EXOPAG SAG 18 Final Report

Metrics for Direct-Imaging with Starshades

Chairs: Tiffany Glassman (tiffany.glassman@ngc.com, Northrop Grumman) and Maggie Turnbull (turnbull.maggie@gmail.com, Global Science Institute)
Community inputs from: Jon Arenberg, Ashley Baldwin, Dominic Benford, Jim Breckinridge, Robert Brown, Eric Cady, Web Cash, Shawn Domagal-Goldman, Anthony Harness, Joe Harrington, Jeffrey Jewell, Charles Lawrence, Doug Leviton, Stefan Martin, Charley Noecker, Aki Roberge, Tyler Robinson, Stuart Shaklan, Ann Shipley, Nick Siegler, Chris Stark, Steve Warwick, Sloane Wiktorowicz, Ben Zeiger

Introduction

The goal of ExoPAG SAG 18 was to define how performance metrics for starshades are being used in the community. SAG18 started with a survey to ask the community what metrics they've been using and how they're being used. We would like to thank everyone who provided inputs to the survey. The questions on the survey were:

Q1. One major goal for SAG 18 is simply to compile a list of ways in which different researchers quantify the observatory performance requirements for detecting planets, and what terms they use to refer to those metrics. These terms are often bandied about in both starshade and coronagraph studies. What term(s) do you typically use to quantify the residual starlight in planet-finding images and/or the minimum brightness of detectable planets (e.g., contrast, suppression, Q, limiting fractional planet brightness)? NOTE: This is purely a terminology question, no definition needed yet.
Q2. Of the metrics you named above, which do you find yourself using most often, and how do you define it? Please describe it qualitatively as best you can. If possible, please also provide quantitative definition. This could be a procedure for how you apply this definition to a test, observation, a modeled image, or an equation defining the metric (you can either point to an equation number in a publication, or type the equation in here and we'll decipher it as best we can).
Q3. Why do you use the metric defined in Q2, instead of others? Please tell us anything you can about why this term is appropriate for your context, ease of use, and/or precedence for using it, that affected your decision to use this metric.
Q4. If possible, please provide a reference to one publication where this metric is used in the way that you've defined it.

In addition to the responses to this survey, we got inputs from a group led by Charles Lawrence of JPL that is tackling similar questions, focused on a plot of testbed performance from Exoplanet office technology appendix (Crill & Siegler 2017, figure 34).

The aim of this report is not to narrow down to one metric – there are many different metrics that may be useful for different situations. The goals is to establish consensus on definitions, differences, and applications of each metric.

The community inputs could be organized into five main categories that define the different ways that people use these metrics in their work. An explanation of each of these terms follows, with a proposed name for the metric, a definition, and a discussion of where it is applicable, and its pros and cons.

Metric Definitions

1) Fractional Planet Brightness

This term defines an astrophysical property of the targets of interest – it's a purely scientific term as opposed to the others which are all based on details of the imaging system used to observe the exoplanet target. The Fractional Planet Brightness is defined as the brightness of the exoplanet relative to the brightness of the star. This is useful, for example, when examining nearby stars and determining the properties of earth-like planets in their habitable zones (for example, Turnbull et al. 2012). However a translation needs to be made when determining the properties of an optical architecture which can observe a star-planet system with e.g. a fractional planet brightness of 10^{-10}. This translation is not always straightforward, which is why it's important to differentiate this astrophysical term from other instrumentation-related terms.

The name Fractional Planet Brightness is taken from Turnbull et al. 2012. Other names that have been used to mean the same thing are "Planet-Star contrast" and "Planet-Star Flux Ratio". Currently, the WFIRST CGI teams are using the term "Planet-Star Flux Ratio," calculated according to Traub & Oppenheimer (2010; Equations 14 and 15) and do not use the word "contrast" when referring to planets or other astrophysical objects. The term "contrast" is reserved for instrument performance, as below.

2) Raw Contrast

The term contrast is most commonly used to define the performance of a direct-imaging system in seeing very faint objects close to very bright stars. We define the term "Raw Contrast" to be the most basic calculation of the average contrast in the focal plane of an imaging system. It is defined as the ratio of the average starlight irradiance in a region of interest to the average irradiance of unblocked starlight in an equivalent aperture centered on the star. The region of interest can be an aperture, pixel, or annulus in any location on the focal plane of the system. There is no correction applied to the measured irradiance in this region.

The advantage of this definition is that it is simple to calculate in test data. However it is hampered by a number of effects that might make it difficult to interpret. First, it doesn't take into consideration the effect of the imaging system on the off-axis point source – the planet. If the imaging system causes a loss of transmission for off-axis sources or degrades the point spread function (PSF) of the planet, then this definition of contrast will over-estimate the ability of the system to detect that planet against the residual starlight in the region of interest. This is mostly a concern for coronagraph systems that can have a large effect on off-axis sources.

For starshade imaging systems, the main limit of this definition is that it is affected by the telescope part of the system, which is unrelated to the performance of the starshade itself. In particular, in a testbed situation, the telescope can be oversized relative to the starshade, resulting in images that are over-resolved compared to flight-like systems. This means that the PSF of the star is tighter than it would otherwise be and residual starlight in the image does not have as much effect on the contrast in off-axis locations (see Figure 1). In this situation, this definition of contrast will again over-estimate the ability of a flight starshade system to enable detection of a planet against the residual starlight in the region of interest.

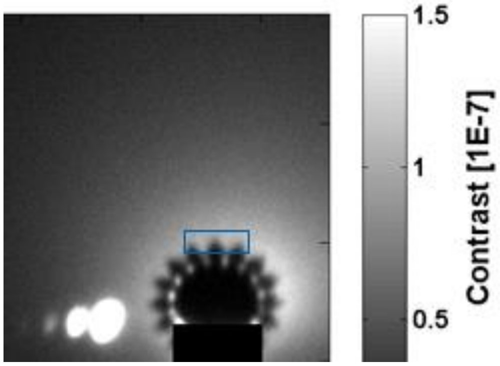
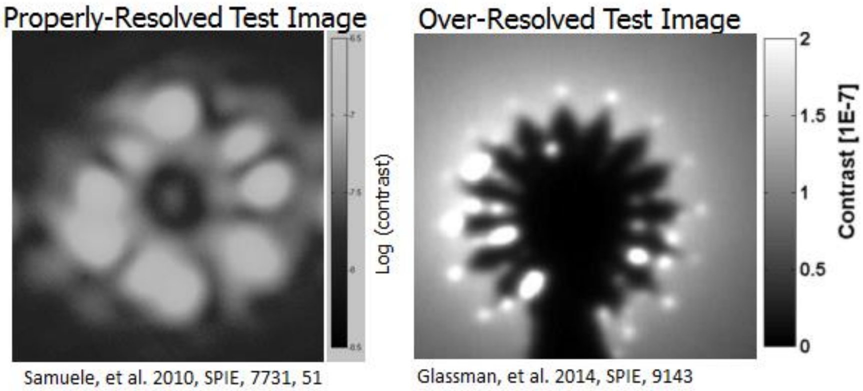


Figure 1: Example of an over-resolved image of a starshade from an outdoor test. The point sources at the bases of the petals don't have as much effect on the residual light outside of the starshade tips because their PSFs are more compact than they would be in a flight system. This also shows the bright background light (in this case from dust in the air) that makes the average in the aperture problematic once the uniform background is subtracted.

Another issue is that if the limiting light in the image comes from a smooth background source (e.g. exozodi light or dust in the air in an outdoor starshade test) then this definition becomes hard to apply. The average of the background across the whole image can be subtracted, which means that the average starlight irradiance in the region of interest becomes just a measure of how well the background was subtracted and may even be negative. In this case, we propose a variant to the raw contrast, which is the RMS raw contrast. In this case, the Nσ standard deviation in the region of interest is calculated instead of the average. This captures the noise in the image, which is what limits the ability to detect a planet against this background. This is also simple to calculate from test data, however it is strongly affected by the background light source rather than the starshade itself. For a more statistically rigorous treatment of contrast in high-contrast imaging, see SAG 19 (Mawet & Jensen-Clem).

3) Aperture-Corrected Contrast

This definition of contrast applies to starshade testbeds where the telescope is larger compared to the starshade than in a flight system and therefore the image is over-resolved. The image is corrected in post-processing to reduce the resolution and determine what the contrast would have been in a correctly-resolved image. The image can be corrected by convolving with a wider PSF. See Figure 2 for the difference between over-resolved and properly-resolved starshade test images.



Samuele, et al. 2010, SPIE, 7731, 51

Glassman, et al. 2014, SPIE, 9143

Figure 2: A starshade test image with a realistic flight-like resolution compared to the size of the starshade (left) compared to an image where the telescope over-resolves the starshade (right).

The advantage of this definition is that it allows a better comparison between a test image and potential flight systems. The problem is that it requires additional post-processing steps that are dependent on models of the “correct” telescope resolution relative to the starshade size. Another issue is that the test system is likely to also be at a different Fresnel number than a flight system (see e.g. Glassman et al. 2014) and this is not something that can be corrected.

4) Q

This term is also a form of contrast but is corrected for the effect of the imaging system on the off-axis point source. This is primarily a concern for coronagraph systems since starshades have little effect on point sources that are located past the tips of the petals.

The effect of the optical system on an off-axis point source could be directly measured in a testbed by placing the star at the off-axis location and directly measuring the irradiance of the star in the off-axis region of interest in the focal plane. For starshade testbeds this is often how contrast is calculated in practice.

For coronagraph testbeds, a post-processing step is often required. In this case the effect of the system on the off-axis planet is modeled and then Q is defined as the brightness of a planet whose light in the region of interest would be equal to the brightness of the residual starlight in the region. For more details, see e.g. Krist et al. (2007) and Brown & Burrows (1990).

5) Suppression

Suppression is different from the other metrics in that it is not defined in the image-plane of an instrument. Instead it measures the total amount of starlight entering the aperture of the telescope. This metric is only meaningful for starshades where the starlight is suppressed before the telescope. The definition is that suppression is the integrated light in the telescope pupil with the starshade in place relative to the integrated light in the telescope pupil without the starshade.

This metric has the advantage of being independent of the size of the telescope, which is useful for assessing and comparing tests with geometries that vary significantly from the flight system and it avoids the problems with measured contrast in over-resolved testbeds. It is a quantitative measure of the total amount of stray light entering the system.

However suppression is difficult to translate into planet detectability in an absolute sense – the ability to see a planet at a given suppression level depends strongly on the distribution of that light in the final image. There is also no meaningful coronagraphic equivalent – suppression must be translated to something like contrast to compare techniques. The total amount of light entering the telescope can also be dominated by background sources, which are difficult to separate without image-plane information, therefore total suppression due to the starshade itself can be difficult to determine in lab or field settings.

6) Starshade Radius

Another parameter related to starshades that is important to define is the radius of the starshade itself. This issue was raised by the Lawrence group at JPL and that group put forth the following definitions.

- $r_{1.0}$ = radius at petal tips
- r_e = radius at the 1-1/e (~63%) transmission point
- $r_{0.5}$ = radius at the 50% transmission point

Relationships between these various radii can vary depending on the starshade design (see e.g. Crill & Siegler 2017), therefore multiple radius values should be provided with each design/ test article if possible. Starshade radius and any other factor derived from that (e.g. Fresnel Number, Inner Working Angle) should always be labeled with which radius was used.

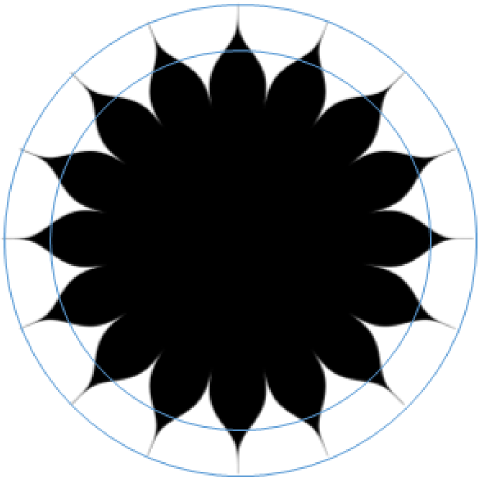


Figure 3: Approximate $r_{1.0}$ (outer) and r_e (inner) for a Hypergaussian starshade. The ratio of $r_{1.0}$ and r_e could look very different for a numerically determined shape.

REFERENCES

Biller, B., et al. 2006, “Contrast Limits with the Simultaneous Differential Extrasolar Planet Imager (SDI) at the VLT and MMT,” Proc. SPIE, 6272, 2D

Brown, R. & Burrows, C. 1990, “On the Feasibility of Detecting Extrasolar Planets by Reflected Starlight Using the Hubble Space Telescope,” ICARUS, 87,484

Brown, R. 2015, “True Masses of Radial-Velocity Exoplanets,” ApJ, 805, 188

Brown, R. 2016, “The Q Metric of Speckle Confusion”

Cady, E., et al. 2015, “Milestone 5 Final Report: Hybrid Lyot and Shaped Pupil Broadband Contrast Testbed Demonstration for WFIRST-AFTA,” wfirst.gsfc.nasa.gov/science/sdt_public/wps/references/WFIRST_CGI_Milestone5_Final_Report.pdf

Crill, B. & Siegler, N. 2017, “EXOPLANET EXPLORATION PROGRAM Technology Plan Appendix 2017,” exoplanets.nasa.gov/system/internal_resources/details/original/566_2017_ExEP_Technology_Plan_Appendix_Rev_B1_Cleared.pdf

Domagal-Goldman, S., Segura, A., Claire, M., Robinson, T. & Meadows, V. 2014, “Abiotic Ozone and Oxygen in Atmospheres Similar to Prebiotic Earth,” ApJ, 792, 90

Glassman, T., Casement, S., Warwick, S., & Novicki, M. 2014, “Measurements of High-Contrast Starshade Performance,” Proc. SPIE, 9143, 20

Glassman, T., et al. 2014, "Demonstration of Starshade Starlight-Suppression Performance in the Field," TDEM-2012 Final Report exoplanets.nasa.gov/exep/files/exep/GlassmanTDEM2012_FinalReport.pdf

Krist, J., et al. 2007, “Hunting Planets and Observing Disks with the JWST NIRCam Coronagraph,” Proc. SPIE, 6693, 0H

Krist, J. 2016, “Contrast Definition for Direct Imaging of Extrasolar Planets”

Mawet, D. & Jensen-Clem, R. SAG19 “Exoplanet Imaging Signal Detection: Theory and Rigorous Contrast Metrics” exoplanets.nasa.gov/system/internal_resources/details/original/314_SAG19exoplanetimagingsignaldetectiontheoryandrigorouscontrastmetrics.pdf

Robinson, T., Stapelfeldt, K., & Marley, M. 2016, “Characterizing Rocky and Gaseous Exoplanets with 2-meter Class Space-based Coronagraphs,” PASP, 128, 025003

Samuele, R., Varshneya, R., Johnson, T., Johnson, A., & Glassman, T. 2010, “Progress at the Starshade Testbed at Northrop Grumman Aerospace Systems -- Comparisons with Computer Simulations,” Proc. SPIE, 7731, 51

Stevenson, K., et al. 2012, “Transit and Eclipse Analyses of the Exoplanet HD149026b using Bliss Mapping,” ApJ, 754, 136

Shaklan, S., et al. 2011, "Stability error budget for an aggressive coronagraph on a 3.8 m telescope," Proc. SPIE, 8151, 09

Shaklan, S., et al. 2015, “Error budgets for the Exoplanet Starshade (Exo-S) Probe-Class Mission study,” Proc. SPIE, 9605, 0Z

Stark, C., Roberge, A., Mandell, A., & Robinson, T. 2014, “Maximizing the ExoEarth Candidate Yield from a Future Direct Imaging Mission,” ApJ, 795, 122

Traub, W. A. & Oppenheimer, B. R. 2010, “Direct Imaging of Exoplanets,” in Exoplanets, edited by S. Seager. Tucson, AZ: University of Arizona Press, 2010, 526 pp. ISBN 978-0-8165-2945-2., p.111-156

Turnbull et al. 2012, “THE SEARCH FOR HABITABLE WORLDS: 1. THE VIABILITY OF A STARSHADE MISSION,” PASP, 124, 418