Starshade exoplanet data challenge

Renyu Hu,[a,b,*] **Sergi R. Hildebrandt**,[a,c] **Mario Damiano**,[a]
Stuart Shaklan,[a] **Stefan Martin**,[a] **and Doug Lisman**[a]
[a]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, United States
[b]California Institute of Technology, Division of Geological and Planetary Sciences, Pasadena, California, United States
[c]California Institute of Technology, Division of Physics, Mathematics and Astronomy, Pasadena, California, United States

Abstract. Starshade in formation flight with a space telescope is a rapidly maturing technology that would enable imaging and spectral characterization of small planets orbiting nearby stars in the not-too-distant future. While performance models of starshade-assisted exoplanet imaging have been developed and used to design future missions, their results have not been verified from the analyses of synthetic images. Following a rich history of using community data challenges to evaluate image-processing capabilities in astronomy and exoplanet fields, the Starshade Technology Development to TRL5 (S5), a focused technology development activity managed by the NASA Exoplanet Exploration Program, is organizing and implementing a starshade exoplanet data challenge. The purpose of the data challenge is to validate the flow down of requirements from science to key instrument performance parameters and to quantify the required accuracy of noisy background calibration with synthetic images. This data challenge distinguishes itself from past efforts in the exoplanet field in that (1) it focuses on the detection and spectral characterization of small planets in the habitable zones of nearby stars, and (2) it develops synthetic images that simultaneously include multiple background noise terms—some observations specific to starshade—including residual starlight, solar glint, exozodiacal light, detector noise, as well as variability resulting from starshade's motion and telescope jitter. We provide an overview of the design and rationale of the data challenge. Working with data challenge participants, we expect to achieve improved understanding of the noise budget and background calibration in starshade-assisted exoplanet observations in the context of both starshade rendezvous with Roman and Habitable Exoplanet Observatory. This activity will thus help NASA prioritize further technology developments and prepare the science community for analyzing starshade exoplanet observations. [DOI: 10.1117/1.JATIS.7.2.021216]

Keywords: starshade; high-contrast imaging; exoplanet; data challenge; Roman space telescope; Habitable Exoplanet Observatory.

Paper 20157SS received Oct. 16, 2020; accepted for publication Mar. 9, 2021; published online Mar. 25, 2021.

1 Introduction

Data challenges have advanced the planning and development of major astronomy facilities both on the ground and in space. Science communities participate in the data challenges to analyze simulated data and gain insight into the detection capabilities of the instrument; in turn, instrument designers learn the precision and noise level needed to reveal objects and phenomena. For example, a series of image analysis challenges have been carried out to develop and test methods to measure weak gravitational lensing from small distortion of galaxies' shapes [e.g., the GRavitational lEnsing Accuracy Testing (GREAT) 3[1]]. These community exercises have been instrumental in consolidating the science plans for space missions such as Euclid. Another

*Address all correspondence to Renyu Hu, renyu.hu@jpl.nasa.gov

prominent example is the data challenges organized by the teams of LIGO-Virgo[2] and LISA,[2] which helped quantify the measurement precision needed to detect gravitational waves and inform the development of these experiments. Data challenges have also advanced many exoplanet projects. For example, a radial velocity fitting challenge was conducted to find the most efficient ways to extract planetary signals embedded in stellar noises.[3,4] For another example, a carefully planned data challenge has helped resolve discrepancies among groups in their data reduction and analysis approaches in exoplanet transit observations with Spitzer.[5]

More recently, community data challenges have been carried out to derive capabilities and inform instrument designs for the exoplanet science with Roman Space Telescope.[6,7] Past efforts include a challenge to efficiently identify and analyze exoplanetary microlensing events from large datasets[8] and the Roman Exoplanet Imaging Data Challenge.[9,10] The latter, probably the first community data challenge for space-based exoplanet imaging,[11] includes efforts to validate models of planetary reflected-light spectra at Roman's Coronagraph Instrument (CGI) wavelengths, to detect planets from simulated Roman images, and to determine planetary orbits from multi-epoch observations. The Roman Exoplanet Imaging Data Challenge has also explored enhanced planet detection and orbit determination capability with a starshade rendezvous at a late phase of the mission[12] while still focusing on detecting Jupiter-sized planets. At the time of writing, the final results of the Roman Exoplanet Imaging Data Challenge have not been published; but the atmospheric modeling effort (as phase I of that effort) succeeded in establishing the ability to retrieve key atmospheric parameters from simulated planetary spectra.[9]

Together with the development of the concepts of starshade rendezvous with Roman[12] and Habitable Exoplanet Observatory (HabEx),[13] both involving a starshade, NASA's Exoplanet Exploration Program (ExEP) is executing the Starshade Technology Development Activity to TRL5 (S5) to rapidly mature the technology and close gaps in optical performance, formation flying, and mechanical precision and stability. Together with S5, ExEP has chartered a Science and Industry Partnership (SIP) to engage the broader science and technology communities during the execution of the S5 activity. A key recommendation that emerged from SIP meetings and discussions is to produce "a flow down of requirements from science to key performance parameters based on synthetic images (rather than scaling formulas only)" and "a plan for the starshade data challenge." Responding to the community recommendation, S5 is now organizing and implementing a Starshade Exoplanet Data Challenge.

The Starshade Exoplanet Data Challenge seeks to verify and improve the exoplanet yield estimates[14–17] by using synthetic images that realistically capture instrumental effects due to the starshade and the telescope. With the completion of most of S5's technology milestones on instrument contrast,[18,19] solar glint,[20] and formation flying,[21] we can now simultaneously include in the images multiple sources of background and noise including residual starlight, solar glint, exozodiacal light, detector noise, as well as variability resulting from starshade's motion in formation flight and telescope's jitter. Many of these terms are specific to starshade observations; while some of them may be included in past exoplanet yield estimates,[14–17] the interplay of these terms of background and their noises can only be revealed and evaluated with the analyses of synthetic images.

A key science question that the Starshade Exoplanet Data Challenge is designed to answer is to what extent the background can be calibrated in the context of starshade-assisted exoplanet imaging. If the background is removed to its photon-noise limit, starshade rendezvous with Roman could provide nearly photon-limited spectroscopy of temperate and Earth-sized planets of F, G, and K stars <4 pc away, and HabEx could extend this capability to many more stars within 8 pc.[17] To achieve these capabilities, the flux of exozodiacal light within the planet's point spread function (PSF) often needs to be calibrated to a precision better than 1% and the solar glint better than 5%.[17] The challenges for photon-limited background calibration may come from the fact that the solar glint varies with the solar angle and the starshade's position and orientation[20] and that an exoplanetary dust disk is likely inclined and may have structures created by dynamical interactions with embedded planets.[22,23] Also, the expected use of slit-prism spectroscopy by Roman CGI may create complexity in spectral extraction together with the background. The Starshade Exoplanet Data Challenge will thus provide the opportunity to quantify the accuracy and precision of noisy background calibration for detection and spectral characterization of small exoplanets.

We expect the outcomes of the data challenge to help NASA identify and prioritize the areas of future technology development. As we assess the abilities to extract planets, dust structures, and their spectra from images that include varying levels of instrumental effects, we will improve our understanding of how critical the instrument performance parameters are. For example, we learn from signal-to-noise ratio calculations that an instrumental contrast of 10^{-10} is likely not needed for many science observations with starshade rendezvous, whereas suppressing solar glint and other stray light sources is of paramount importance.[17] The data challenge will tell us, with fidelity, how much performance would be lost if the contrast was 10^{-9} and the brightness of solar glint was a few times higher than the current best estimate (CBE). To summarize, the Starshade Exoplanet Data Challenge is designed to validate the flow down of requirements from science to key performance parameters, quantify the required accuracy and precision of noisy background calibration, and prepare the science community for analyzing starshade exoplanet observations.

2 Designs and Rationale of the Starshade Exoplanet Data Challenge

2.1 *Overall Structure*

Figure 1 shows the overall structure and workflow of the Starshade Exoplanet Data Challenge. S5 will simulate the images for the data challenge. The images will be generated with the Starshade Imaging Simulation Toolkit for Exoplanet Reconnaissance (SISTER[24,25]), which takes into account the full two-dimensional nature of the astrophysical scene and the spatial variation of the PSF due to the optical diffraction from the starshade. The simulations adopt the nominal performance parameters from current S5 results,[18–21] including the new optical edge coating that reduces the solar glint by a factor of 10.[26] Astrophysical and observational scenarios are selected to represent key science objectives of the well-studied starshade mission concepts, including Roman Rendezvous[12] and HabEx[13] (Table 1). To explore these astrophysical and observational scenarios, as well as key instrument performance parameters (Table 2), approximately 400 images will be simulated.

Participating teams of the data challenge will then develop image-processing algorithms to test the ability to retrieve faint exoplanet signals from the synthetic images and quantify the precision of background calibration. The participating teams will attempt to determine from the images the number of the planets and their locations and brightness, as well as to extract the inclination, density, and potential structures in the exozodiacal dust disk. Estimating the uncertainties of these parameters are essential, because the resulting S/N would be compared with the S/N estimated from idealized exposure time calculators.[17] With the simulated images of slit-prism spectroscopy for Roman and the data cubes of integral field spectroscopy for HabEx, the participating teams will also attempt to extract the planets' spectra. Results from the analyses will determine the detection limit of planets vis-à-vis instrument parameters and indicate how well image-processing algorithms can subtract the background to the photon-noise limit. These results will inform S5 of a realistic noise budget of starshade exoplanet observations and

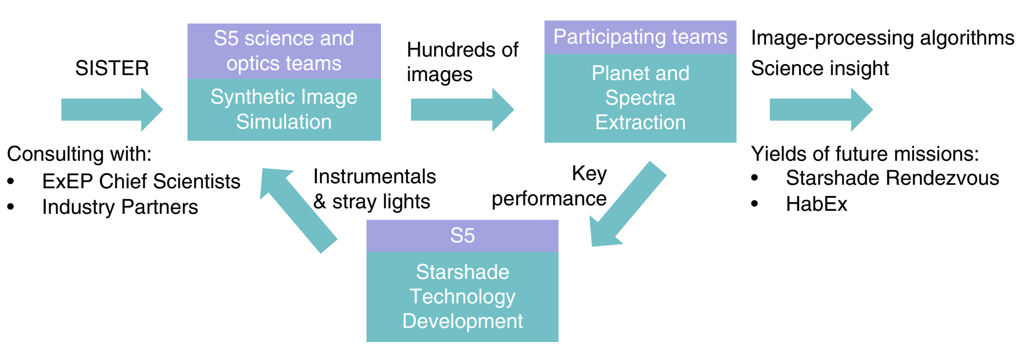

Fig. 1 Overall structure and workflow of the starshade exoplanet data challenge.

Table 1 Astrophysical and observational scenarios adopted by the Starshade Exoplanet Data Challenge.

Star	Exozodi	Planets	Roman Rendezvous	HabEx
τ Ceti	3, 10, and 30 zodis[a], inclination 35 deg, smooth versus clumpy	Two super-Earths, one hypothetical $1.0-R_\oplus$ planet	Imaging in blue[a] and green[a], spectroscopy in green	Spectroscopy 0.3 to 1.0 μm and 1.0 to 1.8 μm
ϵ Indi A	1, 3, and 10 zodis, inclination 30 deg and 80 deg, smooth versus clumpy	Multiple hypothetical $1.0-R_\oplus$ planets	Imaging in blue and green, spectroscopy in green	Spectroscopy 0.3 to 1.0 μm and 1.0 to 1.8 μm
σ Draconis	1, 3, and 10 zodis, inclination 30 deg and 80 deg, smooth versus clumpy	Multiple hypothetical $1.6-R_\oplus$ planets	Imaging in blue and green, spectroscopy in green	N/A
σ Draconis	1, 3, and 10 zodis, inclination 30 deg and 80 deg, smooth versus clumpy	Multiple hypothetical $2.4-R_\oplus$ planets	Imaging in blue and green, spectroscopy in green	N/A
σ Draconis	1, 3, and 10 zodis, inclination 30 deg and 80 deg, smooth versus clumpy	Multiple hypothetical $1.0-R_\oplus$ planets	N/A	Spectroscopy 0.3 to 1.0 μm and 1.0 to 1.8 μm
β CVn	1, 3, and 10 zodis, inclination 30 deg and 80 deg, smooth versus clumpy	Multiple hypothetical $2.4-R_\oplus$ planets	Imaging in blue and green, spectroscopy in green	N/A
β CVn	1, 3, and 10 zodis, inclination 30 deg and 80 deg, smooth versus clumpy	Multiple hypothetical $1.0-R_\oplus$ planets	N/A	Spectroscopy 0.3 to 1.0 μm and 1.0 to 1.8 μm

[a]Blue = 425 to 552 nm, green = 615 to 850 nm (see text). 1 zodi = surface opacity of the exozodiacal dust disk in the habitable zone of the star the same as the surface opacity of the zodiacal disk at 1 AU of the Solar System.[27]

Table 2 Instrumental effects and other backgrounds explored by the starshade exoplanet data challenge.

Item	Description	Variation
Residual starlight	Instrument contrast of 10^{-10} produced by random deviation of the edge of the petals	Contrast level up to 1×10^{-9}
Lateral displacement	Time-dependent lateral shift up to 1 m in formation flight	N/A
Solar glint	Time-dependent sunlight scattered by the optical edge with coating	Up to 3-times brighter than the CBE
Other stray light	Reflection of Milky Way, Earth, Jupiter, and leakage through micrometeoroid holes	N/A
Local zodical light	V band magnitude of 22.5 per arcsec^2	N/A
Telescope jitter	Random pointing error of 14 mas for Roman and 0.3 mas for HabEx	N/A
Integration time	Estimated for S/N per band or spectral element assuming photon-noise background calibration	S/N of 5, 10, and 20

requirements on key instrument performance parameters. The algorithms and science insight gained in the study will be disseminated among science communities via publications and code releases.

The first set of synthetic images has been released to the public in January 2021 and the community participation of the data challenge has started.[28] The data challenge is scheduled to be completed by September 2021.

2.2 *Astrophysical and Observational Scenarios*

Table 1 lists the astrophysical and observational scenarios adopted as the representative cases for the data challenge. These scenarios are chosen to probe the key and limiting science objectives of starshade rendezvous with Roman and HabEx. Four stars will be included in the study, including two stars <4 pc away (τ Ceti and ε Indi A), one star in the 5- to 6-pc distance range (σ Draconis), and one star in the ~8-pc distance range (β CVn). These stars are in the nominal target lists of both starshade rendezvous with Roman and HabEx.[12,13]

While detecting and spectrally characterizing Earth-sized planets in the habitable zones of all these stars belongs to HabEx's science goals, this would only be possible for <4-pc stars with Roman.[17,29] Therefore, we assume hypothetical $1.0 - R_\oplus$ planets when simulating HabEx observations, and larger planets when simulating Roman observing σ Draconis and β CVn. For the larger planets, we consider the two dominant populations of planets discovered by "Kepler":[30] the "super-Earth" population with a representative radius of 1.6 $R_\oplus$ and a larger-radius population with a representative radius of 2.4 $R_\oplus$. The super-Earth populations are likely dominated by large rocky planets,[31] and the larger-radius population can either be rocky planets with massive H_2/He gas envelopes[32,33] or planets with massive water envelopes.[34,35] We use Exo-REL,[36] a well-documented model for exoplanet clouds and reflected-light spectra, to simulate the input spectra of the planets.

The star τ Ceti deserves special attention because of its known planets and outer dust disk. The analyses of radial-velocity data have not reached a consensus, but the two outer planets (the planets e and f) near its habitable zone are consistent between analyses.[37,38] A debris disk has been detected with far-infrared and radio observations,[39–41] with an inner edge at ~6 AU, an outer edge at ~50 AU, and an inclination of ~35 deg. The knowledge motivates us to consider the following in the data challenge for τ Ceti. (1) We adopt the debris disk's inclination as the inclination of the planets and the exozodiacal disk. As such, the true masses of the planets e and f are ~6.9 $M_\oplus$. We further consider the possibility that the planets are either predominantly rocky or with large water envelopes[34] in estimating their radii and simulating the spectra. (2) We include the possibility of a denser exozodiacal disk to test its impact on planet detection (Table 1). The LBTI exozodiacal disk survey did not detect a disk at τ Ceti, but the $1 - \sigma$ upper limit is 44 zodis.[42] (3) We include another planet, with Earth's radius and Earth's mass, between the orbits of the planets e and f. This hypothetical planet is predicted by orbital dynamics and exoplanet population-level information.[43] We verify that the planet would be dynamically stable [44] and would induce a radial-velocity signal amplitude of ~6 cm s^{-1}, which is well below the detection limit of existing data.

In this data challenge, we will test the ability to detect planets embedded in their exozodiacal disk. Because it is not practical to simulate a self-consistent disk with the planets assumed—given the uncertainties in the source of the particles, the existence of outer planets, as well as the particle size distributions—we instead attempt to bound the problem by considering the end-members of a "smooth" disk and a "clumpy" disk. For the smooth disk, we adopt a solar-system disk density profile and use Zodipic[45] to simulate the intensity including the effects of inclination and particle forward scattering.[23] The clumpy disk represents a more challenging scenario for planet detection, where clumps of dust particles are trapped in mean-motion resonance with the planets.[22,23] To our knowledge, this would be the first time that potential structures of the exozodiacal disk are assessed against the imaging and detection of small exoplanets of nearby stars. We envision the data challenge would eventually inform us to what extent a clumpy disk would hinder the revelation of the planets in the system.

Lastly, we design the observational scenarios to mimic the basic ideas of operation outlined by starshade rendezvous with Roman[12] and HabEx.[13] Starshade rendezvous with Roman would perform broadband searches of small planets, and if any feasible planets are detected, conduct spectroscopy immediately following imaging in the green band (615 to 800 nm).[29] We adopt this philosophy by considering two visits per astrophysical scenario, each with both broadband imaging and slit-prism spectroscopy. We will additionally explore two aspects for starshade rendezvous with Roman. (1) We will include broadband imaging in the blue band (425 to 552 nm) and compare it with the green band. The blue band would have less exozodiacal light per resolution element and may thus lead to better planet detection. (2) We will include the green-band spectroscopy up to 850 nm, where the expected instrument contrast degrades from 10^{-10} at 800 nm to 10^{-9} at 850 nm.[12] This is motivated by the desire to extend coverage to longer wavelengths to reduce spectral degeneracies[46] and the realization that an instrument contrast of 10^{-9} may be sufficient.[17] HabEx would conduct most planet searches with its coronagraph and perform integral field spectroscopy with its starshade.[13] We thus focus on spectroscopy for HabEx, also considering two visits per astrophysical scenario. As all scenarios adopted contain planets that fit the definition of "high-interest" by HabEx, we will include near-infrared spectroscopy (1.0 to 1.8 μm) in addition to the 0.3 to 1.0 μm band for all scenarios. It is noteworthy that this entails two spectral integrations per visit, as the starshade needs to be located at a different separation from the telescope for conducting the near-infrared spectroscopy. Because this data challenge focuses on image processing and background calibration, it does not address potential synergy between a starshade and the CGI of Roman, or other operational and logistical constraints of the starshade.

2.3 *Instrument Effects*

Table 2 lists the instrumental and other effects to be included in the synthetic images and explored by the data challenge. We will adopt the CBEs for the residual starlight, solar glint, other stray light, and formation flying performance as summarized by Hu et al.[17] As the residual starlight contrast is the fundamental requirement that controls mechanical precision tolerance, we will vary the contrast level and see how much we could tolerate without adversely impacting planet detection and background calibration. We will adopt the latest estimate of the brightness of the solar glint[26] and also investigate how much brighter we could tolerate. Other stray light sources, such as the reflected light of the Milky Way, Earth, Jupiter, as well as the leakage through micrometeoroid holes, will be dimmer than the solar glint by a factor of at least 2^{16}, and their impact will be studied together with the enhancement factor applied to the solar glint (Table 2).

A new effect that will be simulated for the data challenge is the time-dependent variability of these backgrounds. As the solar glint is a starshade's closest analog to speckles in coronagraphic imaging, it is important to include the temporal variability of the stray light to achieve high-fidelity image simulations. We consider telescope jitter, lateral motion of the starshade in formation flight, as well as the change in Sun's angle during long integration as the main sources of temporal variability. For instance, images will be produced with realistic random pointing errors. This way, the data challenge will be able to assess the precision of background calibration achievable from real observations and how it compares with the photon-noise limit.

In addition to broadband imaging, we will simulate spectroscopy with Roman's slit-prism spectrometer and HabEx's integral field spectrometer. The simulations will mimic the dispersion of the planets and background sources and the recording of the spectra on the detector. The spectral resolution, as well as the width and orientation of the slit when applicable, will follow the specifications of the mission concepts. For Roman, we will study whether a specific slit orientation would be necessary or preferred for extracting planetary spectra from interfering backgrounds such as exozodiacal light. For instance, the slit may be oriented radially along the axis between the star and the planet, perpendicular to that axis, or at some angle to both—this may affect the ability to subtract off the exozodiacal background. To our knowledge, it will be the first time that slit-prism spectroscopy will be simulated and studied in detail for exoplanet direct imaging using a starshade. The knowledge gained in this analysis will help us better understand the background calibration and planet signal extraction in this form of spectroscopy.

In addition to the parameters listed in Table 2, we will adopt the telescope and instrument parameters designed for Roman and HabEx, including the optical throughput and detector properties. The frame rate will be chosen for each observation consistent with the photon counting mode of the electron-multiplying CCDs and the detector noise as the combination of dark current and clock-induced charge will be included accordingly.[16]

3 Summary and Expected Outcomes

To summarize, the Starshade Exoplanet Data Challenge will assess the noise budget of exoplanet observations using a starshade and determine the precision of background calibration achievable with synthetic images. As described in Sec. 2, the data challenge will be based on the hitherto most realistic simulations of starshade-assisted observations that explore the expected diversity of planet types as well as the density, structure, and inclination of the dust disks around the nearby stars. In addition to an ensemble of instrumental effects, we will include temporal variability of residual starlight and stray light due to the starshade's motion and telescope's jitter. The high-fidelity synthetic images will help us better understand the detection limit of planets and their spectra as a function of instrument performance parameters.

Specific outcomes that may be anticipated from the data challenge include (1) estimation of the S/N of planetary parameters (e.g., location, brightness, and spectrum) based on the synthetic images. The estimated S/N, in comparison with the idealized S/N used to set the integration time (Table 2), will tell us how precisely the exozodiacal light and solar glint can be calibrated for planet detection. (2) Detection of exozodiacal disks and constraints of their density, inclination, and possible clumpy structures. We will assess what we could learn about exozodiacal dust disks from direct-imaging observations using a starshade and also evaluate whether the clumpy structures would interfere with planet detection. (3) Extraction of planetary spectra, especially from slit-prism spectroscopy. Using synthetic images, we will evaluate whether it would be feasible to extract planetary spectra when both the planet and the extended background of exozodiacal light are dispersed onto the detector. As a whole, these insights will tell us what we could realistically expect to learn about planets and disks around nearby stars using a starshade, and how these capabilities would depend on the instrument contrast and the suppression of solar glint and other straylight.

The image-processing algorithms and science insight gained in this exercise may also advance high-contrast imaging astronomy in general. Compared to coronagraph direct imaging from the ground and in space, where creative algorithms have been developed to subtract backgrounds and speckles, starshade-assisted direct imaging mainly presents a different set of problems, associated with distinguishing planets from the exozodiacal disk, as well as removing stray light terms specific to the starshade. The data challenge will enable a cross-disciplinary development of techniques to be used to maximize the science yield of the future missions using starshades, ranging from step-by-step (e.g., image subtractions and feature extraction) to more holistic (e.g., Bayesian inversion and deep learning) image analyzing techniques. With the synthetic images, algorithms, science and technology insight, and community partnership, the Starshade Exoplanet Data Challenge will result in an enduring legacy that advances exoplanet astronomy in the years to come.

Acknowledgments

We thank Karl Stapelfeldt, Eric Mamajek, Christopher Stark, and Phil Willems for helpful discussions. The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (No. 80NM0018D0004). @2020 California Institute of Technology. Government sponsorship acknowledged.

References

1. R. Mandelbaum et al., "The Third Gravitational Lensing Accuracy Testing (GREAT3) challenge handbook," *Astrophys. J. Suppl. Ser.* **212**, 5 (2014).

2. D. Meacher et al., "Mock data and science challenge for detecting an astrophysical stochastic gravitational-wave background with Advanced LIGO and Advanced Virgo," *Phys. Rev. D* **92**, 063002 (2015).
3. X. Dumusque, "Radial velocity fitting challenge-I. Simulating the data set including realistic stellar radial-velocity signals," *Astron. Astrophys.* **593**, A5 (2016).
4. X. Dumusque et al., "Radial-velocity fitting challenge-II. First results of the analysis of the data set," *Astron. Astrophys.* **598**, A133 (2017).
5. J. G. Ingalls et al. "Results of the 2015 Spitzer exoplanet data challenge: repeatability and accuracy of exoplanet eclipse depths," in *Am. Astron. Soc. Meeting*, Abstract #228, 401.04 (2016).
6. D. Spergel et al., "Wide-field infrared survey telescope-astrophysics focused telescope assets WFIRST-AFTA 2015 report," arXiv:1503.03757 (2015).
7. R. Akeson et al., "The wide field infrared survey telescope: 100 Hubbles for the 2020s," arXiv:1902.05569 (2019).
8. R. Street and WFIRST Microlensing Science Investigation Team, "WFIRST: microlensing analysis data challenge," in *Am. Astron. Society Meeting,* Abstract #231, 158.06 (2018).
9. S. Hildebrandt and M. TurnbullExoplanet Data Challenge Team, "WFIRST: exoplanet data challenge. Atmospheric retrieval results," in *Am. Astron. Soc. Meeting,* Abstract #231, 158.03 (2018).
10. A. M. Mandell et al., "The WFIRST coronagraph exoplanet data challenge," in *Am. Astron. Soc. Meeting,* Abstract #233, 140.44 (2019).
11. https://www.exoplanetdatachallenge.com.
12. S. Seager et al., "Starshade rendezvous probe mission," *Bull. Am. Astron. Soc.* **51**, 106 (2019).
13. B. S. Gaudi et al., "The Habitable Exoplanet Observatory (HabEx) mission concept study final report," arXiv:2001.06683 (2020).
14. M. C. Turnbull et al., "The search for habitable worlds. 1. The viability of a starshade mission," *Publ. Astron. Soc. Pac.* **124**(915), 418 (2012).
15. C. C. Stark et al., "Maximized exoearth candidate yields for starshades," *J. Astron. Telesc. Instrum. Syst.* **2**(4), 041204 (2016).
16. C. C. Stark et al., "Exoearth yield landscape for future direct imaging space telescopes," *J. Astron. Telesc. Instrum. Syst.* **5**(2), 024009 (2019).
17. R. Hu et al., "Overview and reassessment of noise budget of starshade exoplanet imaging," *J. Astron. Telesc. Instrum. Syst.* **7**(2), 021205 (2021).
18. A. Harness et al., "Demonstration of high contrast in monochromatic light at a flight-like Fresnel number," NASA Starshade Technology Development Activity Milestone Report 1a, 2019, https://exoplanets.nasa.gov/internal_resources/1210/.
19. A. Harness et al., "Demonstration of high contrast in broadband light at a flight-like Fresnel number," NASA Starshade Technology Development Activity Milestone Report 1b, 2019, https://exoplanets.nasa.gov/internal_resources/1211/.
20. E. Hilgemann et al., "Demonstration of solar glint lobe scatter performance," NASA Starshade Technology Development Activity Milestone Report 3, 2019, https://exoplanets.nasa.gov/internal_resources/1544/.
21. T. Flinois et al., "Lateral formation sensing and control," NASA Starshade Technology Development Activity Milestone Report 4, 2018, https://exoplanets.nasa.gov/internal_resources/1077/.
22. C. C. Stark and M. J. Kuchner, "The detectability of exo-earths and super-Earths via resonant signatures in exozodiacal clouds," *Astrophys. J.* **686**(1), 637 (2008).
23. C. C. Stark, "The transit light curve of an exozodiacal dust cloud," *Astronom. J.* **142**(4), 123 (2011).
24. S. R. Hildebrandt et al., "SISTER: Starshade Imaging Simulation Toolkit for Exoplanet Reconnaissance," *J. Astronom. Telesc. Instrum. Syst.* **7**(2) (2021).
25. http://sister.caltech.edu
26. D. McKeithen et al., "Antireflection coatings on starshade optical edges for solar glint suppression," *J. Astron. Telesc. Instrum. Syst.* **7**(2), 021208 (2021).
27. C. C. Stark et al., "Maximizing the exoearth candidate yield from a future direct imaging mission," *Astrophys. J.* **795**, 122 (2014).

28. https://exoplanets.nasa.gov/exep/technology/starshade-data-challenge/.
29. A. Romero-Wolf et al., "Starshade rendezvous: exoplanet sensitivity and observing strategy," *J. Astron. Telesc. Instrum. Syst.* **7**(2), 021210 (2021).
30. B. J. Fulton and E. A. Petigura, "The California-Kepler survey. VII. Precise planet radii leveraging Gaia dr2 reveal the stellar mass dependence of the planet radius gap," *Astron. J.* **156**(6), 264 (2018).
31. L. A. Rogers, "Most 1.6 earth-radius planets are not rocky," *Astrophys. J.* **801**(1), 41 (2015).
32. J. E. Owen and Y. Wu, "The evaporation valley in the Kepler planets," *Astrophys. J.* **847**(1), 29 (2017).
33. S. Jin and C. Mordasini, "Compositional imprints in density–distance–time: a rocky composition for close-in low-mass exoplanets from the location of the valley of evaporation," *Astrophys. J.* **853**(2), 163 (2018).
34. L. Zeng et al., "Growth model interpretation of planet size distribution," *Proc. Natl. Acad. Sci. U. S. A.* **116**(20), 9723–9728 (2019).
35. O. Mousis et al., "Irradiated ocean planets bridge super-Earth and sub-Neptune populations," *Astrophys. J. Lett.* **896**(2), L22 (2020).
36. R. Hu, "Information in the reflected-light spectra of widely separated giant exoplanets," *Astrophys. J.* **887**, 166 (2019).
37. M. Tuomi et al., "Signals embedded in the radial velocity noise-periodic variations in the τ Ceti velocities," *Astron. Astrophys.* **551**, A79 (2013).
38. F. Feng et al., "Color difference makes a difference: four planet candidates around τ Ceti," *Astron. J.* **154**(4), 135 (2017).
39. J. Greaves et al., "The debris disc around τ Ceti: a massive analogue to the Kuiper belt," *Mon. Not. R. Astron. Soc.* **351**(3), L54–L58 (2004).
40. S. Lawler et al., "The debris disc of solar analogue τ Ceti: Herschel observations and dynamical simulations of the proposed multiplanet system," *Mon. Not. R. Astron. Soc.* **444**(3), 2665–2675 (2014).
41. M. A. MacGregor et al., "Alma observations of the debris disk of solar analog τ Ceti," *Astrophys. J.* **828**(2), 113 (2016).
42. S. Ertel et al., "The HOSTS survey for exozodiacal dust: observational results from the complete survey," *Astron. J.* **159**, 177 (2020).
43. J. Dietrich and D. Apai, "An integrated analysis with predictions on the architecture of the Ceti planetary system, including a habitable zone planet," *Astron. J.* **161**(1), 17 (2020).
44. S. R. Kane et al., "A catalog of Kepler habitable zone exoplanet candidates," *Astrophys. J.* **830**(1), 1 (2016).
45. M. Kuchner, "*Zodipic: zodiacal cloud image synthesis*," Astrophysics Source Code Library, ascl–1202 (2012).
46. M. Damiano, R. Hu, and S. R. Hildebrandt, "Multi-orbital-phase and multiband characterization of exoplanetary atmospheres with reflected light spectra," *Astron. J.* **160**(5), 206 (2020).

Renyu Hu is a scientist at the Jet Propulsion Laboratory (JPL) and his research strives to identify and characterize habitable environments in the Solar System and beyond. He is the Starshade scientist of the NASA Exoplanet Exploration Program, providing science leadership to the S5 Starshade Technology Development Activity and managing the Starshade Science and Industry Partnership program. He received his PhD in planetary science from Massachusetts Institute of Technology in 2013.

Sergi R. Hildebrandt is a scientist at the JPL and lecturer at the California Institute of Technology. He received a PhD in theoretical physics at the University of Barcelona. He has worked in the data analysis of the cosmic microwave background, adaptive optics in the visible, and more recently the development of SISTER, a user-friendly, open-source, tool that generates starshade simulations with high fidelity.

Mario Damiano is a postdoc fellow at JPL. He performs data analysis to study and interpret the spectroscopic characteristics of exoplanetary atmospheres to unveil atmospheric composition

and dynamics of these alien worlds. He is also an enthusiast of deep learning and AI. He received his PhD in physics and astronomy from the University College London in 2019.

Stuart Shaklan is the supervisor of the High Contrast Imaging Group in the Optics Section of the JPL. He received his PhD in optics from the University of Arizona in 1989 and has been with JPL since 1991.

Stefan Martin is a senior optical engineer at the JPL. He received his BSc degree in physics from the University of Bristol, United Kingdom, and his PhD in engineering from the University of Wales. At JPL, he has been the leader of the TPF-I Flight Instrument Engineering Team, testbed lead for the TPF-I Planet Detection Testbed, and payload lead for the HabEx Telescope design study. He is currently involved in starshade accommodation on future space telescopes, such as NGRST.

Doug Lisman is the systems engineering lead for starshade technology development at the JPL where he is a member of the Instrument Systems Engineering group. He received his BS degree in mechanical engineering from Washington University in St. Louis in 1984 and has been at JPL since 1984.