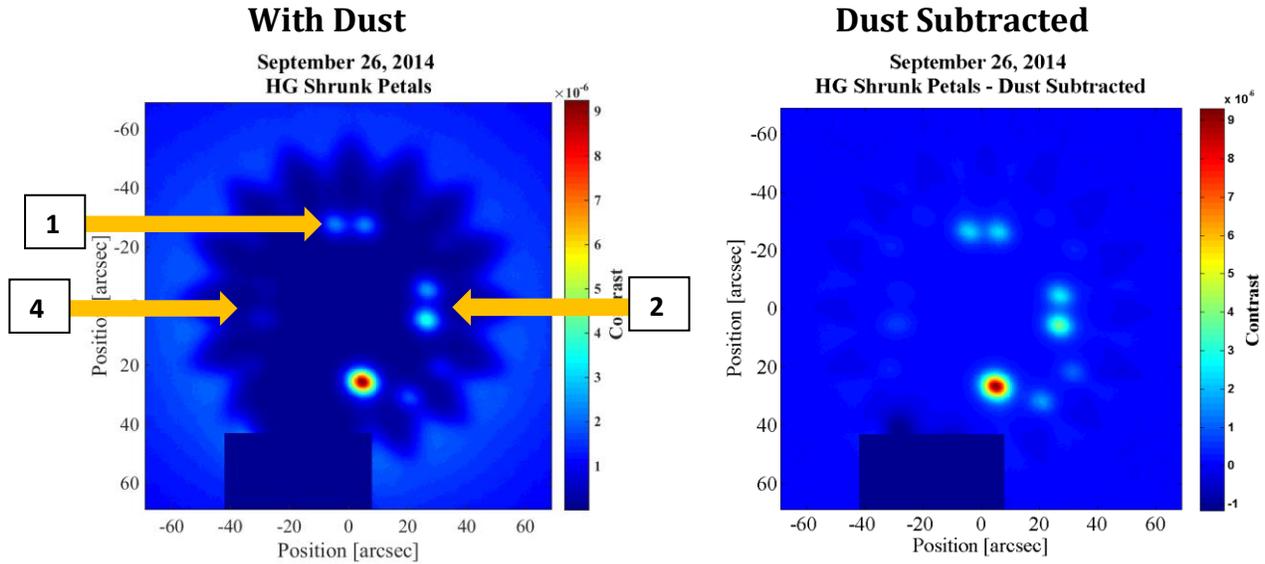
Dust Subtracted

Figure 5-12. Measured contrast maps for the HG edge shape with the petal width variation error, with dust (left) and dust-subtracted (right).

HG Shrunk Petals: Measured Contrast Numbers

The model predictions for the HG shrunk petals Starshade indicate that each pair of points on either side of the shrunk petal would be equal in brightness. For observations at positions 2 and 4, however, these pairs have significantly different brightnesses, with the lower point being brighter by about 30% than the upper point. We can use this information to infer how much the system is out of alignment. The measured peak for each flaw is the average of the two peaks. The bright single peak near the 6:00 position is likely due to an interaction with the stand and was not considered in our analysis.

	Flaw #	Measured Peak	Uncertainty
With Dust	1	3.06E-06	± 6.03E-07
	2	2.93E-06	± 5.40E-07
	3		
	4	3.63E-07	± 7.37E-08
Dust Subtracted	1	3.06E-06	± 5.57E-07
	2	2.93E-06	± 5.34E-07
	3		
	4	3.63E-07	± 7.14E-08

Table 5-10. Measured contrast peaks and comparison to the models for the HG edge shape with the petal width variation error.

HG Shrunk Petals: Model/Experiment Ratios
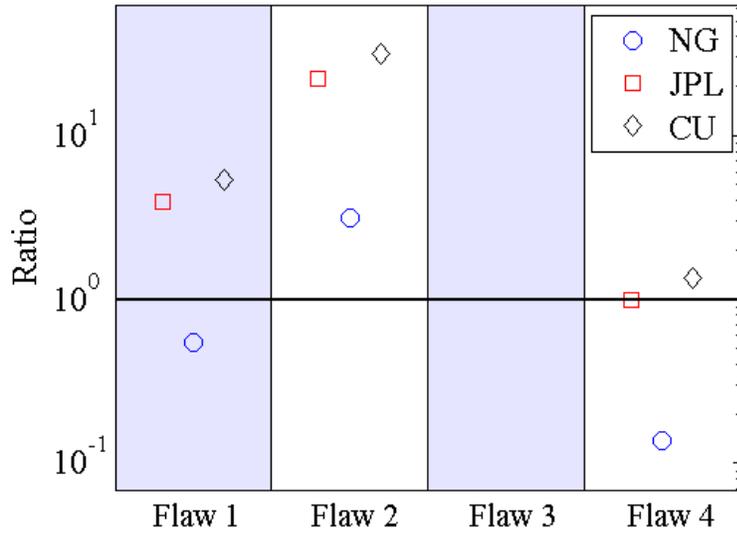
Figure 5-13. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the HG shrunk petal Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

IZ5 Shrunk Petals: Model Predictions
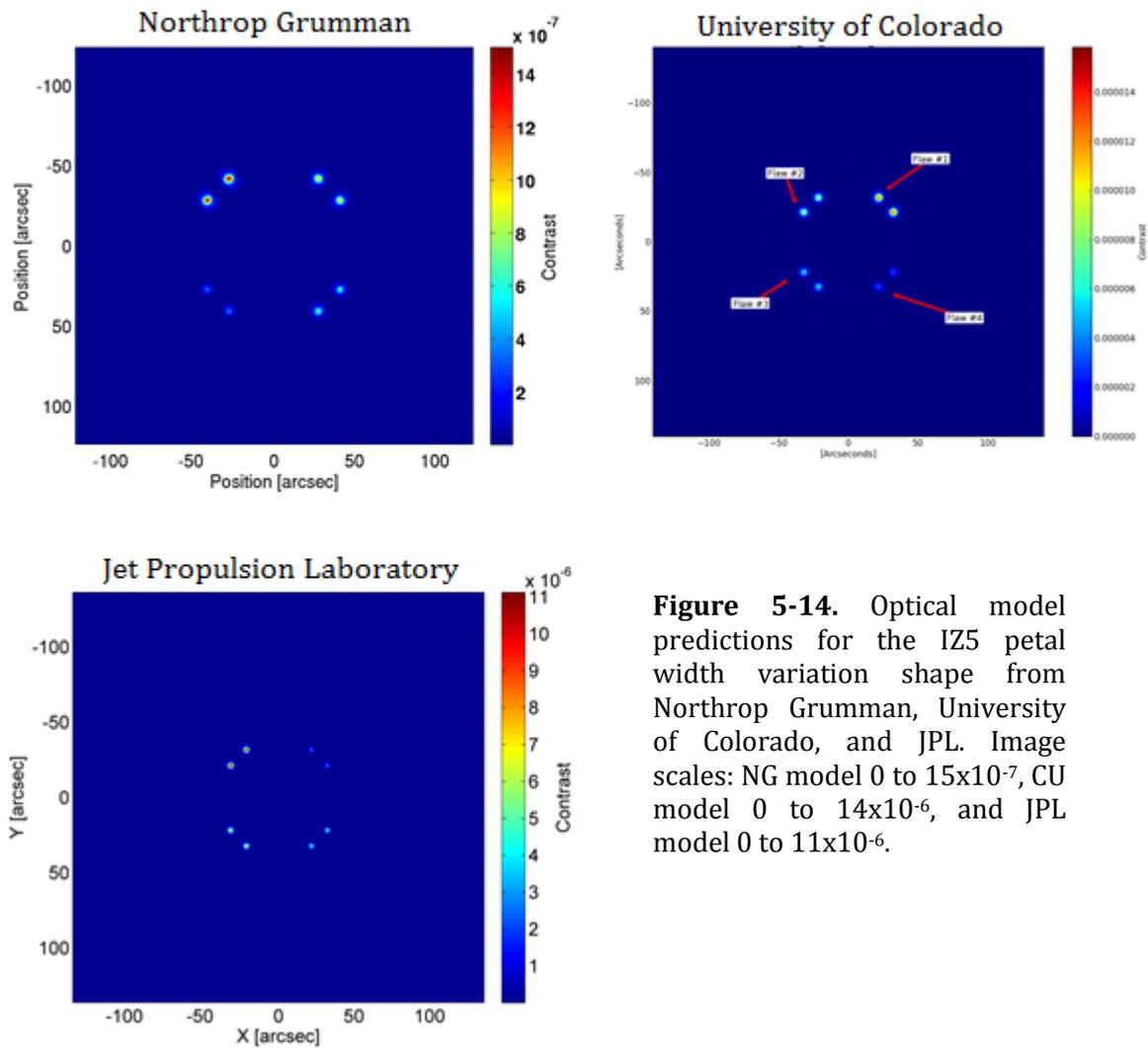
Figure 5-14. Optical model predictions for the IZ5 petal width variation shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 15×10^{-7}, CU model 0 to 14×10^{-6}, and JPL model 0 to 11×10^{-6}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	1.50E-06	1.58E-05	1.10E-05
Flaw 2	1.02E-06	1.07E-05	7.60E-06
Flaw 3	6.50E-07	6.53E-06	4.80E-06
Flaw 4	3.47E-07	3.59E-06	2.60E-06

Table 5-11. IZ5 petal width variation flaw brightnesses computed by each of the models.

IZ5 Shrunk Petals: Measured Contrast Figures
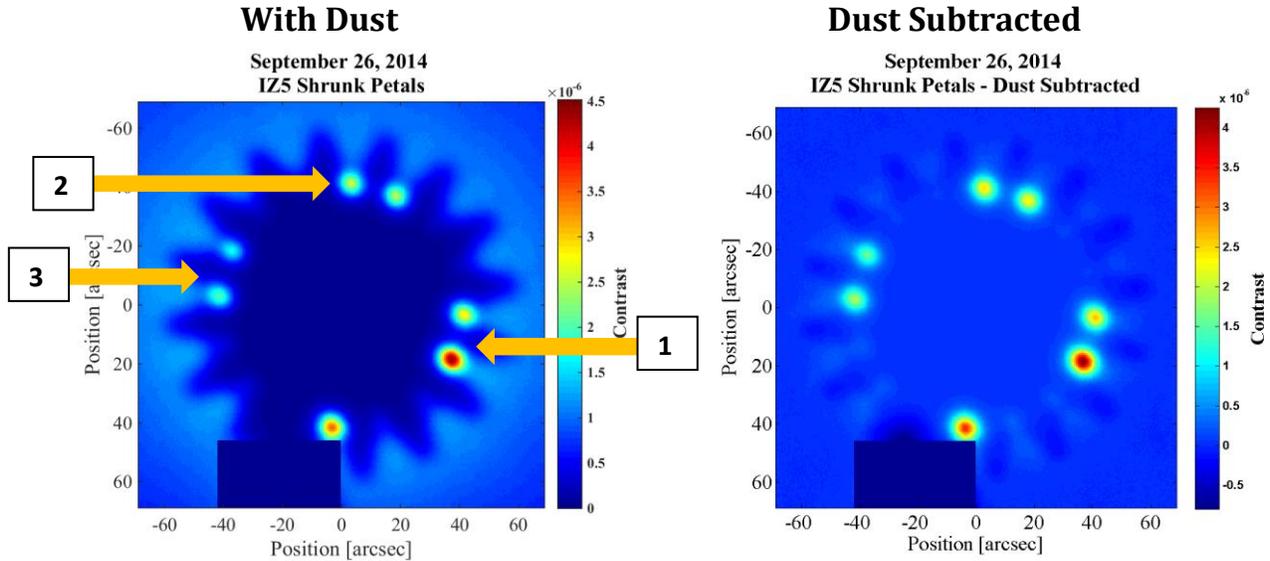
Figure 5-15. Measured contrast maps for the IZ5 edge shape with the petal width variation error, with dust (left) and dust-subtracted (right).

IZ5 Shrunk Petals: Measured Contrast Numbers

We were able to determine peak brightnesses for all four flaws in the IZ5 shrunk petals Starshade. There is likely some influence due to misalignment in this series, which is difficult to estimate due to the slight rotation of the Starshade. The peaks do generally follow the model predictions in that the larger errors resulted in brighter flaws, but the trend deviates from the predictions more as the size of the flaw decreases. The bright spot at 6:00 is likely due to the stand.

	Flaw #	Measured Peak	Uncertainty
With Dust	1	3.09E-06	± 6.61E-07
	2	2.24E-06	± 4.78E-07
	3	1.62E-06	± 3.53E-07
	4	1.33E-07	± 2.95E-07
Dust Subtracted	1	3.09E-06	± 6.52E-07
	2	2.49E-06	± 5.36E-07
	3	2.24E-06	± 4.80E-07
	4	1.80E-06	± 3.90E-07

Table 5-12. Measured contrast peaks and comparison to the models for the IZ5 edge shape with the petal width variation error.

IZ5 Shrunk Petals: Model/Experiment Ratios
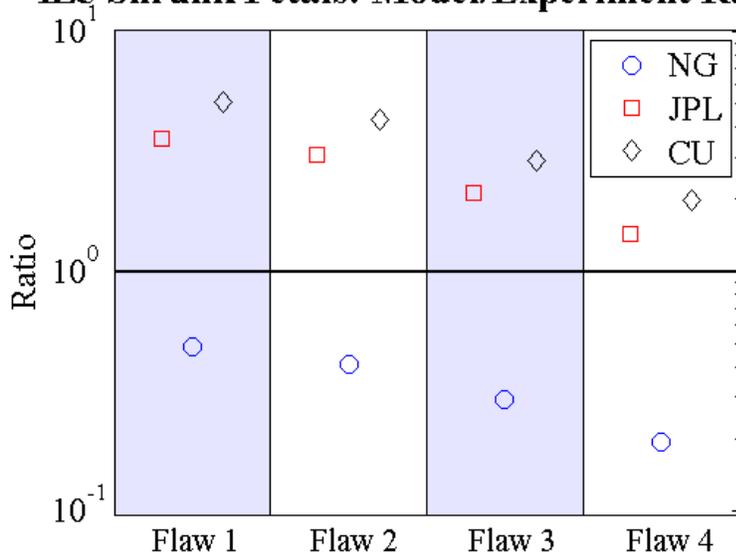
Figure 5-16. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the IZ5 shrunk petal Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

5.4.3 PETAL CLOCKING:

A petal is shifted in plane, with the valleys fixed and the tip moving the most. The motion varies linearly with radius from the valley to the tip. The amplitude is given as the change in angle of the petal with respect to the center of the Starshade.
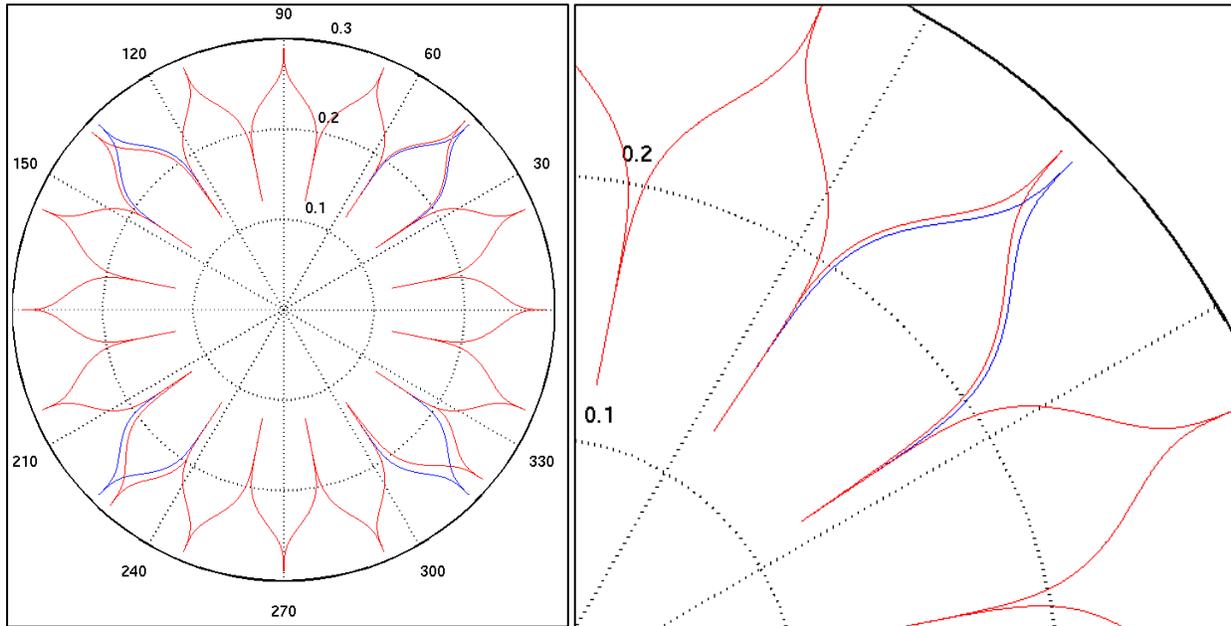
Figure 5-17: Clocking of Four Petals from the Base HG Design (blue). The shape with errors is in red, but is only different from the base shape at the four error locations. The right-hand figure is zoomed in to one petal to show details of the deviation.

HG Petal Clocking: Model Predictions
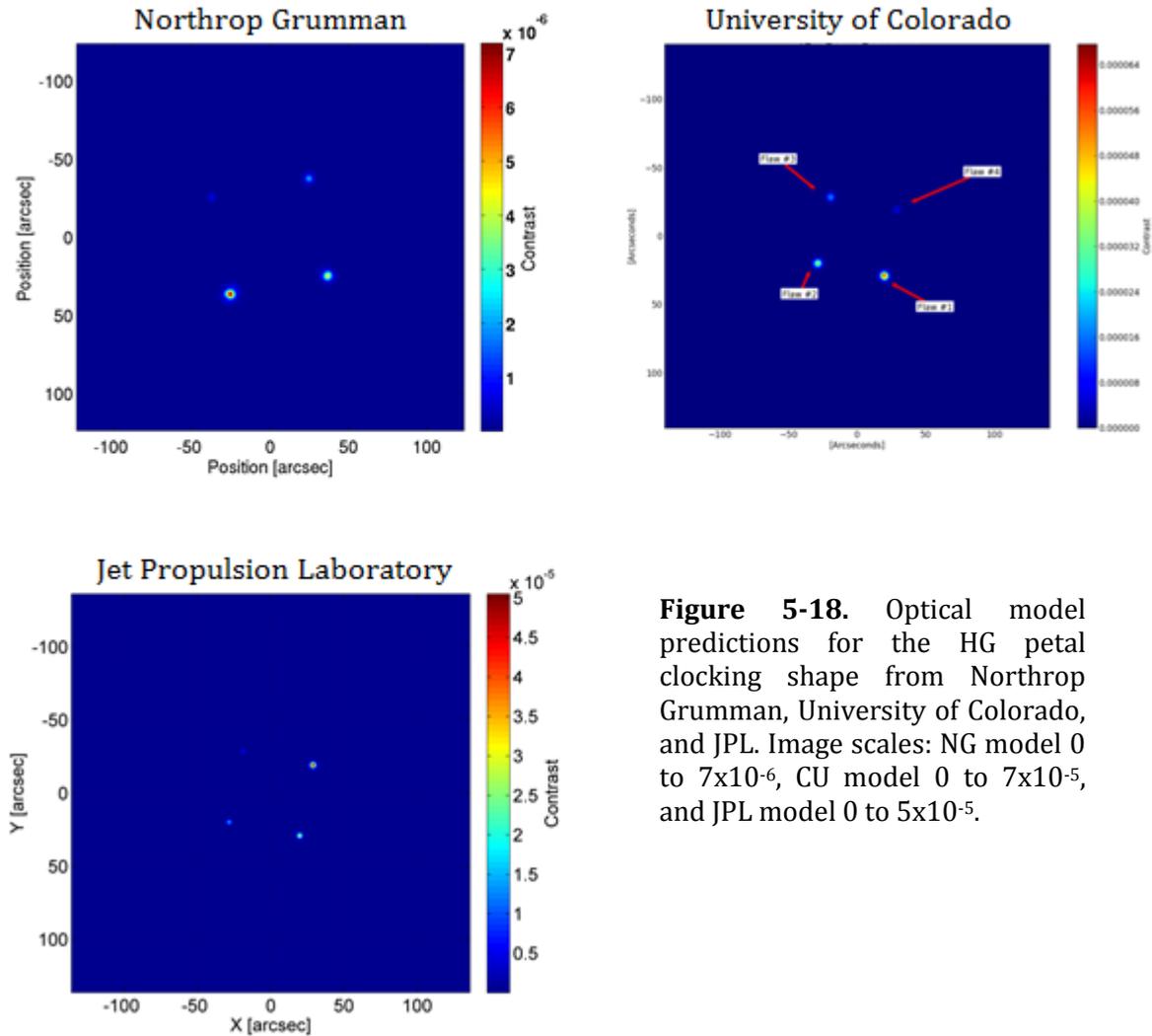
Figure 5-18. Optical model predictions for the HG petal clocking shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 7x10^{-6}, CU model 0 to 7x10^{-5}, and JPL model 0 to 5x10^{-5}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	7.18E-06	6.76E-05	5.00E-05
Flaw 2	4.32E-06	4.13E-05	3.00E-05
Flaw 3	2.09E-06	2.07E-05	1.50E-05
Flaw 4	6.92E-07	6.71E-06	4.80E-06

Table 5-13. HG petal clocking flaw brightnesses computed by each of the models.

HG Petal Clocking: Measured Contrast Figures

With Dust	Dust Subtracted
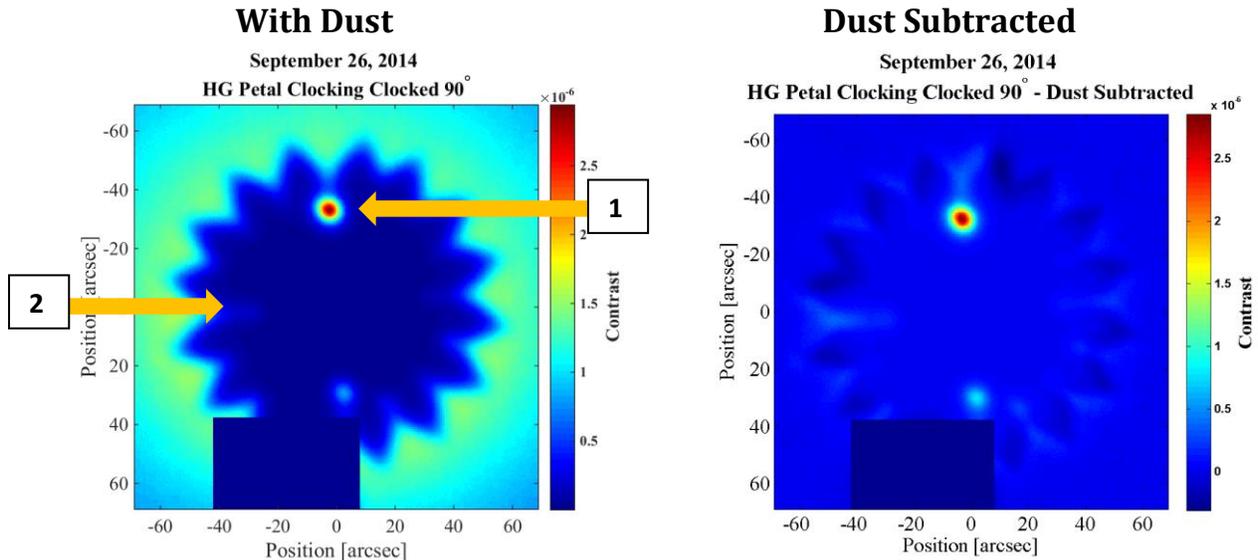
Figure 5-19. Measured contrast maps for the HG edge shape with the petal clocking error, with dust (left) and dust-subtracted (right).

HG Petal Clocking: Measured Contrast Numbers

Flaw 3 in the image is most likely interacting with the stand and was not considered in our analysis. Our data for flaw 3 is taken from a separate set of images with the Starshade rotated by 90 degrees. All models over-predicted the peaks of these flaws.

	Flaw #	Measured Peak	Uncertainty
With Dust	1	1.46E-06	± 2.82E-07
	2	1.54E-07	± 3E-08
	3	3.1E-07	± 1.01E-08
	4		
Dust Subtracted	1	1.46E-06	± 2.82E-07
	2	1.54E-07	± 4.09E-08
	3	5.26E-07	± 1.01E-07
	4	1.46E-06	± 2.82E-07

Table 5-14. Measured contrast peaks and comparison to the models for the HG edge shape with the petal clocking error.

HG Clocked Petals: Model/Experiment Ratios
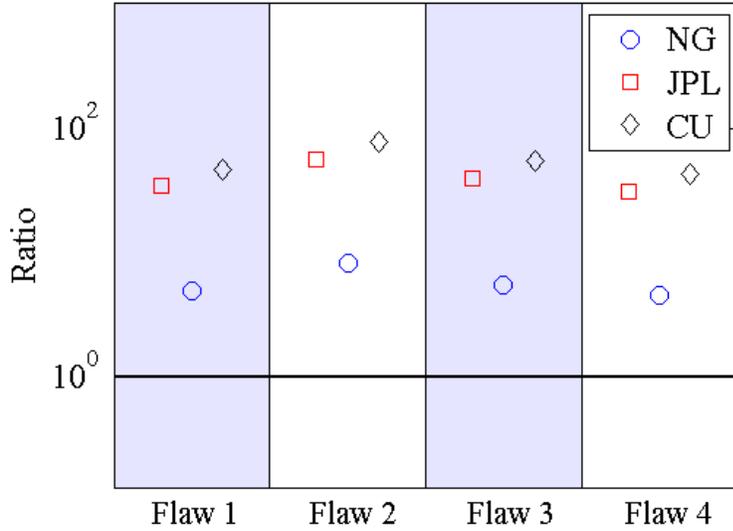
Figure 5-20. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the HG clocked petals Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

IZ5 Petal Clocking: Model Predictions
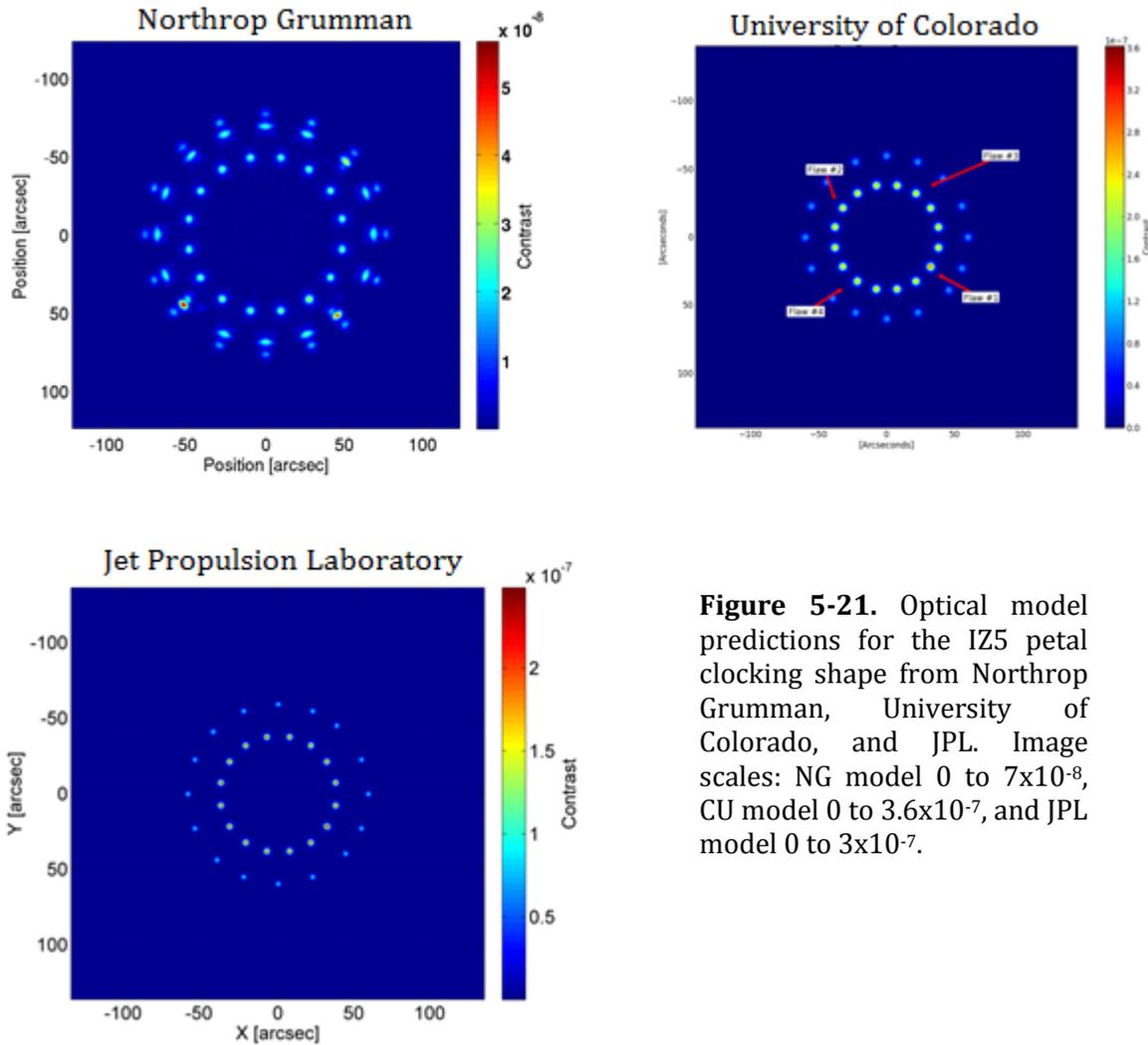
Figure 5-21. Optical model predictions for the IZ5 petal clocking shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 7×10^{-8}, CU model 0 to 3.6×10^{-7}, and JPL model 0 to 3×10^{-7}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	7.18E-06	6.76E-05	5E-05
Flaw 2	4.32E-06	4.13E-05	3E-05
Flaw 3	2.1E-06	2.07E-05	1.5E-05
Flaw 4	6.92E-07	6.71E-06	4.8E-05

Table 5-15. IZ5 petal clocking flaw brightnesses computed by each of the models.

IZ5 Petal Clocking: Measured Contrast Figures

With Dust   **Dust Subtracted**
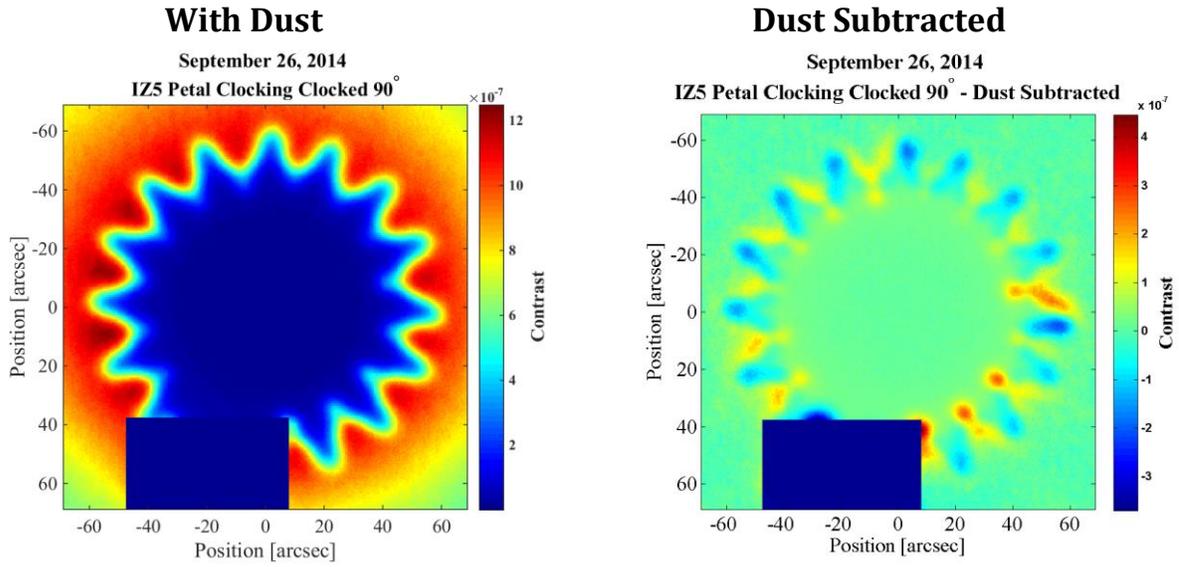
Figure 5-22. Measured contrast maps for the IZ5 edge shape with the petal clocking error, with dust (left) and dust-subtracted (right).

*Our post-processing was unable to extract any flaw peaks from the observational data for this shape.

5.4.4 EDGE SINES:

A sine wave with a specified frequency and amplitude is added to the position of the petal along the adjacent edges of two petals (both sides of one valley). This effectively makes the width of the petal vary sinusoidally versus radius compared to the base petal width. The parameters specify both the frequency (N in cycles per petal edge) and the amplitude (in cm).
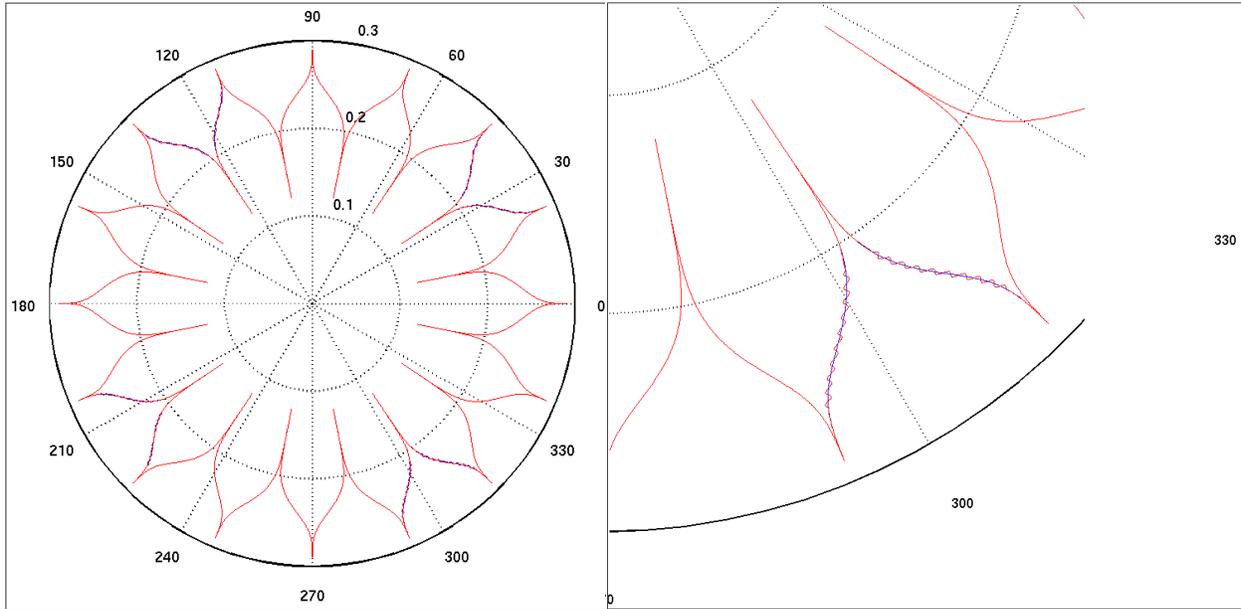
Figure 5-23: Sine wave variations from the base HG design (blue) along the edges of four petals. The shape with errors is in red, but is only different from the base shape at the four error locations. The right-hand figure is zoomed in to one petal to show details of the deviation.

HG Sines: Model Predictions

Northrop Grumman

University of Colorado

Jet Propulsion Laboratory
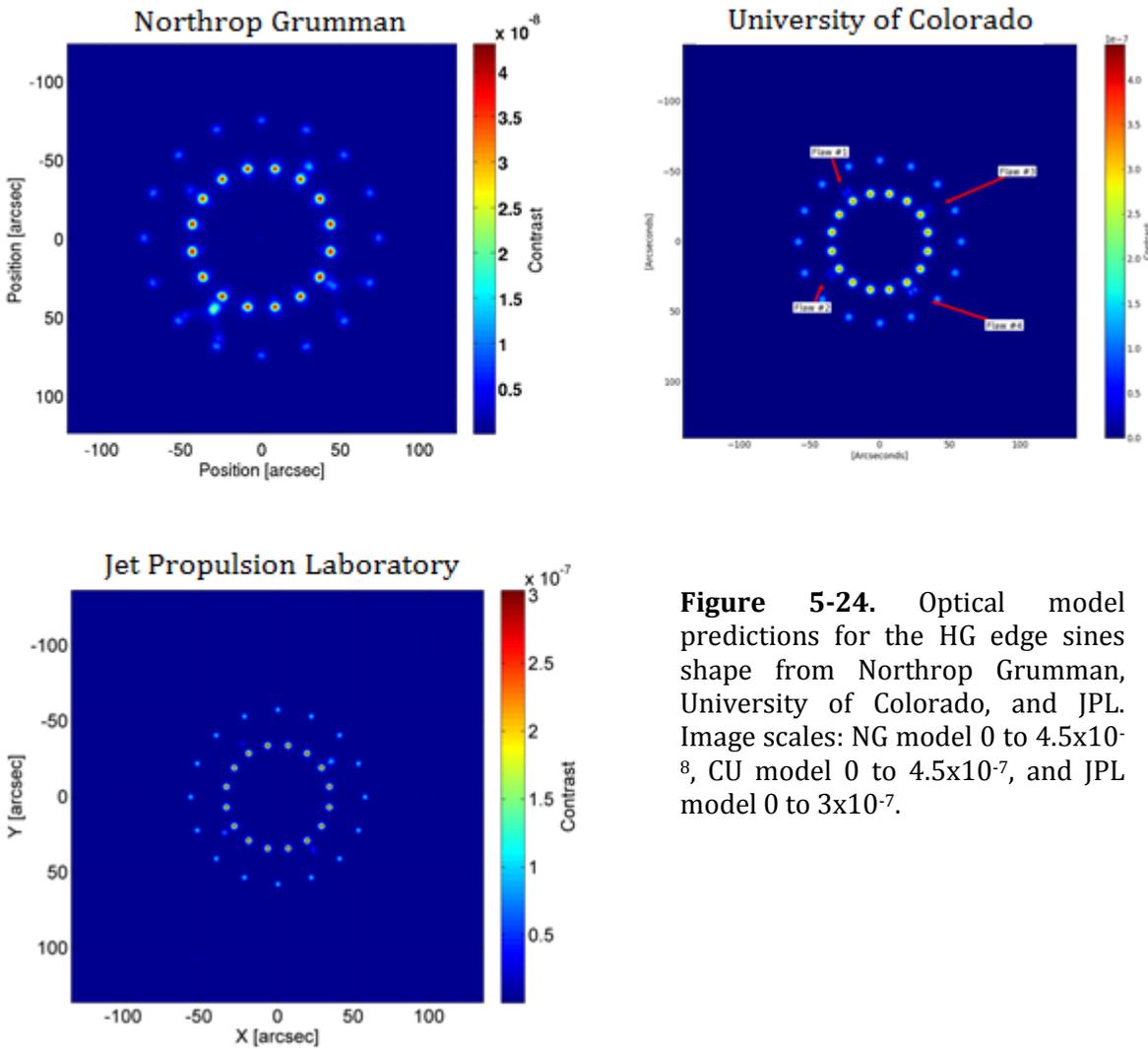
Figure 5-24. Optical model predictions for the HG edge sines shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 4.5×10^{-8}, CU model 0 to 4.5×10^{-7}, and JPL model 0 to 3×10^{-7}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	6.96E-09	8.57E-08	9.53E-08
Flaw 2	1.35E-08	4.72E-08	7.03E-08
Flaw 3	1.04E-08	4.43E-08	4.41E-08
Flaw 4	2.31E-08	3.67E-08	3.44E-08

Table 5-16. HG edge sines flaw brightnesses computed by each of the models.

HG Sines: Measured Contrast Figures

With Dust
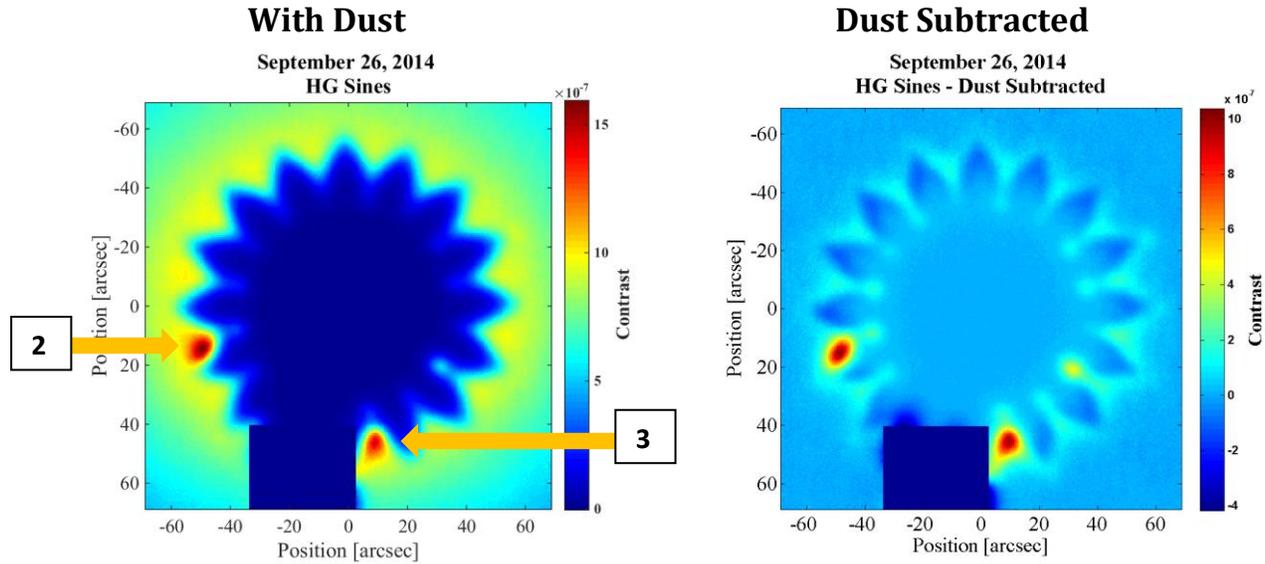
Dust Subtracted

Figure 5-25. Measured contrast maps for the HG edge shape with the edge sines error, with dust (left) and dust-subtracted (right).

HG Sines: Measured Contrast Numbers

The observations of the sine wave flaws produced surprising results. The flaws were predicted to have a number of close together features, which is consistent with the non-circular shapes of the peaks. Gaussian fitting to these peaks is less reliable due to this shape. In addition, the flaws are significantly (by 2 orders of magnitude) brighter than all of the model predictions. The two measurable flaws correspond to the sine waves with the small amplitude but two different frequencies. The larger amplitude flaws were not detectable in our observations (positions 1 and 4).

	Flaw #	Measured Peak	Uncertainty
With Dust	1		
	2	1.08E-06	± 1.86E-07
	3	5.06E-06	± 6.66E-07
	4		
Dust Subtracted	1		
	2	9.60E-07	± 1.85E-07
	3	3.21E-06	± 6.66E-07
	4		

Table 5-17. Measured contrast peaks and comparison to the models for the HG edge shape with the edge sines error.

HG Sines: Model/Experiment Ratios
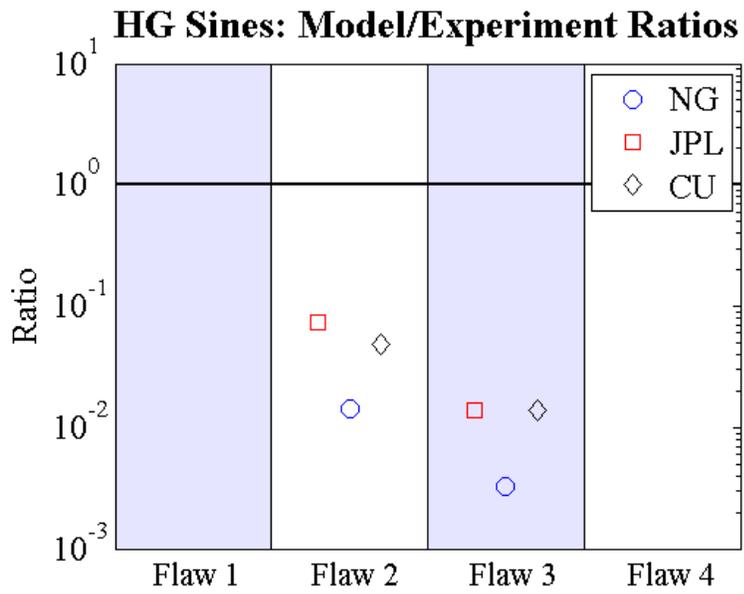
Figure 5-26. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the HG edge sines Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

IZ5 Sines: Model Predictions

Northrop Grumman

University of Colorado

Jet Propulsion Laboratory
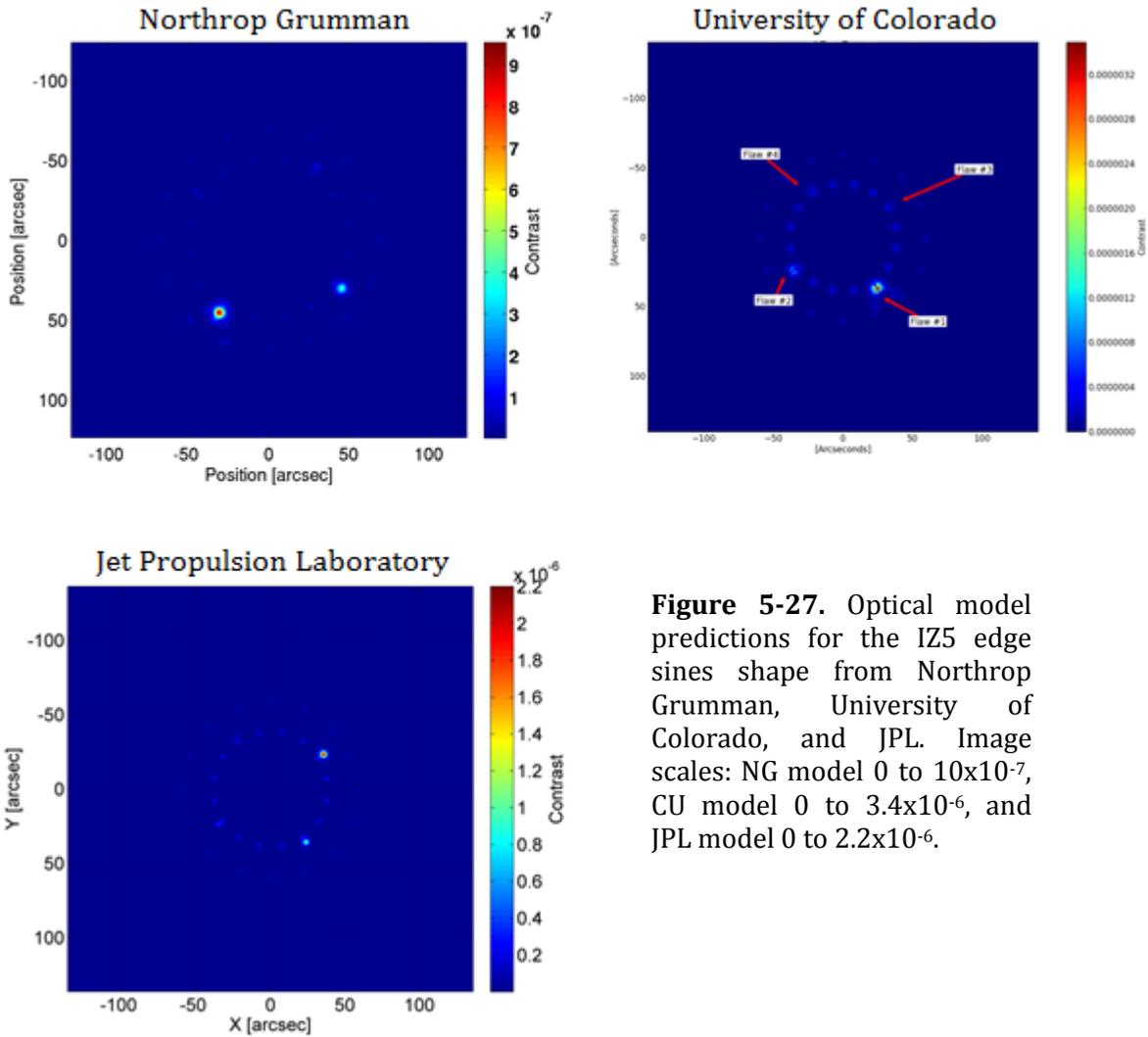
Figure 5-27. Optical model predictions for the IZ5 edge sines shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 10×10^{-7}, CU model 0 to 3.4×10^{-6}, and JPL model 0 to 2.2×10^{-6}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	3.56E-08	3.49E-06	2.20E-06
Flaw 2	7.65E-08	1.08E-06	1.00E-06
Flaw 3	4.04E-07	3.19E-07	3.40E-07
Flaw 4	9.54E-07	3.18E-07	1.58E-07

Table 5-18. IZ5 edge sines flaw brightnesses computed by each of the models.

IZ5 Sines: Measured Contrast Figures
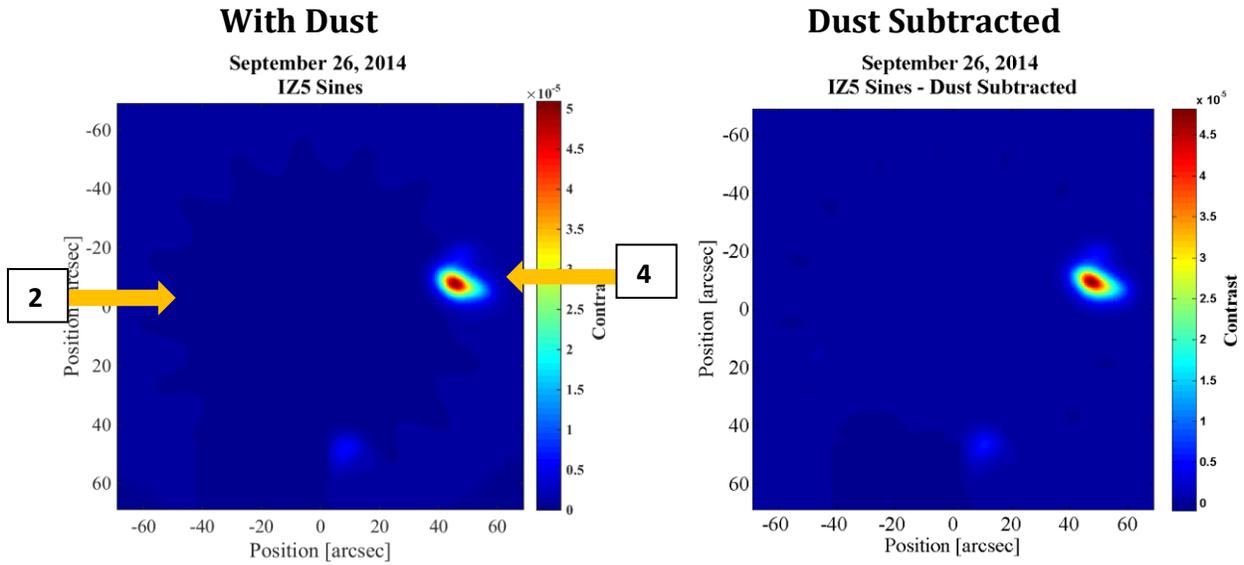
Figure 5-28. Measured contrast maps for the IZ5 edge shape with the edge sines error, with dust (left) and dust-subtracted (right).

IZ5 Sines: Measured Contrast Numbers

The observations of the IZ5 sines Starshade were very similar to those for the HG sines, in that the models under-predicted the peak values significantly. They also had a similar challenge in that they were non-circular and therefore were difficult to match with any one individual point in the model predictions.

	Flaw #	Measured Peak	Uncertainty
With Dust	1		
	2	1.26E-06	± 5.28E-06
	3	4.82E-05	± 6.06E-06
	4	5.23E-05	± 1.61E-05
Dust Subtracted	1		
	2	2.08E-05	± 5.28E-06
	3	2.44E-05	± 5.96E-06
	4	5.23E-05	± 1.60E-05

Table 5-19. Measured contrast peaks and comparison to the models for the IZ5 edge shape with the edge sines error.
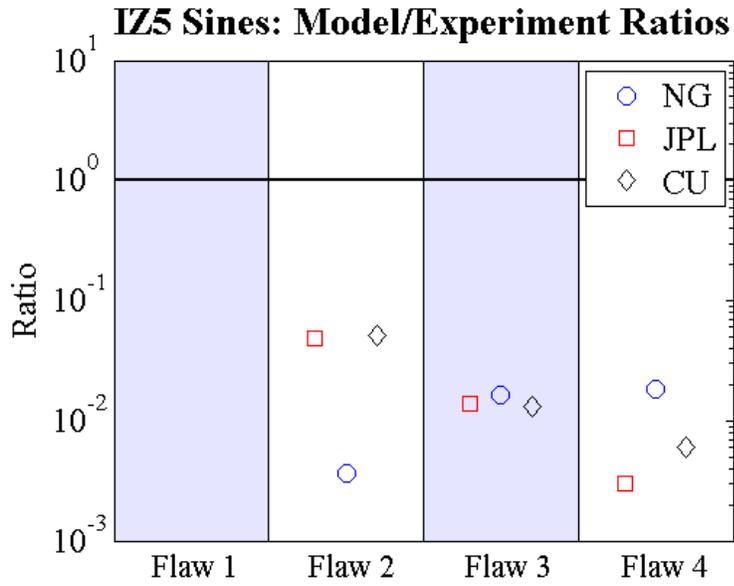
Figure 5-29. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the IZ5 edge sines Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

5.4.5 DISPLACED EDGES:

This case is included to represent an offset error in the edge of a Starshade.   Full-scale shades will likely need to have segmented edges, and there exists the possibility of misalignment of edge segments during assembly or a perturbation during launch or deployment. For this error shape, a region of length L along one petal edge is shifted outwards, with the new edge path remaining parallel to the base design. The length of the offset, L, is about 13 mm for all 4 added errors. The offset distance values, w, are 0.3, 0.6, 0.9, and 1.2 mm moving counter-clockwise around the Starshade.
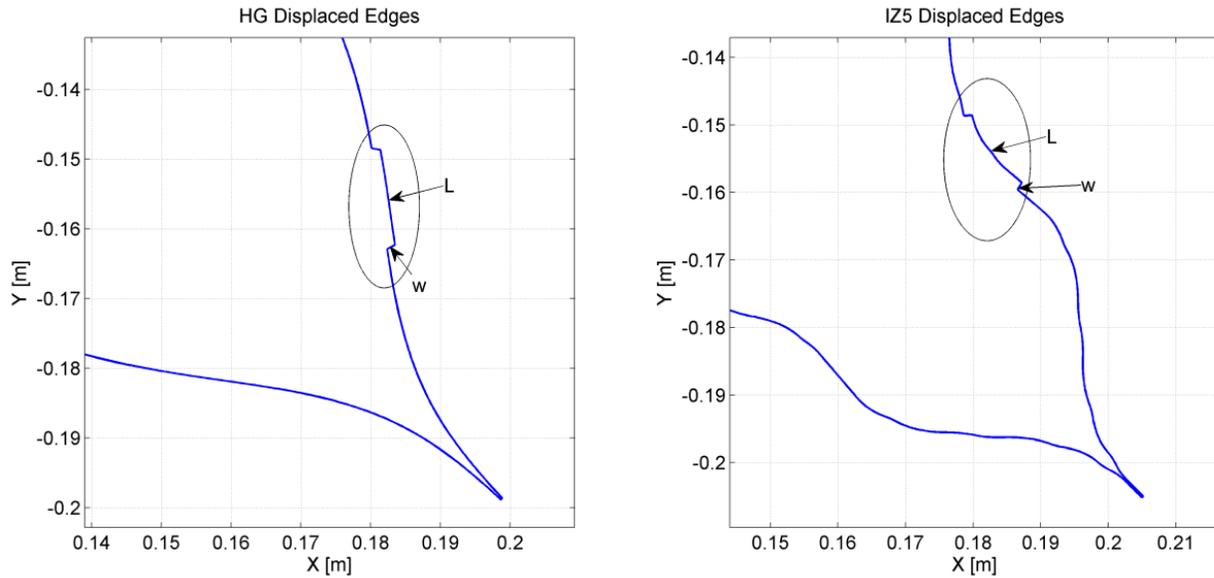
Figure 5-30: Examples of displaced petal errors on HG (left) and IZ5 (right) base shapes. L is the length of the displaced region, along the edge of the petal, and w is the amplitude of the displacement.

HG Displaced Edges: Model Predictions
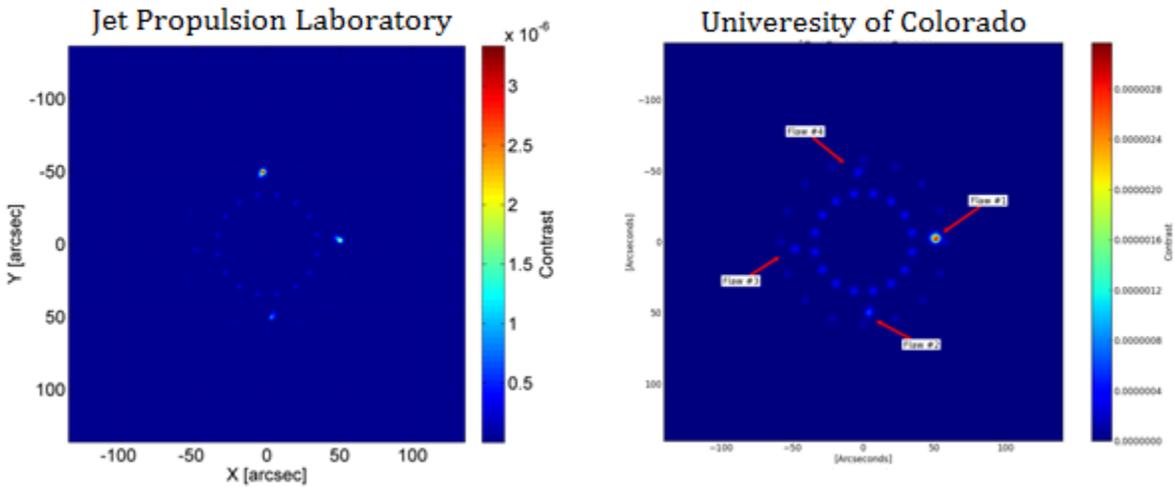
Figure 5-31. Optical model predictions for the HG displaced edges shape from University of Colorado and JPL. Image scales: CU model 0 to 3×10^{-6}, and JPL model 0 to 4×10^{-6}.

*The NG model capability of simulating non-smooth Starshades is in beta mode and was therefore not used to make predictions for this shape.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1		3.17E-06	3.30E-06
Flaw 2		6.32E-07	2.00E-06
Flaw 3		4.58E-07	9.10E-07
Flaw 4		3.66E-07	2.10E-07

Table 5-20. HG displaced edges flaw brightnesses computed by each of the models.

HG Displaced Edges: Measured Contrast Figures

With Dust
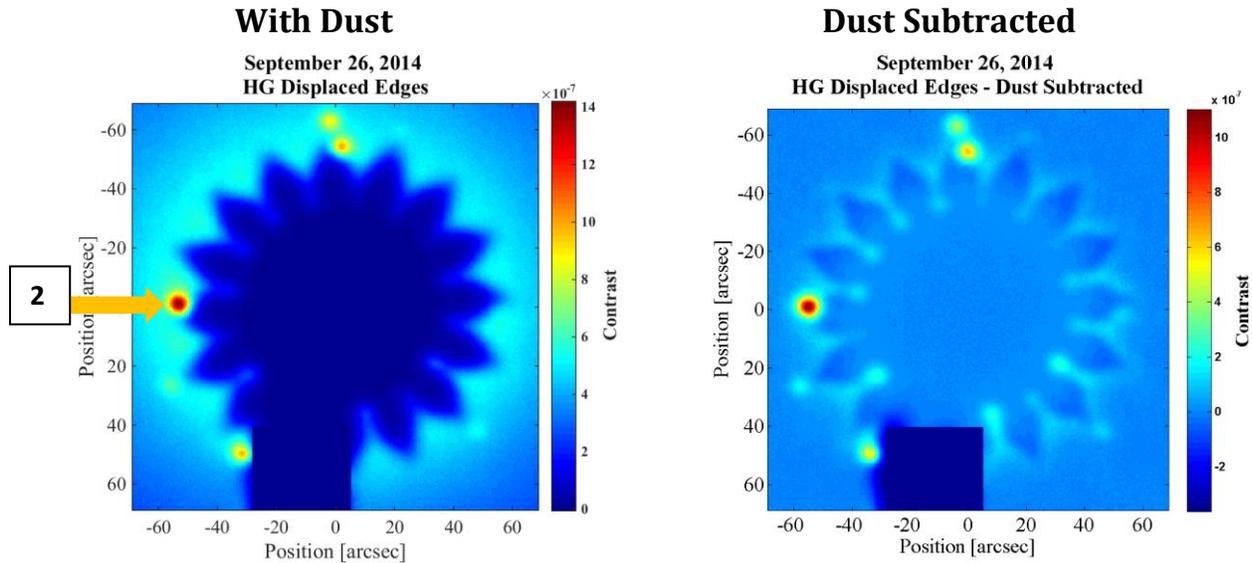
Dust Subtracted

Figure 5-32. Measured contrast maps for the HG edge shape with the displaced edges error, with dust (left) and dust-subtracted (right).

HG Displaced Edges: Measured Contrast Numbers

Model predictions for the HG displaced edges Starshade indicated that there would be a pair of bright peaks for each flaw at the corners of the displaced edge. Only one point for each observed flaw was visible, which is likely due to atmospheric blurring. The bright pair of points at the 12:00 position are the flaw and a broken tip. The point at 7:00 is likely due to the Starshade stand. In the dust-subtracted image the tips and valleys of the Starshade are clearly visible.

	Flaw #	Measured Peak	Uncertainty
With Dust	1		
	2	9.42E-07	± 1.73E-07
	3	3.17E-07	± 7.35E-08
	4		
Dust Subtracted	1		
	2	7.61E-07	± 1.73E-07
	3	3.17E-07	± 7.24E-08
	4		

Table 5-21. Measured contrast peaks and comparison to the models for the HG edge shape with the displaced edges error.

Figure 5-33. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the HG displaced edges Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

IZ5 Displaced Edges: Model Predictions
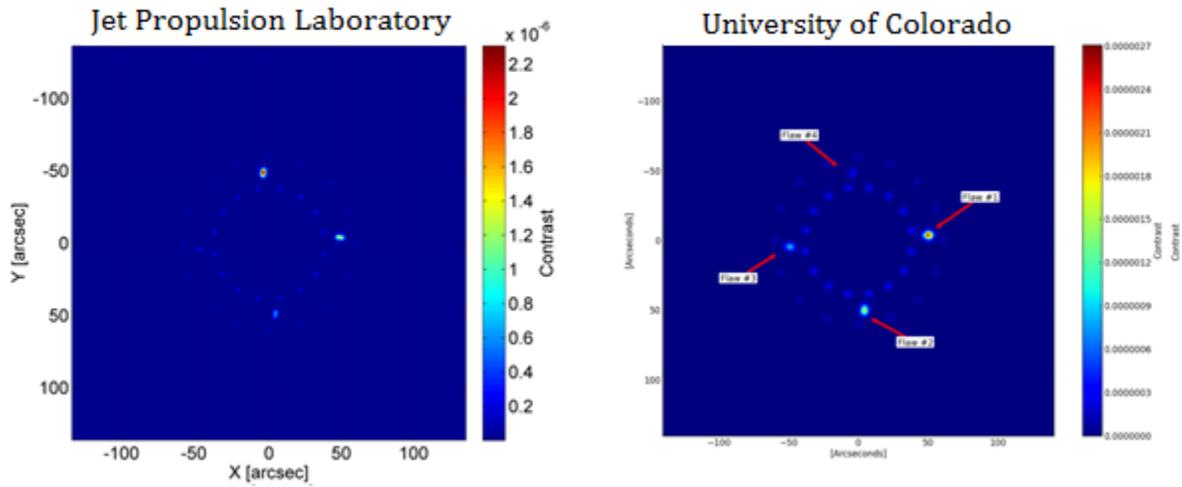
Figure 5-34. Optical model predictions for the IZ5 displaced edges shape from University of Colorado and JPL. Image scales: CU model 0 to $2x10^{-6}$, and JPL model 0 to $2.3x10^{-6}$.

*The NG model capability of simulating non-smooth Starshades is in beta mode and was therefore not used to make predictions for this shape.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1		2.71E-06	2.30E-06
Flaw 2		1.74E-06	1.30E-06
Flaw 3		8.95E-07	6.70E-07
Flaw 4		2.98E-07	1.90E-07

Table 5-22. IZ5 displaced edges flaw brightnesses computed by each of the models.

IZ5 Displaced Edges: Measured Contrast Figures

With Dust
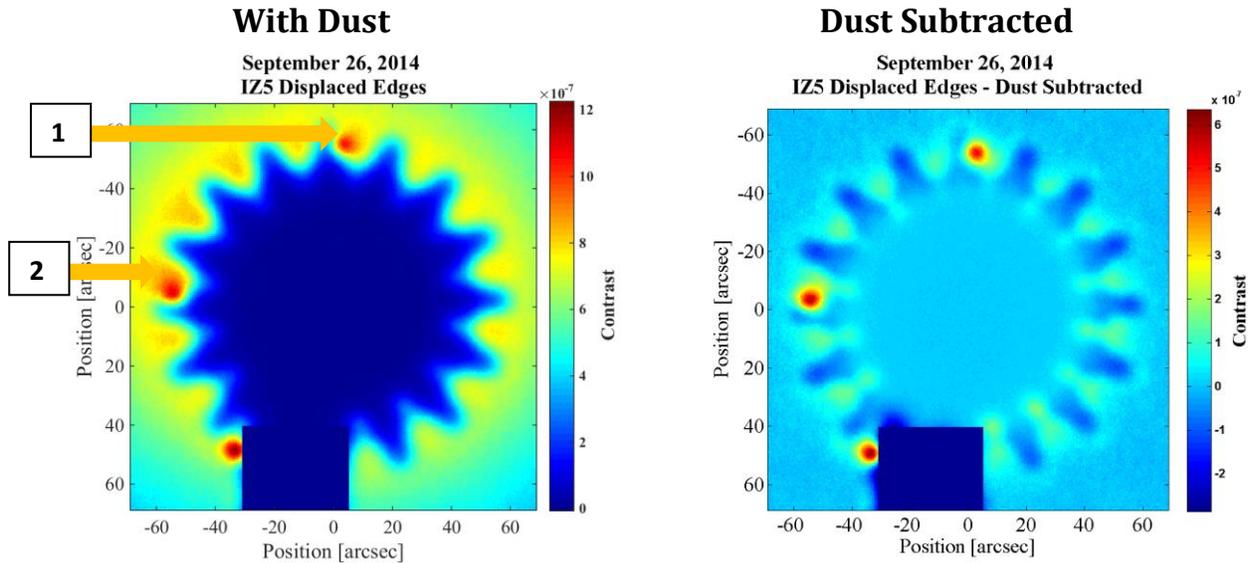
Dust Subtracted

Figure 5-35. Measured contrast maps for the IZ5 edge shape with the displaced edges error, with dust (left) and dust-subtracted (right).

IZ5 Displaced Edges: Measured Contrast Numbers

Model predictions for the IZ5 displaced edges Starshade indicated that there would be a pair of bright peaks for each flaw at the corners of the displaced edge. Only one point for each observed flaw was visible, which is likely due to atmospheric blurring. The point at 7:00 is likely due to the Starshade stand.

	Flaw #	Measured Peak	Uncertainty
With Dust	1	8.08E-07	± 1.89E-07
	2	7.05E-07	± 1.65E-07
	3	4.92E-07	± 1.16E-07
	4		
Dust Subtracted	1	8.08E-07	± 1.89E-07
	2	7.05E-07	± 1.65E-07
	3	4.92E-07	± 1.15E-07
	4		

Table 5-23. Measured contrast peaks and comparison to the models for the IZ5 edge shape with the displaced edges error.

IZ5 Displaced Edges: Model/Experiment Ratio
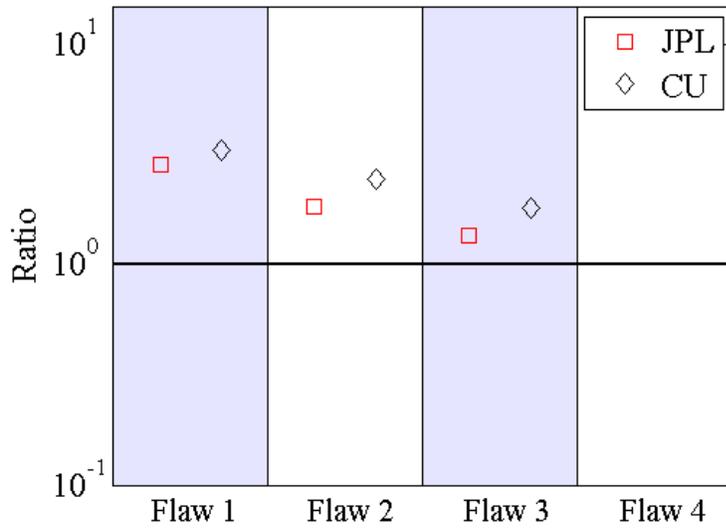
Figure 5-36. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the IZ5 displaced edges Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

5.4.6 *VALLEY TRUNCATION:*

The valley between a pair of petals is closed from the center to a specified radius. The amplitude given is the minimum radius of the petal for that valley. With no error the minimum petal radius is 16.91 cm for the base HG and 18.81 cm for the base IZ5.
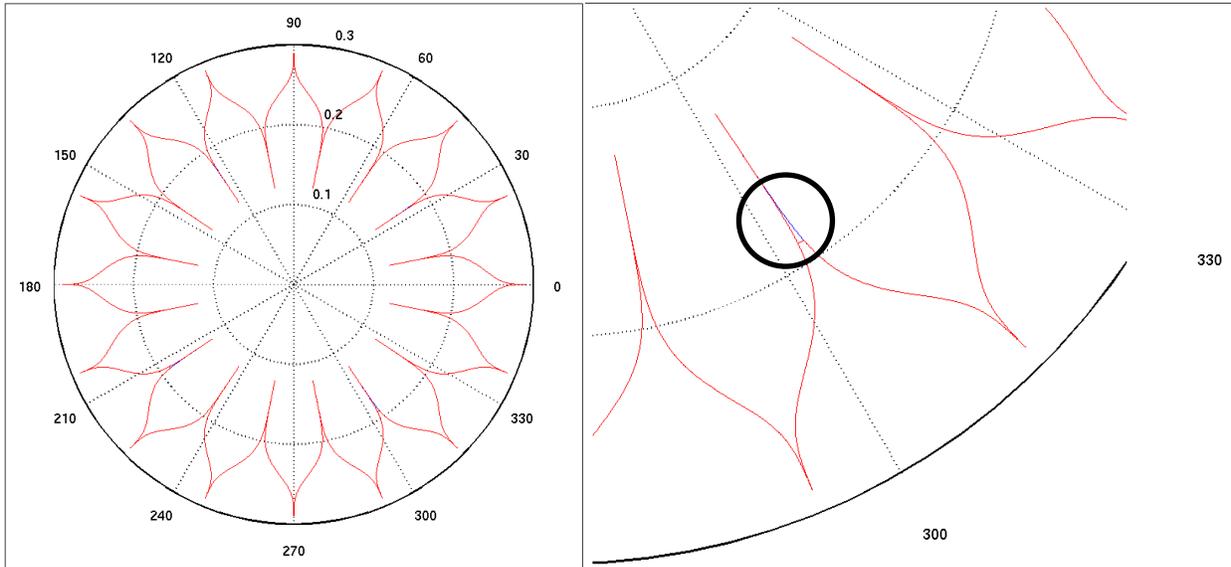
Figure 5-37: Truncation of the valleys at four locations relative to the base HG design (blue). The shape with errors is in red, but is only different from the base shape at the four error locations. The right-hand figure is zoomed in to one petal to show details of the deviation. The circle indicates the location of the truncated valley.

HG Valley Truncation: Model Predictions
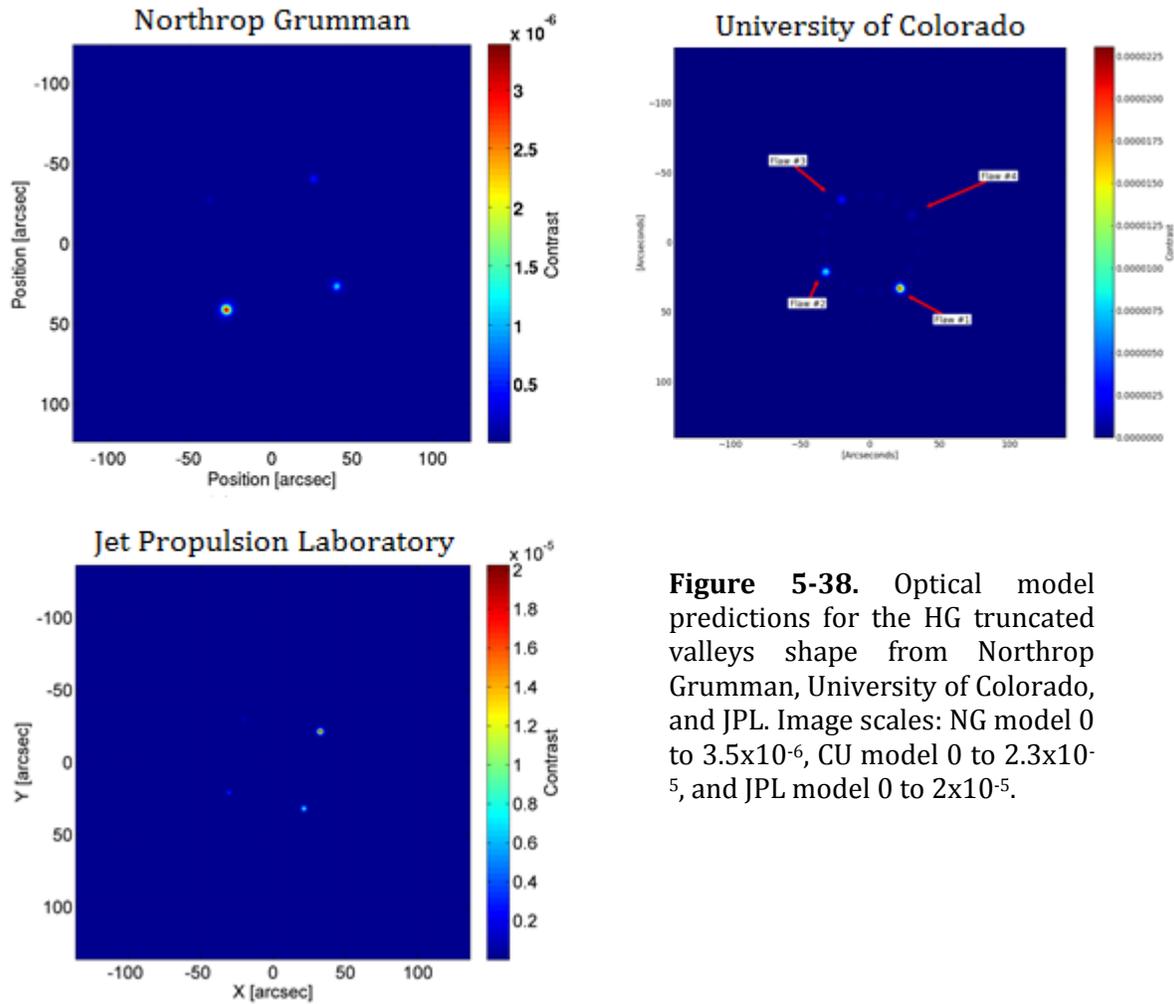
Figure 5-38. Optical model predictions for the HG truncated valleys shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 3.5x10^{-6}, CU model 0 to 2.3x10^{-5}, and JPL model 0 to 2x10^{-5}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	3.38E-06	2.31E-05	2.00E-05
Flaw 2	8.63E-07	1.03E-05	8.80E-06
Flaw 3	4.59E-07	3.57E-06	3.20E-06
Flaw 4	2.12E-08	1.42E-06	1.20E-06

Table 5-24. HG truncated valleys flaw brightnesses computed by each of the models.

HG Valley Truncation: Measured Contrast Figures

With Dust
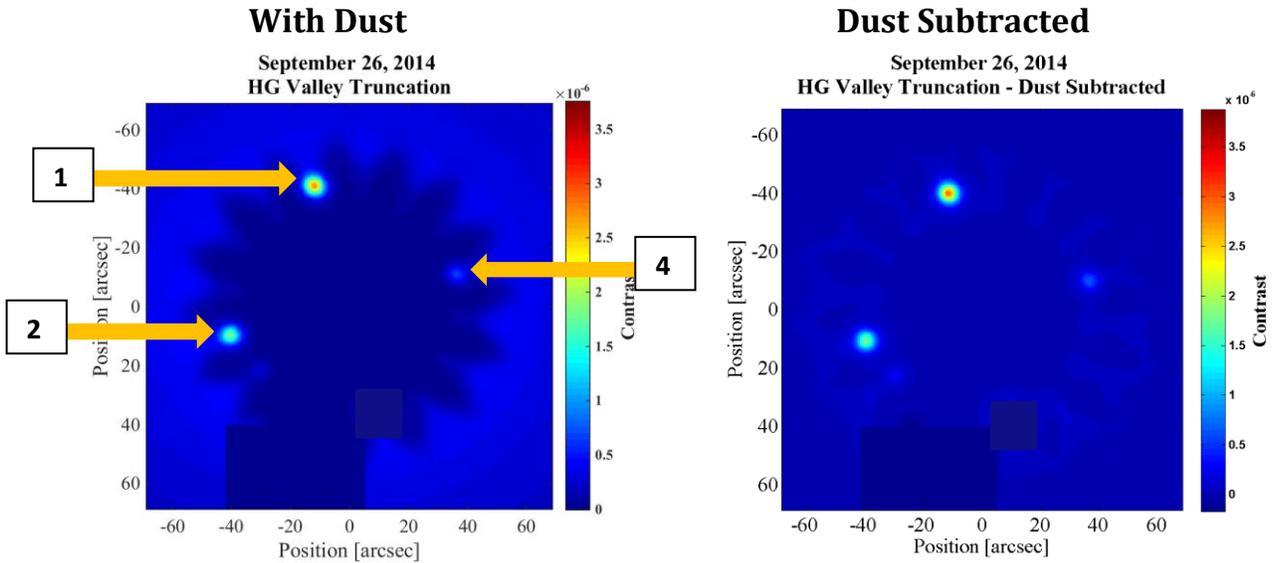
Dust Subtracted

Figure 5-39. Measured contrast maps for the HG edge shape with the valley truncation error, with dust (left) and dust-subtracted (right).

HG Valley Truncation: Measured Contrast Numbers

All four valley truncation flaws were measured for the HG valley truncation Starshade. The flaw at 6:00 had interference from the Starshade stand and therefore is masked out in the above image. That flaw was measured from the clocked set of images. The disagreement with the models increases with decreasing flaw size. (A larger flaw is a larger deviation from the base Starshade.)

	Flaw #	Measured Peak	Uncertainty
With Dust	1	1.88E-06	± 1.61E-06
	2	1.68E-06	± 8.83E-07
	3	1.61E-06	± 4.06E-07
	4	4.66E-07	± 7.68E-08
Dust Subtracted	1	1.88E-06	± 1.61E-06
	2	5.45E-06	± 8.83E-07
	3	2.51E-06	± 4.06E-07
	4	4.66E-07	± 7.69E-08

Table 5-25. Measured contrast peaks and comparison to the models for the HG edge shape with the valley truncation error.

HG Valley Trunc: Model/Experiment Ratios
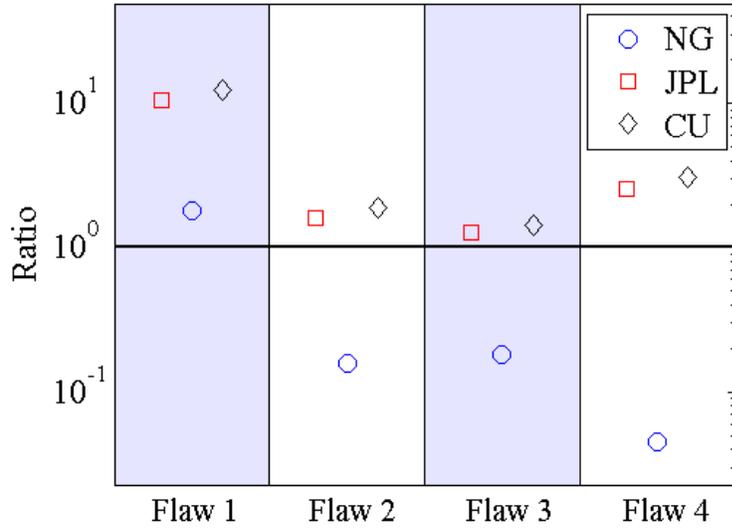
Figure 5-40. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the HG truncated valleys Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

IZ5 Valley Truncation: Model Predictions
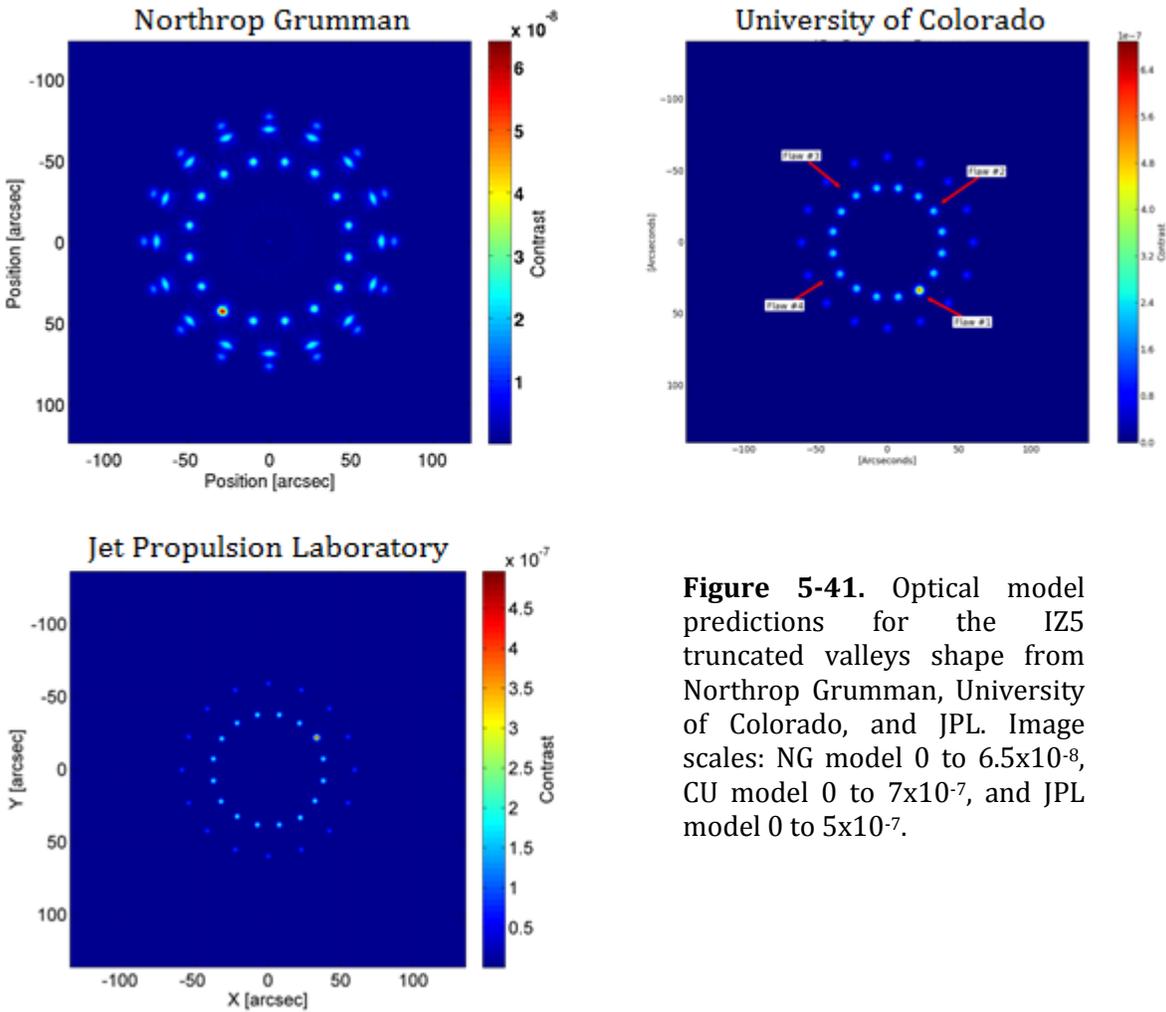
Figure 5-41. Optical model predictions for the IZ5 truncated valleys shape from Northrop Grumman, University of Colorado, and JPL. Image scales: NG model 0 to 6.5×10^{-8}, CU model 0 to 7×10^{-7}, and JPL model 0 to 5×10^{-7}.

	Peak Contrast (NG)	Peak Contrast (CU)	Peak Contrast (JPL)
Flaw 1	6.42E-08	6.89E-07	5.00E-07
Flaw 2	3.45E-08	3.10E-07	2.20E-07
Flaw 3	2.82E-08	3.07E-07	2.20E-07
Flaw 4	3.19E-08	2.67E-07	2.10E-07

Table 5-26. IZ5 valley truncation flaw brightnesses computed by each of the models.

IZ5 Valley Truncation: Measured Contrast Figures

With Dust
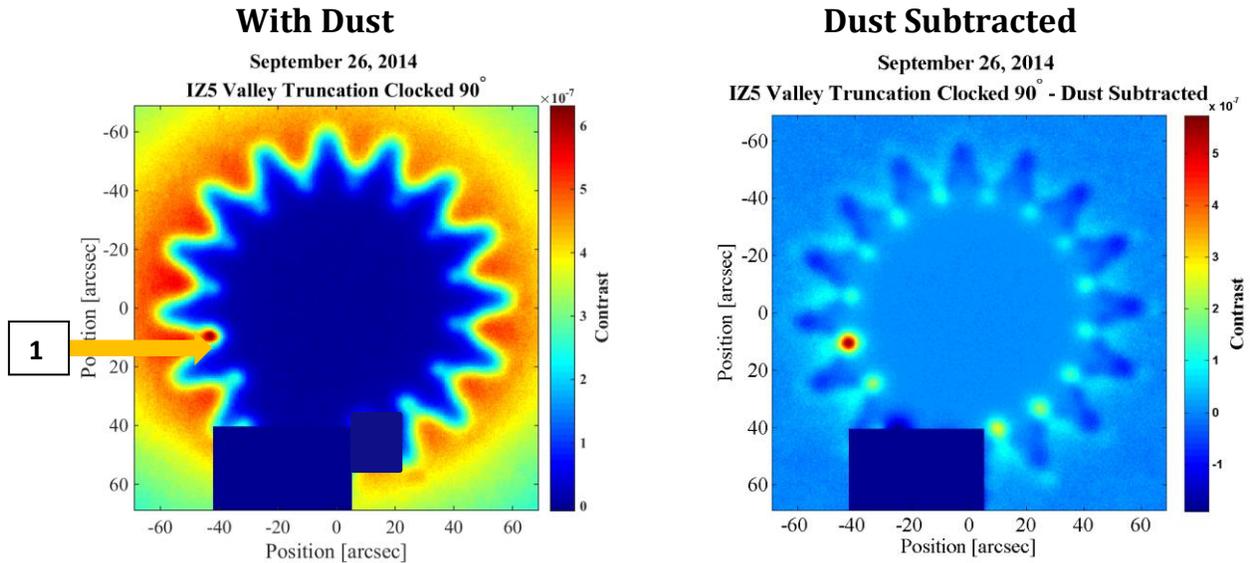
Dust Subtracted

Figure 5-42. Measured contrast maps for the IZ5 edge shape with the valley truncation error, with dust (left) and dust-subtracted (right).

IZ5 Valley Truncation: Measured Contrast Numbers

The flaws on the IZ5 valley truncation Starshade were extremely faint, and only one intentional flaw was measured. All flaws for this Starshade were expected to be faint, but more should have been visible based on the model predictions.

	Flaw #	Measured Peak	Uncertainty
With Dust	1	3.91E-07	± 9.57E-08
	2		
	3		
	4		
Dust Subtracted	1	3.91E-07	± 9.57E-08
	2		
	3		
	4		

Table 5-27. Measured contrast peaks and comparison to the models for the IZ5 edge shape with the valley truncation error.

IZ5 Valley Trunc: Model/Experiment Ratios
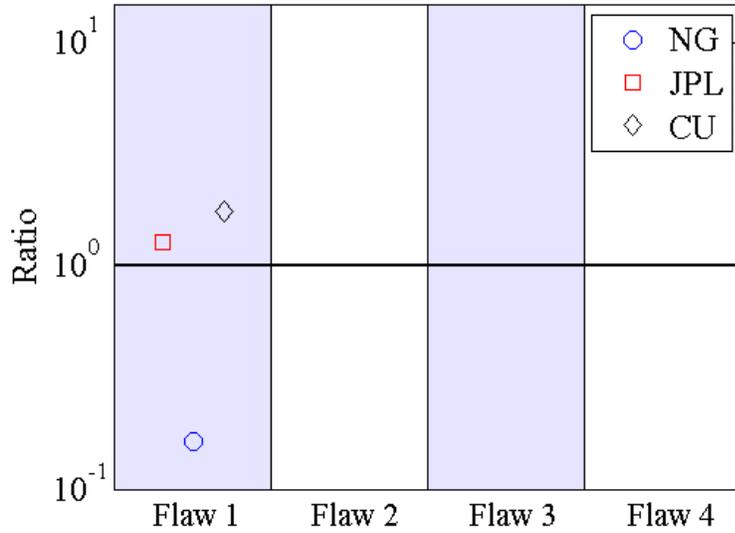
Figure 5-43. Ratios of flaw peaks modeled independently by NG, JPL, and CU to the peaks measured in the field for the IZ5 truncated valleys Starshade. We show the ratios of the models to the dust subtracted peaks. (Points above the line indicate the model predicted a brighter response than was measured)

5.5 DISCUSSION OF RESULTS AND RELATION TO MILESTONE #2

In Milestone #2 we aimed to "demonstrate agreement between the measured and predicted contrast resulting from a range of Starshade shapes". Our measurements are generally within a factor of five of the model predictions (See ratio figures in Section 5.4), with the exception being the edge sines flaw, for which all three of our models severely under-predicted the peaks. The edge sines flaws were predicted to have a number of close-together features, which we believe to have been blurred together into one peak by the turbulent atmosphere during our measurements. Because of this issue, unique to this flaw, Gaussian fitting to these peaks was much less reliable, and we therefore have not included these peaks in our consideration of the data set as a whole. Other outlying points of comparison such as flaw #2 on the IZ5 tip truncation Starshade (Section 5.4.1, Figures 5-8, 5-9) could be caused by an uncontrolled variation to the Starshade shape (a damaged tip, or a spider).

As detailed in the uncertainty analysis (Section 5.3), the average uncertainty on our measurements is around ± 25%. This is largely due to the variability of the peak of the unblocked source (due to seeing), which we use to convert our images to contrast units. The flaw peaks that we measure for comparison to models are susceptible to the same atmospheric variability. On top of that, misalignment models (Section 5.3.1) suggest that misalignment in our setup could impact our measurements by up to ± 100%. With these uncertainties in mind, general agreement to model and experiment, and the general trends of between different scales of flaws is encouraging.

Worth noting is the disagreement between the model predictions from NG, CU, and JPL. NG consistently predicts dimmer peaks than both CU and JPL, while CU and JPL are generally close to each other, but for some classes of flaws, do not show consistent difference between the two models, indicating some variation in how the models are operating.

In previous comparisons between the three models we have seen agreement of within 5%. These comparisons were based on modeling a flight-scaled system using parallel light, i.e. the light source is at infinity.

Since then the models have been modified to handle spherical wavefronts in order to model the field test setup where the light source is at 1km from the Starshade. It is likely that the disagreement we see in the models is due to these modifications. Future resolution of these model discrepancies is important to the use of these models as predictors of Starshade performance in flight.

6 CONCLUSION

The 2012 Starshade Field Testing TDEM was awarded to address two scientific milestones:

> **Milestone #1: Demonstrate, using a Starshade, contrast better than 10^{-9}, at all radii past the Starshade tips, in 50% bandwidth light.**

> **Milestone #2: Demonstrate agreement between the measured and predicted contrast resulting from a range of Starshade shapes.**

The status of these two milestones is as follows:

Milestone #1

During our April 2015 test, we were able to take an extended series of images using two Starshade shapes with excellent atmospheric conditions. After subtracting the dust from these images, we were able to combine data from a single night into an image with an effective exposure time of more than 45 minutes. This combined image had a 3σ standard deviation of less than 10^{-9} in a 40x40 pixel box just outside the tips of the IZ5 Starshade. Similar results were obtained using a HG Starshade.

Without dust subtraction, the best contrast we were able to obtain was just shy of the contrast goal set in milestone #1. This indicates that to further improve contrast measurements we will likely need to seek an alternate site with lower dust in the atmosphere. When imaging the scene with no main source and therefore little to no dust contribution, we obtain comparable contrast to our dust-subtracted data increasing the confidence in the performance of the dust subtraction algorithms.

Milestone #2

During our September 2014 test, we were able to take images of a wide variety of Starshades with intentional flaws. These flaws had all been previously modeled using an optical model that had been updated to include expanding source beam. We did a comprehensive comparison between the observations and the model predictions from three different optical models. The comparison was ultimately not conclusive due to the error bars on the measurements and the disagreement between the different model predictions.

7 FUTURE WORK

Previous testing of sub-scale Starshades by this team was done in a controlled environment (see for example Samuele, et al. 2010, SPIE, 7731, 51 and Samuele, et al. 2009, SPIE, 7440, 2). The biggest issue with that test setup was that the limited width of the vacuum tube led to the need for a collimated light source. Achieving a collimated beam big enough to cover the Starshade (with a clean enough wavefront) proved to be the limiting factor in that test. The available length of the test facility limited the Starshades to a few cm in diameter but a bigger Starshade would make the issue of the collimated beam even harder. With the outdoor test environment, we are able to use an expanding beam and not worry about it reflecting from any nearby objects. Kasdin et al. have pursued a different strategy to managing this problem using a controlled test environment (see for example Kim et al. 2014, SPIE, 9605, 84 and Sirbu et al. 2014, SPIE, 9143, 2P). We believe these tests are complimentary and should both continue to be pursued.

Prior to and during this TDEM, Northrop Grumman has developed the equipment and team necessary to carry out long distance optical testing of Starshades. The process, team, and equipment should now be seen as a testing asset that can be deployed to test any future Starshade optics of interest to the community. Tests that have been suggested include:

- Expanding the baseline to 4 km and using a smaller Starshade to take observations that more closely match flight-like conditions for both Fresnel Number and Resolution Element (See Section 7.1).
- Re-imaging the scenes with color filters to explore the effects of color on the different Starshade shapes in a more comprehensive manner.
- Carrying out detailed alignment measurements, measuring the optical effects at the edge of the Starshade as the Starshade is moved away from perfect alignment.
- Repeating observations from the September 2014 test (flawed Starshades) in better desert conditions and setup. As one possible cause of discrepancy between models and observations is misalignment between the source and the Starshade, developing a method to minimize misalignment would likely drastically improve the reliability of the measurements.
- Re-imaging the Starshades with sine wave errors along their edges in better atmospheric conditions. Since periodic flaws are possible in a flight system, it is important that we understand why we measured them to produce such bright contrast features and why our measurements disagree with the models by two orders of magnitude. Given the nature of the predicted sine wave error features (many close-together contrast peaks that could easily blur together), imaging in more stable conditions will likely yield significantly different results.
- Repeating the experiments at a new location that has less dust and more stable atmospheric conditions. This should lead to better contrast ratio values at flight like Fresnel numbers and resolution elements.
- Testing in detail the impact of tilt of the Starshade (Section 3.3). In particular the suppression between the petals and the tips for different tilt angles.

Other possible future work includes:

- Refining our post-processing methodology to assure we are extracting optimum numbers from the data. Examples of this include examining our data reduction methods, optimizing our image-rejection criteria, considering a larger area (a ring rather than a small box) when calculating 3σ contrast levels at various radii, and experimenting with the use of aperture photometry to define source brightness rather than Gaussian peaks.
- Resolving the differences between the three different optical models. In previous model comparisons the models were in agreement with the assumption of a point source at infinity (i.e. flat wavefronts interacting with the Starshade), show divergence when the expanding beam option was implemented. A detailed examination of the methods used in each of the models, particularly as it relates to the spherical wavefront modelling may provide some insight into the differences between them. Noticeably, differences between the models themselves do not appear to be consistent even within a single class of flaw.
- Including dust and atmospheric effects in the simulations to better simulate the test environment

7.1 DISCUSSION OF FUTURE 4 KM BASELINE TEST

Our current test setup uses a much greater Fresnel number and has a greater number of resolution elements across the Starshade than a flight-like scenario. One way to mitigate this discrepancy is to increase our baseline from 2km to 4km and decrease the size of our Starshade. In order to investigate the feasibility of a test with these parameters, we placed LEDs at heights of approximately 5' and 11' on a vertical stand 4km away from the telescope for the duration of tests #2 and #3. Doubling the length of our path from the light sources to the telescope leads to an increase in all of the major limiting factors for our current test: more refraction, more blurring, and more dust. As dust subtraction is a good solution for mitigating the dust contribution, we analyzed the characteristics of the 4km sources with respect to overall movement and PSF variability. Shown in Figures 7-1 and 7-2 are the 4km FWHM values for the upper and lower sources along with the FWHM of one of the planets at 2km for reference. For 20cm and 29cm Starshades (⅓ and ½ the size of our current test articles) we estimate that we need seeing conditions that are respectively better than 8.3 and 12.4 arcseconds. For the September 2014 test, the overall seeing conditions were too poor for the ⅓ size Starshade, but were sufficient for the second half of the night to use a ½ size Starshade. The seeing conditions were even better for April 2015 test, when conditions were sufficient for a ⅓ size Starshade using source lights at 4km.
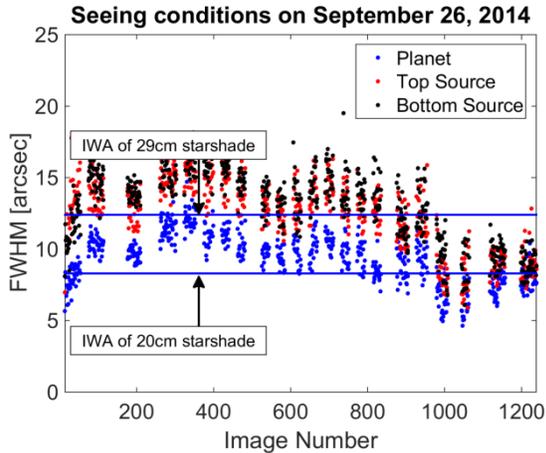
Figure 7-1: Seeing Conditions Support Longer Range Observations. The seeing, as measured by the FWHM of the sources in the images, shows that on a nominal night the source FWHM can be smaller than the IWA of a 29 cm Starshade at 4km. The black and red points are LEDs 4km from the telescope at heights of 6' and 11'. The blue points are from a 'planet' 2km from the telescope.
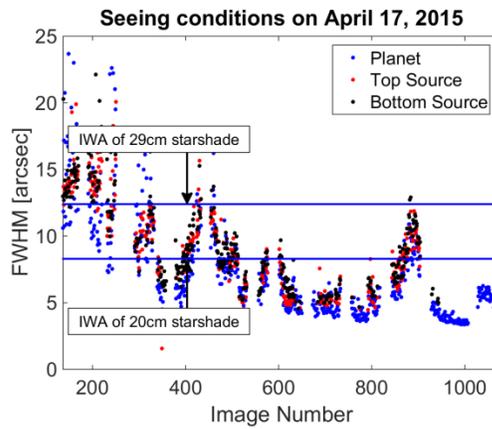
Figure 7-2: Seeing Conditions Support Longer Range Observations. Similar observations from April 2015 are consistent with those from September 2014. The black and red points are LEDs 4km from the telescope at heights of 5' and 11'. The blue points are from a 'planet' 2km from the telescope. Seeing conditions become more consistent as time progressed.

The refraction of the 4km sources by the atmosphere also poses a greater challenge for the longer baseline. We can track vertical movement of the two sources in addition to the variation in seeing. A movement of greater than half the size of the IWA of a given Starshade size would lead to the main source 'peeking' out from behind the Starshade and eliminate the benefits of using a Starshade. Figure 7-3 shows the vertical movement of both the top and bottom sources relative to the mean vertical position of the sources for a given set of time-adjacent images. The deviation from the mean location is often larger than the IWA of a 29cm diameter Starshade, but would still allow for 5-20% of the images taken to be in adequate alignment.
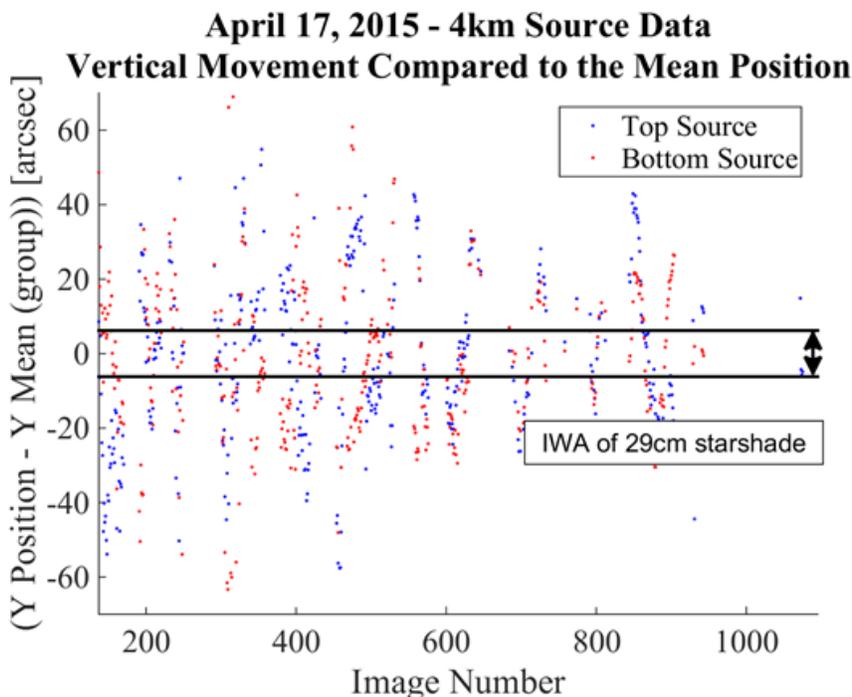
Figure 7-3: Refraction Challenges for Longer Range Observations. The vertical movement of the two sources at 4km on April 17, 2015 compared with the mean vertical position of those sources in time-adjacent images. The large vertical spread shows that only a small fraction of the data would be in alignment using a 29cm Starshade, indicating the need for a test dedicated solely to this type of observation to obtain sufficient exposure time in order to achieve contrast goals of 10^{-9}.

To better control alignment during future tests we will follow a 3 step process.

1) Determine the best alignment position for all Starshades by using the small bright points around the tips or valleys of a non-flawed Starshade. When this Starshade is aligned, tips and valleys are equally bright all around the Starshade.
2) By finding an image where this occurs in post processing, we note the relative position of the planet LEDs to the bright glints on the Starshade stand in the image.
3) We can then look at the flawed Starshade images in post processing and select the images that have the same relative planet and Starshade stand position.

There is a 1:1 mapping of distortions that move main source and planet LEDs. Similarly there is a 1:1 mapping of distortions that move the Starshade and the Starshade stand. This technique uses the Planet LEDs and the Starshade stand as surrogates for the distortion movement (and therefore misalignment) of the source and the Starshade. This will significantly reduce the alignment errors due to misalignment as outlined in Section 5.3.1

While a test of a smaller Starshade over a 4km range will be challenging, the combination of better alignment selection techniques and rejection of poor quality images due to seeing and alignment lead us to believe that a successful test can be conducted over 5 nights of observations.

8 APPENDIX

8.1 DUST SUBTRACTION APPROACH

As the main purpose of April 2015 test was aimed toward Milestone #1, to demonstrate contrast better than 10^{-9}, large numbers of images needed to be precisely co-aligned.

In each image the dust in the atmosphere scattered the light produced by the main source. This scattering was seen as a halo of light around the Starshade that was brightest near the edge of Starshade and tapered off as the distance from the Starshade increased. The dust halo varied in shape and brightness significantly between sets of images due to changing air conditions. The images in Figure 8-1 taken within a few hours of each other illustrate this difference. These images are rendered in the same color scale to emphasize the shape and size of the dust halo.
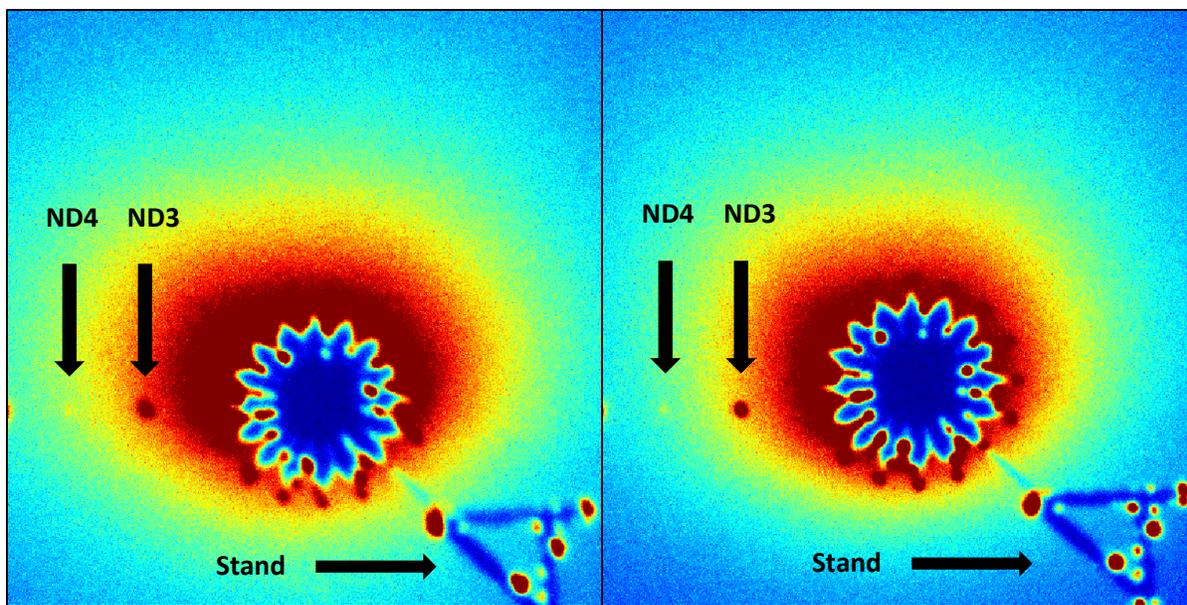
Figure 8-1: Typical difference in dust halo shape and size (HG Starshade). In the left image, the red band of the dust halo completely engulfs the ND3 Planet, while it only clips it in the right image. More of the Starshade tips can be seen in this scale in the right image than in the left image.

Due to this variation, our process of co-aligning and combining multiple images in order to calculate the background standard deviation needed to be adjusted in order to combine images over the course of an entire night. Our solution was to model the dust halo in each image and subtract it out before we combined the images (hereafter referred to as dust subtraction). Our previous methods of dust subtraction were not designed to handle the amount of precision needed to cover a large area of the image, so we needed to implement a more robust fitting technique.

We made one major assumption when designing our current fitting method: in each image we assumed the dust halo was a smoothly varying function of pixel position. We believe this is a valid assumption as the dust particles were evenly distributed throughout the atmosphere captured by our image. Taking

this assumption allowed us to approximate the dust halo using a large degree polynomial function. The process we used to implement this approximation method is as follows:

1. Choose a consistent frame in each image on which to perform the dust fit.
2. Mask off parts of the image containing significant contributions from sources other than the dust (i.e. Starshade, Starshade stand, and planet sources).
3. Choose a subset of pixels not masked off on which to perform the dust fit.
4. Choose a degree for the polynomial to fit the chosen pixels.
5. Perform a linear least squares regression on the chosen pixels.

Each step of our process is described in more detail below.

1. There were two important considerations regarding the choice of frame. Firstly, it was critical that the size of the frame was large enough to compare an area far from the Starshade to areas close to the Starshade. The larger the frame however, the more pixels would be considered, and ultimately the more computing time it would take to run the curve fitting routine. We decided on a frame 600 by 1200 pixels to leave an adequate distance from the furthest corner to the Starshade while keeping a reasonable total number of pixels. The second consideration was on which feature to co-align the images. Due to the issue of atmospheric turbulence and the additional distance between the telescope and planet sources compared to the distance between the telescope and the Starshade, the Starshade and the planet sources were not relatively aligned in all of our images. We chose to align on the planets for our stacking process, as it allowed the images to be registered on bright point sources, making our co-alignment easier and more accurate. This approach also benefits from being more flight-like, as the on-orbit Starshade may move slightly against the image for the period of integration.
2. As we were only fitting the dust halo, we had to decide which pixels were significantly influenced by other sources. This required finding and disregarding the Starshade, Starshade stand, and planet sources. These pixels were masked off to not influence the fit.
3. For computational efficiency we subsampled the image. Our subsampling method was based on the necessity of a higher degree of accuracy close to the Starshade. To achieve this, the image was broken into subsections. In each subsection, the ratio of pixels that were chosen to use in the fit was based off of the distance of that subsection to the center of the Starshade. This produced a mesh as seen in Figure 8-2.
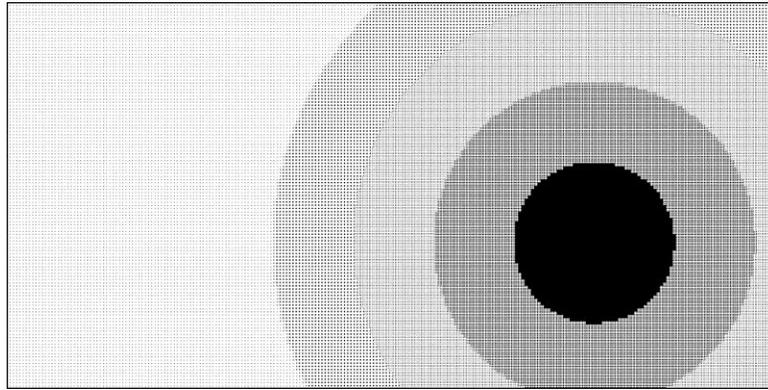
Figure 8-2: A visualization of our subsampling methodology. Black pixels represent pixels chosen for the fit.

4. The factor that contributed the most to the goodness of the fit was the choice of degree for the polynomial approximation. To determine an adequate degree, we chose a typical image from each night of data and fit a polynomial of over-estimated degree (50). We ran a statistical F-test on each group of terms of the same degree to measure the probability that the coefficients of these terms differ from 0. The overall conclusion of this test was at the $\alpha = 0.05$ confidence level, we failed to find a significant influence from the terms of degree 30 or higher; however, there were instances of a few more degrees being necessary. To be well within tolerance, we decided on using a 40 degree polynomial to fit each image. Typical plots of the p-values from the F-tests for days 1, 2, 4 and 5 are shown in figure 8-3. (We did no dust subtraction for day 3).
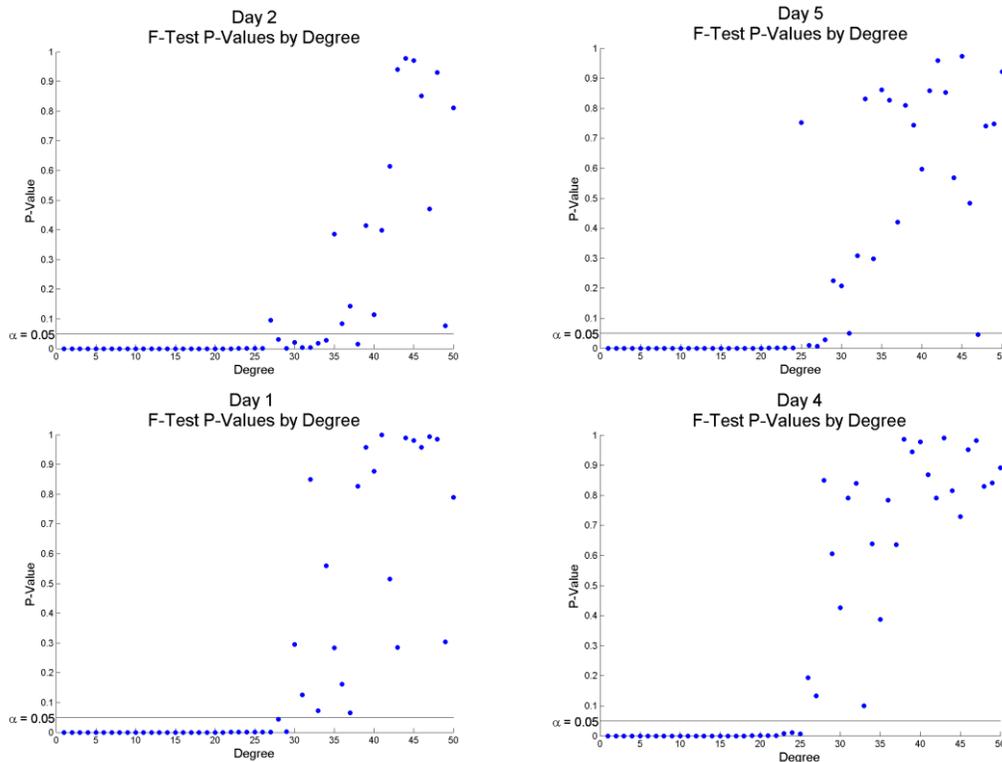
Figure 8-3: Typical P-Values obtained from running an F-Test by degree in the polynomial fit to the dust. Values close to zero indicate a large influence on the fit.

Based on the parameters chosen above, the linear least squares regression algorithm resulted in a polynomial approximation for the dust halo valid within the chosen frame. While polynomial fits are notoriously poor for extrapolation, the purpose of the dust subtraction was solely to remove the dust's effect on the standard deviation of the background. We did rely on the interpolative ability of the polynomial fit due to our subsampling method, but no extrapolation was needed or utilized. The effects of the chosen polynomial degree and our use of subsampling are charted in figure 8-4. The general result from this comparison was at a 20th degree polynomial fit or higher the difference in the overall standard deviation of the background was on the scale of tenths of a percent, whether subsampled or not. The use of subsampling did produce a visual "rippling" effect as seen in the lower right images in Figure 8-4 as a byproduct of the sparse sampling on the left side of each image. The rippling however was on a small enough scale and far enough away from the Starshade to not influence any of our results. Ultimately, we concluded that our subsampling technique did not significantly affect the standard deviation of the background for large degree polynomial fits, and its use was warranted as it allowed us to use a large enough degree polynomial (as outlined in our F-Test) to fit the dust halo.
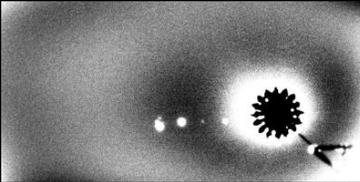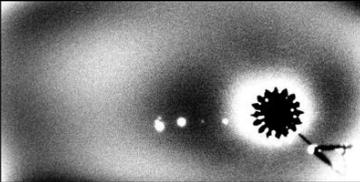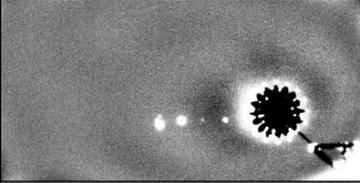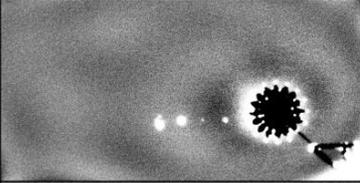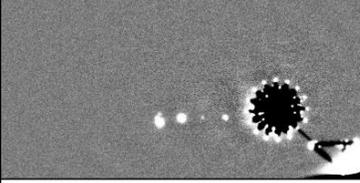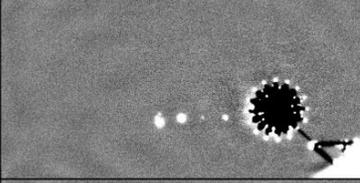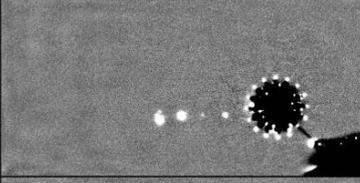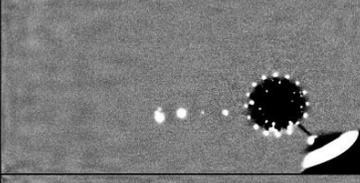
Figure 8-4: Table of dust subtractions for various polynomial degree fits, with and without subsampling.

8.2 CONVOLUTION FILTER

Here we present a method we investigated to improve upon our proposed nominal three-sigma detection level for planet sources. A three-sigma detection level alone works on a pixel-by-pixel basis; it tells us if a particular pixel is likely to have come from a source other than noise. However, the image of a planet source spans multiple pixels, and assuming the planet is a point source its image is a point spread function (PSF). If we know the form of the point spread function, we can therefore utilize the entire planet signal for detection.

The particular method we investigated to this end was the use of an approximation of the PSF as a convolution filter on the image. In general, this method is used to increase the signal to noise ratio (SNR) of images, which is defined to be the signal divided by the standard deviation of the noise floor. For the simplest case of the image being comprised only of the planet source and noise, this drives down the standard deviation of the noise more than the planet source, effectively increasing the SNR of the image. In the filtered image, each pixel represents how well the corresponding group of pixels in the original image matches with the PSF. Therefore a three-sigma detection in the filtered image references the entire planet source against the noise.

The mathematical basis for this process comes from our approximation of the PSF as a Gaussian curve, and our assumption that the noise is random and normally distributed (Gaussian noise). We assumed the PSF was a normalized Gaussian of the form

$$PSF = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{x^2}{2\sigma_x{}^2}-\frac{y^2}{2\sigma_y{}^2}\right)$$

where σ_x and σ_y are characteristics of the PSF shape in the x and y dimensions respectively. Filtering an image consisting of a planet source and noise with a function of this form reduces the standard deviation of the noise by a factor of $\sqrt{4\pi\sigma_x\sigma_y}$, while only reducing the peak of the planet source by a factor of 2. Effectively this increases the SNR of the image by a factor of $\sqrt{\pi\sigma_x\sigma_y}$, theoretically allowing us to detect planet sources this factor fainter.

In practice, the improvement factor is weaker, mostly due to the fact that the background of our images is not perfect Gaussian noise. In order to effectively estimate the actual benefit of employing this method we needed the mean of the background to be essentially zero. As our images are significantly affected by dust (as explained in the first appendix), we relied on our dust subtraction routine to achieve this. We ran this method on our best contrast dust-subtracted images from the April 2015 test. The results are summarized in Figure 8-5.
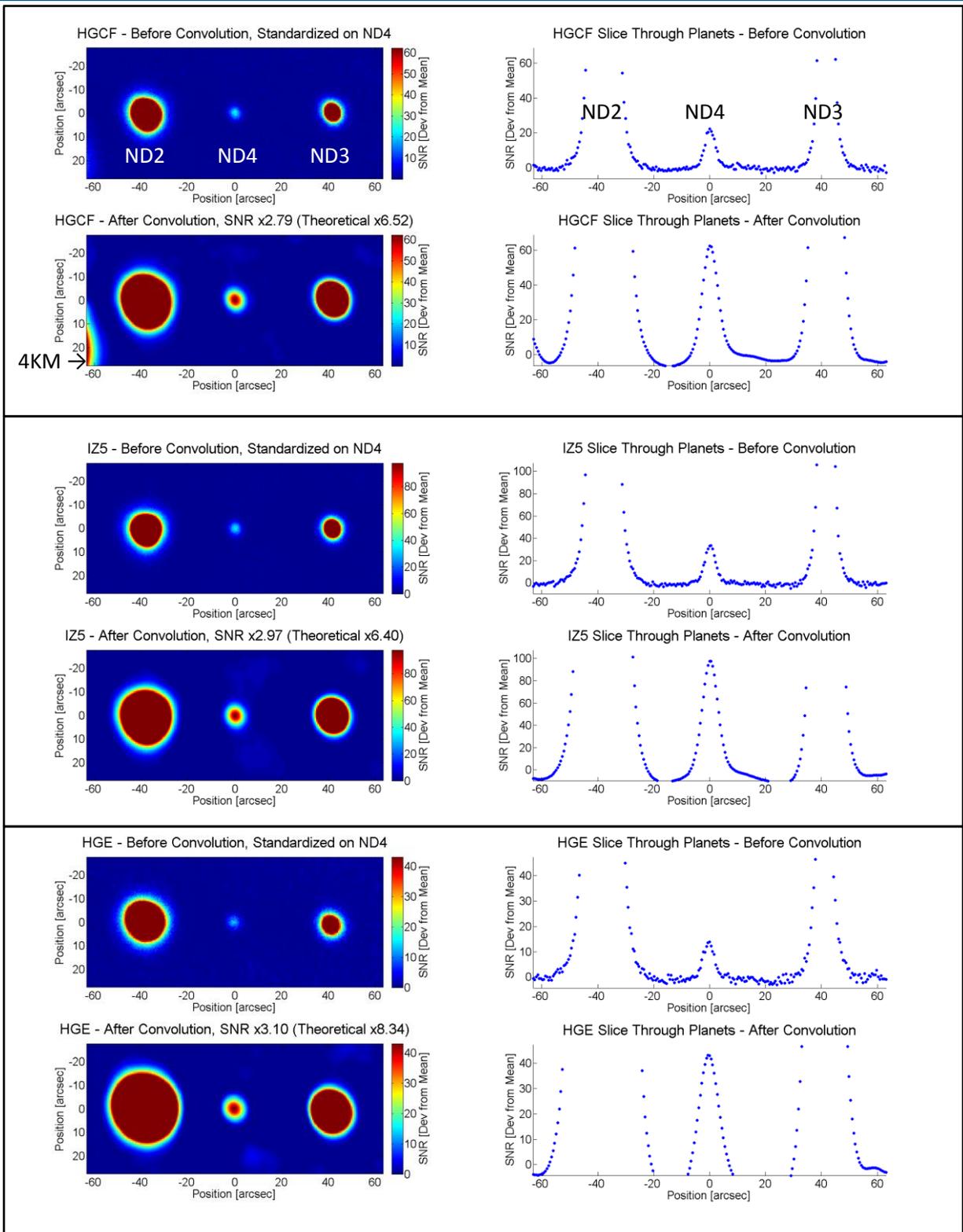
Figure 8-5: Best contrast results planets for Hypergaussian Carbon Fiber (HGCF), IZ5, and Hypergaussian Etched (HGE) Starshades before and after convolution filter is applied. Shown on the left are crops of the images in standard deviation units, and shown on the right are one-dimension slices through the peak of the ND4 planet sources. Also shown are actual SNR increases versus theoretical SNR increases.

9 REFERENCES

[1] Glassman, T., Casement, S., Warwick, S., Armagan, O. 2013, "Achieving High Contrast Ratios with a 60 cm Starshade," Proc. SPIE, 8864, 42.

2 Cady, E., Kasdin, J., Vanderbei, R., & Belikov, R. 2007, "Optimal design of petal-shaped occulters for extra-solar planet detection," Proc. SPIE, 6693, 04.

[3] Dumont, P., Shaklan, S., Cady, E., Kasdin, J., & Vanderbei, RDumont et al. 2009, "Analysis of External Occulters in the Presence of Defects," Proc. SPIE, 7440, 6.

[4] Cash, W. 2011, "Analytic Modeling of Starshades," Astrophysical Journal, 738, 76.

[5] Shaklan, S., Noecker, C., Glassman, T., Lo, A., Dumont, P., Kasdin, J., Cady, E., Vanderbei, R., & Lawson, P. 2010, "Error Budgeting and Tolerancing of Starshades for Exoplanet Detection," Proc. SPIE, 7731, 75.

[6] Bruno, O., Chaubell, J. 2013, "Theoretical Basis of the CWI 1.00 Code"

[7] Schindhelm, E., Shipley, A., Oakley, P., Leviton, D., Cash, W., & Card, G. 2007, "Laboratory Studies of Petal-Shaped Occulters," Proc. SPIE, 6693, 5.