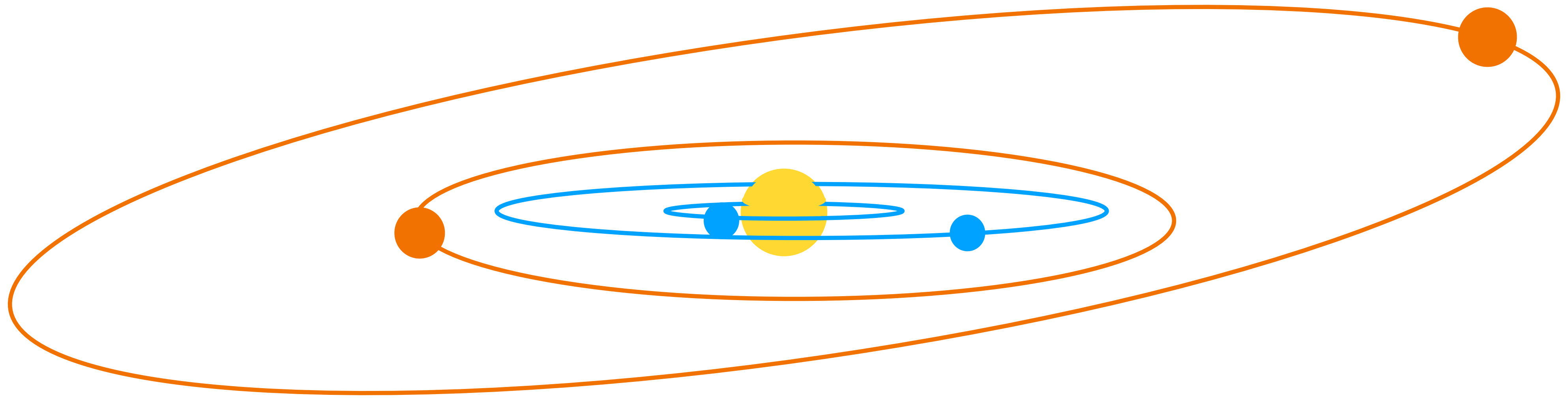
Upstream knowledge of targets with radial velocities

Nathan Hara
Laboratoire d’astrophysique de Marseille

HALO workshop
5 December 2024

Outline

(1) The output of *HWO, PCS* and *LIFE* would be greatly enhanced by a pre-detection of suitable targets (precursor survey)

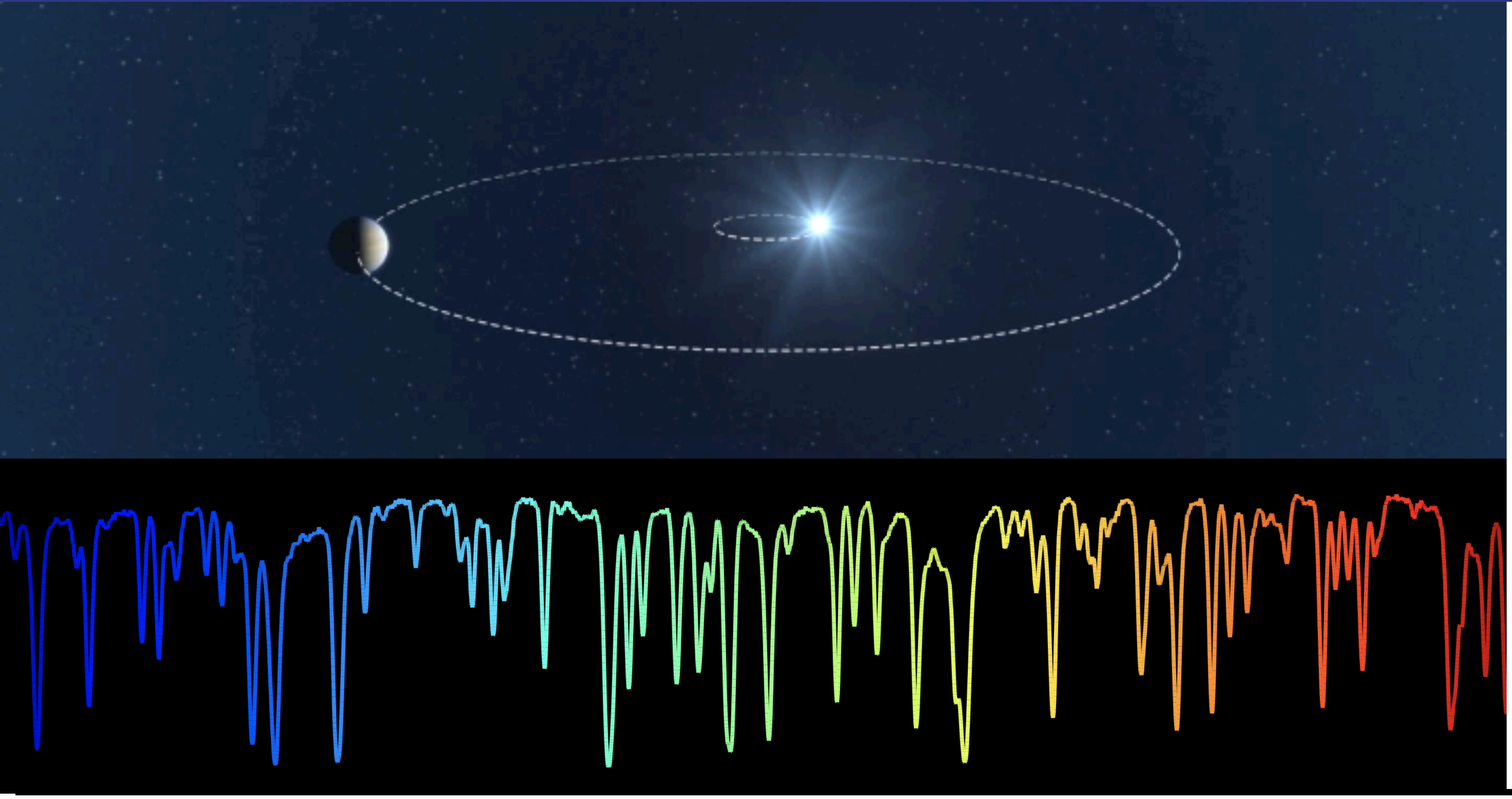


Transits would help very marginally
(2) **Radial velocity** or **astrometry** could do it

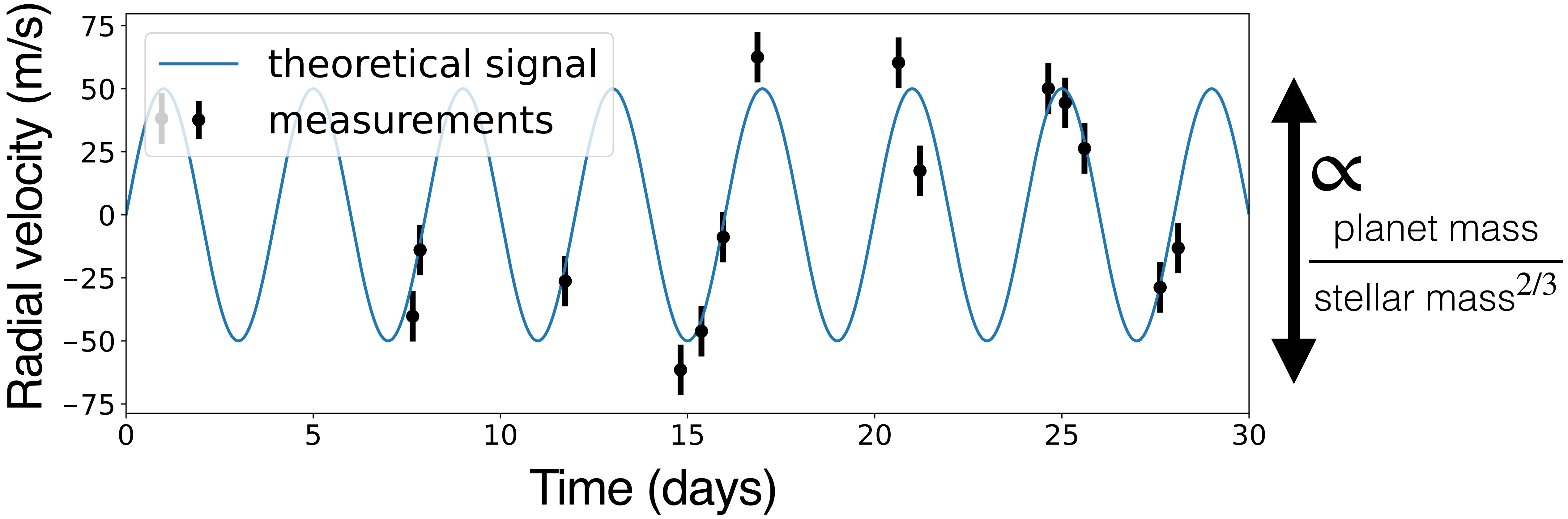
(3) Earth-analog detection with radial velocity is hindered by stellar variability and systematics: mostly not a hardware problem, needs **new data analysis methods** and appropriate **observation strategy**
Detection with astrometry would require a new space mission (concept in study)

(4) A community effort is needed to make precursor surveys happen

The Radial Velocity (RV) technique



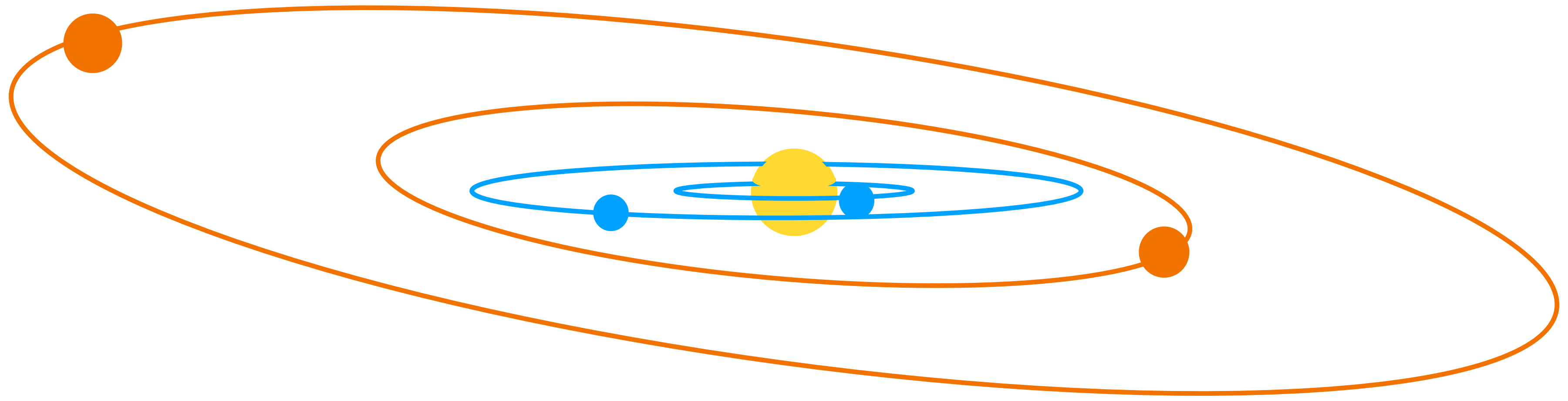
Radial velocity (RV) = velocity of the star projected onto the line of sight



$\propto \frac{\text{planet mass}}{\text{stellar mass}^{2/3}}$

Radial velocities (RVs) are essential

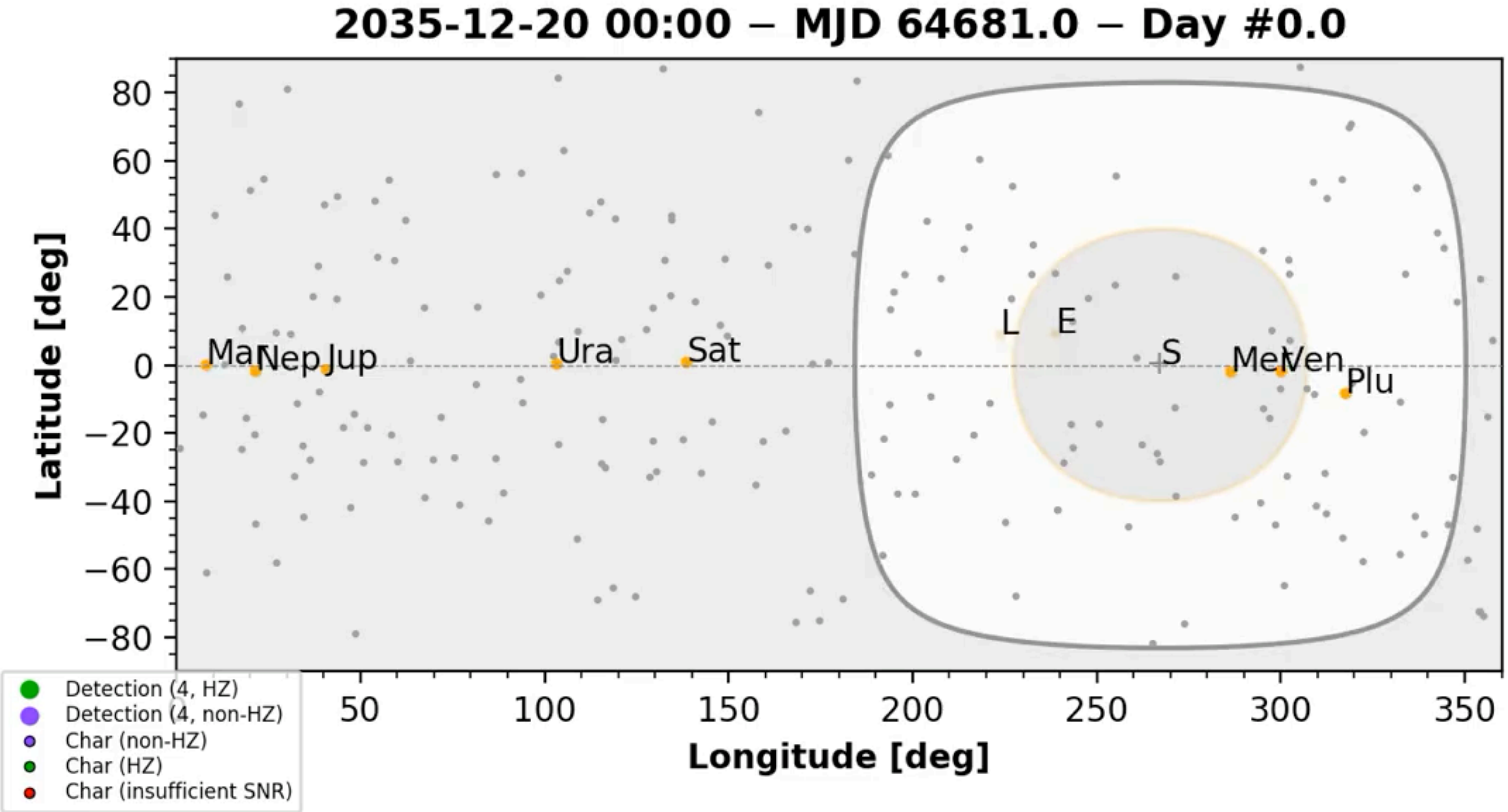
To probe outer regions of planetary systems with low transit probability



To measure exoplanetary masses: their most fundamental parameter

Impact of a precursor survey

Simulating the yield of *HWO*



Morgan et al. 2021, Savransky et al. 2020

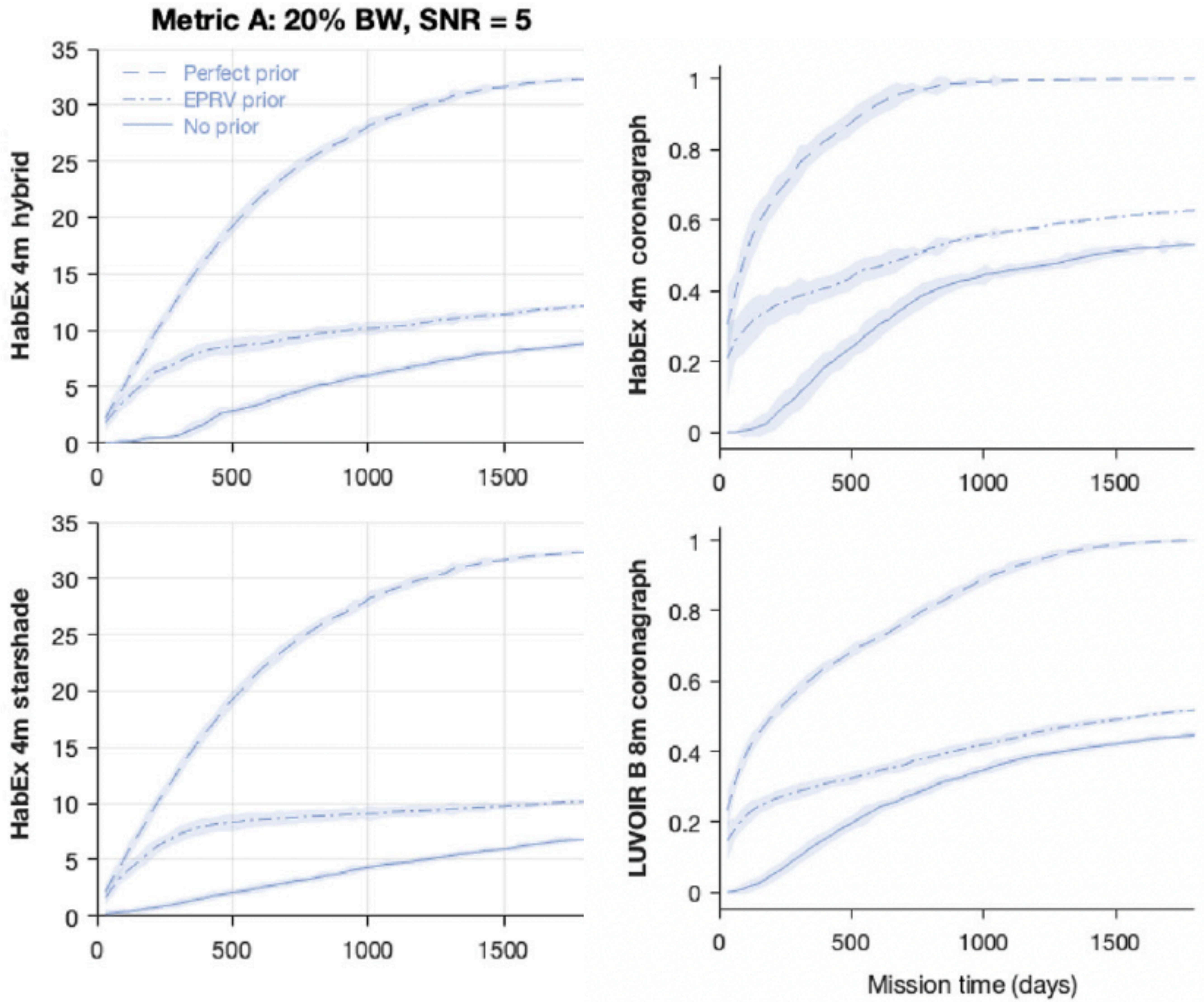
Impact of a radial velocity precursor survey (Morgan et al. 2021)

HWO yield is greatly improved if targets are detected in advance

Universe + mission simulated with EXOSIMS (Delacroix et al. 2015, Savransky et al. 2020)

With a RV survey

- Results obtained faster, more time for characterisation
- Mitigates the risk of the missions



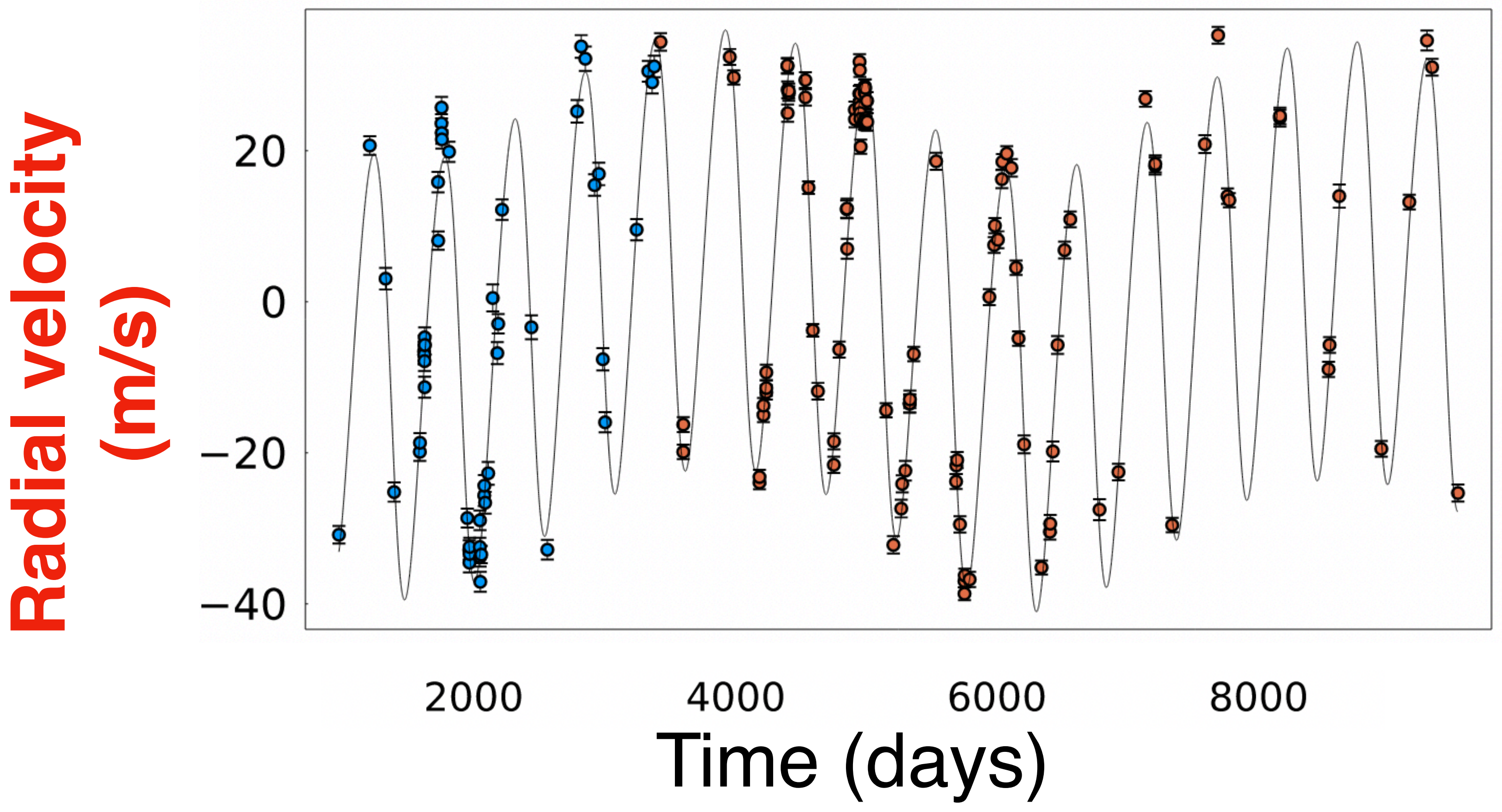
(3 cm/s) RV instrument on a 10-m class telescope surveying ~53 HabEx targets in a 5 year, 25% time survey

RV precursor survey is feasible in principle

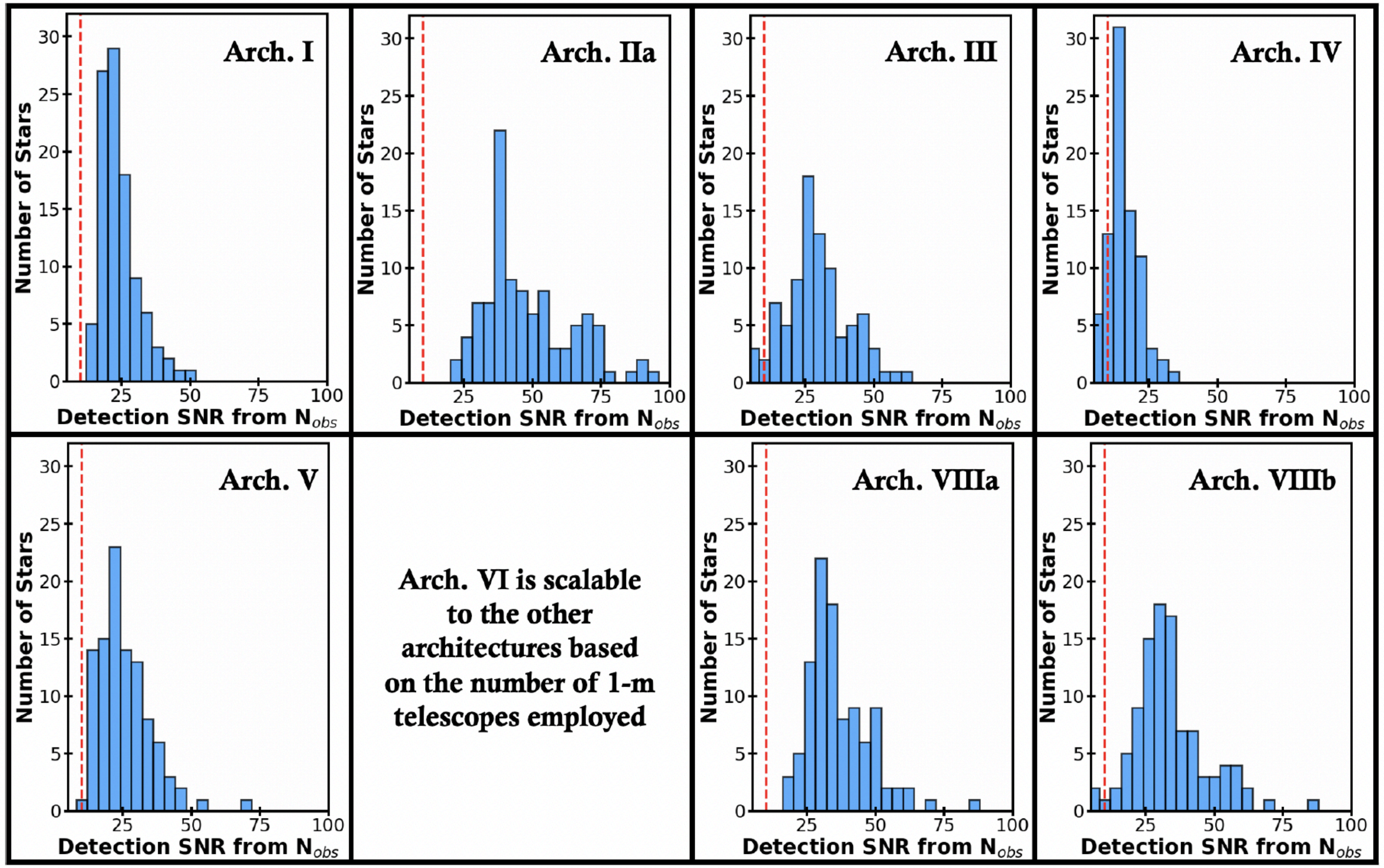
Simulation with different telescope/instrument associations

In principle the 100 prioritised targets can all be characterised

Simulation done in the best case scenario: White noise



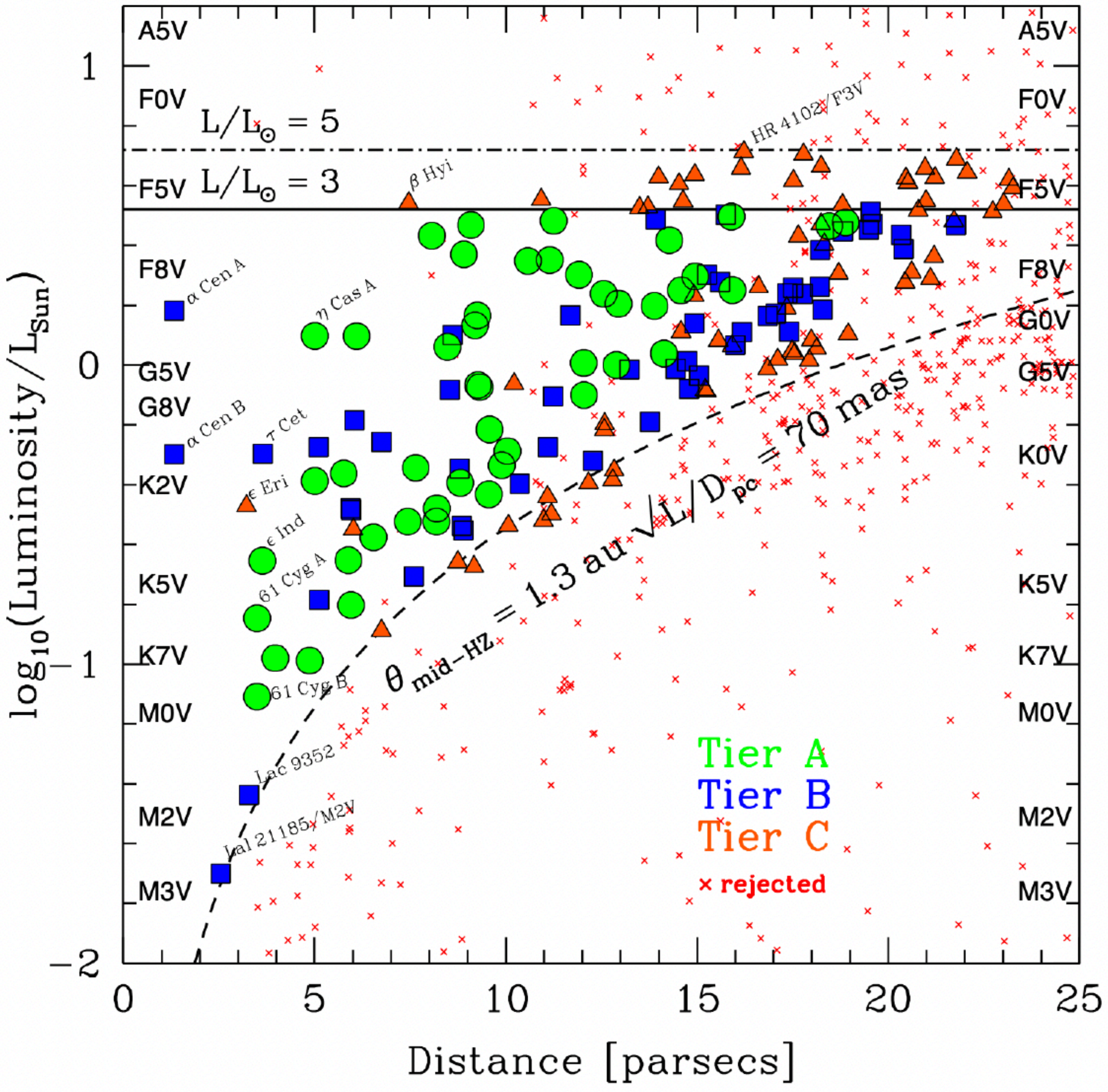
Histogram of obtained signal to noise ratio for 100 priority stars assuming they all have an Earth



Crass et al. 2021

HWO prioritised targets (Mamajek et al. 2024)

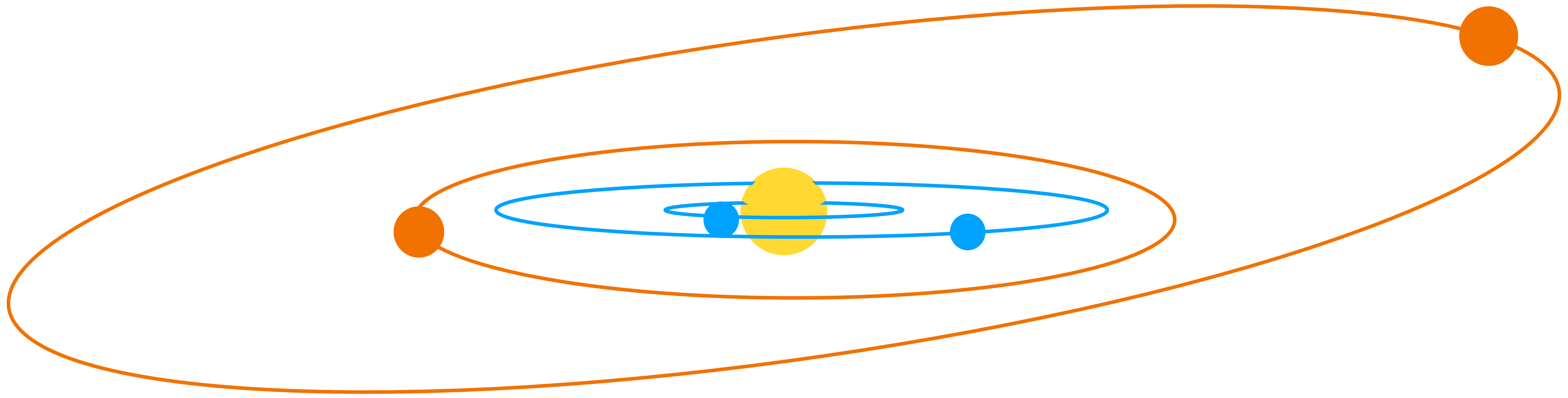
Parameter	**Tier A**	**Tier B**	**Tier C**
IWA constraint	83 mas	72 mas	65 mas
Exoplanet brightness limit (Rc)	30.5 mag	31.0 mag	31.0 mag
Exoplanet-star Brightness ratio limit	4e-11	4e-11	2.5e-11
Disk criterion	No known dust disks of any kind	No disk, or KB disks OK if $L_{disk}/L_* <= 10^{-4}$	All disks OK, even if $L_{disk}/L_* >= 10^{-4}$ or detected HZ warm dust disk
Treatment of binaries	Single or binary companion > 10″ sep	Single or binary companion 5″ - 10″ sep	Single or binary companion 3″ - 5″ sep
Number of Stars	**47**	**51**	**66**



Sample	**F**	**G**	**K**	**M**
Tier A	14	15	17	1
Tier B	15	23	11	2
Tier C	37	17	12	0
Total (A+B+C)	66	55	40	3

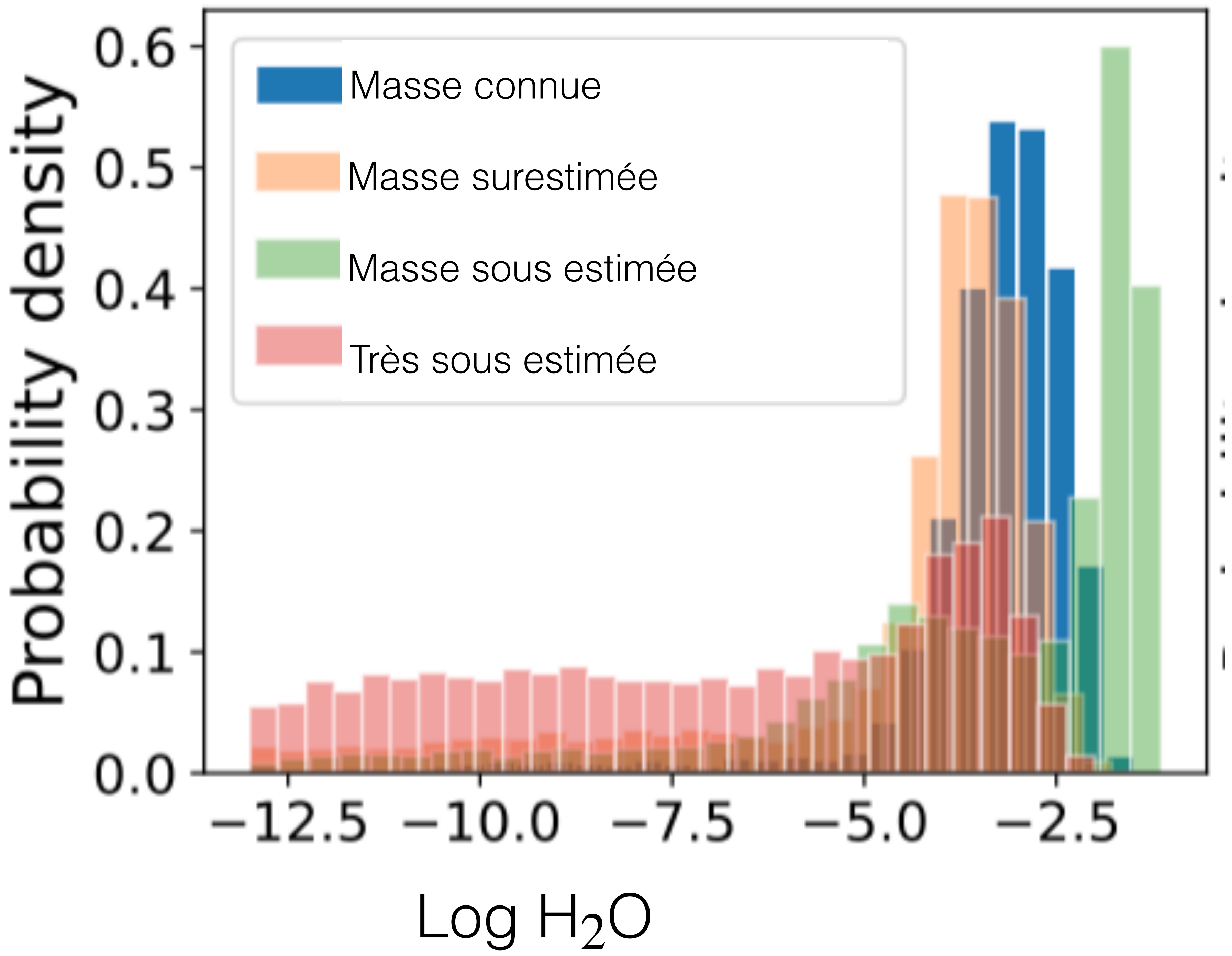
Further arguments

For a RV survey of nearby stars



- Even if Earth-like planets are not detected, information on the system architecture
- Presence more or less likely of certain

For extreme precision RVs



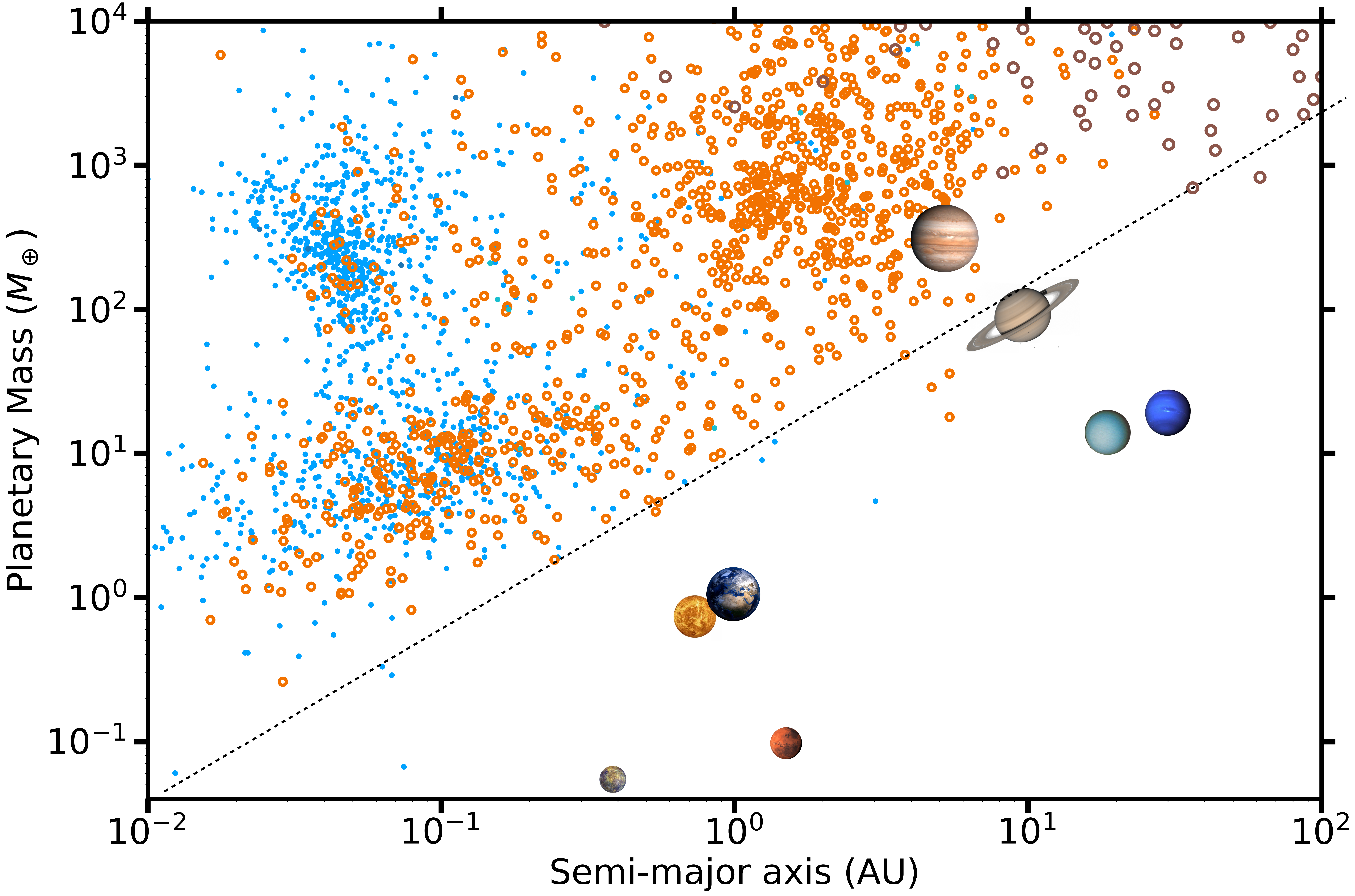
- Direct measurements of the mass is critical (Batalha+ 2019, Kempton+ 2018)
- Number of measurements needed proportional to individual RV uncertainty squared $N \propto \sigma_{\text{RV}}^2$

NASA Extreme precision radial velocity report (Crass+ 2021)

« There exist multiple plausible system architectures in terms of telescope size, longitude and latitude distribution, and dedication that could, **if stellar variability mitigation, telluric mitigation, and instrumental precision goals are met**, successfully acquire a set of measurements with the statistical precision required to detect Earth analogs. »

Difficulties of RV data analysis

Known exoplanets



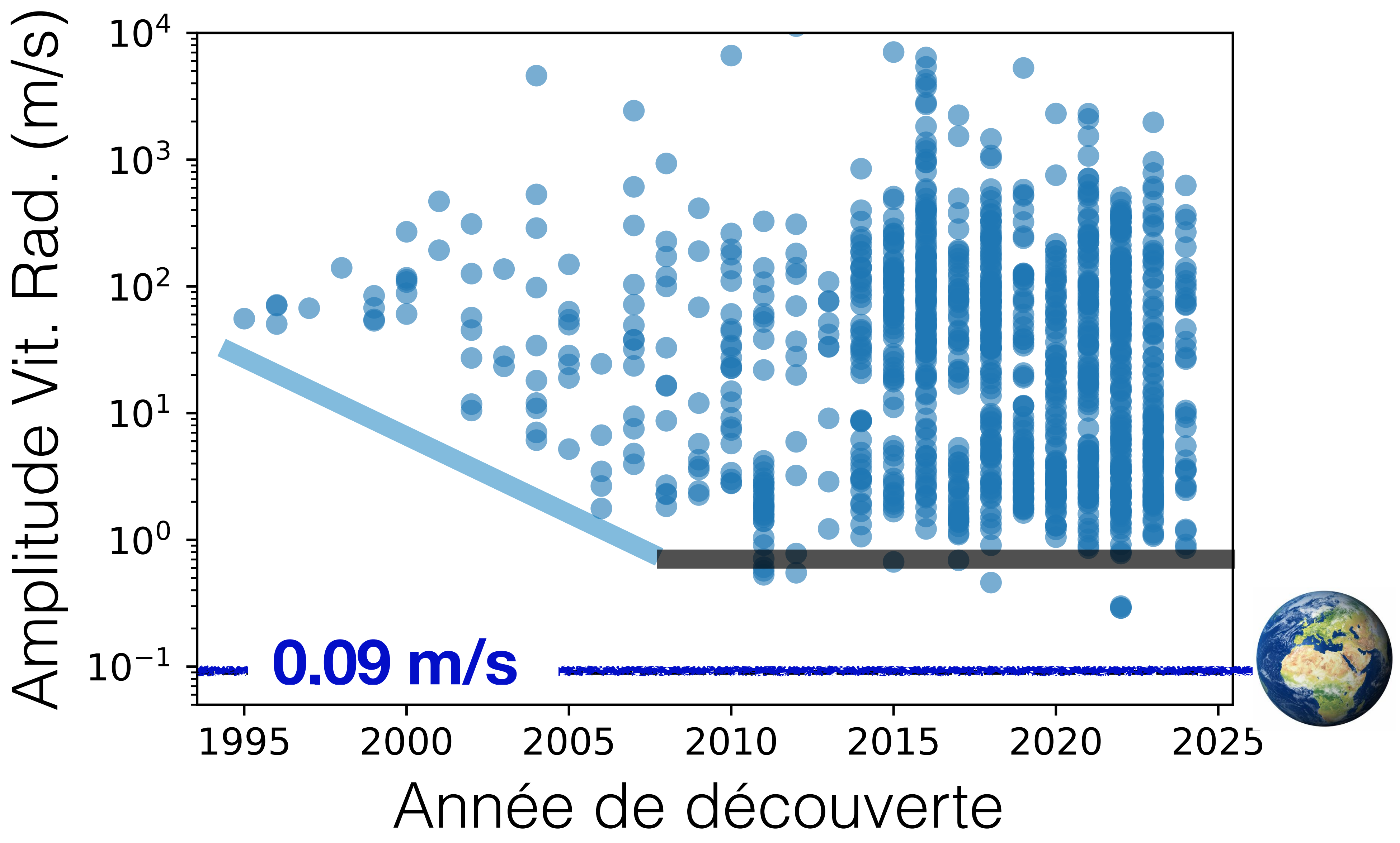
Transits

Radial velocity

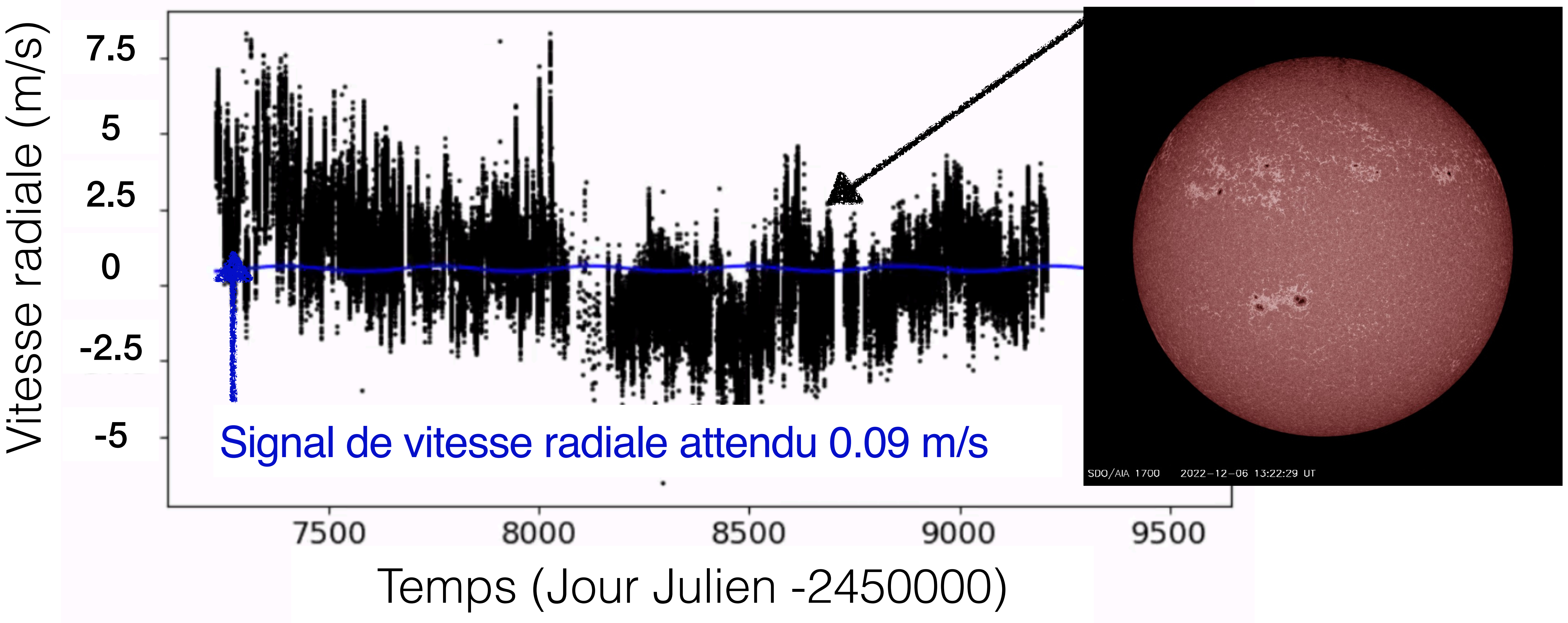
Imaging

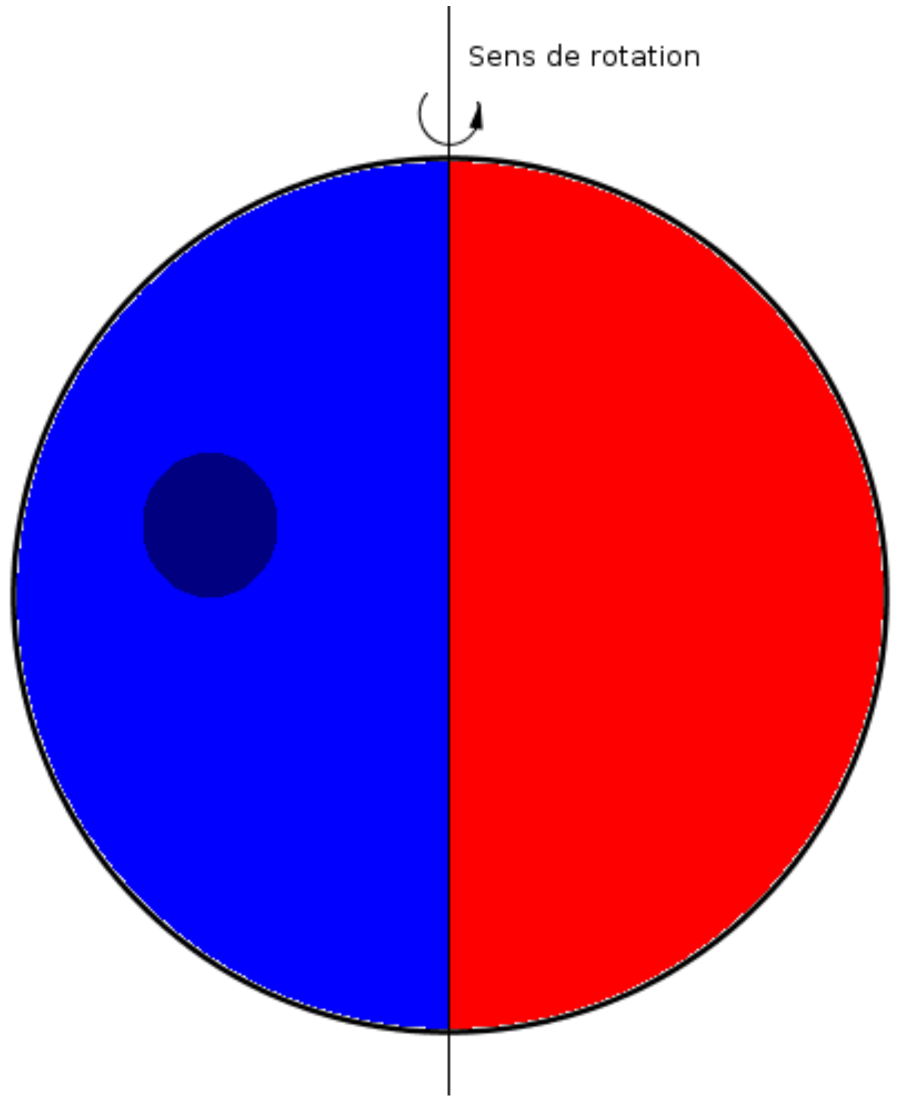
Source: exoplanet.eu
April 2024

The stellar variability problem



Vitesse radiale mesurée du Soleil par HARPS-N





No exo-Earth detection possible with radial velocities so far

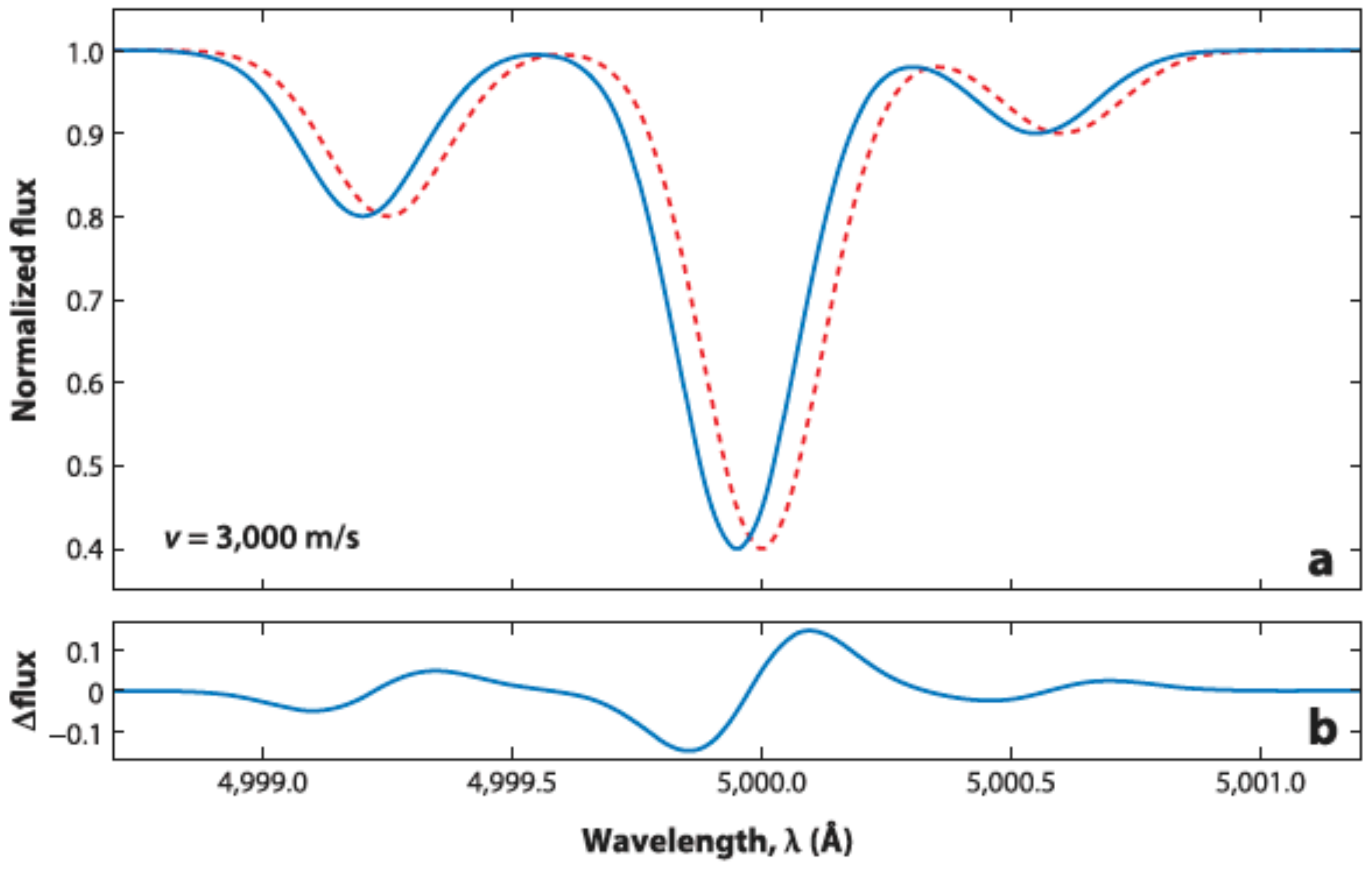
Calibration is still to be improved but instrument precision is not the main limitation

Only solution: new data analysis methods

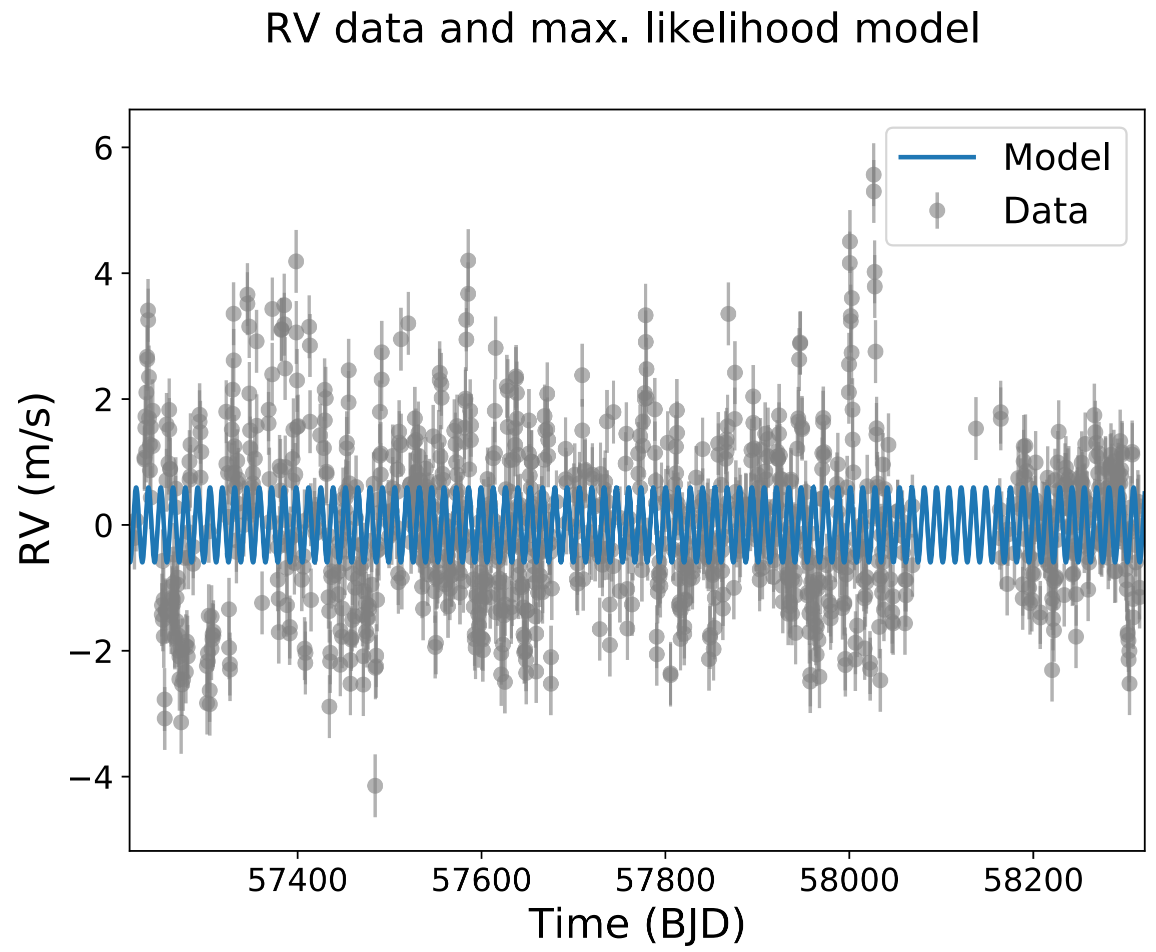
(>250 ref. Hara & Ford 2023, *Annual Reviews*)

Key ideas to correct stellar variability

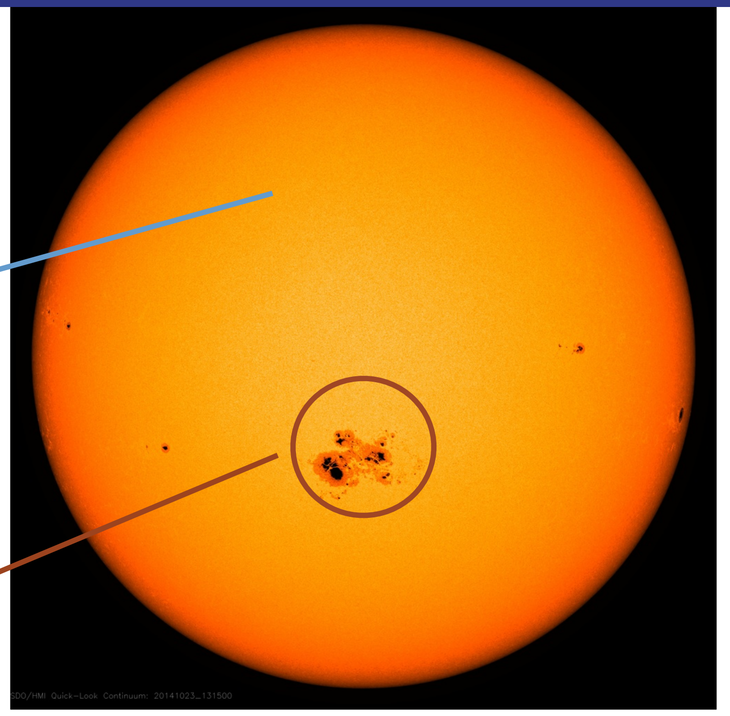
Planets induce a pure Doppler shift

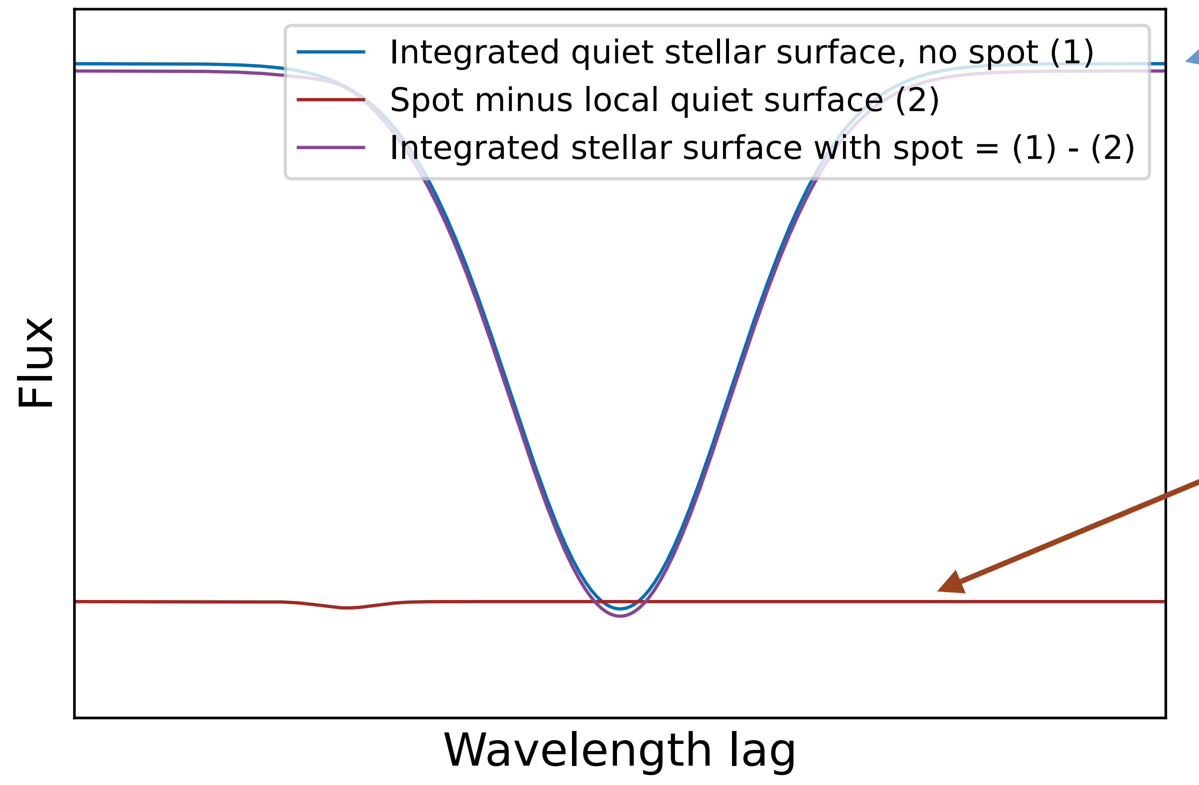


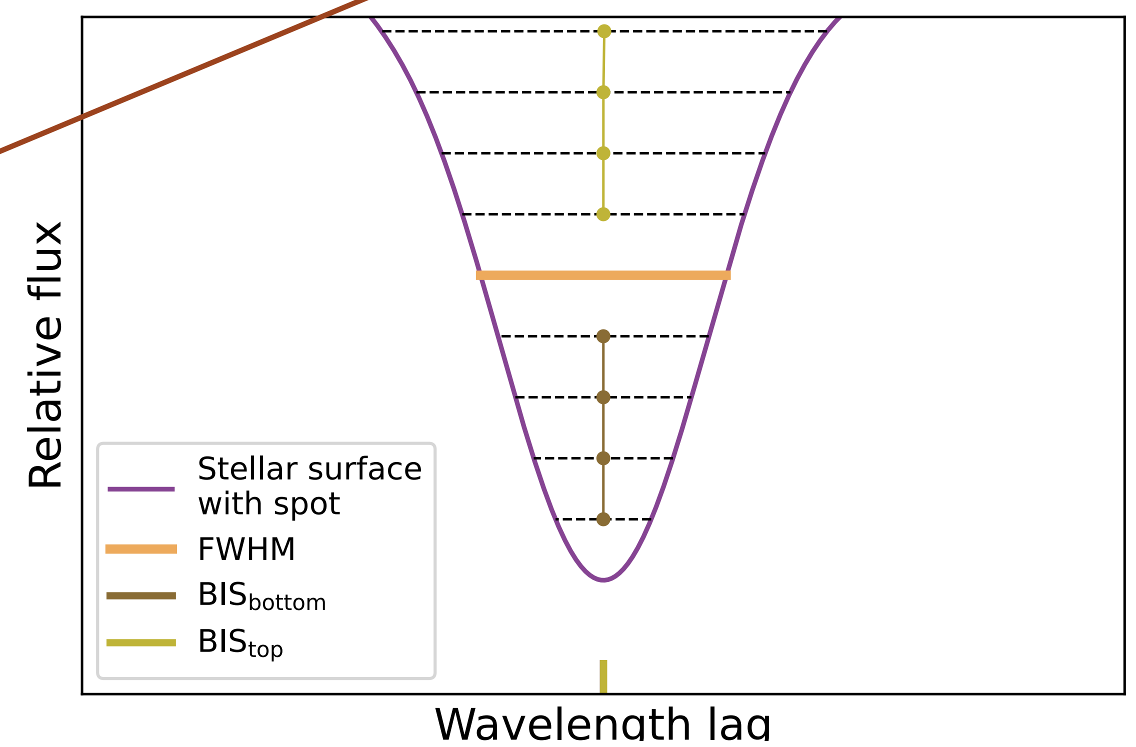
Planets induce a periodic signal



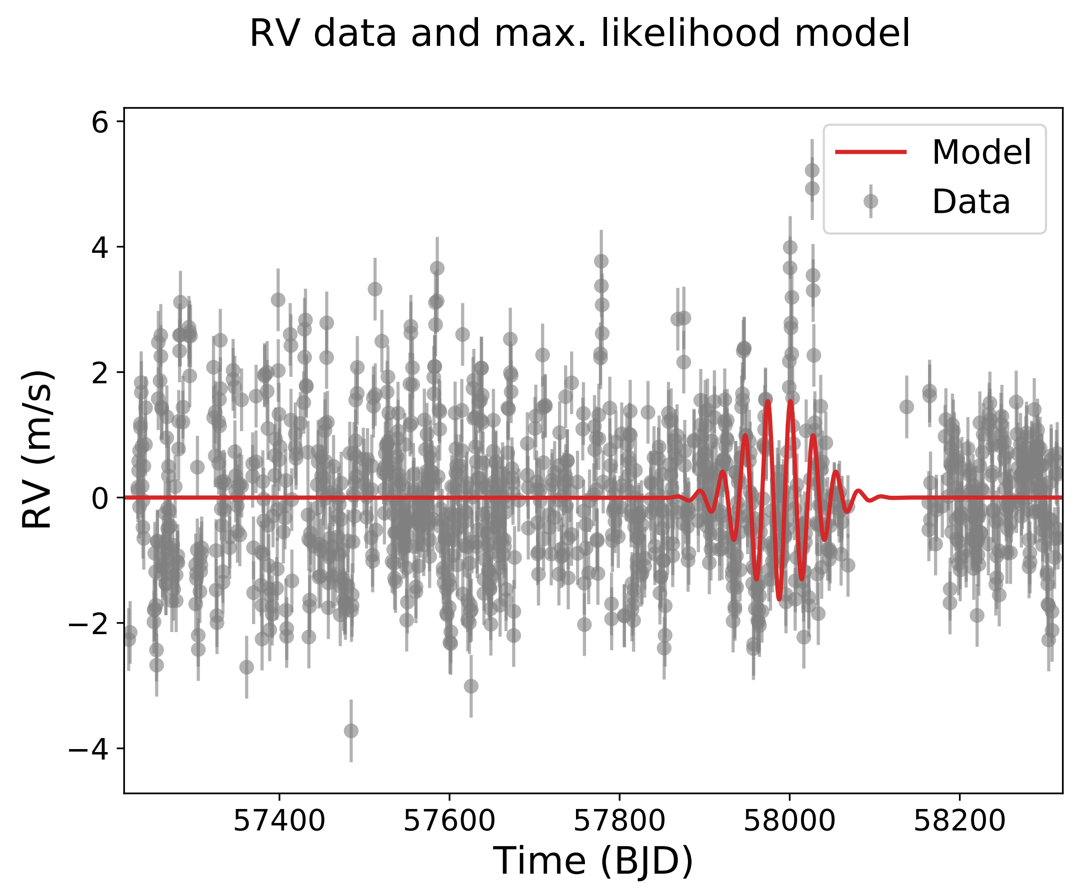
stellar and instrumental effects change the shape of the spectrum







Stellar and instrumental effects are (usually) not strictly periodic



The extreme precision RV problem in a nutshell

$$RV_{\text{measured}} = RV_{\text{center of mass}} + RV_{\text{contam}}$$

We want

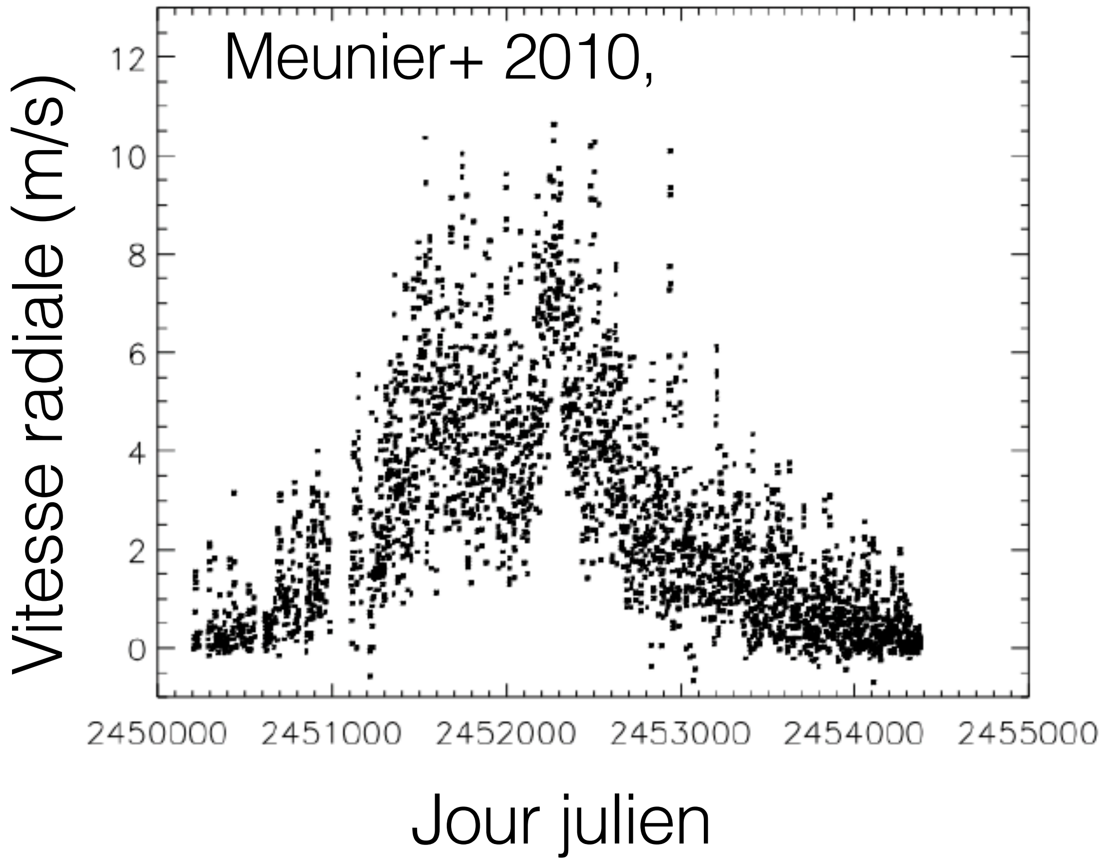
1) An RV whose definition is as close as possible as a global Doppler shift
2) To leverage the shape of the spectra to estimate the contamination
3) To estimate the uncertainties on RV_{contam}

Can be seen as a statistical or machine learning problem

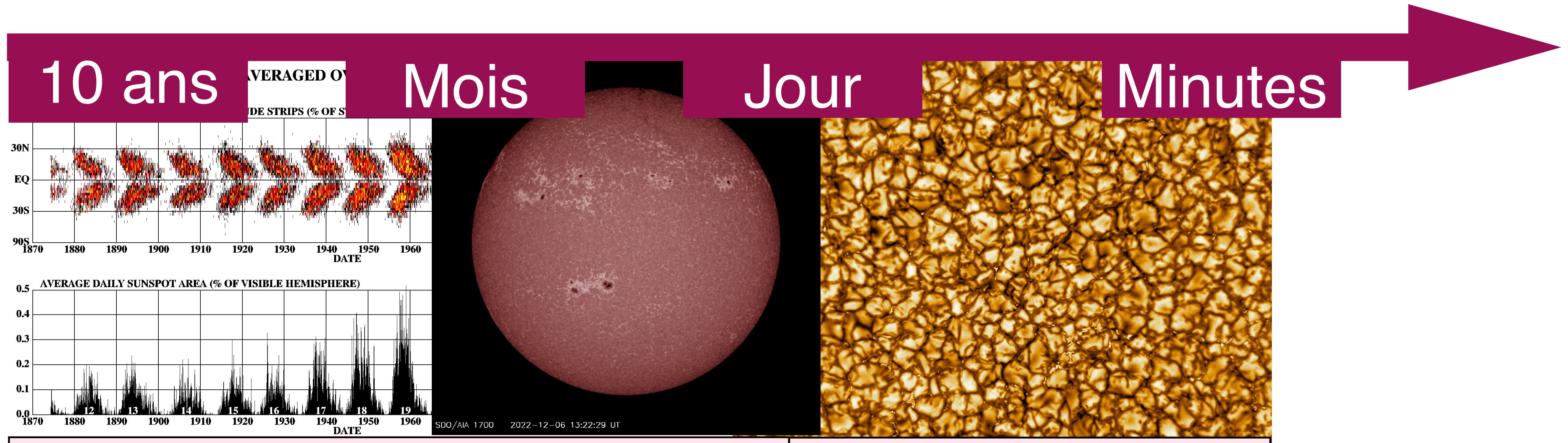
Understanding stellar variability and modelling it in practice

Detailed physical models
Analysis of solar data

Likelihood supposed Gaussian and stationary with qualitative parameters



Numerous processes at
Different timescales

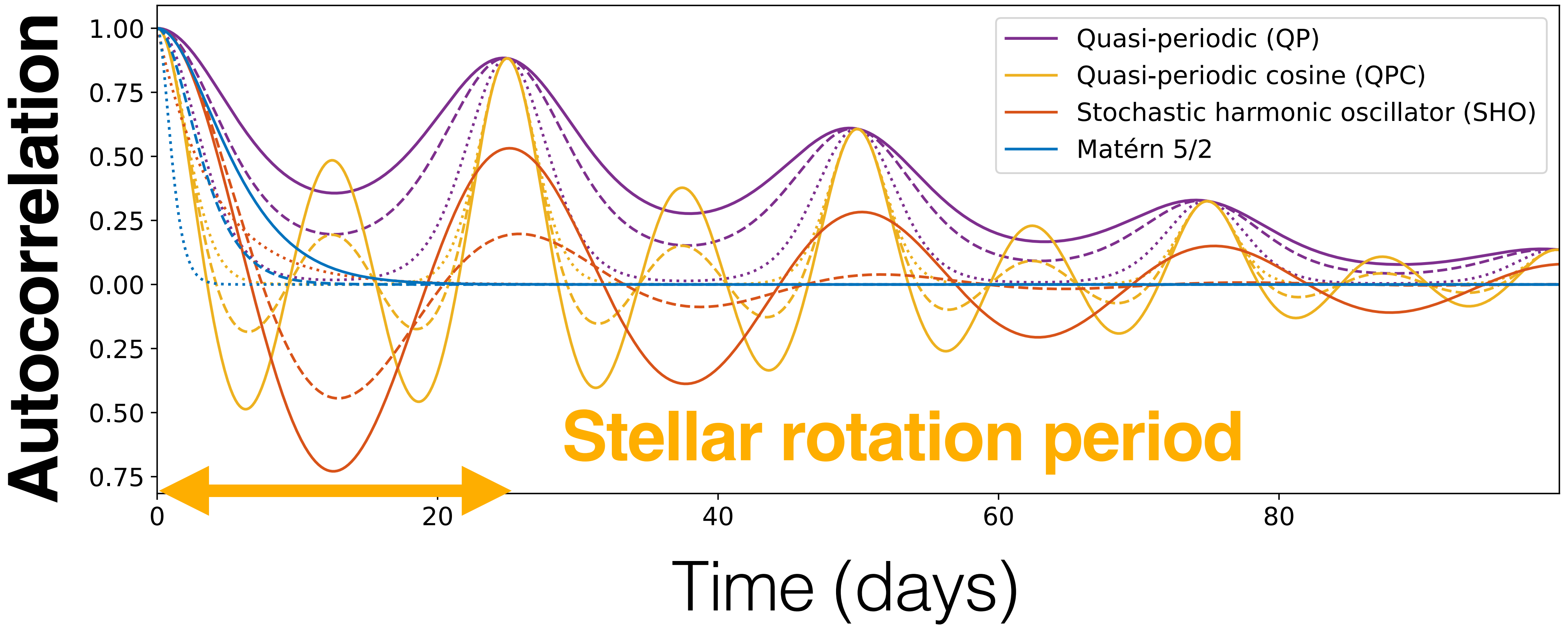


Magnetic activity
p.ex. Meunier+ 2010, 2012, 2019
Boisse+ 2012, Dumusque+ 2014,
Haywood+ 2016, Al Moulla 2023

Granulation and super granulation
Cegla+2013, 2019
Dravins+ 2021,

+ oscillations, winds, gravitational redshift…

$$p\left(\text{data}\mid\theta,\eta\right)=\frac{1}{\sqrt{2\pi}^{N}\sqrt{|V(\eta)|}}\,e^{-\frac{1}{2}\left(\text{data}-f(\theta,\eta)\right)^{T}V(\eta)^{-1}\left(\text{data}-f(\theta,\eta)\right)}$$



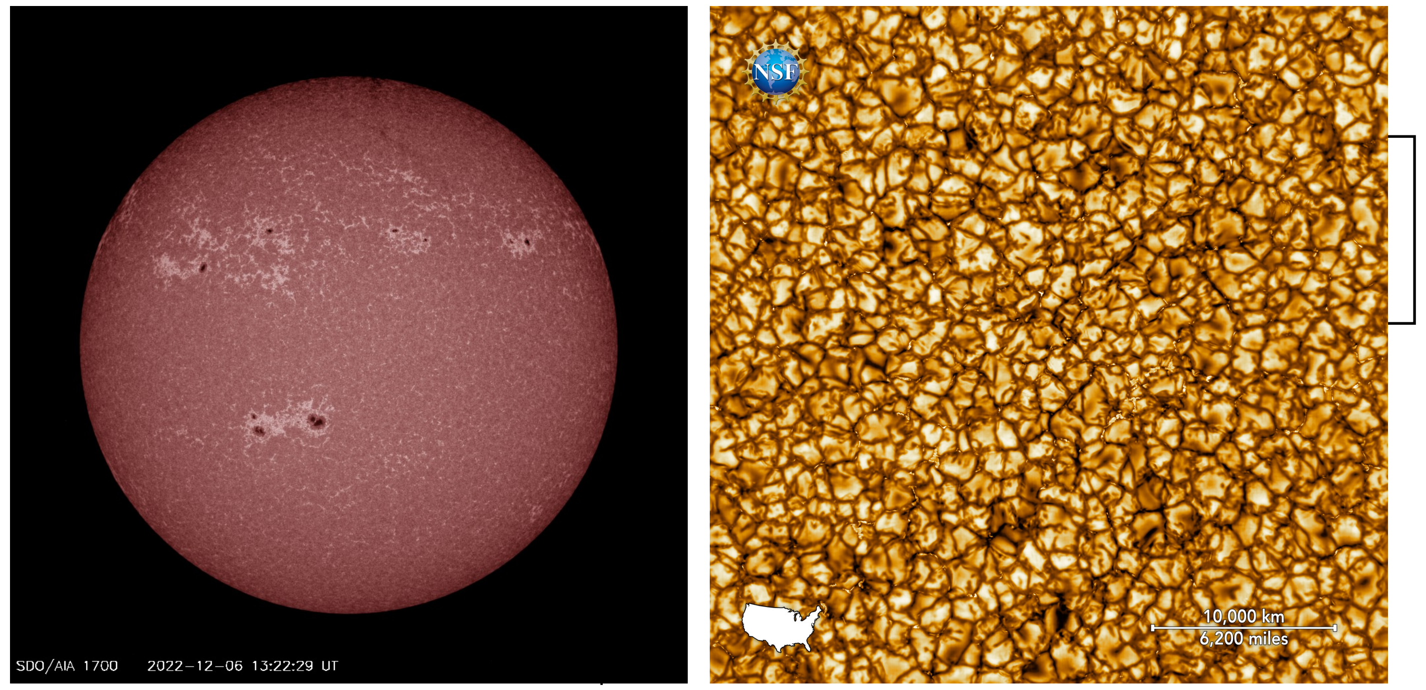
Aigrain et al. 2012, Haywood et al. 2014, Foreman Mackey et al. 2017, Rajpaul et al. 2015, Perger et al. 2021. Jones et al. 2022…

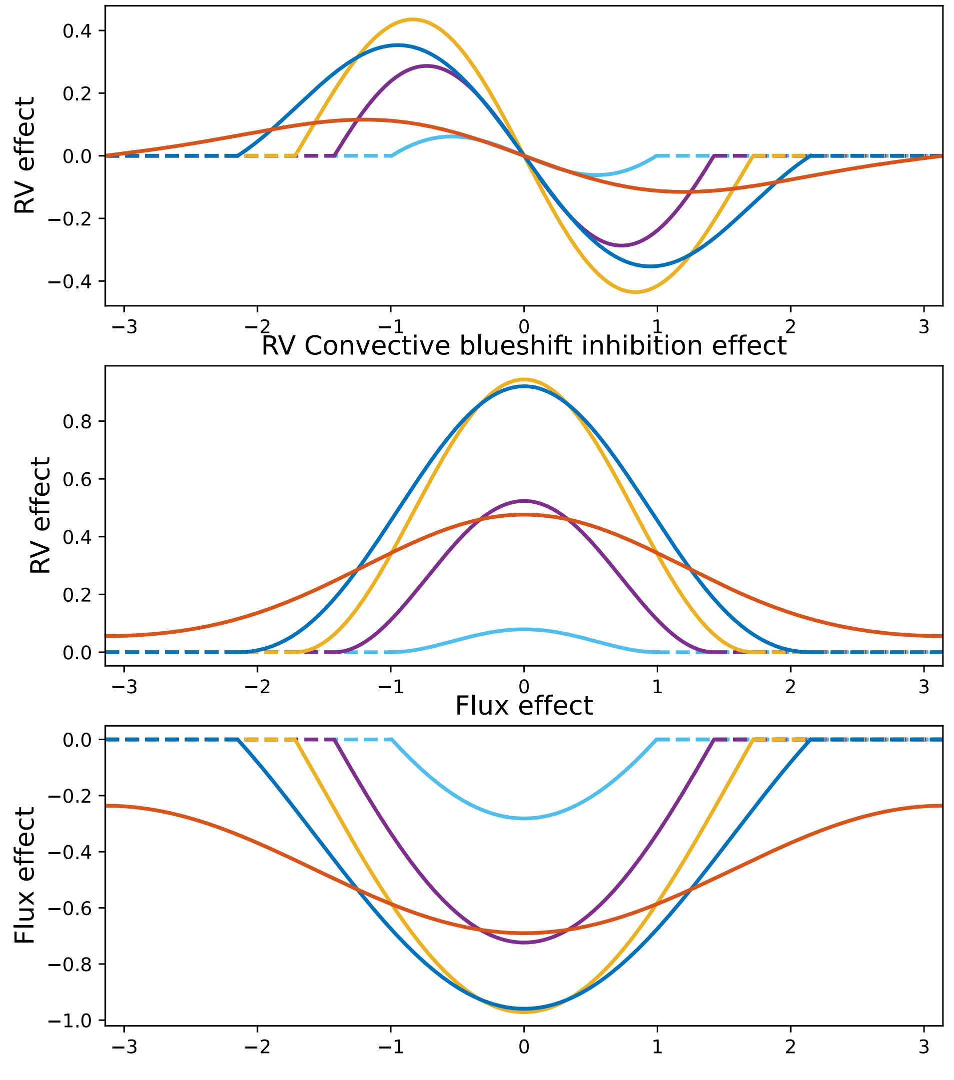
Statistical models are not quantitatively related to the physical model

From the physical model, we compute the covariance

$$\iint \boxed{g(t-t_1,\gamma)}\boxed{h(t-t_1',\gamma)}\boxed{\lambda(t)}\boxed{p(\gamma \mid t,\eta)}\,\mathrm{d}t\,\mathrm{d}\gamma = \kappa_\eta(P(t_1), RV(t_1'))$$

Hara & Delisle
2023

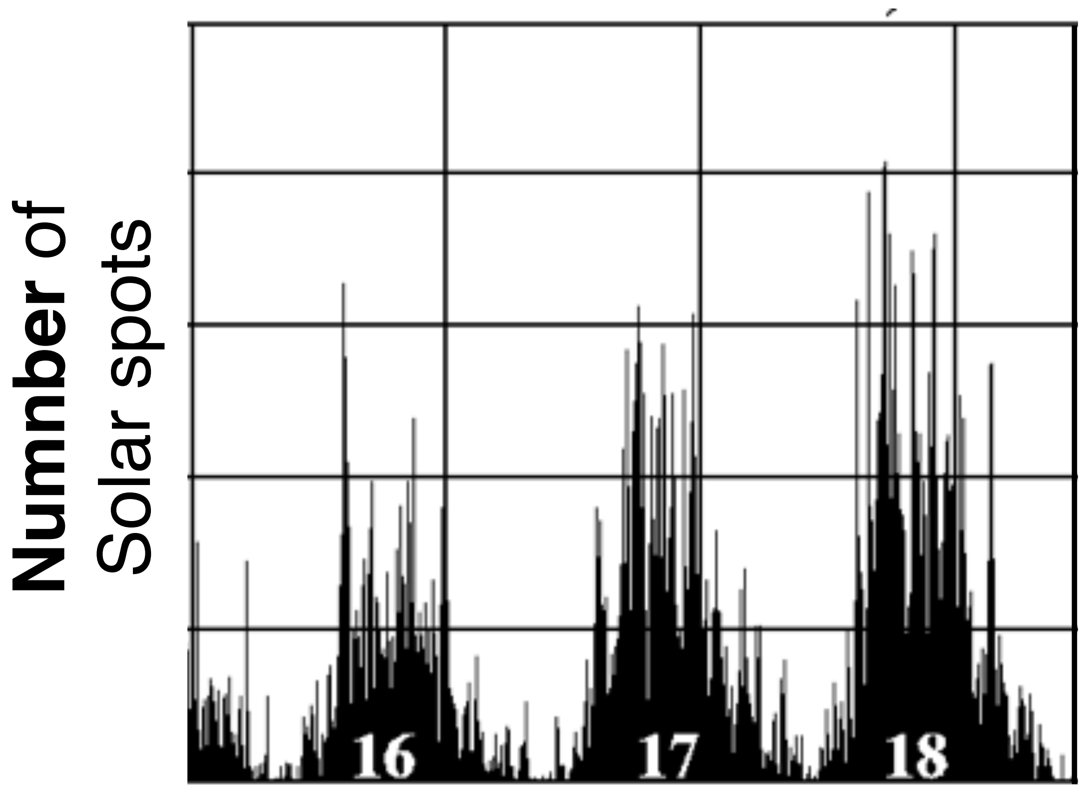




Time

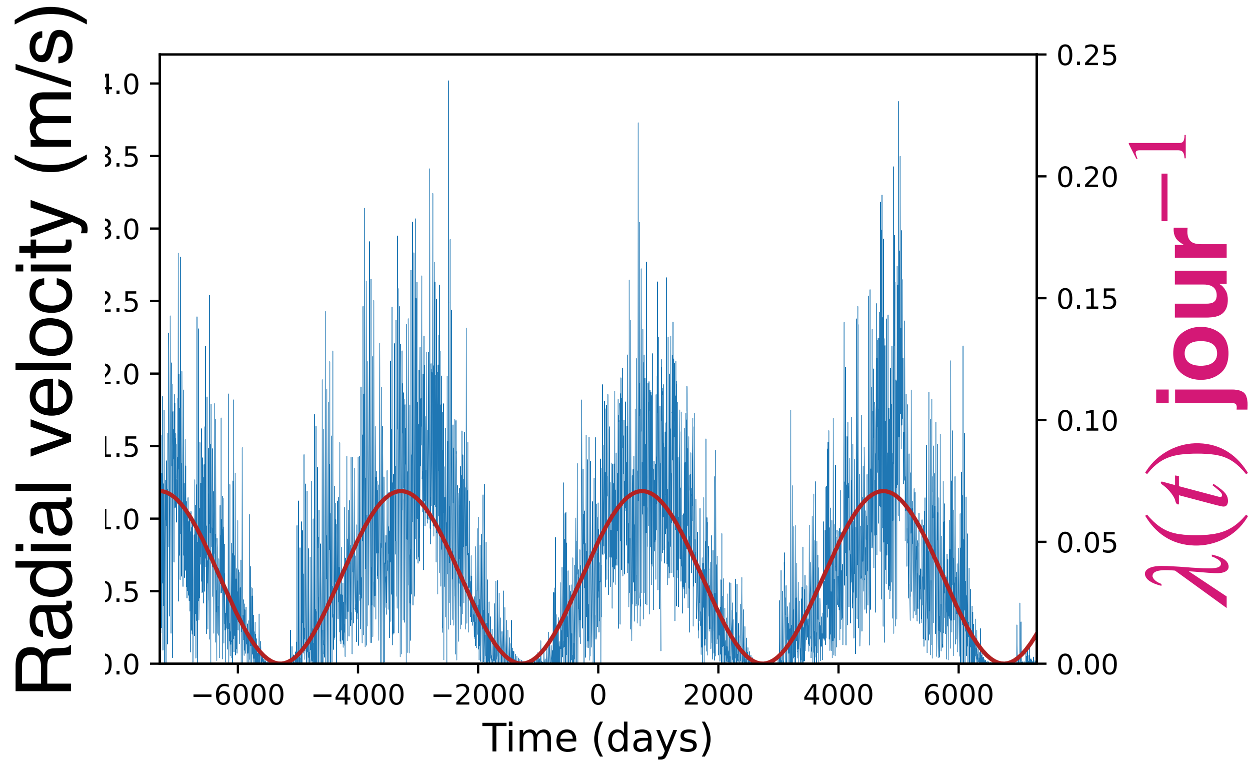
Effect on RV
$h(t-t_1,\gamma)$

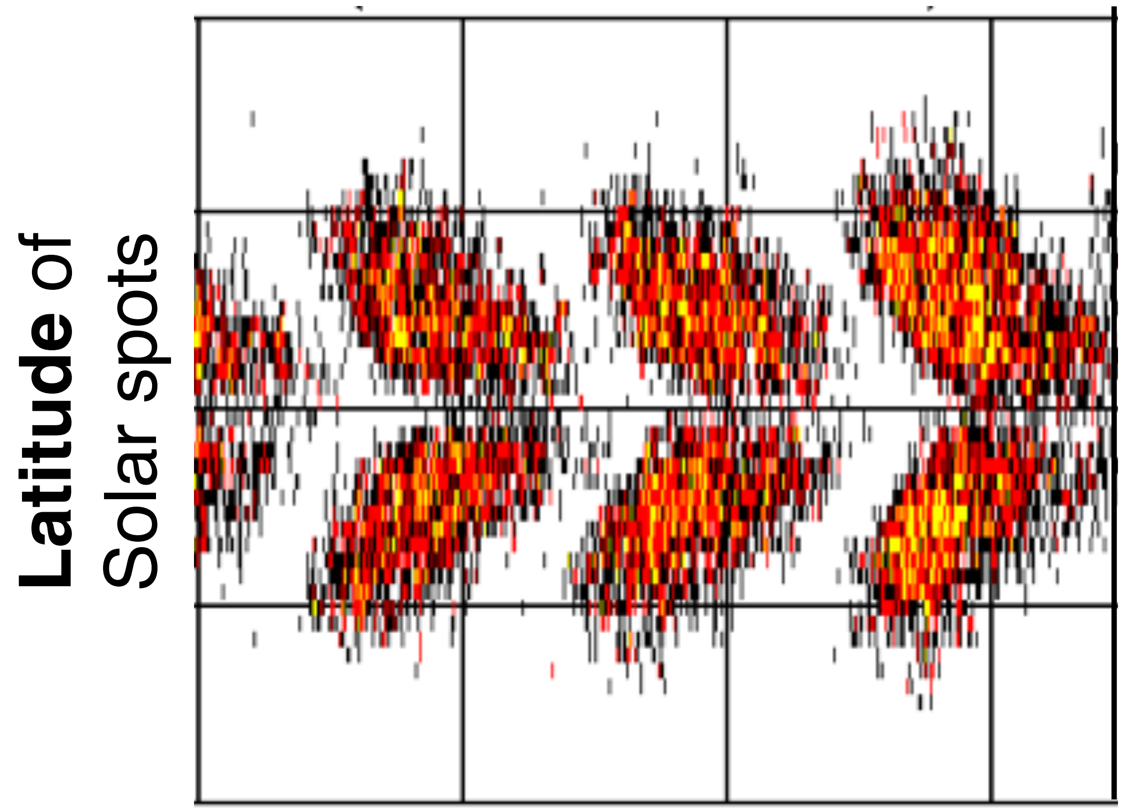
Effect on photometry
$g(t-t_1,\gamma)$,



Numnber of Solar spots

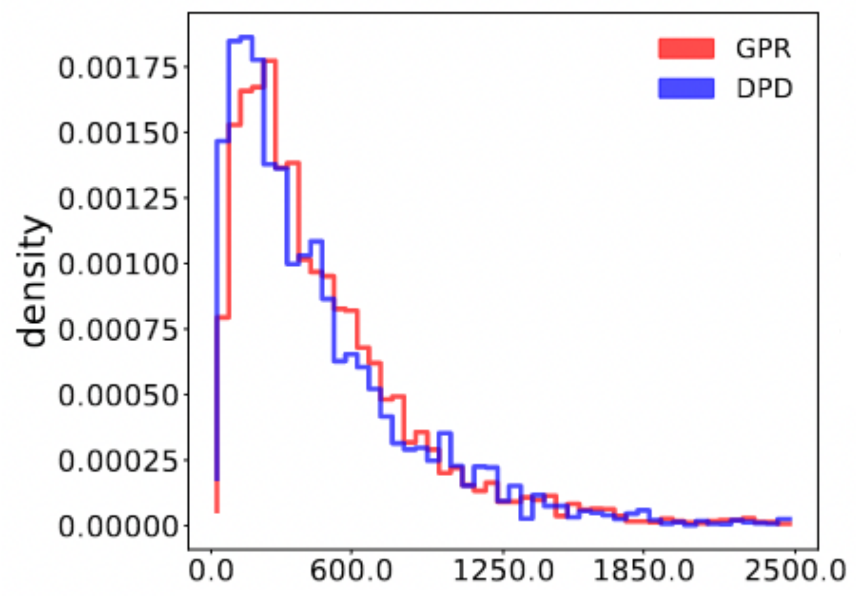
Time





Latitude of Solar spots

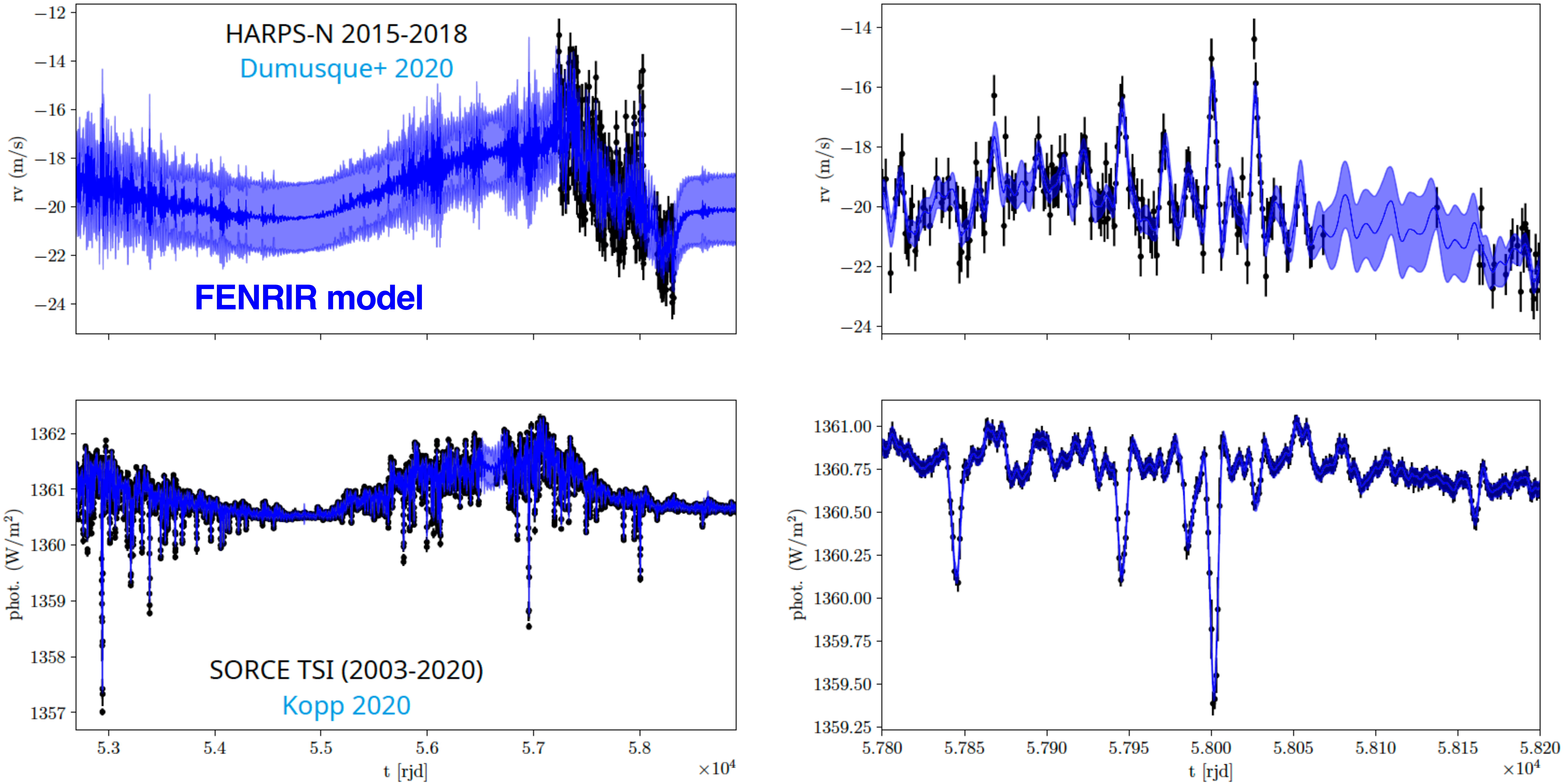
Time



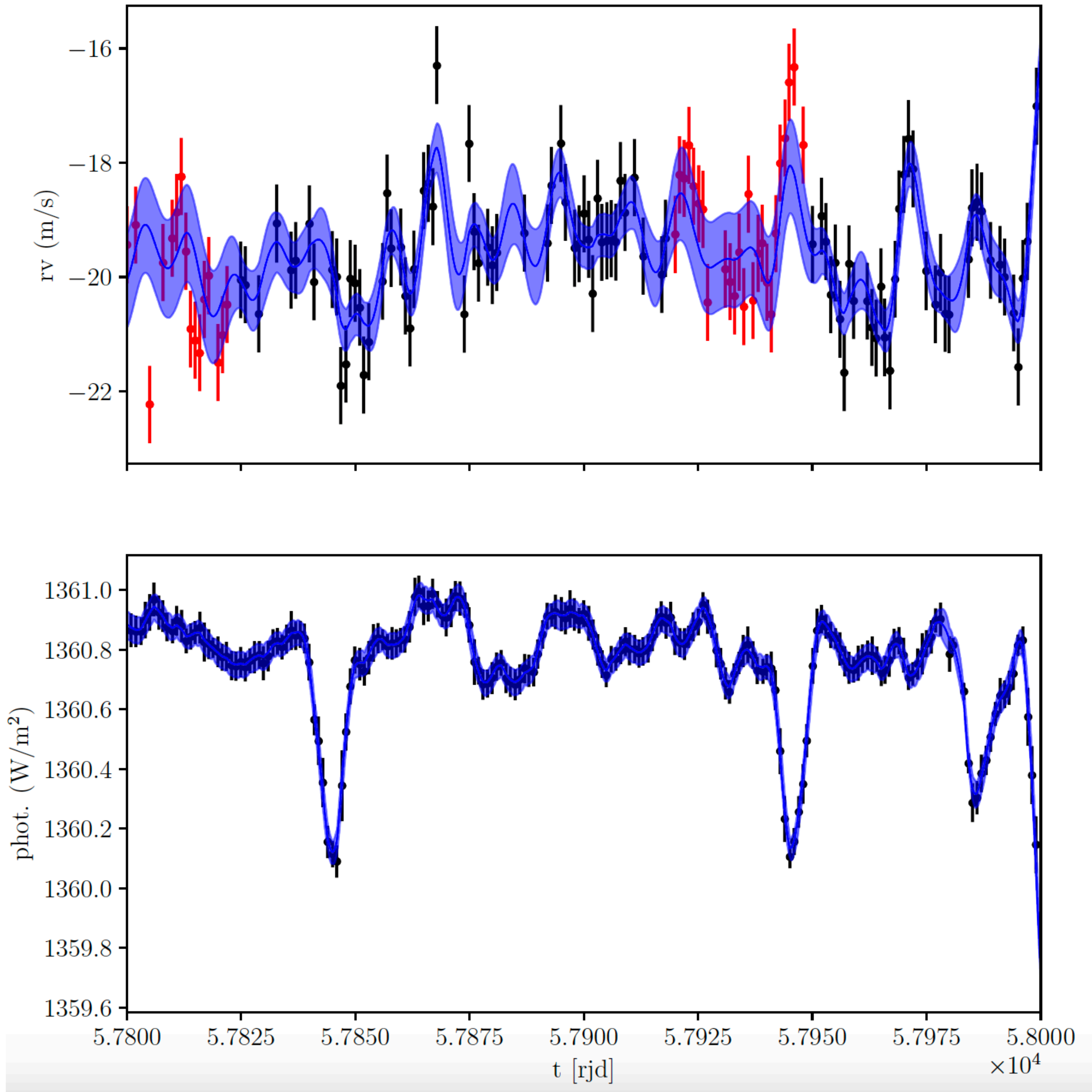
Area of spots
(μHemisphere)

$p(\gamma \mid \eta, t)$

Performance of the stellar activity model



Performance of the stellar activity model

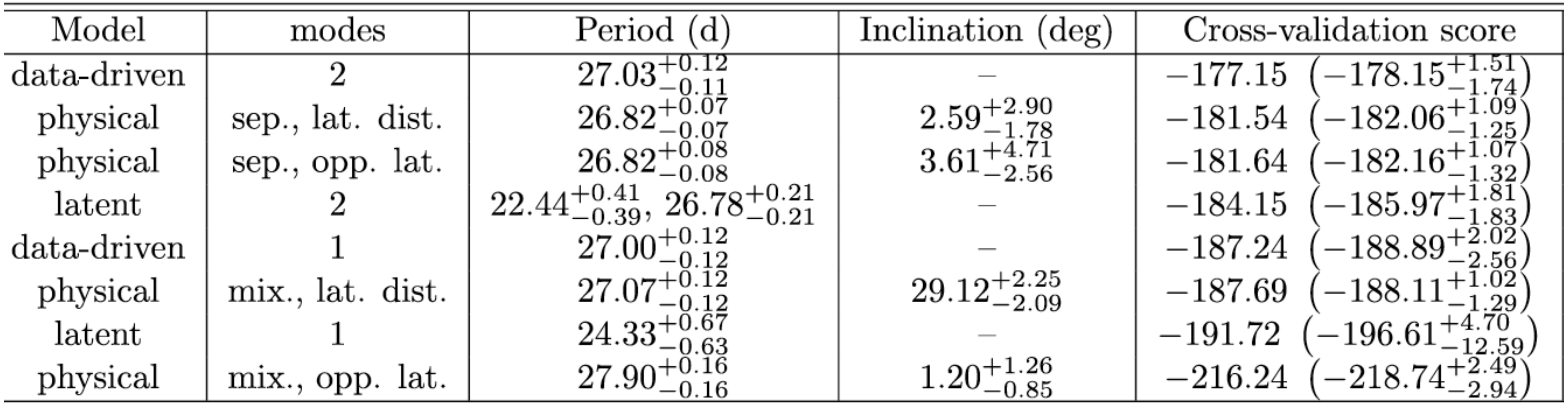


We remove a the **test set**

We fit the parameters to the remaining points (the **training set**)

We compare the prediction of the **model** with the **test set**

We compute the likelihood of the predictive distribution

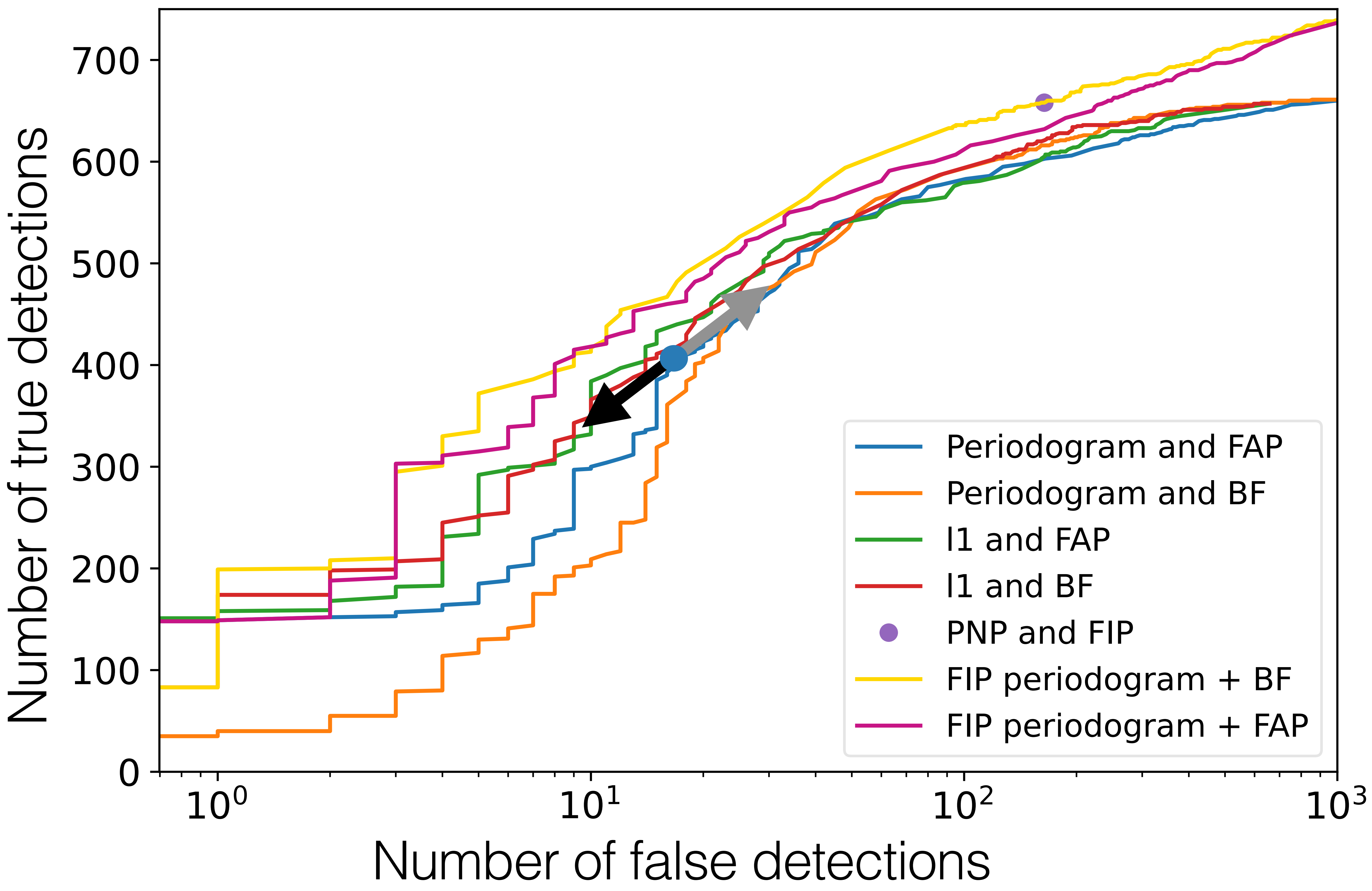


Model	modes	Period (d)	Inclination (deg)	Cross-validation score
data-driven	2	$27.03^{+0.12}_{-0.11}$	–	-177.15 $\left(-178.15^{+1.51}_{-1.74}\right)$
physical	sep., lat. dist.	$26.82^{+0.07}_{-0.07}$	$2.59^{+2.90}_{-1.78}$	-181.54 $\left(-182.06^{+1.09}_{-1.25}\right)$
physical	sep., opp. lat.	$26.82^{+0.08}_{-0.08}$	$3.61^{+4.71}_{-2.56}$	-181.64 $\left(-182.16^{+1.07}_{-1.32}\right)$
latent	2	$22.44^{+0.41}_{-0.39}$, $26.78^{+0.21}_{-0.21}$	–	-184.15 $\left(-185.97^{+1.81}_{-1.83}\right)$
data-driven	1	$27.00^{+0.12}_{-0.12}$	–	-187.24 $\left(-188.89^{+2.02}_{-2.56}\right)$
physical	mix., lat. dist.	$27.07^{+0.12}_{-0.12}$	$29.12^{+2.25}_{-2.09}$	-187.69 $\left(-188.11^{+1.02}_{-1.29}\right)$
latent	1	$24.33^{+0.67}_{-0.63}$	–	-191.72 $\left(-196.61^{+4.70}_{-12.59}\right)$
physical	mix., opp. lat.	$27.90^{+0.16}_{-0.16}$	$1.20^{+1.26}_{-0.85}$	-216.24 $\left(-218.74^{+2.49}_{-2.94}\right)$

Optimal criterion to detect exoplanets (1)

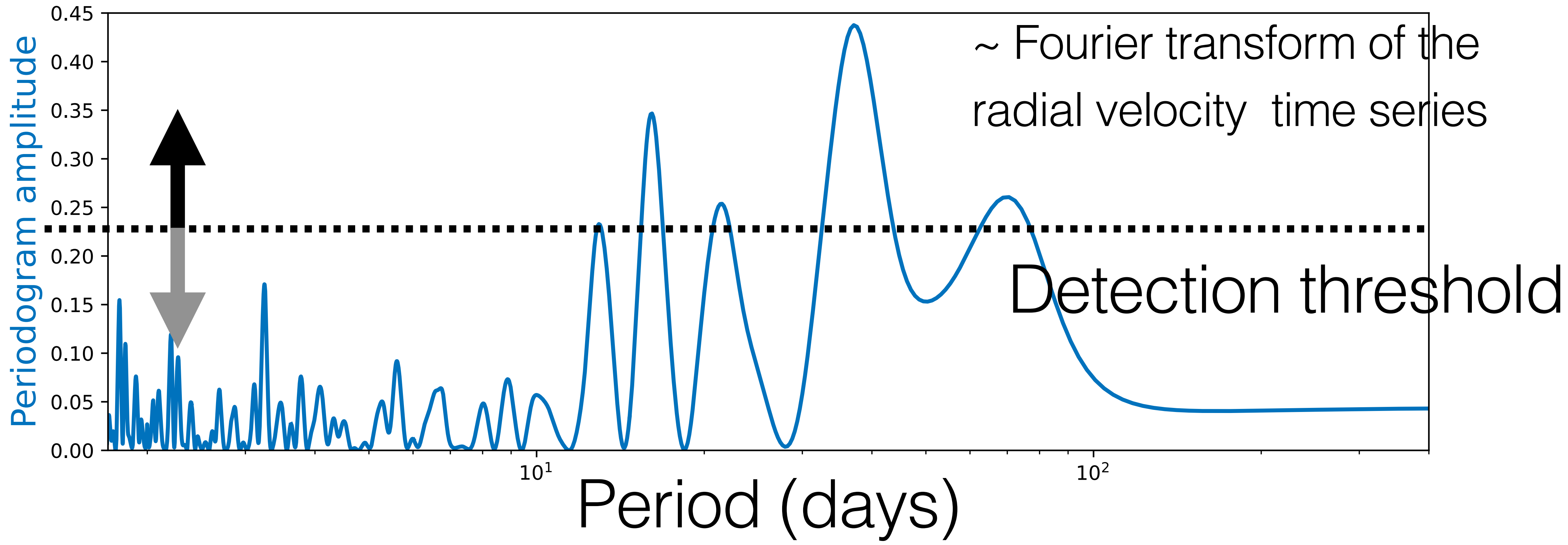
1000 radial velocity datasets with 0,1 or 2 planets

- Analysed with the same model
- With different detection criteria



Bayes factors and FAPs

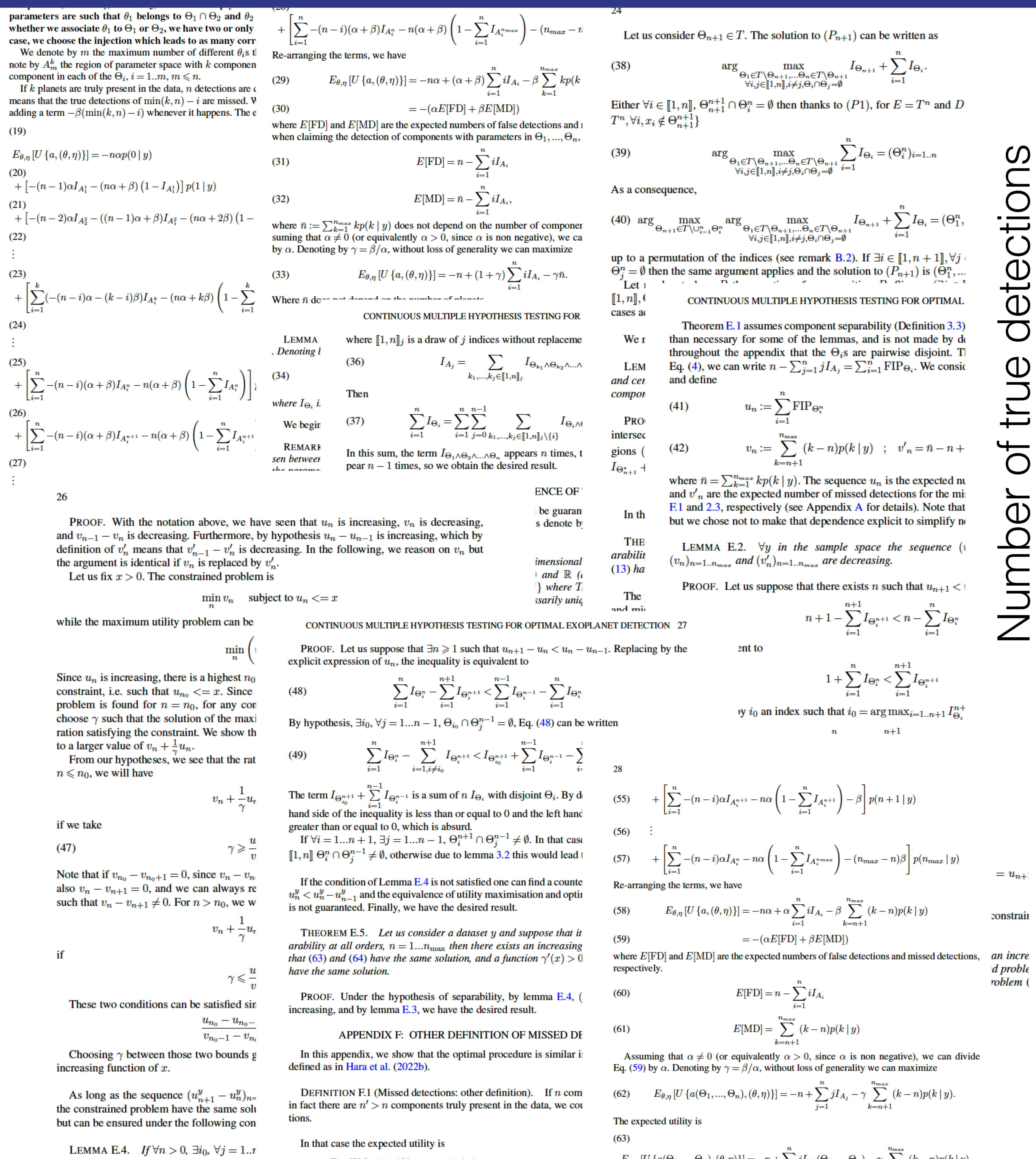
- **Optimal?**
  Which criterion maximises true detections?
- Do not encode where the planet is
- Are not defined on a very intuitive scale

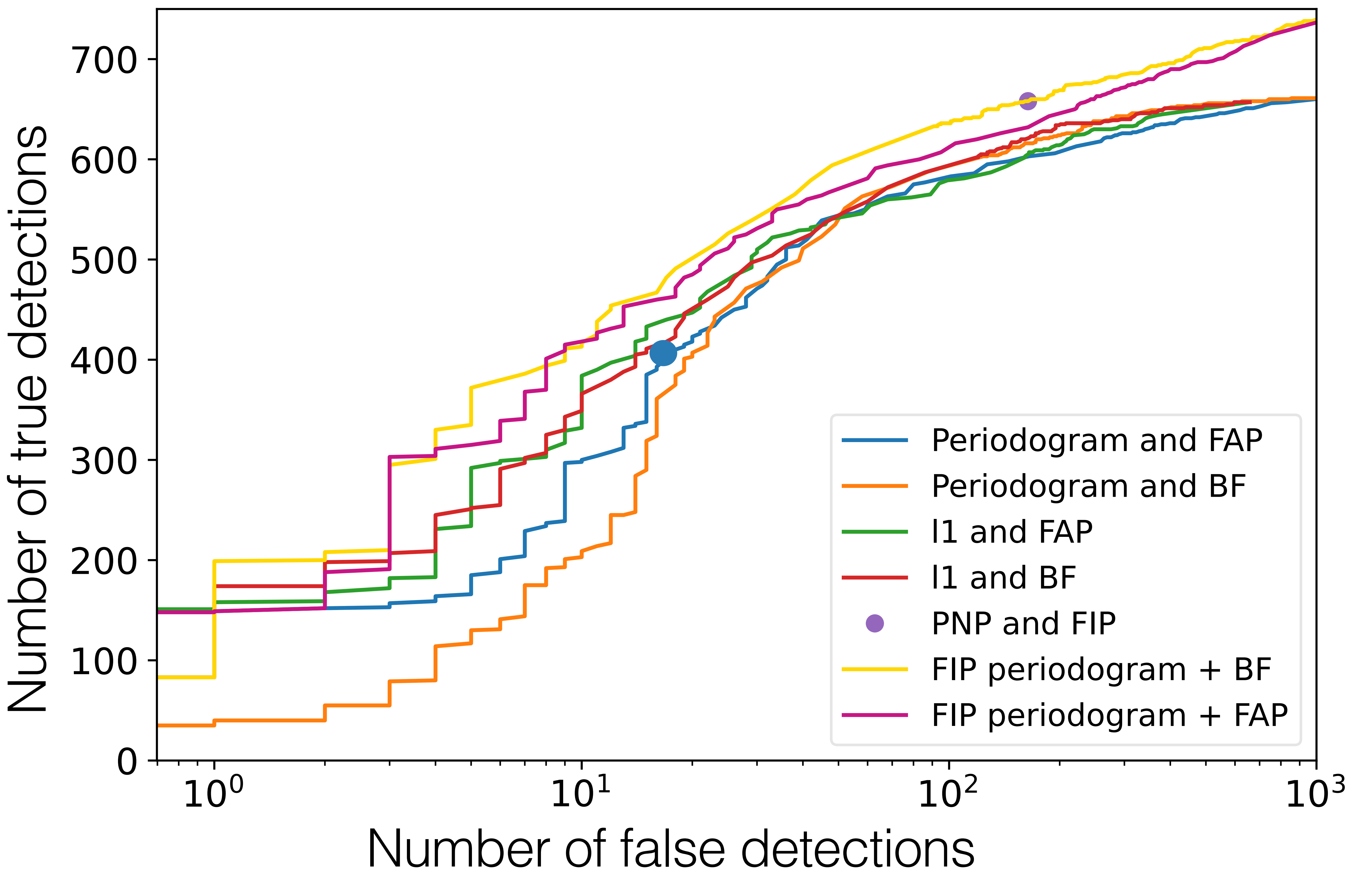


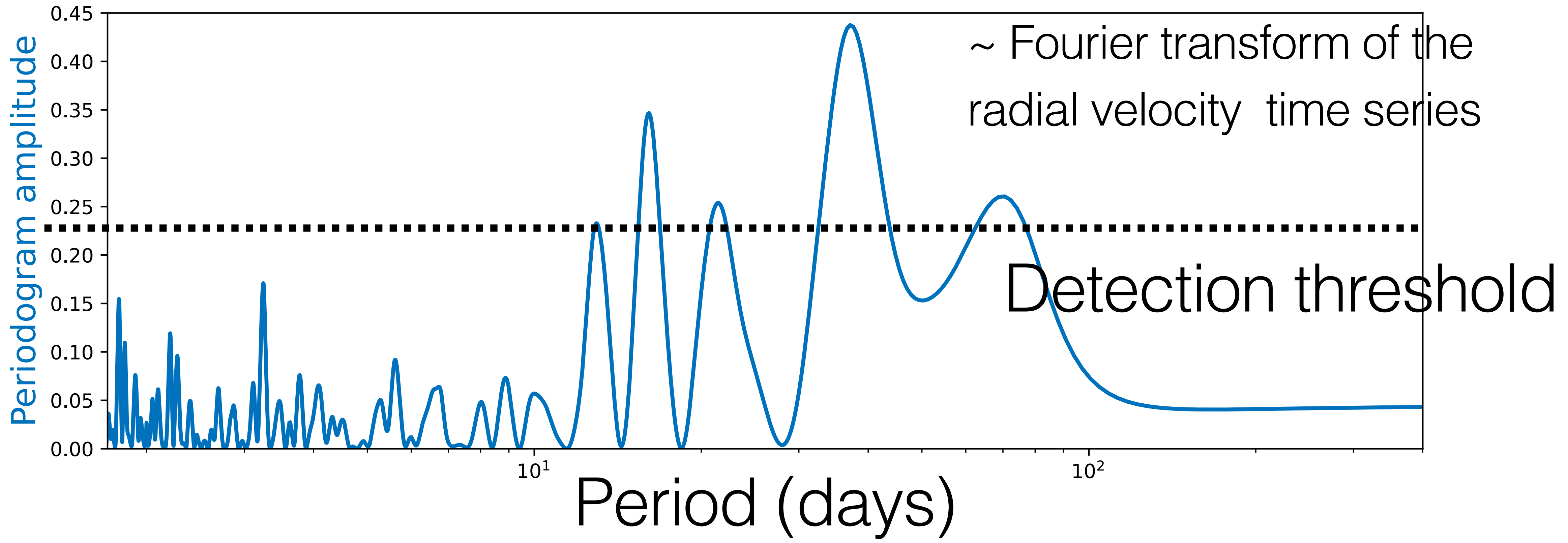
~ Fourier transform of the radial velocity time series

Detection threshold

Optimal criterion to detect exoplanets (2)







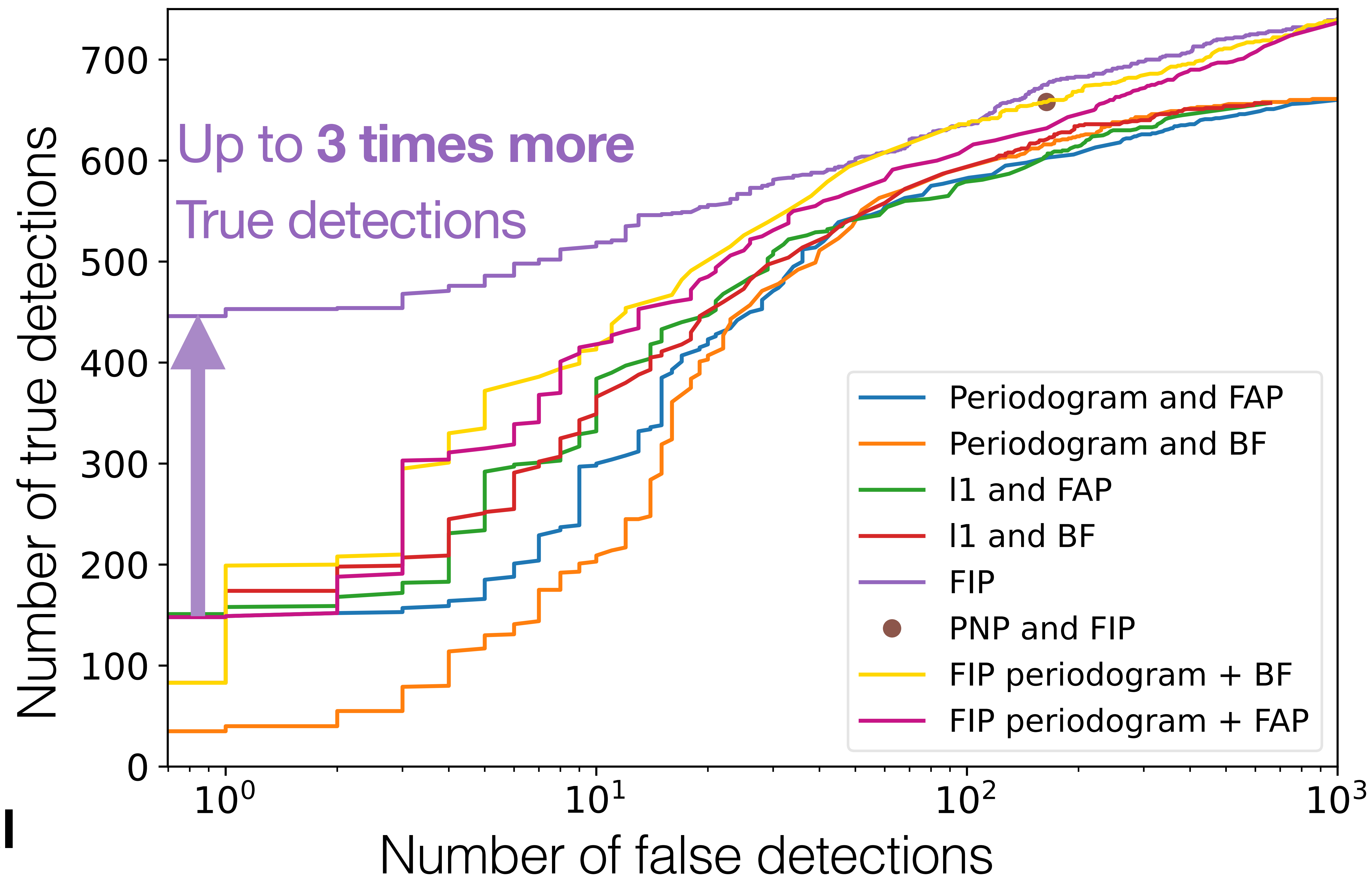
~ Fourier transform of the radial velocity time series

Detection threshold

Optimal criterion to detect exoplanets (3)

- **Mathematical proof of optimality** of a new detection criterion called « True inclusion probability » (TIP)
- Optimal in a general case

Hara et al. 2023, Annals of Applied Statistics (in revision)
Hara, Unger, Delisle, Díaz, Ségransan 2022

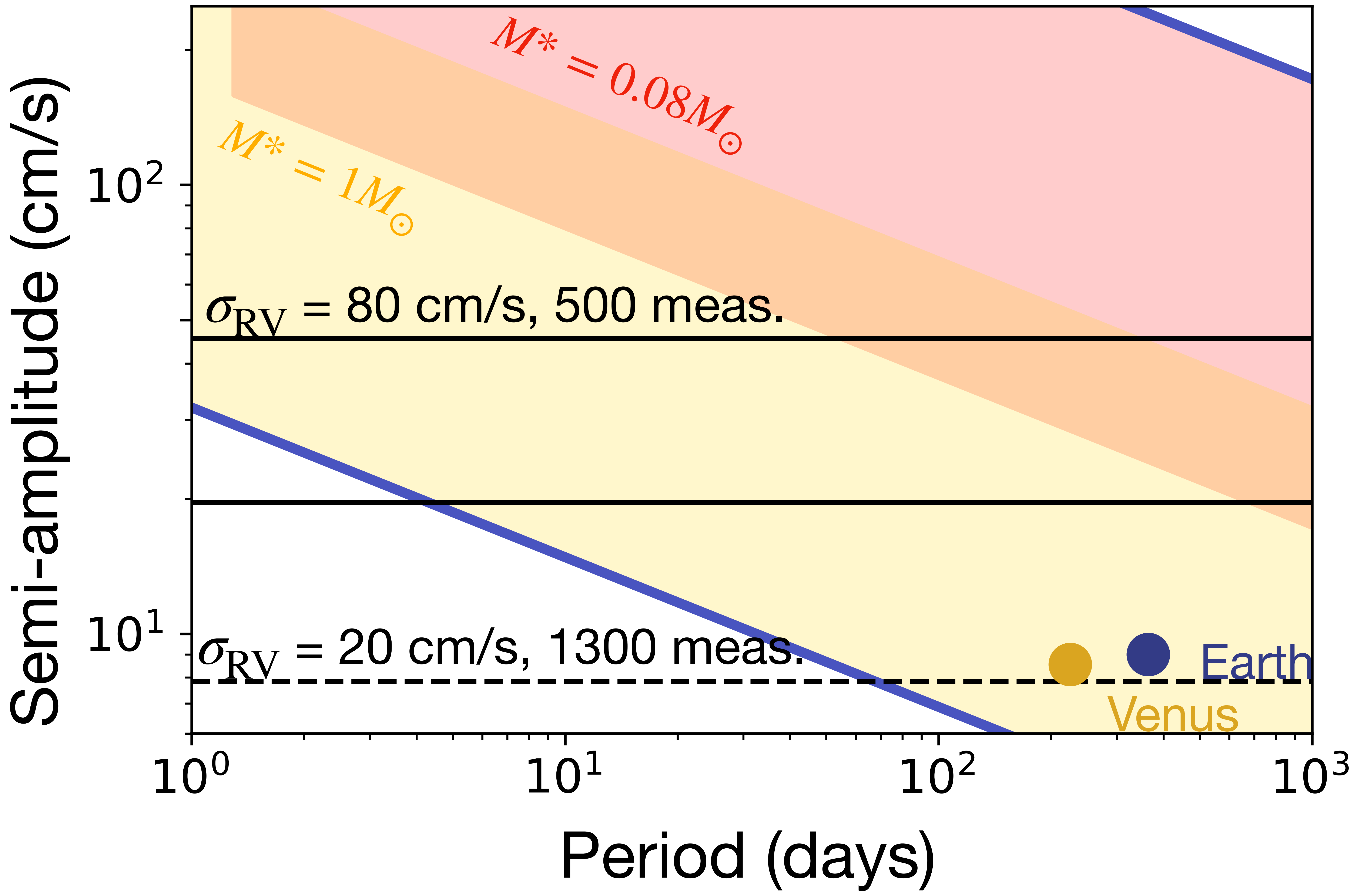


Bayes factors and FAPs

- Optimal?
  → **New criterion demonstrably optimal**
- Do not encode where the planet is
  → **Encoded in new criterion**
- Are not defined on a very intuitive scale
  → **New criterion is an actual probability**
- **Optimal for all exoplanet detection methods**

In a collection of independent detections made
with TIP 99%, on average 99% are correct
90% 90%
50% 50%

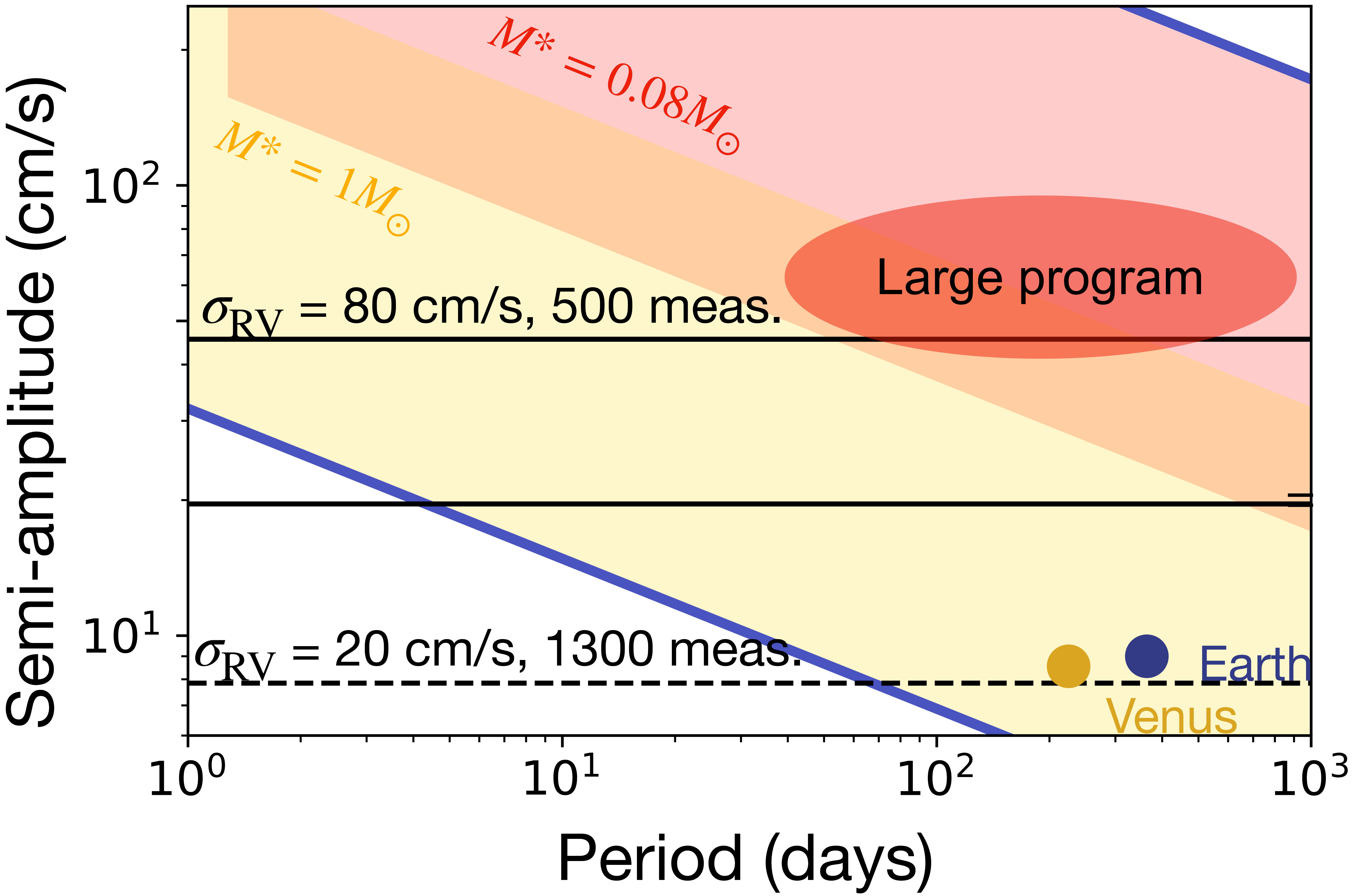
Motivation for a precursor survey



Shaded areas:
semi amplitude of RV signals
planets from 0.5 to 5 Earth masses, 0.01-10 AU for a $1M_\odot$ star in yellow, and a $0.08\ M_\odot$ in red.

Horizontal lines:
semi amplitudes measurable with a 10% precision on mass assuming **uncorrelated,** Gaussian error of amplitude σ_{RV}

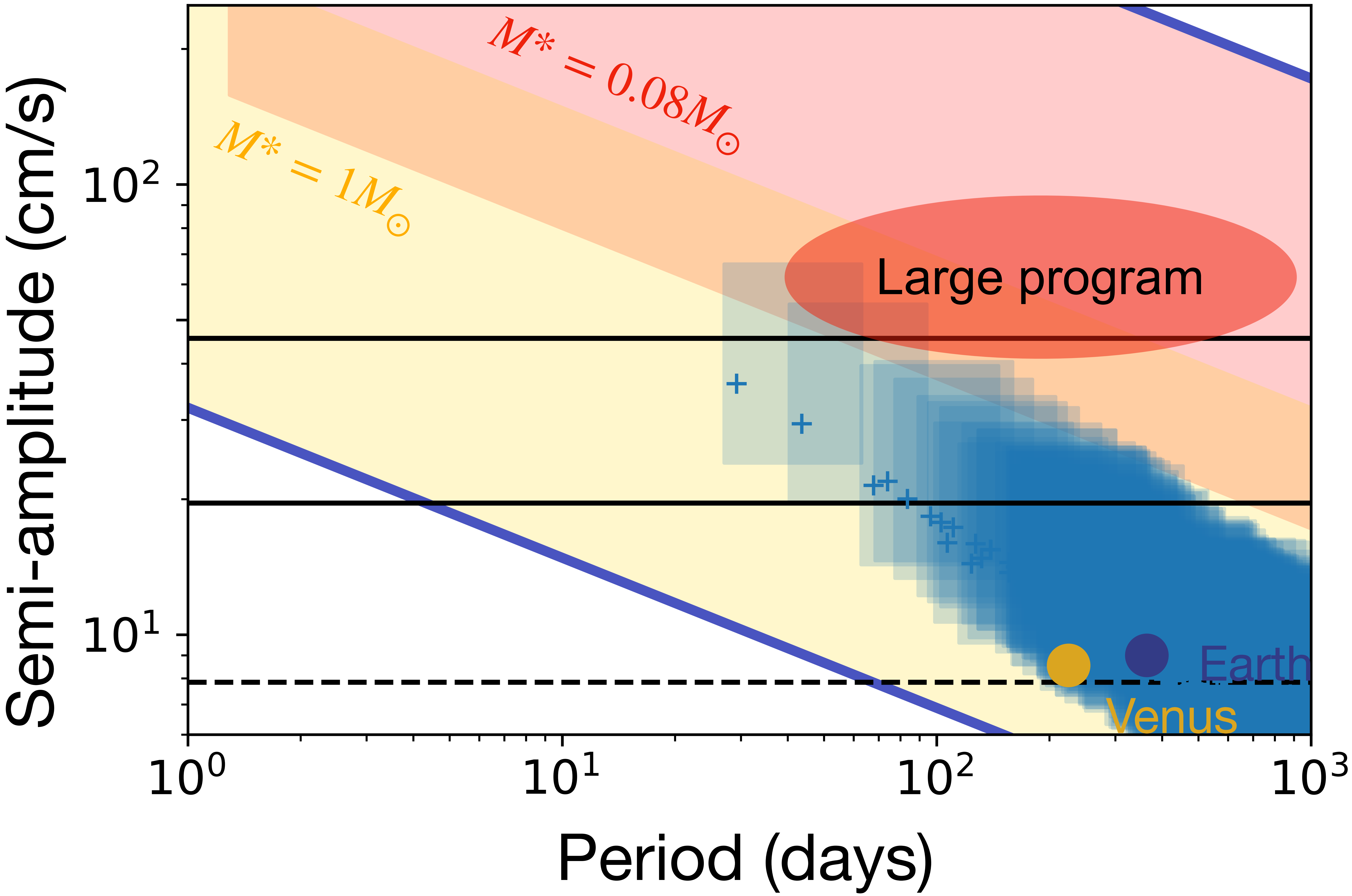
Motivation for a precursor survey



Shaded areas:
semi amplitude of RV signals
planets from 0.5 to 5 Earth masses, 0.01-10 AU for a $1M_\odot$ star in yellow, and a $0.08\ M_\odot$ in red.

Horizontal lines:
semi amplitudes measurable with a 10% accuracy assuming **uncorrelated,** Gaussian error of amplitude σ_{RV}

Motivation for a precursor survey



Expected RV amplitude for the Mamajek+ 2024 *HWO* target list
[0.95, 1.67] AU, [0.8, 1.4] $M_\oplus$

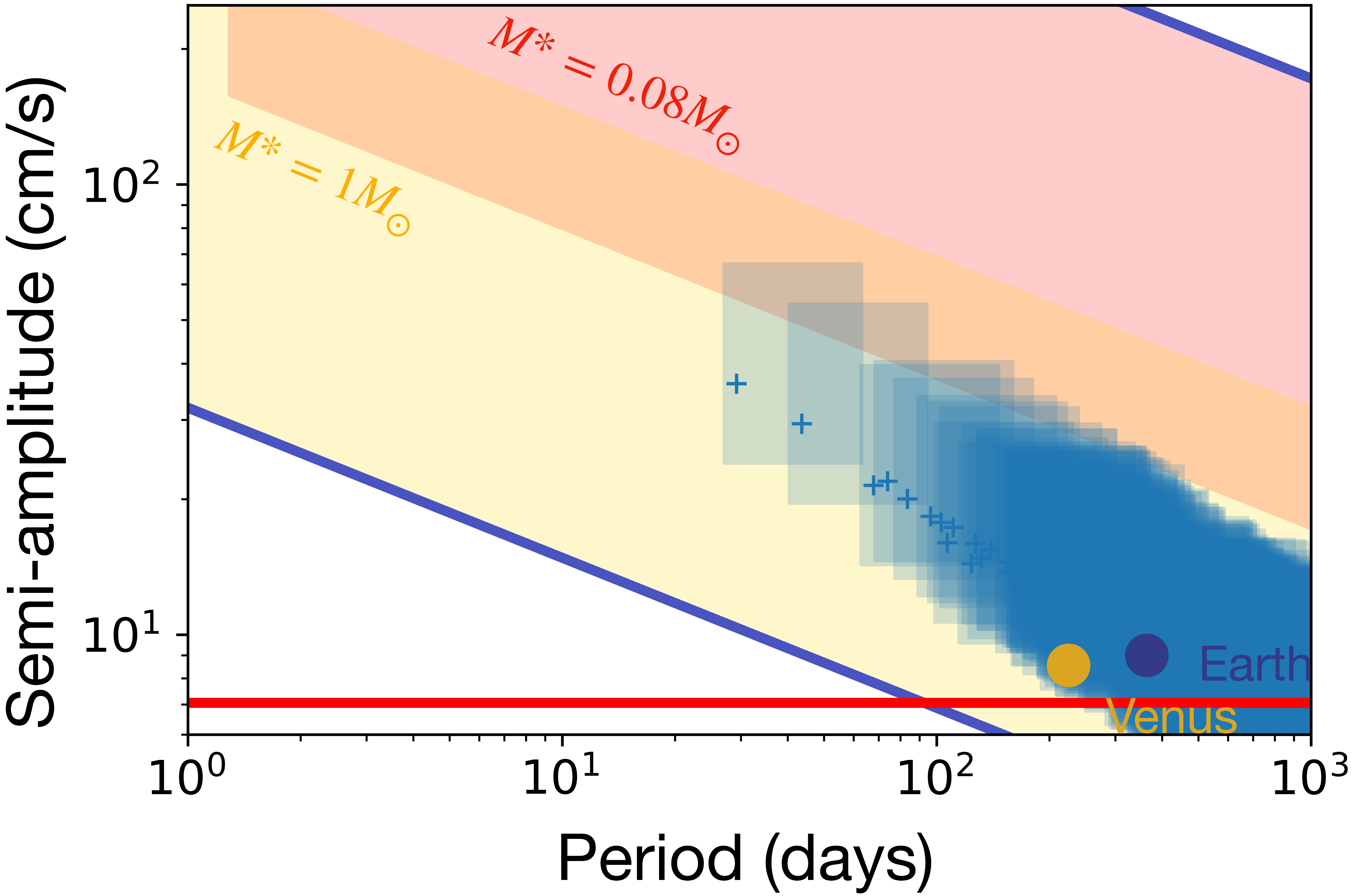
σ_{RV} = 80 cm/s, 500 meas.

σ_{RV} = 50 cm/s, 1300 meas.
or σ_{RV} = 20 cm/s, 208 meas.

σ_{RV} = 20 cm/s, 1300 meas.

10% precision on mass =
10 sigma detection

Motivation for a precursor survey



Expected RV amplitude for the Mamajek+ 2024 *HWO* target list

[0.95, 1.67] AU, [0.8, 1.4] $M_\oplus$

σ_{RV} = 30 cm/s, 1300 meas, 6 sigma detection

NASA Extreme precision radial velocity report (Crass+ 2021)

Existing high precision spectrograph:

HARPS, HARPS-N, **ESPRESSO**, NIRPS (for M dwarves)/ SOPHIE, SPIRou (Europe-led)

NEID, KPF (US)

Existing network: NASA Extreme precision RV initiative

Upcoming:

HARPS3 (Cambridge), G-Clef (Harvard)

Some key findings of the EPRV report:

- Establishment of a NASA EPRV Research Coordination Network and Standing Advisory
- More interdisciplinary collaboration (with stellar physicists in particular)
- Collaborative work on standard datasets (especially Sun-fed spectrographs, coordinate observations with major instruments on a small set of bright standard stars)

Other topics

Achieving <1 cm/s calibration accuracy with laser frequency combs

Observation strategy: how to most efficiently spend observation time?

Telluric contamination: how to remove it?

Is spectropolarimetry interesting to correct stellar variability?

Is there an advantage to using adaptive optics?

Target list for PCS: NIRPS spectrograph in the infrared

HARPS3 spectrograph dedicated to HWO target scanning

Impact of RV survey on *LIFE*

Conclusion

RV precursor surveys could greatly improve the yield of *HWO, LIFE,* PCS

Needs work on

- High resolution spectrograph hardware (better calibration source, adaptive optics)
- Data analysis methods (telluric, systematics, stellar variability)
- Lobbying: getting funds to push precursor surveys
- Quantifying the impact on *LIFE*
- What is the required precision on mass measurements to interpret atmosphere observations?
- Are there data analysis techniques that could be transferred between RV, transits and imaging?

References

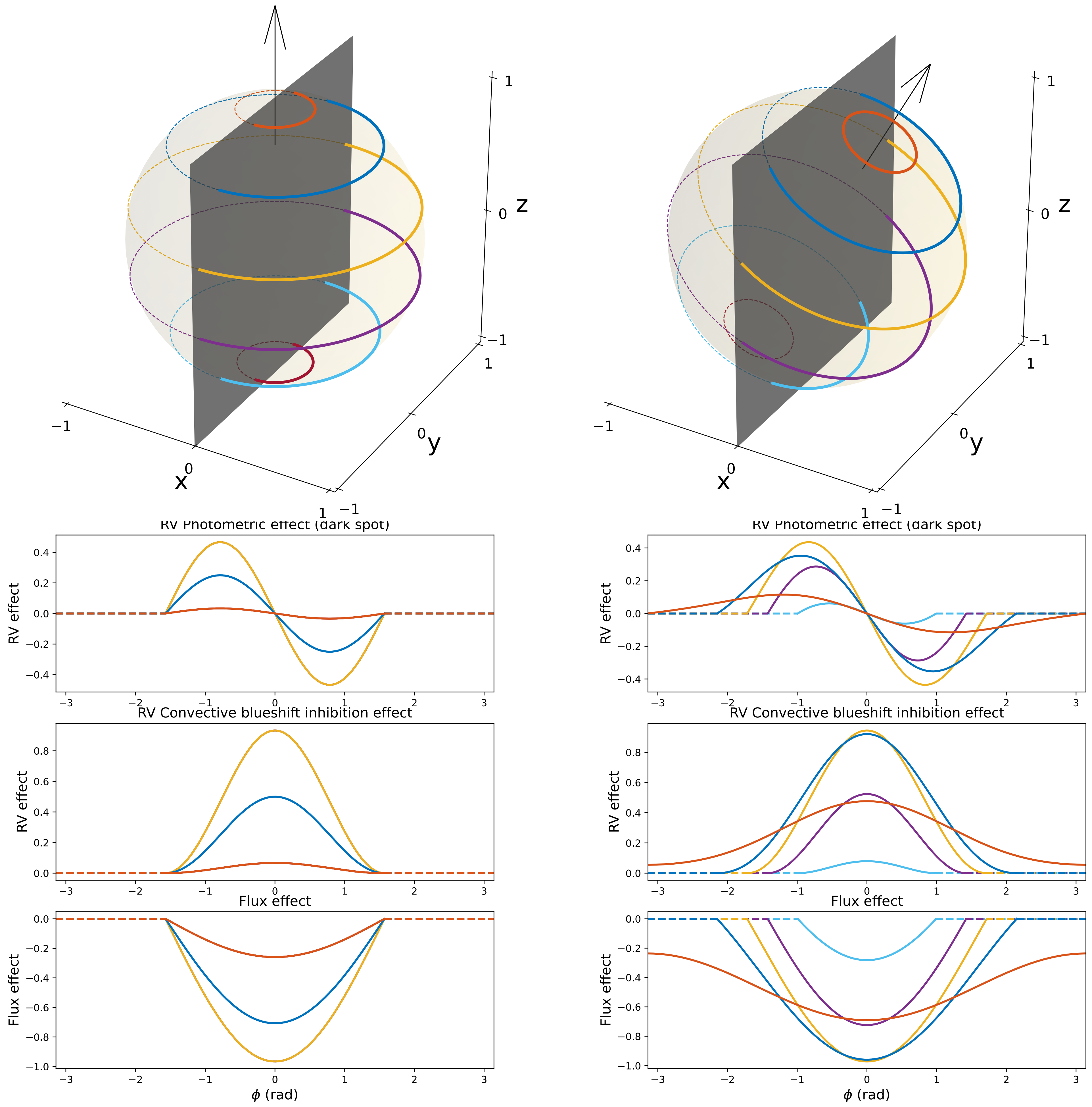
Target list: Mamajek et al. 2024**,** Gaudi et al. 2020

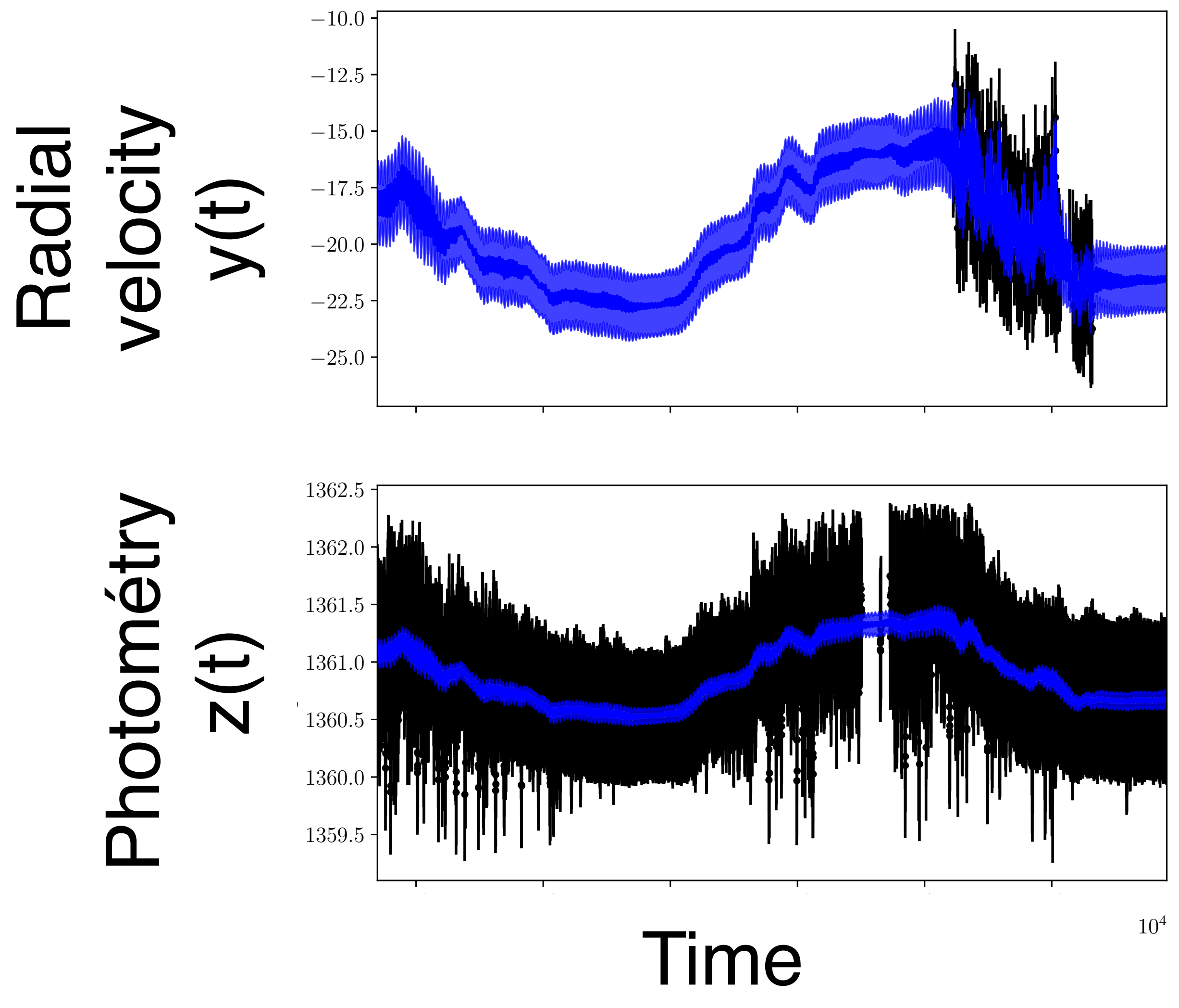
Impact of RV survey, Morgan et al. 2021, Crass et al. 2021

Roadmap to Extreme Precision Radial Velocity: Crass et al. 2021

Statistical methods for exoplanet detection with radial velocities: Hara & Ford 2023

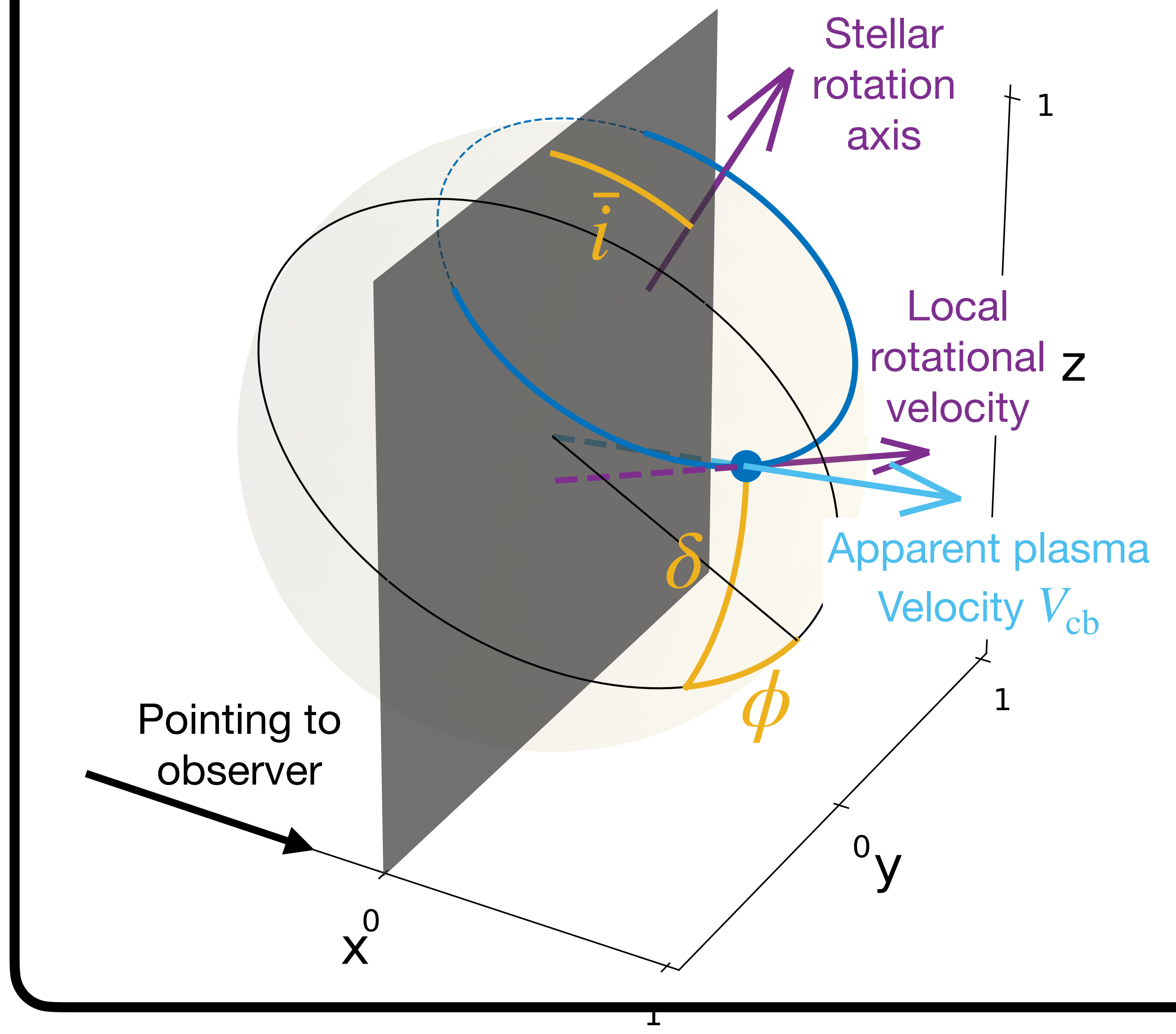
By-product: statistical Doppler imaging





Solar observations

RVs (HARPS-N) and photometry (SORCE)



Finding the solar inclination from the statistical distribution of the signal