EXOPLANET IMAGING
COMMUNITY
DATA CHALLENGE

www.exoplanetdatachallenge.com

Julien Girard & Turnbull SIT
@djulik

STScI Liaison for the Coronagraph Instrument

Roman Coronagraph Instrument Information Sessions - October 26 & 28 2021 - IPAC/Virtual
To *broaden* and *deepen* our knowledge as exoplanet community

To get the community acquainted with the *Coronagraph Instrument* data’s *new contrast regime* and astrophysics that will be enabled:

- **giant planets in reflected light**

To *develop*, *use* and *improve* data simulation and analysis *tools*

To foster collaborations and *train* future exoplanet scientists!

[www.exoplanetdatachallenge.com](http://www.exoplanetdatachallenge.com)
M. Turnbull leads one of the 2 Science Investigation Teams (SIT) for the **Coronagraph Instrument** (form. CGI)

Data Challenge Team, ∈ Turnbull SIT

- Neil Zimmerman
- Eli Bogat
- Julien Girard
- Junellie Gonzalez
- Sergi Hildebrandt
- Neil Zimmerman
- Stephen Kane
- Chris Stark
- Tiffany Meshkat
- Zhexing Li
Roman Exoplanet Imaging Data Challenge: Organization

Ell Bogat (GSFC)
Julien Girard (STScI)
Junellie Gonzalez-Quiles (JHU)
Sergi Hildebrandt (JPL)
Stephen Kane (UCR)
Zhexing Li (UCR)
Avi Mandell (GSFC)
Tiffany Meshkat (IPAC)
Chris Stark (GSFC)
Maggie Turnbull (SETI, SIT PI)
Neil Zimmerman (GSFC)
A NEW CONTRAST REGIME, A NEW TYPE OF SCIENCE

GIANT EXOPLANETS IN REFLECTED LIGHT!
Roman Exoplanet Imaging DC & Coronagraph Science Cases

Self-luminous, Young Super Jupiters: Atmospheric Properties

- WFIRST CGI Passbands
- Imaging
- Spectroscopy

Mark Marley (Ames)

Circumstellar disks:
Protoplanetary (young)
Debris (mature)
Exozodi (mature, HZ)

Blind search for exoplanets

Garret, Savransky & Macintosh (2017)

Mature Jupiter Analogues in Reflected Light: Atmospheric Properties

- 1x Solar Metallicity
- Water Clouds
- 30x Solar Metallicity
- Alkali Clouds

Natasha Batalha & Mark Marley (Ames)

Orbital Solution and Mass Measurement

- HD 32207
- HD 141492
- TW Hya

John Debes (STScI)

Natasha Batalha & Mark Marley (Ames)

Credit: JPL ROMAN Project Science Team
Roman Exoplanet Imaging DC: Coronagraph Modes

Narrow field of view mode
- Full 360 deg
- Inner working angle (IWA): 3 lambda/D (0.15")
- Outer working angle (OWA): 9 lambda/D (0.45")
- Band 1: 575 nm, 10.1% bandwidth

Wide field of view mode
- Full 360 deg
- IWA: 6.5 lambda/D (0.43")
- OWA: 20 lambda/D (1.45")
- Band 4: 825 nm, 9.9% bandwidth

Spectroscopy mode
- Prism+ slit
- Band 3: 730 nm, 16.7% bandwidth

+ Starshade 360° large FoV @425-552nm (assuming a rendezvous)

Credit: JPL / 2020
An original Data Challenge
- **6 imaging epochs** of the same target throughout mission: 47 UMa
- 3 planet with matching and realistic **radial velocity data**
- Extract sources, compute relative photometry & astrometry, disentangle from background sources, exozodiacal light
- Compute orbital solution using all the information available

Variability: Phase curve!

Junellie Gonzalez
Orbitize/OFTI
(Blunt 2017)
Roman Exoplanet Imaging DC: HLC & Precursor RV data

Talk by Stephen Kane & Zhexing Li on RV precursor work

“Truth” (hackathon nº2)
4 planets, 2 detectable with RV

Target Star
(47 UMa: G1V, V=5.0)

Reference Star
(η UMa: B3V, V=1.9)

Precursor RV Data provided (hackathon nº2)

2006 - 2020 ~1 m/s accuracy (e.g. Keck)
2020 - 2024* ~0.3 m/s accuracy (e.g. NEID)

* assuming cross-instrument calibrations are ok
ROMAN Exoplanet Imaging DC: Starshade Rendezvous Simulations

Sergi Hildebrandt (JPL)

www.sister.caltech.edu
Roman Exoplanet Imaging DC: Contrast Regime (OS6-OS9)

For $5 < V < 6$ stars

The expected contrast is

$< 10^{-7}$ (required)

$\sim 10^{-9}$ (predicted)

100 to 1,000 times better than current facilities.

Optimistically, image "Jupiters" to "Neptunes" @50pc in reflected light!
Data Challenge Design: 3-Jupiter analogs + disk

Exozodiacal Cloud

Dynamical Simulations

Exozodiacal Debris Disk Model

Stark & Kuchner et al. 2008, 2009

Junellie Gonzalez-Quiles (GSFC)
Chris Stark (STScI)
Neil Zimmerman (GSFC)
Sergi Hildebrandt (JPL)

ΔT = 0.00 yr

ZIMMERMAN+ (IN PREP)
In-house Analysis
The Roman Exoplanet Imaging DC: 4 Steps

**DATA**

6 imaging epochs throughout the mission

- Realistic simulations: OS6 Speckle field time series, detector model, background contamination sources, exozodiacal light

- Hybrid Lyot Coronagraph
  - 4 epochs, 2 rolls
  - + Calibrations

- Star Shade
  - 2 epochs
  - + Calibrations

& 15 years of precursor RV data

**CHALLENGE**

1. Extract & identify point sources in 4 HLC epochs, disentangle from background sources, provide census and rough astrometry

2. Compute orbital parameters & masses with those 4 epochs, use priors from RV data

3. Refine orbital parameters & masses using additional 2 SS epochs, all the information available

4. For a given planet, measure the phase curve assuming it is Lambertian, provide radius & albedo given mass-radius relationship

www.exoplanetdatachallenge.com
4 Tutorial Events: a Young & Diverse Crowd

CALTECH/IPAC
JUNE ‘19

Tokyo
OCTOBER ‘19

www.exoplanetdatachallenge.com
~ 70 people participated in person to our four "hack events"

Diverse crowd in age, seniority, gender and country of origin / workplace

8 teams (1 to 4 persons) entered the competition

4 of them have completed the step 3 and have access to the star shade data

One of these “top” participants did not attend our events
DC: HLC PSF Subtraction / Processing

Neil Zimmerman

4 HLC epochs

Noiseless “truth”

With OS6 Speckle Field (co-added frames)

PSF subtracted / processed

The “science grade data product”

We have to work with

ZIMMERMAN+ IN PREP

(DC DESIGN, IN HOUSE ANALYSIS FOR PLANET C & CODE)
In-house Analysis: PSF Subtracted Images
Uses cross-correlation to a high resolution PSF to obtain accurate astrometry for each signal.
In-house Analysis: Final Orbital Fit

Ell Bogat & Neil Zimmerman

Finds initial orbital constraints from radial velocity data using RadVel

Uses orbitize! (parallel MCMC chain fit) to fit astrometry data for each planet with prior constraints from RadVel

Truth versus In-House Analysis
Metrics, Results & Prize
Step 1 results: quick astrometry / identification

7 Teams
Step 1 Astrometry: Epoch 1 (T + 0.00 yr), HLC

- Host Star
- HLC OWA
- Truth Values
- Team Pentagon
- Team Square
- Team Hexagon
- Team Triangle
- Team Plus
- Team Diamond
- Team X

- planet_d
- planet_c
- planet_b
- bg_object
Final astrometry / identification by epoch

5 Teams
Final Astrometry: Epoch 3 (T + 1.00 yr), HLC
Final Astrometry: Epoch 4 (T + 2.00 yr), HLC

- **Host Star**
- **HLC OWA**
- **Planet b**
- **Planet c**
- **Planet d**
- **Background star (not visible)**
- **Truth Values**
- **Team Triangle**
- **Team Diamond**
- **Team Pentagon**
- **Team Square**
- **Team Hexagon**

[Graph showing various celestial objects and their positions]
Final Astrometry: Epoch 6 (T + 4.00 yr), Starshade

- **Host Star**
- **Planet b**
- **Planet c**
- **Planet d**
- **Background star** (not visible)

- **Truth Values**
- **Team Triangle**
- **Team Diamond**
- **Team Pentagon**
- **Team Square**
- **Team Hexagon**

**Axes:**
- Dec (mas)
- RA (mas)

**Color Scale:**
- $10^{-1}$
- $10^{-2}$

**Units:**
- photons/sec

**Scale:**
- 50 mas
Final Planet Count for the 3 Best Teams

**TEAM PENTAGON**
13/18 detections  0 false positives
Mass: 0.3% error  Radius/Albedo: No estimate

**TEAM TRIANGLE**
10/18 detections  0 false positives
Mass: 0.5% error  Radius/Albedo: 21% error

**TEAM DIAMOND**
10/18 detections  1 false positive
Mass: 11% error  Radius/Albedo: 49% error

FINALIST!  FINALIST!  FINALIST!
Final Planet Count for the 2 Next Teams (4th-5th)

Mis-matching of the planets caused significant errors

False positives and mis-matching of the planets caused significant errors
# Challenge Results: Planet detection matrix

<table>
<thead>
<tr>
<th>Time Elapsed (years)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Planet b</td>
<td>Flux Ratio</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(HLC) 0.00</td>
<td>3.58e-10</td>
<td>4.54e-09</td>
</tr>
<tr>
<td></td>
<td>□□□□□□□□</td>
<td>▲▼▼▼▼▼▼</td>
</tr>
<tr>
<td>(HLC) 0.15</td>
<td>6.9e-10</td>
<td>4.52e-09</td>
</tr>
<tr>
<td></td>
<td>□□□□□□□□</td>
<td>▲▼▼▼▼▼▼</td>
</tr>
<tr>
<td>(HLC) 1.00</td>
<td>5.19e-09</td>
<td>3.61e-09</td>
</tr>
<tr>
<td></td>
<td>□▼▼▼▼▼▼</td>
<td>▲▼▼▼▼▼▼</td>
</tr>
<tr>
<td>(HLC) 2.00</td>
<td>4.82e-09</td>
<td>1.83e-09</td>
</tr>
<tr>
<td></td>
<td>□▼▼▼▼▼▼</td>
<td>▲▼▼▼▼▼▼</td>
</tr>
<tr>
<td>(Starshade) 3.00</td>
<td>4.96e-10</td>
<td>5.6e-10</td>
</tr>
<tr>
<td></td>
<td>□□□□□□□□</td>
<td>▲▼▼▼▼▼▼</td>
</tr>
<tr>
<td>(Starshade) 4.00</td>
<td>2.4e-09</td>
<td>1.21e-10</td>
</tr>
<tr>
<td></td>
<td>□▼▼▼▼▼▼</td>
<td>▲▼▼▼▼▼▼</td>
</tr>
</tbody>
</table>

**Legend**
- ● Detection
- ○ Non-Detection
- Easy
- Medium
- Difficult

- Planet c is the easiest
- Team ● detected Planet d for very challenging epochs (experience & post-processing skills)
Step 2 & 3: Refined Astrometry / orbital fit - Nominal Planet c

Final Astrometry Results: Planet c

![Graphs showing the final astrometry results for Planet c. The plots display the shifted positions of point sources in the sky, with different symbols representing various teams and time points. The x and y axes represent ΔRA and ΔDec in milliarcseconds (mas).]
Step 2 & 3: Refined Astrometry / orbital fit - Outer Planet d
Step 2 & 3: Refined Astrometry / orbital fit - Inner Planet b

Final Astrometry Results: Planet b

ΔDec [mas] vs. ΔRA [mas]

- Planet Orbit
- HLC IMA/OWA
- Truth Values
- Team Square
- Team Triangle
- Team Hexagon
- Team Pentagon
### SAT style scoring:

- 1 point awarded for a planet detection
- 0.25 points subtracted for a false positive
- etc.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Step 1</th>
<th>Astrometry</th>
<th>Photometry</th>
<th>Orbit</th>
<th>Mass</th>
<th>Step 4</th>
<th>Total</th>
<th>Ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼</td>
<td>15</td>
<td>75</td>
<td>40</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>230</td>
<td>W</td>
</tr>
<tr>
<td>▼</td>
<td>4.29</td>
<td>60</td>
<td>30</td>
<td>40</td>
<td>40</td>
<td>25</td>
<td>199.29</td>
<td></td>
</tr>
<tr>
<td>▼</td>
<td>12.86</td>
<td>45</td>
<td>50</td>
<td>30</td>
<td>10</td>
<td>12.5</td>
<td>160.36</td>
<td></td>
</tr>
<tr>
<td>▼</td>
<td>10.71</td>
<td>30</td>
<td>10</td>
<td>20</td>
<td>20</td>
<td>0</td>
<td>90.71</td>
<td>C</td>
</tr>
<tr>
<td>▼</td>
<td>8.57</td>
<td>15</td>
<td>20</td>
<td>10</td>
<td>30</td>
<td>0</td>
<td>83.57</td>
<td></td>
</tr>
<tr>
<td>▼</td>
<td>6.43</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.43</td>
<td>-</td>
</tr>
<tr>
<td>▼</td>
<td>2.14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.14</td>
<td>-</td>
</tr>
</tbody>
</table>
Team Identities: Who’s Who?

Check the jamboree page presentations from 3 finalists!
exoplanetdatachallenge.com/home/jamboree

<table>
<thead>
<tr>
<th>Team</th>
<th>Symbol</th>
<th>Team Identities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang</td>
<td>⬤</td>
<td>Jason Wang</td>
</tr>
<tr>
<td>Princeton</td>
<td>▼</td>
<td>Leonid Pogorelyuk &amp; Brianna Lacy</td>
</tr>
<tr>
<td>Wagner</td>
<td>♂</td>
<td>Kevin Wagner</td>
</tr>
<tr>
<td>Tanner</td>
<td>♂</td>
<td>Angelle Tanner</td>
</tr>
<tr>
<td>Planet Hunters</td>
<td>■</td>
<td>Mia Hu, Jonathan Brande, Tomás Silva, Taichi Uyama</td>
</tr>
<tr>
<td>Milou</td>
<td>✤</td>
<td>Julien Milli</td>
</tr>
<tr>
<td>Agrawal</td>
<td>✗</td>
<td>Shubham Agrawal</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team</th>
<th>Symbol</th>
<th>Rank</th>
<th>Prize</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang</td>
<td>⬤</td>
<td>15</td>
<td>75</td>
</tr>
<tr>
<td>Princeton</td>
<td>▼</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Wagner</td>
<td>♂</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Tanner</td>
<td>♂</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Planet Hunters</td>
<td>■</td>
<td>10</td>
<td>25</td>
</tr>
<tr>
<td>Milou</td>
<td>✤</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Agrawal</td>
<td>✗</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

W = 50
C = 40
Metrics, Results & Prizes

3 winning teams

2 challenger teams

2 other teams
Data Challenge: Winners

Data Challenge Results

exoplanetdatachallenge.com/home/results

3 Winning Teams

Team Princeton
Team Wang
Team Wagner
The CGI contrast regime and capabilities offers exciting science demonstration prospects in addition to the technological pathways: imaging giant planets in reflected light.

We are better prepared to exploit the real Roman Coronagraph data even if the DC data (OS6) was quite optimistic.

We are able to perform orbital retrieval on HLC simulations (> 2-3 epochs) using real exoplanetary systems with RV trends (not yet Gaia).
Engaging the community is not too difficult but getting (young) people to commit and submit (results) is not easy as they are already pressured with paper writing, graduate school applications, etc. It could be a good idea to involve their supervisors early on. We encouraged people to team up and it has given some positive results in the engagement.

Several teams have found decent astrometric solutions for at least one planet with or without the priors from RV precursor data.

All participants struggled with calibrating photometry: we improved our tutorial on this matter.
One team (Wang) has been able to recover a challenging planet in some epochs for which we thought it was not possible! Post-processing and experience on precursor data helps! It might be determinant for OS9, OS11 and the real data.

Some participants preferred to develop their own tools rather than use the publicly available packages (potential added value).

A few bugs have been found and it has been rewarding for the public package developers (e.g. for orbitize!) which suddenly get many avid testers.
Papers on the Data Challenge itself

**GIRARD+**
Online (2020)
SPIE Proceedings
- General paper on the Data Challenge

**TURNBUL+**
Online (2021)
JATIS Special issue on Star Shades
- Focus on star shade
- Focus on planet d

**ZIMMERMAN+**
In prep (2021)
Astronomical Journal
- Challenge design
- & HLC simulations
- In-house analysis
- & code for planet c

**GIRARD+**
In prep (2021)
Astronomical Journal
- Challenge organization
- In-house analysis for all 3 planets
- Lessons learned
- Participating teams feedback

www.exoplanetdatachallenge.com
Papers from SIT members related to the DC

HILDEBRANDT+
Online (2020)
AAS Proceedings

• SISTER, Starshade Imaging Simulation Toolkit

LI (ZHEXING)+
Online (2021), on arxiv/ADS
Astronomical Journal
“Direct Imaging of Exoplanets Beyond the Radial Velocity Limit”

• Compares Roman Coronagraph + Star shade rendezvous with HabEx

SAXENA+
Online (2021), on arxiv/ADS
Astronomical Journal
"Simulating Reflected Light Exoplanet Spectra of the Promising Direct Imaging Target, \( \upsilon \) Andromedae d, with a New, Fast Sampling Method Using the Planetary Spectrum Generator"

www.exoplanetdatachallenge.com
Legacy Tutorial suite (Jupyter notebooks)
Roman Exoplanet Imaging Data Challenge: Website

Data Challenge
Team, Jamboree/Results, Timeline, Events

Turnbull SIT
Final Report, Publications, etc.

Data & Tutorial
Legacy Tutorial, older versions

Links & Resources
New section for the SIT

www.exoplanetdatachallenge.com