Predicting Earth-like-planets-hosting systems

Jeanne Davoult - DLR

PLATO-ESP2025: Planets throughout the Habitable Zone

24.06.2025

Diversity and correlations in exoplanet demographic and architecture

- Positive relation between stellar metallicity/mass and giant planets

e.g. Santos+2001, Johnson+2010, Bonfils+2013

Diversity and correlations in exoplanet demographic and architecture

- Positive relation between stellar metallicity/mass and giant planets
- Lower-mass planets (Super-Earths, Mini-Neptune) are more common than giant planets

e.g. Mayor+2011, Mulders+2015

Diversity and correlations in exoplanet demographic and architecture

- Positive relation between stellar metallicity/mass and giant planets
- Lower-mass planets (Super-Earths, Mini-Neptune) are more common than giant planets
- Abundance of packed and regular inner systems

  ➔ Peas-in-Pod architecture

e.g. Lissauer+2011, Millholland+2017, Weiss+2018

Diversity and correlations in exoplanet demographic and architecture

- Positive relation between stellar metallicity/mass and giant planets
- Lower-mass planets (Super-Earths, Mini-Neptune) are more common than giant planets
- Abundance of packed and regular inner systems

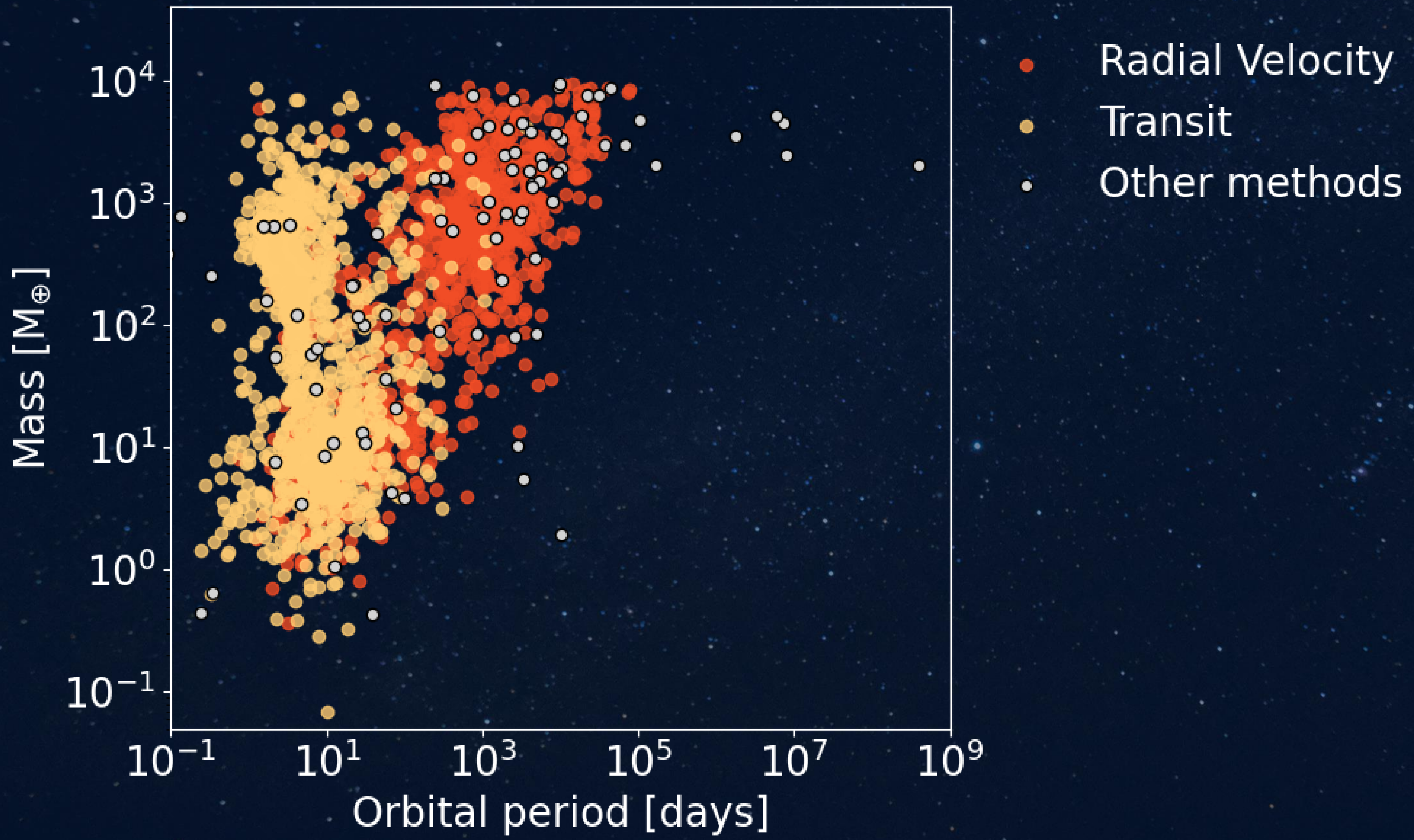
  ➔ Peas-in-Pod architecture

- Correlation between outer giant planets and inner terrestrial planets

  ➔ linked with the metallicity

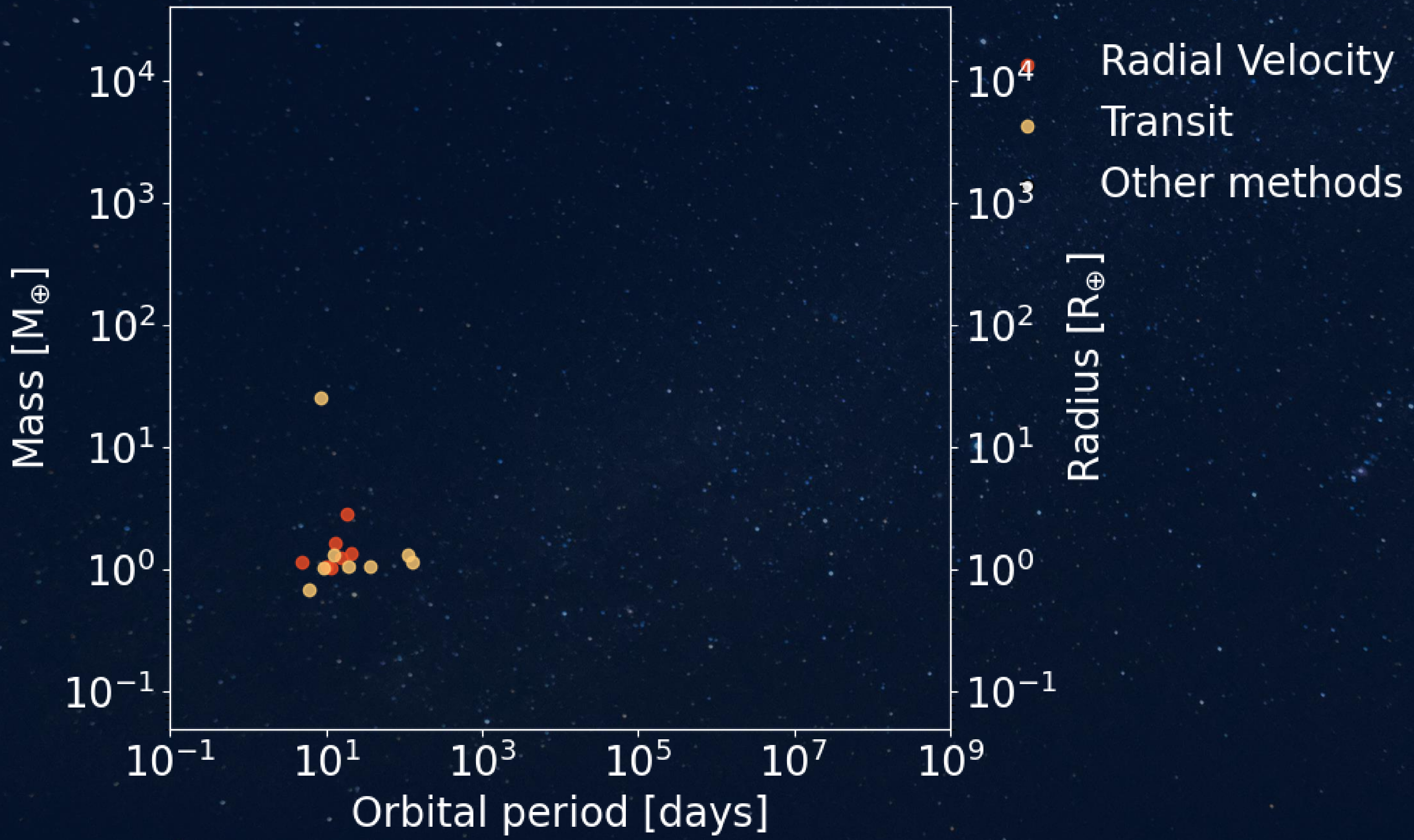
e.g. Zhu&wu2018, Bryan+2019

Exoplanet demographic



Earth-like: 0.5 < Rp < 1.5 Rearth OR 0.5 < Mp < 3 Mearth
0.25 < Sp < 1.1 Searth

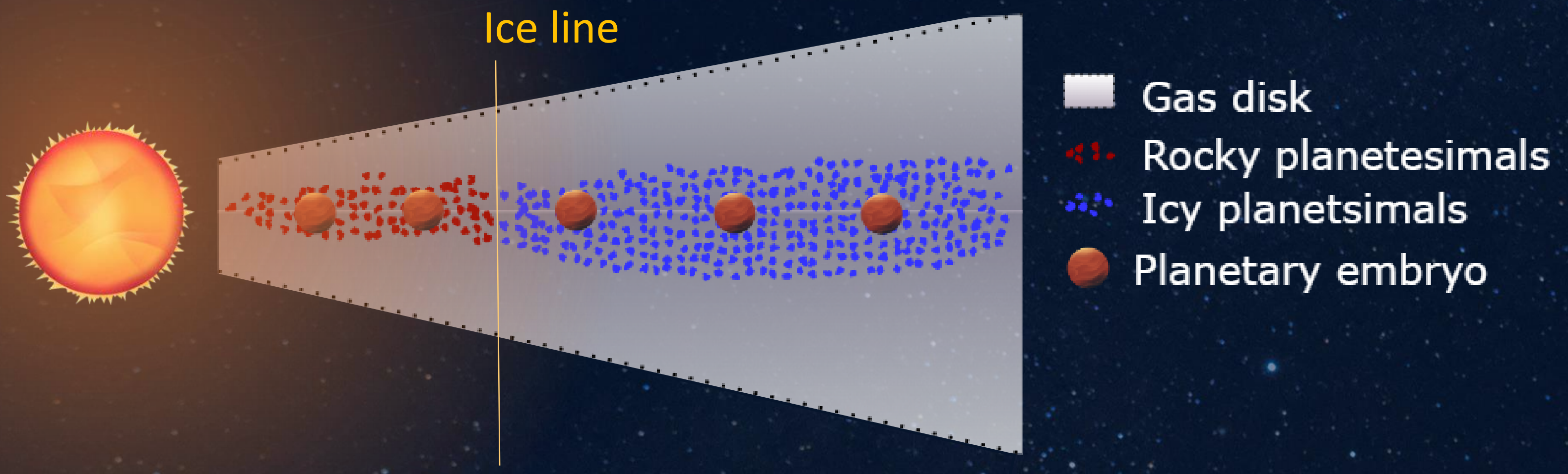
Exoplanet demographic



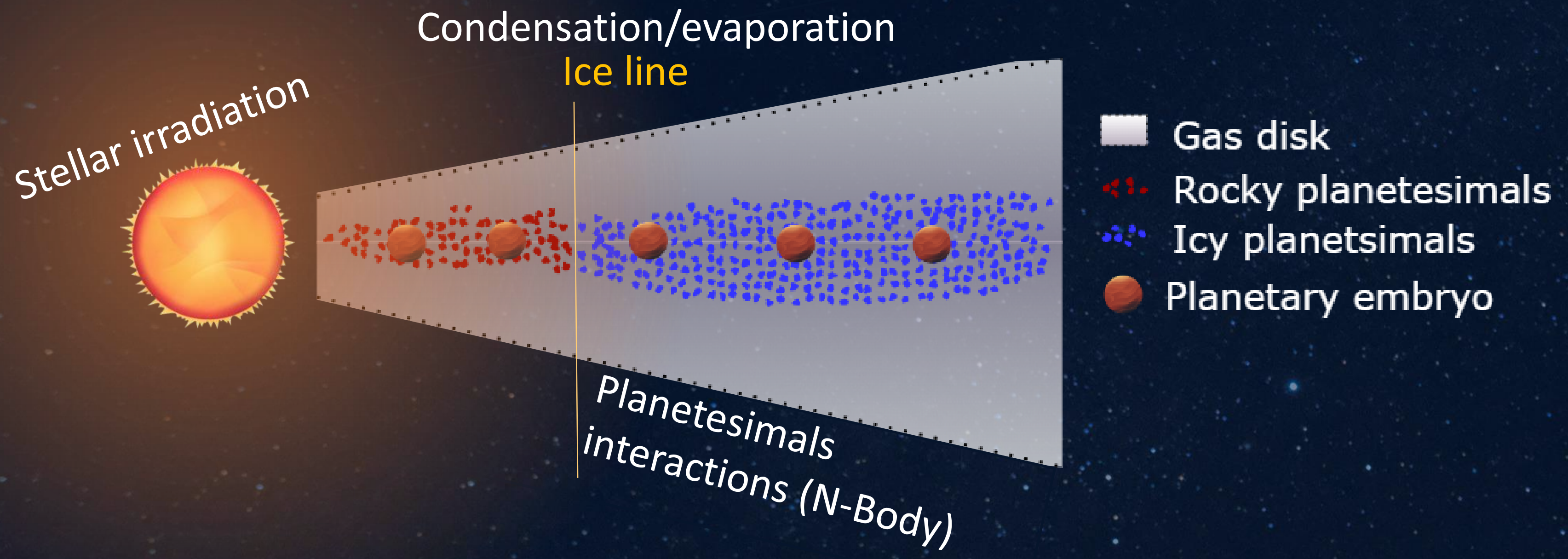
The Bern model

A population synthesis model for the formation and evolution of planetary systems

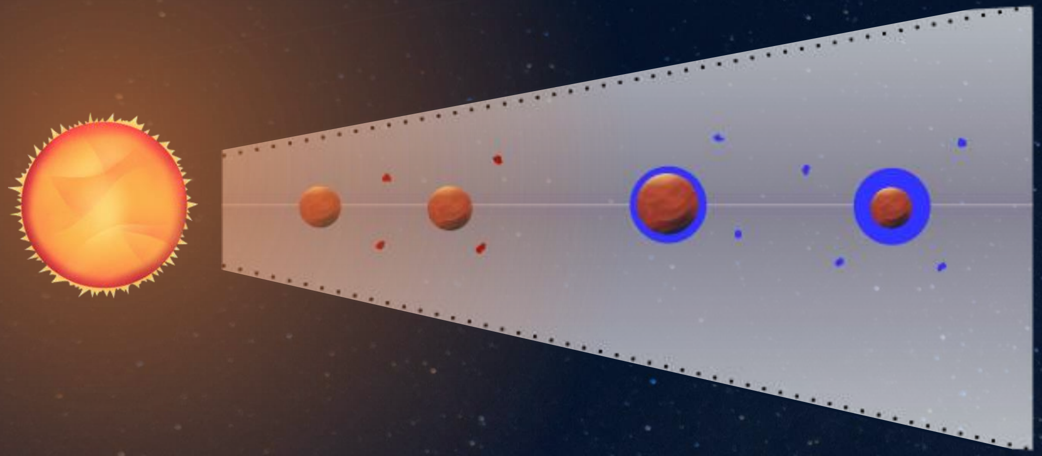
The Bern model



The Bern model



The Bern model



Core accretion paradigm: solid and gas accretion

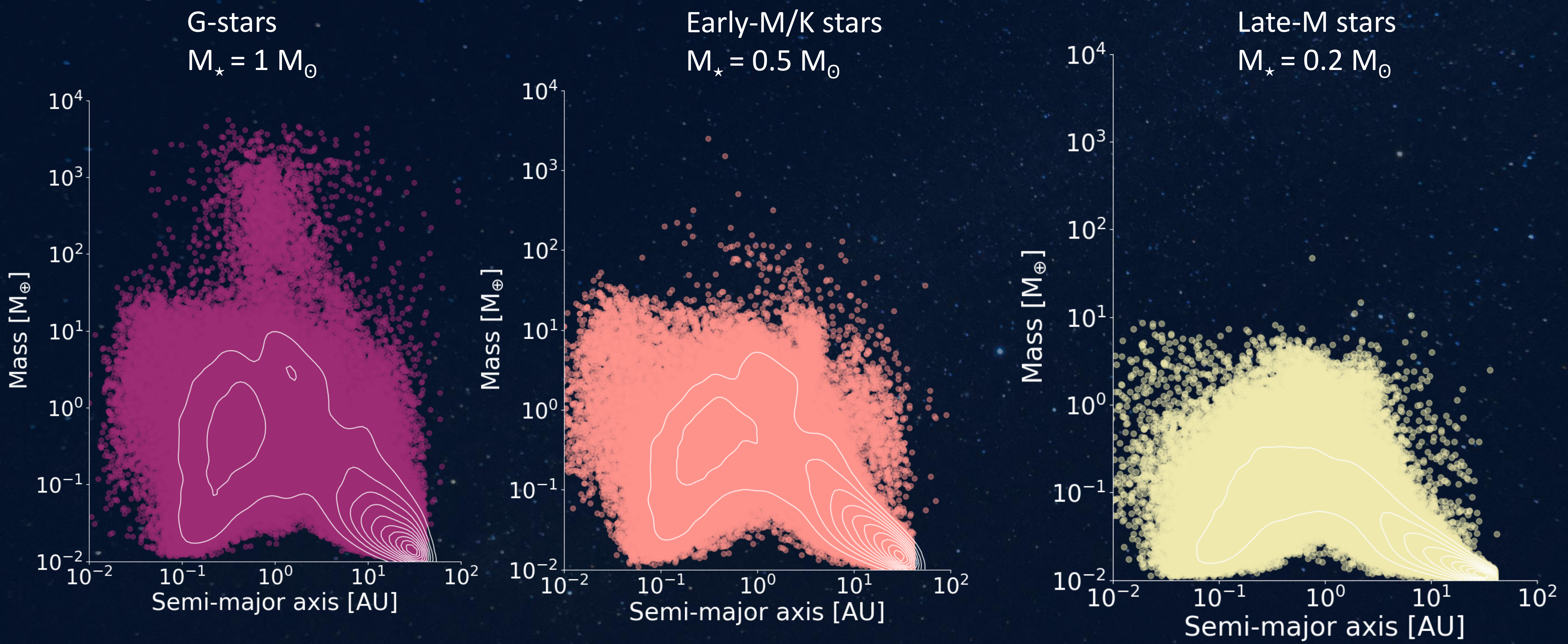
The Bern model

The Bern model

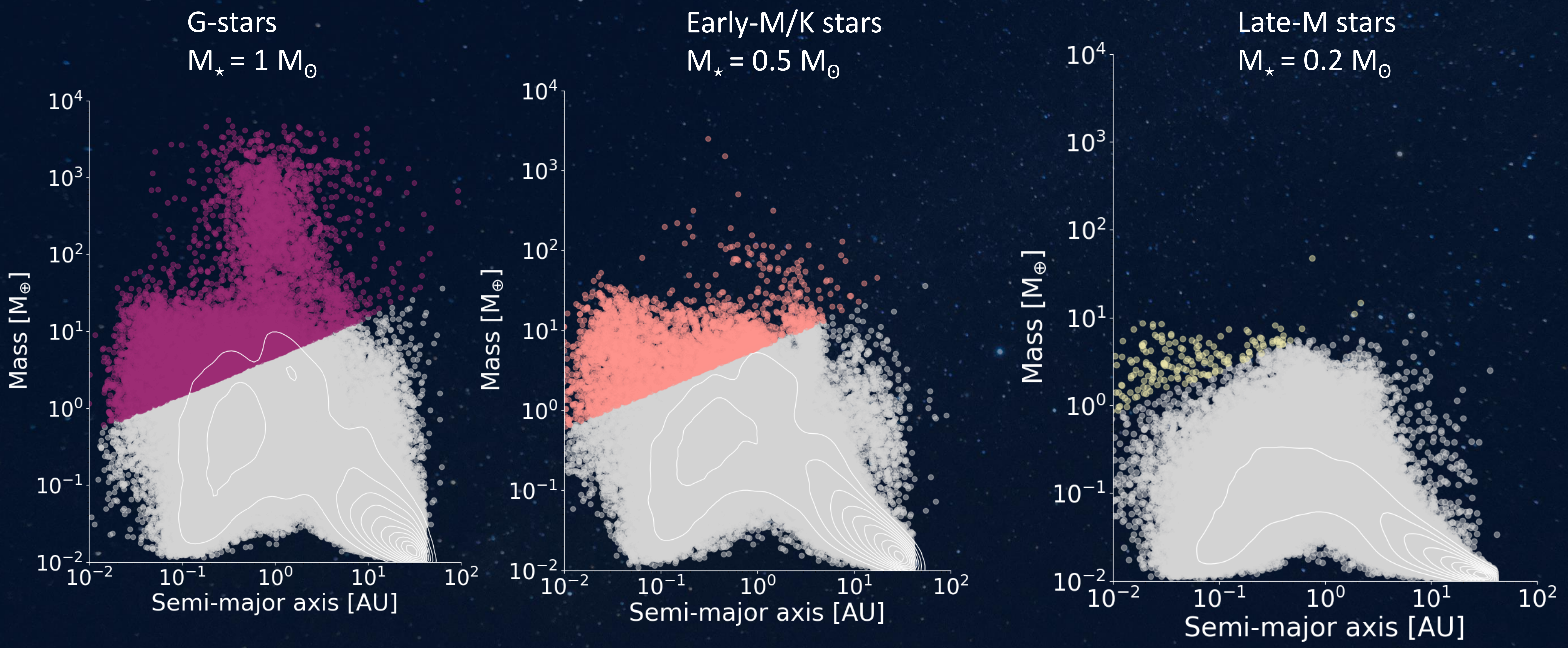
The Bern model

The Bern model

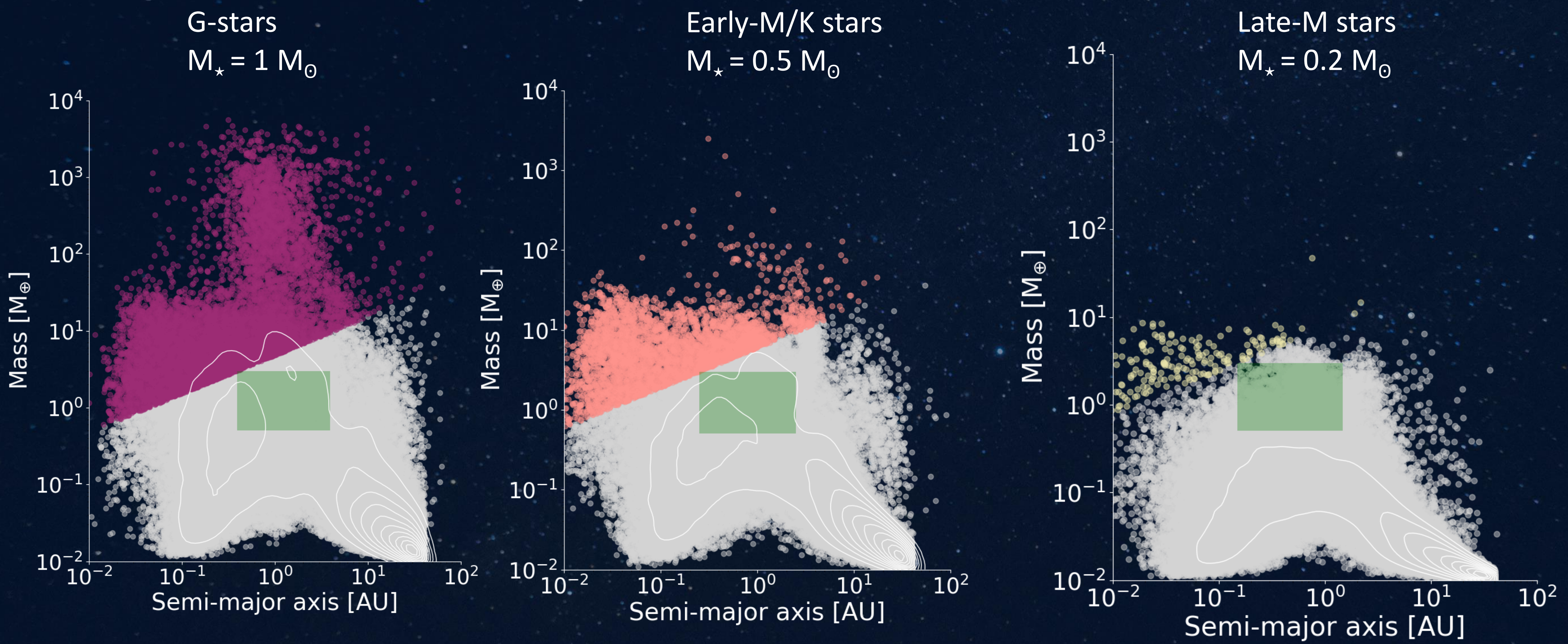
Populations of synthetic planetary systems



Populations of synthetic planetary systems

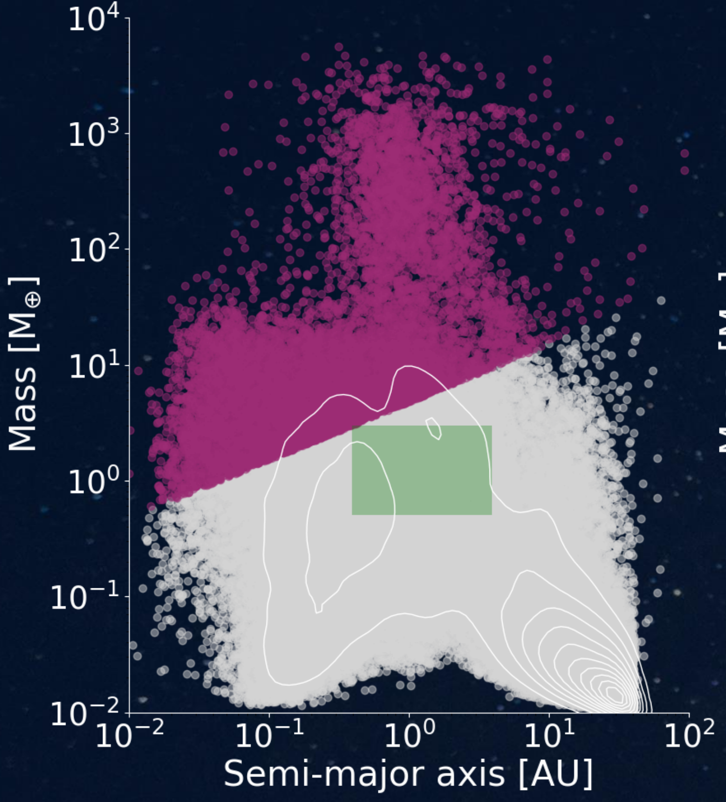


Populations of synthetic planetary systems



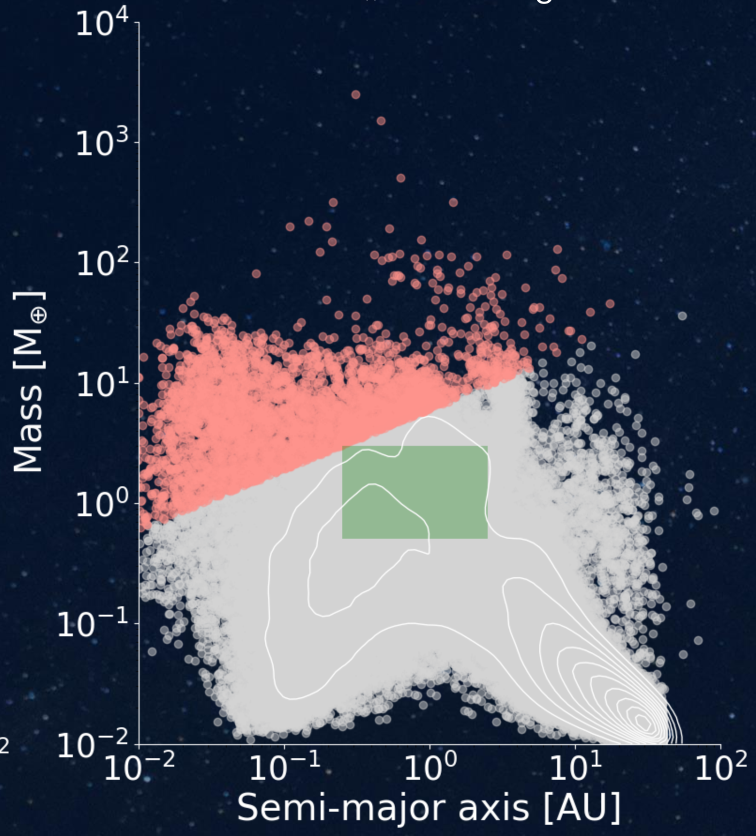
Populations of synthetic planetary systems

G-stars
$M_\star = 1\ M_\odot$



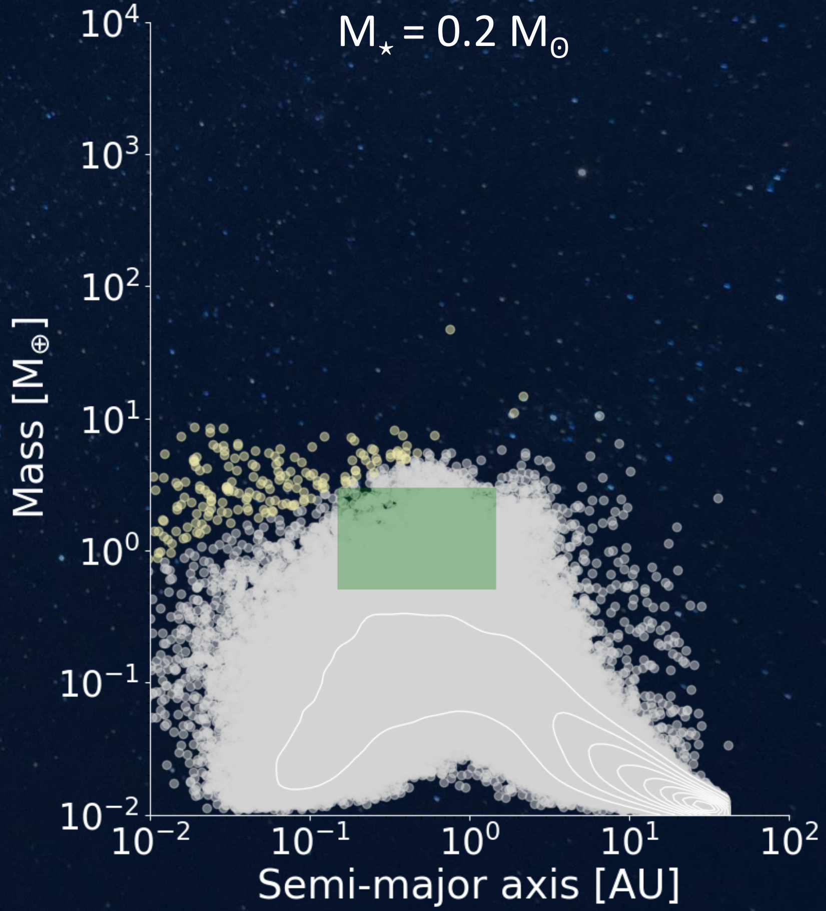
$\eta_{syst,ELP} = 0.6$

Early-M/K stars
$M_\star = 0.5\ M_\odot$



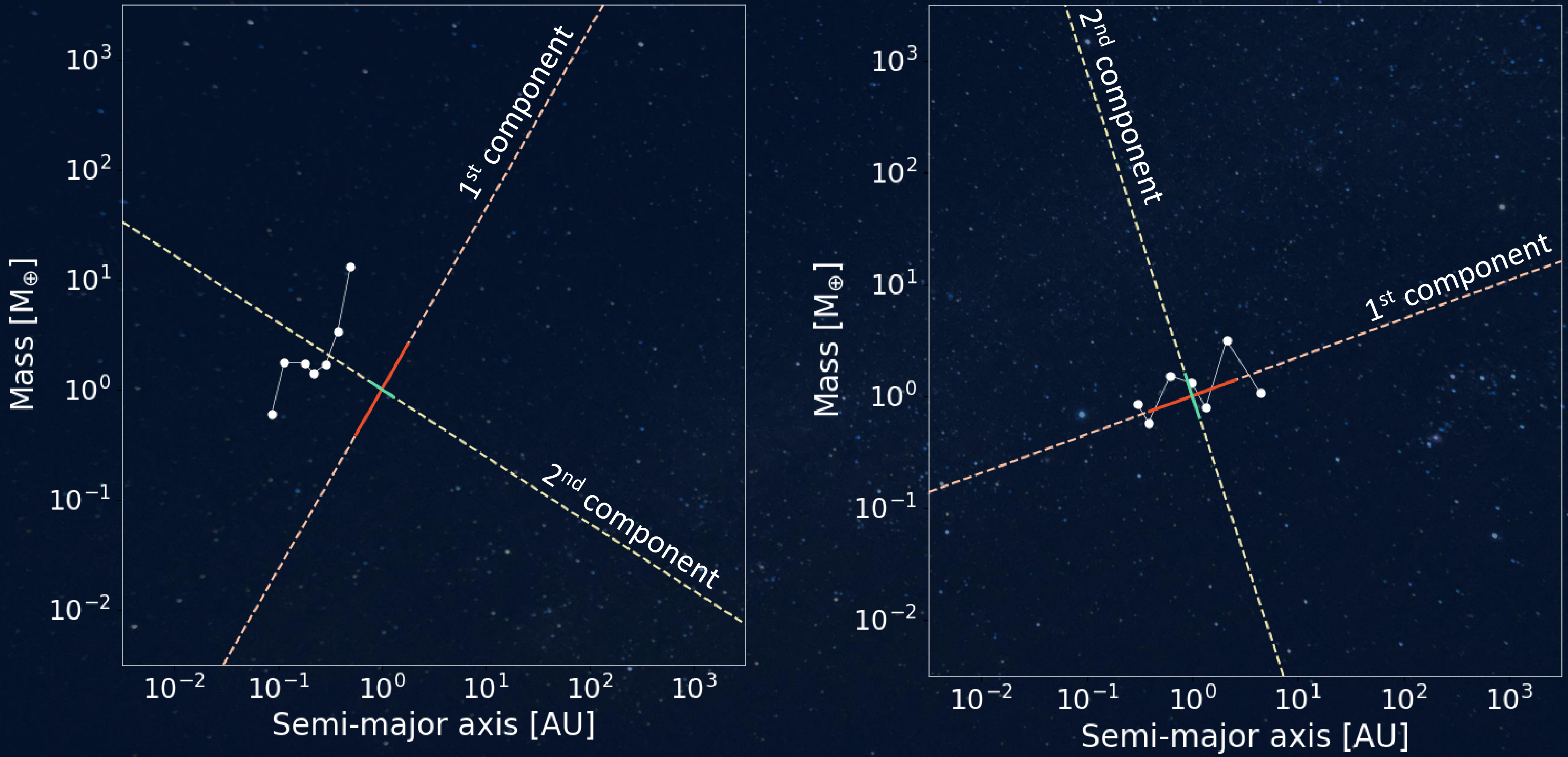
$\eta_{syst,ELP} = 0.74$

Late-M stars
$M_\star = 0.2\ M_\odot$

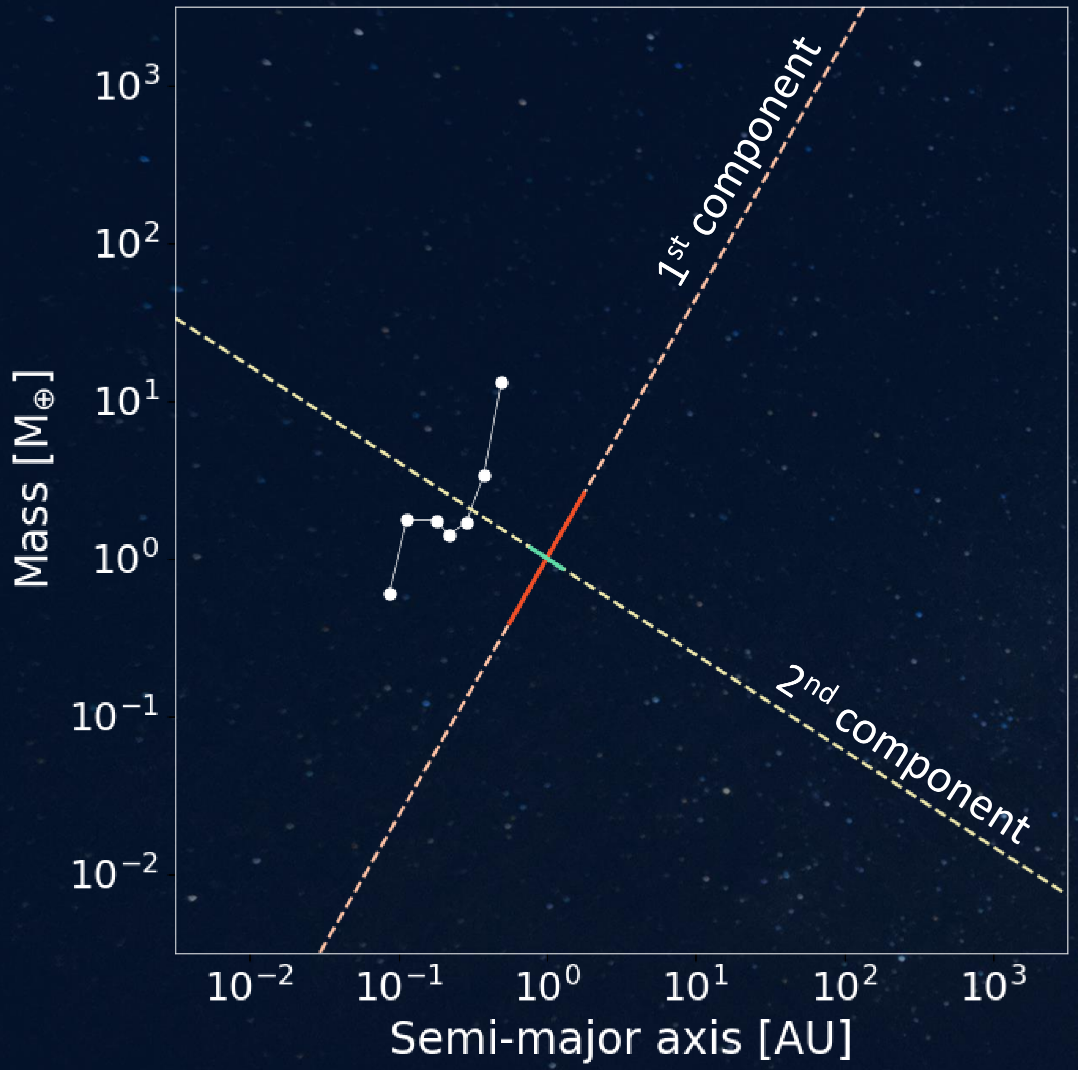


$\eta_{syst,ELP} = 0.4$

Principal component analysis (PCA)

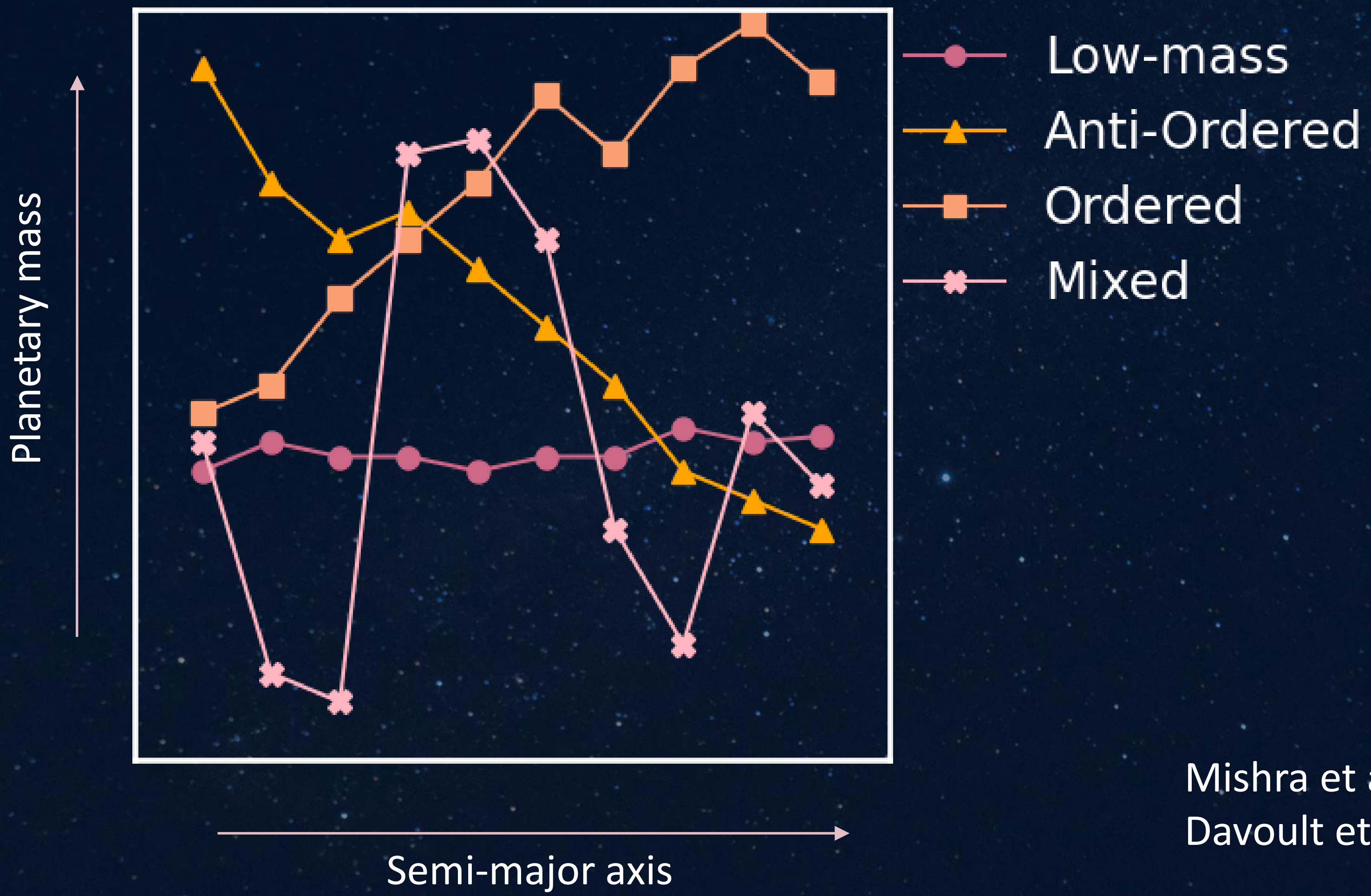


Principal component analysis (PCA)



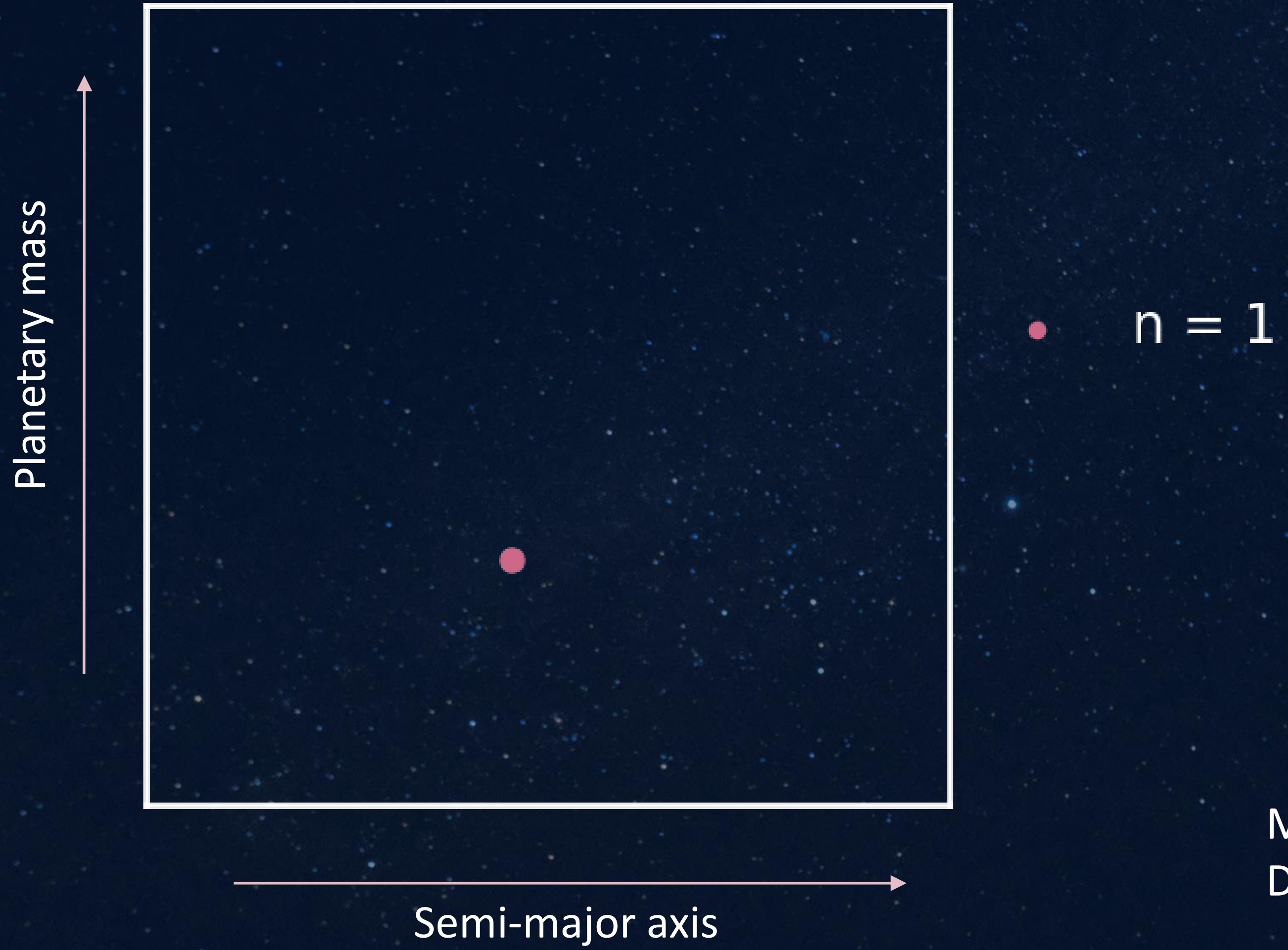
- Slope of the 1st Component: $S(C_1)$
- Variance of the 2nd Component: $V(C_2)$
- Mass of the most massive planet
- Total number of planets in the system

System architecture



Mishra et al. (2023)
Davoult et al. (2024)

System architecture



Mishra et al. (2023)
Davoult et al. (2024)

Profile of ELP-hosting systems – Davoult et al. (2024)

	$M_\star = 1\ M_\odot$	$M_\star = 0.5\ M_\odot$	$M_\star = 0.2\ M_\odot$
Low-mass	$R_{IDP} < 2.5\ R_\oplus$ → 55.8% $R_{IDP} > 2.5\ R_\oplus$ → 88% $P_{IDP} < 10$ days → 38% $P_{IDP} > 10$ days → 83%	$R_{IDP} < 2.75\ R_\oplus$ → 64% $R_{IDP} > 2.75\ R_\oplus$ → 95% $P_{IDP} < 10$ days → 60% $P_{IDP} > 10$ days → 79%	88%
Anti-Ordered	$M_{IDP} < 100\ M_\oplus$ → 38% $M_{IDP} > 100\ M_\oplus$ → 6% $R_{IDP} < 10\ R_\oplus$ → 34% $R_{IDP} > 10\ R_\oplus$ → 5%	N.A.	N.A.
Ordered	$M_{IDP} < 10\ M_\oplus$ → 31% $M_{IDP} > 10\ M_\oplus$ → 7% $R_{IDP} < 6\ R_\oplus$ → 30% $R_{IDP} > 6\ R_\oplus$ → 3%	$M_{IDP} < 10\ M_\oplus$ → 50% $M_{IDP} > 10\ M_\oplus$ → 22% $R_{IDP} < 2\ R_\oplus$ → 50% $R_{IDP} > 2\ R_\oplus$ → 34%	N.A.
Mixed	$M_{IDP} < 10\ M_\oplus$ → 32% $M_{IDP} > 10\ M_\oplus$ → 13% $R_{IDP} < 2.5\ R_\oplus$ → 27% $R_{IDP} > 2.5\ R_\oplus$ → 8%	N.A.	N.A.
n = 1	$M_{IDP} < 100\ M_\oplus$ → 95% $M_{IDP} > 100\ M_\oplus$ → 4% $R_{IDP} < 8\ R_\oplus$ → 95% $R_{IDP} > 8\ R_\oplus$ → 8% $P_{IDP} < 30$ days → 37% $R_{IDP} > 30$ days → 91%	$M_{IDP} < 10\ M_\oplus$ → 93% $M_{IDP} > 10\ M_\oplus$ → 90% $R_{IDP} < 2.75\ R_\oplus$ → 75% $R_{IDP} > 2.75\ R_\oplus$ → 97% $P_{IDP} < 10$ days → 50% $P_{IDP} > 10$ days → 96%	94%

- IDP = Innermost Detectable Planet
- N.A. : Not Applicable

Profile of ELP-hosting systems – Davoult et al. (2024)

	$M_\star = 1\,M_\odot$	$M_\star = 0.5\,M_\odot$	$M_\star = 0.2\,M_\odot$
Low-mass	$R_{IDP} < 2.5\,R_\oplus$ → 55.8% **$R_{IDP} > 2.5\,R_\oplus$ → 88%** $P_{IDP} <$ 10 days → 38% **$P_{IDP} >$ 10 days → 83%**	$R_{IDP} < 2.75\,R_\oplus$ → 64% **$R_{IDP} > 2.75\,R_\oplus$ → 95%** $P_{IDP} <$ 10 days → 60% **$P_{IDP} >$ 10 days → 79%**	**88%**
Anti-Ordered	$M_{IDP} < 100\,M_\oplus$ → 38% $M_{IDP} > 100\,M_\oplus$ → 6% $R_{IDP} < 10\,R_\oplus$ → 34% $R_{IDP} > 10\,R_\oplus$ → 5%	N.A.	N.A.
Ordered	$M_{IDP} < 10\,M_\oplus$ → 31% $M_{IDP} > 10\,M_\oplus$ → 7% $R_{IDP} < 6\,R_\oplus$ → 30% $R_{IDP} > 6\,R_\oplus$ → 3%	$M_{IDP} < 10\,M_\oplus$ → 50% $M_{IDP} > 10\,M_\oplus$ → 22% $R_{IDP} < 2\,R_\oplus$ → 50% $R_{IDP} > 2\,R_\oplus$ → 34%	N.A.
Mixed	$M_{IDP} < 10\,M_\oplus$ → 32% $M_{IDP} > 10\,M_\oplus$ → 13% $R_{IDP} < 2.5\,R_\oplus$ → 27% $R_{IDP} > 2.5\,R_\oplus$ → 8%	N.A.	N.A.
n = 1	**$M_{IDP} < 100\,M_\oplus$ → 95%** $M_{IDP} > 100\,M_\oplus$ → 4% **$R_{IDP} < 8\,R_\oplus$ → 95%** $R_{IDP} > 8\,R_\oplus$ → 8% $P_{IDP} <$ 30 days → 37% **$R_{IDP} >$ 30 days → 91%**	**$M_{IDP} < 10\,M_\oplus$ → 93%** $M_{IDP} > 10\,M_\oplus$ → 90% $R_{IDP} < 2.75\,R_\oplus$ → 75% **$R_{IDP} > 2.75\,R_\oplus$ → 97%** $P_{IDP} <$ 10 days → 50% **$P_{IDP} >$ 10 days → 96%**	**94%**

- IDP = Innermost Detectable Planet
- N.A. : Not Applicable

Earth-like planet Predictor – Davoult et al. (2025)

A Machine Learning approach

- $M_\star$
- Architecture
- $M_{IDP}, R_{IDP}, P_{IDP}$

➔ Correlated with ELPs

Random Forest

Decision Tree
max_depth = 5

x500

Earth-like planet Predictor – Davoult et al. (2025)

A Machine Learning approach

- $M_\star$
- Architecture
- $M_{IDP}, R_{IDP}, P_{IDP}$

➔ Correlated with ELPs

Random Forest

Decision Tree
max_depth = 5

x500

➔

Performance metrics:

- Precision score = $\frac{TP}{TP+FP}$
- Recall score = $\frac{TP}{TP+FN}$

Earth-like planet Predictor – Davoult et al. (2025)

A Machine Learning approach

- $M_\star$
- Architecture
- $M_{IDP}, R_{IDP}, P_{IDP}$

➔ Correlated with ELPs

Random Forest

Decision Tree
max_depth = 5

x500

➔

Voting rate threshold	Precision score
> 70%	0.85
> 80%	0.98
> 90%	0.99

Earth-like planet Predictor: 44 results

G stars

Systems	Voting rates
HD 103949	96%
HD 42618	95%
HD 85390	93%
HIP 41378	**97%**
Kepler-22	92%
Kepler-538	96%
KMT-2021-BLG-0171L	94%

K stars

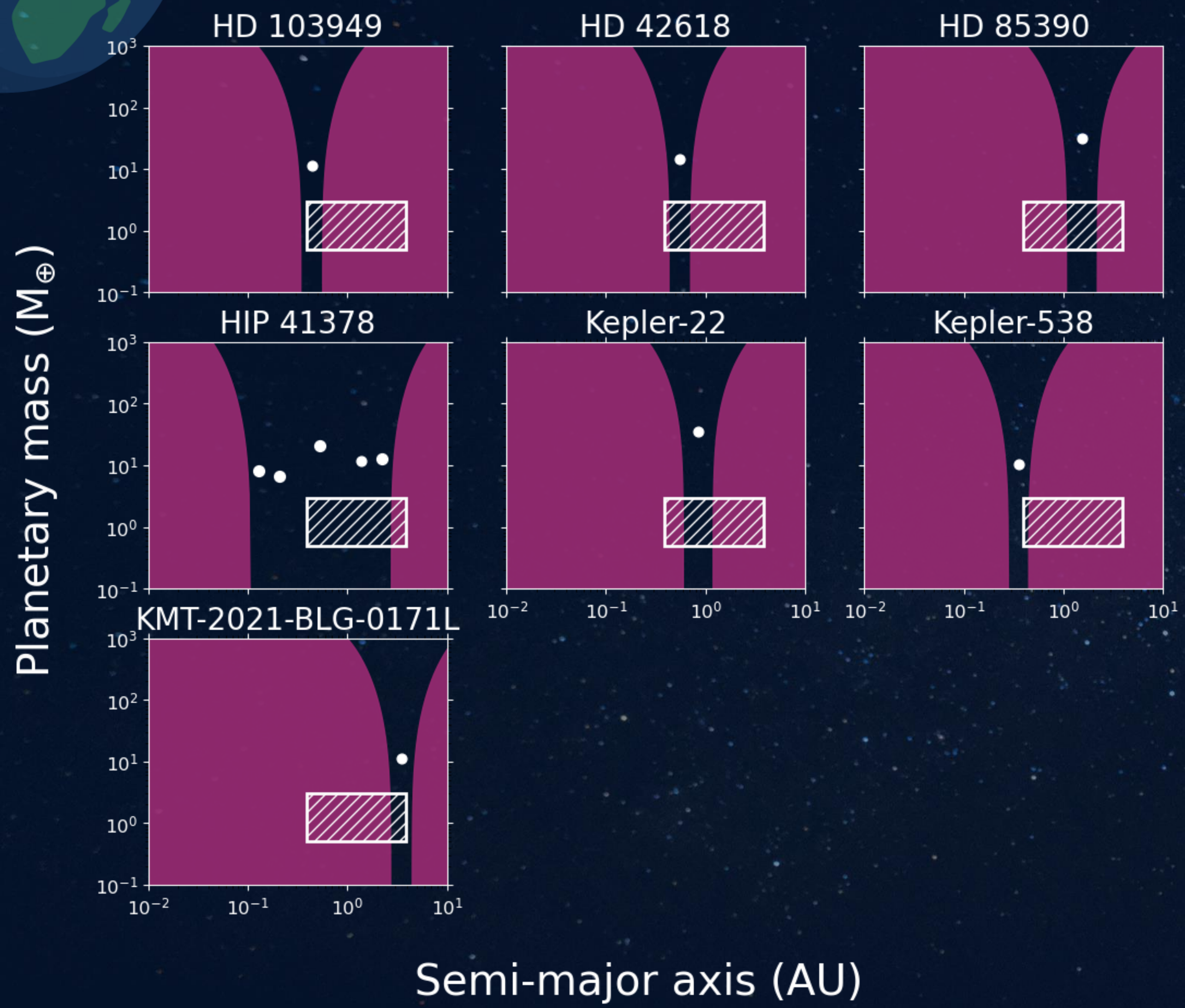
Systems	Voting rates
OGLE-2017-BLG-1691L	**97%**
GJ 685	92%
Gl 514	95%
HD 147379	96%
HD 211970	92%
HIP 71135	95%
K2-286	94%
KMT-2022-BLG-0440L	**98%**
KMT-2022-BLG-0475	92%
OGLE-2007-BLG-368L	96%
OGLE-2015-BLG-0966L	96%
OGLE-2015-BLG-1670L	96%
OGLE-2018-BLG-0506	96%
OGLE-2018-BLG-0516	95%
OGLE-2018-BLG-1126	96%
OGLE-2018-BLG-1185	96%
TCP J050742+244755	**97%**
TOI-1231	91%
TOI-2285	91%

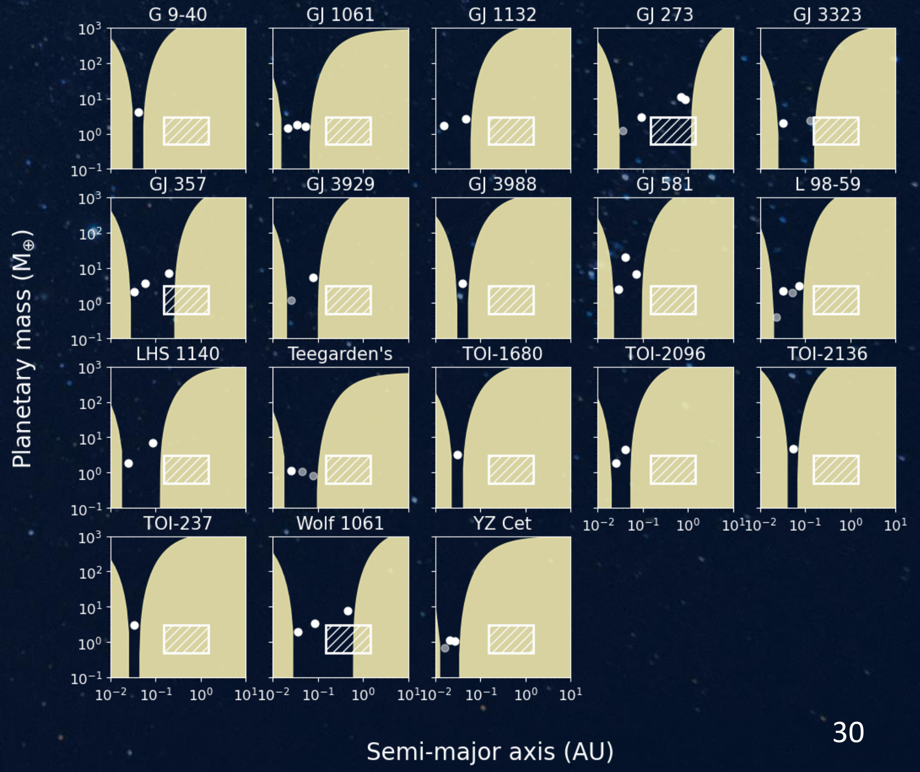
M-stars

Systems	Voting rates
G 9-40	91%
GJ 1061	**98%**
GJ 1132	94%
GJ 273	92%
GJ 3323	95%
GJ 357	**98%**
GJ 3929	91%
GJ 3988	91%
GJ 581	**98%**
L 98-59	**98%**
LHS 1140	95%
Teegarden's	95%
TOI-1680	91%
TOI-2096	**98%**
TOI-2136	91%
TOI-237	92%
Wolf 1061	**98%**
YZ Cet	**98%**

Stability criterion from Fabrycky+2014, Chen+2024

Earth-like planet Predictor: G stars





Take-away messages

- In the Bern model, we found correlations between the presence of Earth-like planets and observable properties of their systems such as their architecture, the mass, radius and period of the innermost detectable planet

- A Machine Learning model with very high performance during the training phase identified 44 systems as the most likely to host an Earth-like planet, and a study of their stability confirmed this possibility

- But one of the caveat of these studies, is how we deal with observational bias which is something I address in upcoming works (**Eltschinger, Davoult et al. (in prep.)** for RV with HARPS or ESPRESSO and **Davoult et al. (in prep.)** for photometry with PLATO)

- PLATO will populate an unreachable region of planet demographic so far, allowing a better comparison between models and data