Constraining Exoplanet Structure and Composition: Linking Theory & Observations

Fernandez Otegi, Jon

Université de Genève
Département d'Astronomie

Faculté des Sciences
Professeur François Bouchy
et Professeur Ravit Helled

Constraining Exoplanet Structure and Composition:

Linking Theory & Observations

Thèse

présentée à la Faculté des Sciences de l'Université de Genève
pour obtenir le grade de Docteur ès sciences,
mention Astronomie et Astrophysique

par

Jon Fernandez OTEGI

de

Villanueva de la Cañada (Espagne)

Thèse N° 5698

Genève
Observatoire Astronomique de l'Université de Genève
2022

Acknowledgements

My main thanks must necessarily go to Ravit and Francois. Professionally, I have always felt that they maintained a perfect balance between guiding me and allowing me freedom. Whenever I felt that I started to get frustrated or blocked with some task, I knew that meeting with either of them would resolve my doubts, and they would make me see "the way forward" again. Not least, they have always managed to keep my motivation high. I believe that in every meeting with the two of them I have always ended up more motivated with my work at the end than at the beginning, which says a lot about them. Personally, they always made me feel very calm and secure. They are both protective and caring, which gives a lot of confidence. I don't want to fall into clichés, but I really feel very fortunate to have had them as bosses, and I think that a very large part of what I have been able to do well during this Ph.D. is thanks to them. Thank you very much for everything!

Otherwise, it has been a very nice four years of my life. Both in Zurich and Geneva I have found very pleasant working environments with nice people. Special thanks to Claudio, who in spite of being a pain in the ass, you end up loving him a little, and to Sandra for being an unbeatable office companion. In Geneva, I must dedicate special thanks to Nolan, Jules, and, especially, Angelica for making my day to day life much more enjoyable both at work and outside. And lastly, of course, thanks to Altair for showing me how much you can love a silly poop machine.

Résumé

Depuis la découverte de la première exoplanète à la fin du vingtième siècle, les progrès technologiques ont considérablement élargi notre compréhension des systèmes planétaires. Le grand nombre d'exoplanètes découvertes a révélé une diversité stupéfiante dans leur nature, ce qui a conduit à la conclusion surprenante que notre système solaire pourrait ne pas être un système typique. L'un des principaux objectifs de la science planétaire est de comprendre la composition et la structure interne des planètes, qui peuvent ensuite être liées à leur formation et à leur évolution. La caractérisation de l'intérieur des exoplanètes est un défi, et nécessite une forte interaction entre la théorie et les observations. Une combinaison de plusieurs techniques de détection est nécessaire pour déduire les paramètres fondamentaux des planètes tels que la masse, le rayon, ou la température d'équlibrium, et ces observables permettent de mettre des contraintes sur les modèles théoriques. D'autre part, les modèles théoriques sont cruciaux pour guider les campagnes d'observation. Cette thèse représente un effort pour avancer dans l'interaction entre les observations et la théorie pour comprendre la nature sous-jacente des planètes.

Une partie importante de ce travail est consacrée à l'étude de la démographie des exoplanètes. Nous entrons dans une ère où nous pouvons comparer et classer les exoplanètes en populations distinctes en fonction de nombreux paramètres comme la masse, le rayon, le rayonnement entrant, le type stellaire, etc. Nous présentons un nouveau catalogue d'exoplanètes aussi fiable que possible, basé sur des mesures robustes de masse et de rayon. Ce catalogue revisité permet d'identifier des tendances et des sous-populations d'exoplanètes qui n'avaient pas été vues auparavant. La grande diversité des exoplanètes s'étend également à l'architecture des systèmes multi-planétaires. Nous utilisons également le catalogue d'exoplanètes revisité pour effectuer une analyse approfondie de la similarité de la masse, du rayon et de la densité des systèmes multi-planétaires.

Outre les études démographiques, nous nous sommes également concentrés sur la caractérisation de la structure interne des planètes individuelles. En nous appuyant sur un modèle de structure antérieur, nous avons mis à jour plusieurs parties des modèles de noyau, d'eau et d'enveloppe volatile, et nous l'utilisons pour caractériser diverses exoplanètes découvertes.

Nous utilisons également ces modèles internes pour étudier plusieurs aspects qui affectent la caractérisation des super-Terres et des sub-Neptunes, tels que les incertitudes observationnelles, la localisation dans le diagramme masse-rayon, ou les hypothèses du modèle.

Cette thèse se veut être une étape supplémentaire vers une compréhension plus profonde de la nature exoplanétaire.

Abstract

Since the discovery of first exoplanet by the end of the twentieth century, technological advancements have hugely extended our understanding of planetary systems. The large number of discovered exoplanets has unveiled a staggering diversity in their nature, leading to the surprising conclusion that our Solar System may not be a typical system. One of the key objective of planetary science is to understand the composition and internal structure of planets, which can then be linked to their formation and evolution path. The characterization of exoplanet interior results challenging, and requires a strong interaction between theory and observations. A combination of several detection techniques is needed to infer planet fundamental parameters as the mass, radius, or equlibrium temperature, and these observables allow to put constraints on theoretical models. On the other hand, theoretical models are crucial to guide observational campaigns. This thesis represents an effort to advance in the interplay between observations and theory to understand the underlying nature of planets.

A substantial part of this work is dedicated to the investigation of the exoplanet demography. We are entering an era in which we can compare and classify exoplanet in distinct populations based on many parameters as mass, radius, incoming radiation, stellar type, etc. We present a new as reliable as possible exoplanet catalog based on robust mass and radius measurements. This revisited catalog allows to identify trends and exoplanet subpopulations that were not seen before. The large exoplanet diversity also extend to the architecture of multi-planetary systems. We also use the revisited exoplanet catalog to perform an in-depth analysis of the similarity in mass, radius, and density of multi-planetary systems.

Besides the demographic studies, we have also focused on the characterization of the internal structure of individual planets. Building on a previous structure model, we updated several parts of the core, water and volatile envelope models, and use it to characterize various discovered exoplanets. We also use these internal models to study several aspects that affect the characterization of super-Earths and sub-Neptunes, such as observational uncertainties, location in the mass-radius diagram, or model assumptions.

This thesis aims to be an additional step towards a deeper understanding of the exoplanetary nature.

Contents

CHAPTER 1

INTRODUCTION

In 1992 the first-ever discovery of an exoplanet was announced (Wolszczan & Frail 1992). They detected three terrestrial planets around the pulsar PSR 1257+12. Three years later, on October 6 1995, another milestone in science was achieved. Michel Mayor and Didier Queloz discovered the first exoplanet around a main-sequence star (Mayor & Queloz 1995). Located about 50 light years from us, 51 Pegasi b is a gas giant with a mass about 150 times that of Earth. This has been one of the most significant planets ever discovered, so important that in 2019 the Nobel Prize in Physics was awarded to the astronomers who found it. Indeed, with this discovery, the field of exoplanets was born.

Since then the number of discovered exoplanets has grown tremendously. The exoplanet research developed rapidly during the last two decades. As of writing this thesis 4884 exoplanets have been discovered to date (considering the data from the NASA Exoplanet Archive in January 2022), and we are now aware of the large existing diversity of exoplanets. We have also realized that the properties of the identified exoplanets show a noticeable contrast when compared to the planets in the Solar System, leading us to the surprising conclusion that our Solar System may not be a typical planetary system. Nowadays, we are able to perform demographic studies comparing and classifying exoplanets in different populations. Figure 1.1 shows the mass or minimum mass against the semi-mayor axis of the whole known exoplanet population from the NASA Exoplanet Archive[a]. We can see that exoplanets tend to group into the following populations:

[a] https://exoplanetarchive.ipac.caltech.edu/

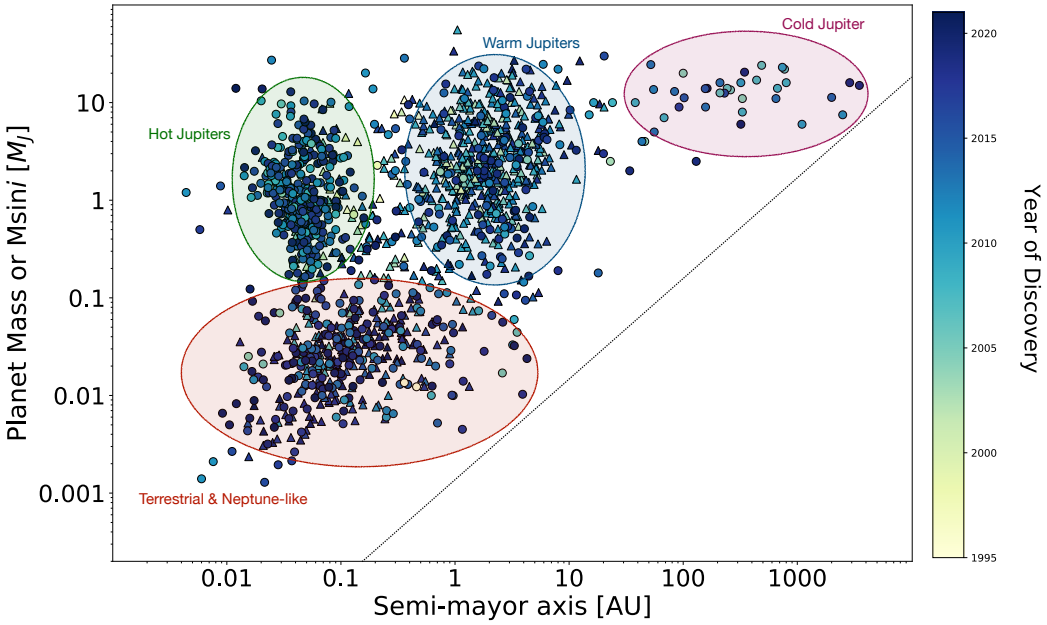

Figure 1.1: Mass against semi-mayor axis of the known exoplanet population from the NASA Exoplanet Archive. The color represents the year of their discovery. Triangle markers indicate planets with minimum mass. Below the dashed line there are no detected planets due to observation biases. We indicate the different populations of exoplanets.

- *Hot Jupiters:* These exoplanets are gas giants with masses comparable or larger than Saturn orbiting very close to their host stars (<0.1AU). Usually they are tidally locked, and they are likely to have extreme atmospheres due to their proximity to the host star. These exoplanets are the easiest to find with radial velocity (RV) technique and transit method.
- *Warm and Cold Jupiters:* Warm Jupiters are located at separations larger than 0.5AU, cold Jupiters beyond 10AU. The observed separation between the groups of giant planets is due to different detection techniques. Cold Jupiters are mostly discovered through direct imaging, most of the warm Jupiters are minimum mass measurements obtained by radial velocity technique, and the majority of hot Jupiters have been detected through transit technique.
- *Terrestrial and Neptune-like planets:* Terrestrial planets are planets that are mainly composed of metals and silicate rocks. Neptune-like exoplanets are expected to have a similar core as terrestrial planets, but also posses a significant atmosphere that typically is hydrogen- and helium-dominated.

The observed distribution of planets is strongly affected by observational and detection

biases. The absence of planets below the dashed line in Figure 1.1 is caused by the limitations of the current detection techniques. We can see that most of the planets discovered until 2010 were giant planets. Detecting small planets is technically challenging, and most of the terrestrial and Neptune-like exoplanets have been discovered after 2015. Continuing the search to discover more worlds is crucial to understand the origin of planetary systems in general, and will eventually allow us to understand what conditions are necessary for planets to be capable of hosting life.

1.1 Detection Methods

Due to the difference in brightness between an exoplanet and its host star, it is very challenging to detect it via direct imaging. Most of the observed exoplanets have been discovered through indirect detection methods, i.e., methods in which the exoplanet's effects on its host star are used to infer its presence. Among all detection techniques, transit photometry and radial velocimetry have been the most successful methods to discover exoplanets, accounting for ~95% of the detections.

Figure 1.2 shows the number of discovered planets every year, with different colors representing different discovery methods. Note that we use the NASA Exoplanet Archive. After (Mayor & Queloz 1995) discovered the first exoplanet around a main-sequence star, the radial velocity technique was the most important detection method for almost twenty years. The number of planets with measured radii was very low in the 2010s, until the Kepler mission. Kepler was launched in 2009 to explore the structure and diversity of planetary systems searching transiting planets. However, these planets were often too dim to perform radial velocity follow-up, so the number of planets with measured mass and radius was still relatively low. The loss of a reaction wheel on the spacecraft in 2013 finished Kepler's four-year science mission. In its second phase, the K2 mission continuously monitored more than 150,000 stars near the ecliptic plane (Howell et al. 2014a) and lead to a strong increase in the number of detected planets. Compared to Kepler, the K2 mission observed more diverse stellar populations and focused on brighter targets, which are more amenable for radial velocity follow-up observations. Later in 2018, the Transiting Exoplanet Survey Satellite (TESS) was launched to survey 85% of the sky for transiting exoplanets around bright stars (Ricker et al. 2015). RV follow-up programs have triggered a rapid expansion of the number of confirmed planets with a mass measurement, which has allowed to perform an in-depth analysis of the demography of exoplanets. In this section, I will go through some main exoplanet detection techniques.

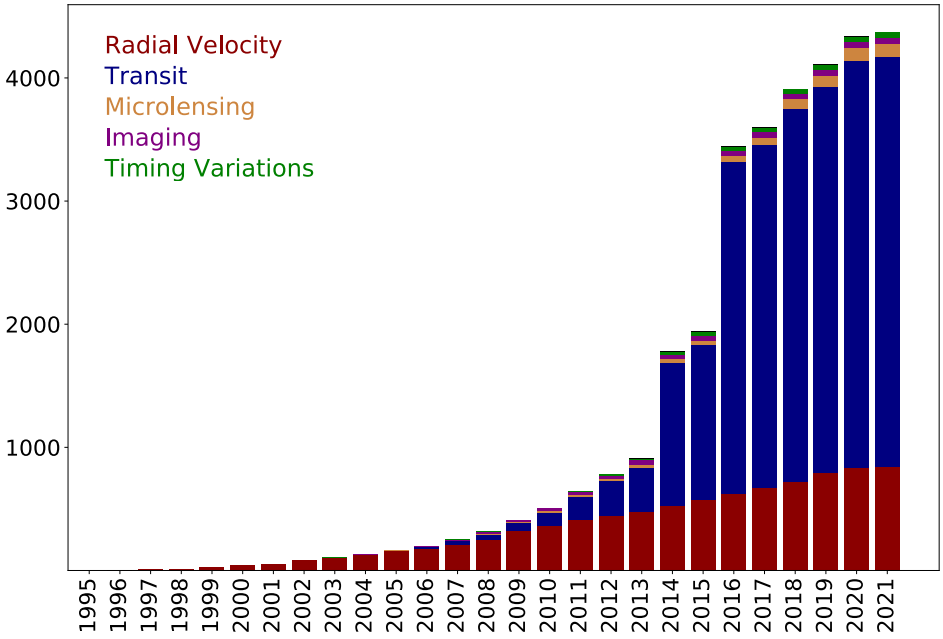

Figure 1.2: Number of discovered exoplanets every year to date. Different colors represent different discovery methods.

1.1.1 Transit Photometry

Transit photometry is currently the most effective and sensitive method to detect exoplanets at short period. If a planet crosses in front of its host star, then the observed flux drops by a small amount. Figure 1.3 provides a schematic view of the transit photometry. Assuming that the flux coming from the planet is negligible, the drop in flux (ΔF) is called transit depth, and it is determined by the ratio of the areas of the planet and the star:

$$\Delta F \approx \left(\frac{R_p}{R_s}\right)^2 \tag{1.1}$$

where R_p and R_s are the planetary and stellar radii respectively. Jupiter and the Earth are nearly 10 and 100 times smaller than the Sun respectively, so their transit depths are 1% and 0.01%. We see that very precise photometry is needed to detect small planets. Using basic geometry, Seager & Mallen-Ornelas (2003) derived analytic expressions relating the transit observables to the orbital parameters:

$$b \equiv \frac{a}{Rs}\cos i = \left(\frac{(1-\frac{R_p}{R_s})^2 - [\sin^2(t_F\pi/P)/\sin^2(t_T\pi/P)](1-\frac{R_p}{R_s})^2}{\cos^2(t_F\pi/P)/\cos^2(t_T\pi/P)}\right)^{1/2} \tag{1.2}$$

where b is the impact parameter, a the semi-mayor axis, i the inclination, t_T the total transit duration and t_F the time in which the whole planet is in front of the star. This equation assumes a circular orbit. These expressions allow estimating the semi-mayor axis of the orbit, the star density, the planet/star radius ratio, and inclination.

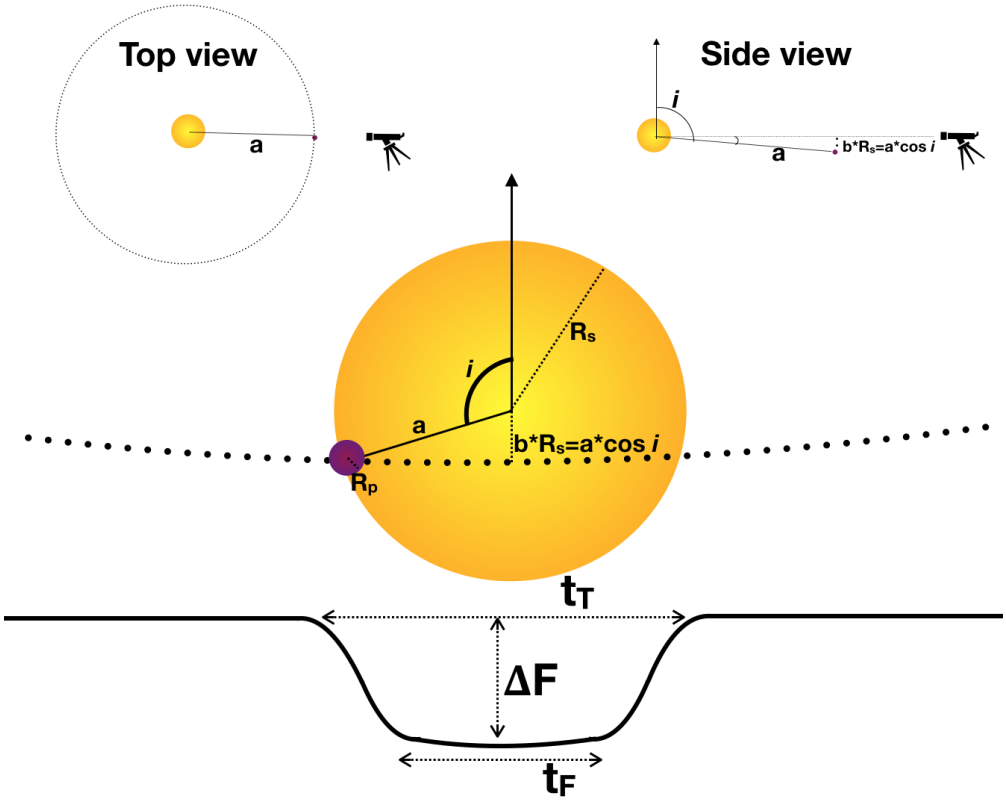

Figure 1.3: Schematic view of a transit. The lower solid curve represents the observed light curve, the middle planet illustrates the model from the observer's point of view, and the top panel from other points. Figure taken from Deeg & Alonso (2018).

This method has two main disadvantages. First, the transits are observable only when the planet's orbit is almost perfectly aligned with the vantage point. The probability of observing a planet in a transit is provided by Winn (2010):

$$p_{\rm tr} = \left(\frac{R_s \pm R_p}{a}\right)\left(\frac{1 + e\sin\omega}{1 - e^2}\right) \tag{1.3}$$

where the plus and minus signs refer to grazing and full transits, respectively. Nearly 10% of planets with small orbits have such alignment, and the fraction decreases for planets with larger orbits. For a hot Jupiter with semi-major axis of 0.05 AU the probability is on the order of 10%, and for an Earth-size planet at 1 AU it is 0.5%. Since the majority of planet systems will

not display transits, transit surveys generally scan large areas of the sky, monitoring thousands of stars at once. This way ways transit surveys have managed to find more exoplanets than any other method.

The second main disadvantage of the transit method is the high rate of false detections. There are some astrophysical configurations that can lead to similar transit-shaped light curves. For example, a stellar eclipsing binary in a sky position that is so close to a brighter star that the light of both objects is in the same pixel of the detector. Also, some false positives may be caused by artifacts of the red noise or other instrumental effects. Santerne et al. (2012) found that false positives for transits observed by the Kepler mission could be larger than ~40% for giant planet candidates in single-planet systems. Therefore, follow-up observations are required to confirm the detection. These follow-up observations include multicolor photometry, spectroscopic observations or orbital brightness modulation.

1.1.1.1 Transit Timing Variations

Transit timing variations (TTVs) is a highly sensitive technique to detect additional planets in a system. It is based on the gravitational pull of planets in tightly packed systems among themselves, which causes some planets to accelerate and some others to decelerate. The Kepler mission enabled detailed statistical analysis of multi-planet systems (Fabrycky et al. 2014; Lissauer et al. 2014), suggesting that at least half of the planets are in multi-planet systems, and a significant proportion of them are not easy to detect though transit (i.e. some may not transit) or radial velocity (i.e. some may not be massive enough). Analyzing the variations on the transit timing is then a powerful tool to reveal additional planets in the system with very low masses.

For a single planet, the time of the transit will be the following:

$$t_{\mathrm{cte}}(n) = nT + t_0 \tag{1.4}$$

where T is an interval, n is an integer, and t_0 is the time at the transit n=0. An additional planet in the system will cause variation in the transit interval of the planet, so the TTVs of the nth transit can be written as:

$$\delta t(n) = t(n) - t_{\mathrm{cte}}(n) = t(n) - nT - t_0 \tag{1.5}$$

T and t_0 can then be computed by fitting a linear regression to the data. This method not only reveals additional objects interfering with the motion of the star and the planet, it also

allows calculating the masses of the planet and the additional object. There are several ways to make this estimation, depending on the orbit characteristics. A common method is based on making an n-body simulation, varying the masses of the planets until the observations are fitted. If the system has two planets in approximate resonance, the following approximation from Agol et al. (2005) can be applied:

$$\delta t_{\max} \approx \frac{T}{4.5j} \frac{M_2}{M_1 + M_2} \tag{1.6}$$

where M_1 and M_2 are the masses of the planets, $\delta t_{\max}$ the maximum transit timing variation, T the average time between transits and j the order of the resonance.

Like all the other detection methods, TTVs more effective to detect some specific exoplanets. According to Steffen (2016), the sensitivity of TTVs can be expressed as follows:

$$\mathrm{SNR}_{\mathrm{TTV}} \sim \frac{M_p R_p^{3/2} P^{5/6}}{\sigma_{TTV}} \tag{1.7}$$

where σ_{TTV} is the intrinsic uncertainty of a measurement. This means that the TTV technique can determine masses more easily for more massive, bigger, and longer period exoplanets.

Despite being a much more powerful to measure masses and eccentricities in small planets at relatively long orbital periods than RVs, using TTVs to infer planet properties requires fitting a large number of parameters that normally are degenerate. In order to overcome this inference problem, Marcov Chain Monte Carlo (MCMC) simulations are usually used. However, interpreting the MCMC results is often difficult because of strong parameter correlations and sensitivity to priors. In Hadden & Lithwick (2017) they used both N-body and analytic fitting in a complementary approach in order to identify the degeneracies inherent to TTVs and classify inferred planet masses as robust or not.

A highly remarkable multi-planetary system robustly characterized by TTVs is TRAPPIST-1. With a mass of $0.09 M_\odot$, a radius slightly larger than Jupiter and a surface temperature of 2560K, TRAPPIST-1 hosts seven terrestrial planets in orbital resonance. Four of these planets are thought to be located withing the habitable zone of the star, with temperatures suitable to the presence of liquid water.

1.1.1.2 Space transit surveys

After the first observed planet transit (Charbonneau et al. 2000) several ground based surveys have been developed to search for transiting planets. The Hungarian Automated Telescope Network (HATNet Bakos et al. 2004) and the Wide Angle Search for Planets (WASP Pollacco

et al. 2006) were the two ground-based transit surveys, and they have been responsible for the majority of transiting planets discovered from the ground. However, these missions were mainly sensitive to giants planets. CoRoT (Convection, Rotation et Transits planétaires Baglin 2003; Auvergne et al. 2009; Moutou et al. 2013) was the first space mission looking for transits. Its goal was to search for smaller radii of rocky worlds, and its most remarkable discovery was CoRoT-7b (Léger et al. 2009; Queloz et al. 2009), which became the first detected exoplanet with a rock or metal-dominated composition.

The Kepler space telescope (Borucki et al. 2010) was the successor of CoRoT, which aimed to discover Earth-size exoplanets transiting their host stars. Kepler had a photometer that continuously monitored the brightness of 150 000 main sequence stars in a fixed field of view of 115 degree^2 in the Cygnus constellation. During the four years the mission lasted, Kepler performed photometry of unprecedented precision, observed 530 506 stars and detected ~2700 exoplanets. Most of the discovered planets range by Kepler range between the size of Earth and Neptune, and it also found some Mercury-size planets and planets in the habitable zone. The impact of Kepler in the field was huge, not only because of the number of discovered planets, but also because of the richness and the unexpected nature of the planetary system architectures discovered. In addition, thanks to Kepler, we knew that there are more planets in the Milky Way than stars (Ballard & Johnson 2016). In 2013 a reaction wheel stopped functioning, leading to the K2 phase Howell et al. (2014b). Compared to Kepler, the K2 mission covered more sky and observed more diverse stellar populations. Even if the second phase of the Kepler mission focused on brighter stars than the first one, their relatively faint magnitudes made difficult radial velocity (RV) follow-up.

Later, the Transiting Exoplanet Survey Satellite (TESS) was launched in 2018 to search for exoplanet in an area 400 times larger than that covered by the Kepler mission. TESS is an all sky survey which searches for transits around the brightest stars around the Earth. The satellite consists of an array of four wide-field cameras, and the primary mission had the goal of covering the whole sky in two years to find planets smaller than $4R_\oplus$. It targeted stars that are significantly brighter than the ones observed in the Kepler and K2 missions, amenable to perform RV follow-up. In Figure 1.4 shows the mass-radius diagram of known exoplanets with measured mass and radius from the PlanetS Catalog[b], highlighting the ones detected by Kepler, the K2 mission, and TESS. 12% of exoplanets in this plot have been discovered by Kepler, 8% by the K2 mission, and 17% by TESS. In addition, we note that the majority of planets detected by Kepler and K2 are in the gas giant regime, while most of the planets found

[b] https://dace.unige.ch/exoplanets/

by TESS are located below $4R_{\oplus}$. We see that TESS has triggered a rapid expansion of the number of well characterized sub-Neptunes, which has allowed to perform in-depth analysis of the demographic patterns of this population.

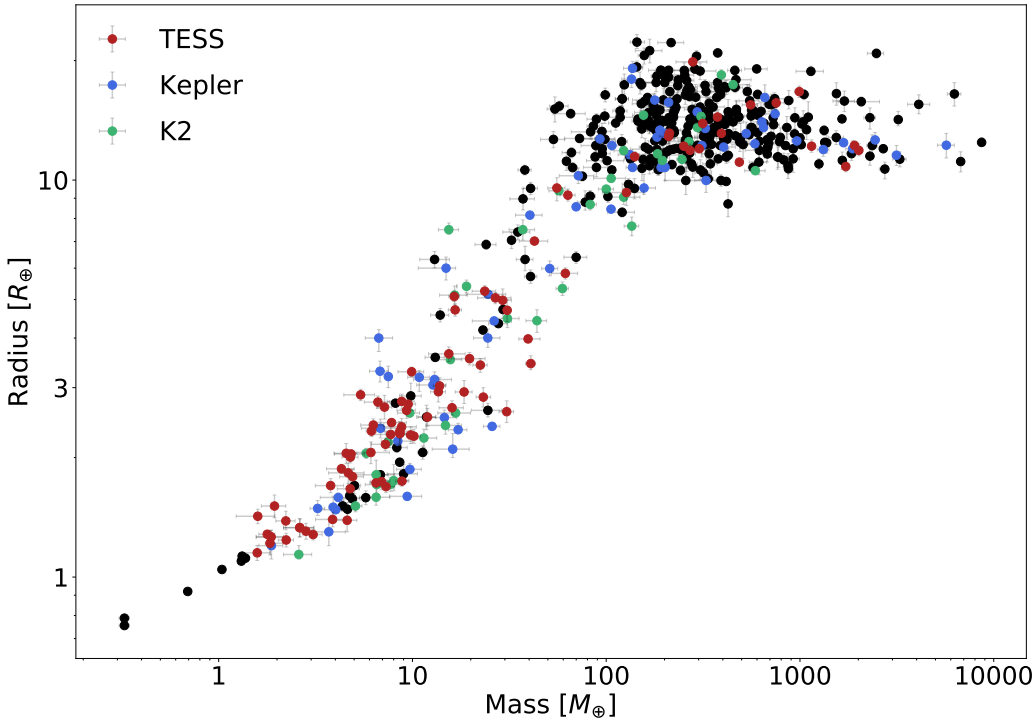

Figure 1.4: Mass-radius diagram with exoplanets from the PlanetS Catalog. We highlight exoplanets detected by the primary Kepler mission (blue), K2 (green), and TESS (red).

1.1.2 Radial Velocimetry

The radial velocity (RV) method was responsible for the first discovery of a planet orbiting a main-sequence star (Mayor & Queloz 1995). A planet orbiting a star exerts a gravitational force, making the star to orbit around the center of mass of the two-body system. The radial velocity describes the velocity variations of the star projected to the line of sight, and it can be detected through the Doppler effect. Using high-resolution spectrographs, the stellar perturbations caused by the planet can be detected through Doppler red shifts and blue shifts of the spectral lines. The RV can then be measured as follows:

$$RV = c\frac{\lambda_{\text{obs}} - \lambda_{\text{rest}}}{\lambda_{\text{rest}}} \tag{1.8}$$

where c is the speed of light, and λ_{obs} and λ_{rest} are the observed and rest wavelengths, respectively. On the other hand, the RV of the star can be described with a Keplerian orbital

solution:

$$RV = \gamma_0 + K[\cos(\nu(t, P, T_0) + \omega) + e\cos(\omega)] \tag{1.9}$$

where γ_0 is the systemic velocity of the star relative to the observer, K is the RV semi-amplitude, P the orbital period, T_0 the reference epoch, e the orbital eccentricity, ω the argument of periastron, and $\nu(t, P, T_0)$ the true anomaly, which determines the position of the star with respect to the argument of periastron. The semi-amplitude can then be related to the mass and semi-mayor axis using the following relation presented in Lovis & Fischer (2010):

$$K = \sqrt{\frac{G}{1-e^2}} M_p \sin i (M_s + M_p)^{-1/2} a^{-1/2} \tag{1.10}$$

where G is Newton's gravitational constant, i the inclination of the orbit, a the semi-mayor axis, and M_p and M_s are the planetary and stellar masses, respectively. This expression can be expressed in more convenient astronomer-units as follows:

$$K = \frac{28.4329 \text{ m s}^{-1}}{\sqrt{1-e^2}} \frac{M_p \sin i}{M_{Jup}} \left(\frac{M_s + M_p}{M_\odot}\right)^{-1/2} \left(\frac{a}{1 \text{ AU}}\right)^{-1/2} \tag{1.11}$$

We note that there is an inherent degeneracy in the mass determination through this method, and only the minimum mass ($M_p \sin i$) can be estimated. In consequence, coupling RV measurements with transits results extremely useful not only because it allows to constrain the true mass of the planets, but it also allows deriving the full architecture of the system and minimize false positives.

A major difficulty when analyzing RV data comes from the stellar activity, since several types of stellar activity can induce RV signals. For instance, some stars have dark spots and bright places that can interfere with the RVs. These spots can distort the absorption lines in the stellar spectra as it rotates in and out of the view, creating additional signals in the RV time series (Saar & Donahue 1997). A common method to deal with this is to tune Gaussian process (GP) hyperparameters on photometric data, and then use them to model the stellar effect on the RVs (Aigrain et al. 2012). Also, the GP can be trained on one or various activity indicators or on the RVs themselves (Haywood et al. 2014).

The sensitivity for the RV method derived in Steffen (2016) is the following:

$$\text{SNR}_{\text{RV}} \sim \frac{M_p}{\sigma_{RV} P^{1/3}} \tag{1.12}$$

where σ_{RV} is the intrinsic uncertainty of a measurement. If we compare this expression to Equation 1.7 we realize that the RV and the TTV methods have different observational biases, since it is easier for the RV technique to derive the mass of a short period planet, while the TTV technique can determine masses more easily for longer periods. Several studies have claimed the sensitivity of the TTVs is more uniform, and that for a given radius TTVs can detect a wider range of masses than RVs, probing smaller densities (e.g. Wolfgang et al. 2015; Steffen 2016).

1.1.3 Other detection techniques

Apart from the transit photometry and radial velocity methods, there are other techniques that have their own advantages. Combining all the techniques allows discovering a large range of planets distributed around different types of stars. The main detection techniques other than transit and RV are the following:

- **Direct Imaging:** This is the most direct method to detect a planet. It is a difficult technique because of the large contrast between the light emitted by the star and the light emitted/reflected from the planet, and the proximity of the two bodies on the sky. It is more sensitive for massive, young exoplanets located far from the star.

- **Gravitational Microlensing:** This technique is based on Einstein's General Relativity theory, where the light passing close to a massive object is bent by its gravitational field. The gravitational field acts like a lens that magnifies the light of a background star, and a planet around the lensing star can further induce a magnification that can be detectable. This technique is sensitive to planets with masses down to the Earth, located up to several AU from the star.

- **Pulsar Timing Variations:** The fast rotation and misaligned magnetic fields with respect to the rotation axis of the pulsars make them emit periodic pulses of synchrotron radiation. These pulse rates range from a few milliseconds to various seconds, depending on the speed of the star's rotation. Variations in the timing of the pulses indicate that the pulsar is moving back and forth, because of the gravitational pull of one or various planets.

- **Astrometry:** This technique detects the motion of a star by making precise measurements of its position on the sky. Similarly to the RV method, measuring the star's motion around the common barycenter of the star and the exoplanet allows determining the exoplanet's mass.

These detection methods combined to transits and RVs have allowed to retrieve many parameters as planetary mass, radius, orbital period, insolation, eccentricity, inclination, host

star parameters, etc. Studying these parameters in individual systems or in a statistical way is fundamental to understand the nature of exoplanets, what are they made of, and their formation and evolution histories.

1.2 Characterization of Exoplanet Interiors

The large number of discovered exoplanets over the last decades has provided a new perspective on planetary structure. They cover a very wide range physical conditions, from extremely puffy gas giants to small compact rocky planets with densities close to that of iron. Solar System planets as reference, one can naively divide the exoplanet population into three subgroups attending to their interiors:

- *Gas Giants:* Gas giants are essentially made of hydrogen and helium, with small contributions of heavier elements. Most of the known observed gas giants have similar masses and radius to Jupiter and Saturn, and probably also have a similar bulk structure. A schematic representation of the interior of Jupiter is shown in Figure 1.5. Jupiter's outer ~20% in radius is mostly made of molecular hydrogen, while the bulk of Jupiter consists of metallic hydrogen. High in the atmosphere at several kilo bars there is the radiative-convective boundary, which separates the outer thin layer (that is approximately of the same relative depth as the skin of a grapefruit) and a deep envelope with higher opacities in which heat is transported through convection. Gas giants may also have heavy elements cores, although this is one of the biggest uncertainties for many of them. Some known transiting gas giants have smaller radii than what H-He can produce, and therefore these planets must have a significant amount of heavy elements. However, for inflated planets, it is nearly impossible to know whether they have cores.
- *Neptune-like planets:* The so-called 'Neptune-like' planets still have a thick H-He envelope, but it does not account for most of the planet mass. Figure 1.5 shows how the interiors of these planets differs from those of the gas giants and terrestrial planets. Most of the structure models studying Uranus and Neptune find that most of the planet mass is in a deep fluid ionic sea consisting of water, ammonia and methane. In addition, their atmospheres are strongly enriched in metals. The carbon abundance of both Uranus and Neptune has been measured via spectroscopy (Fletcher et al. 2010), and they found a carbon enhancement of ~50x with respect to the Sun, implying that Neptune-like planets may have enriched atmospheres compared to their host stars. Whether these planets are primarily made of water and other fluid planetary ices, or mainly of rock and iron, is still an open question. If they were formed beyond the ice line, where the water has condensed, we would expect them to be true ice giants. However, if they are formed

within the ice line, they would probably possess rock and iron rich interiors with an accreted gaseous envelope.

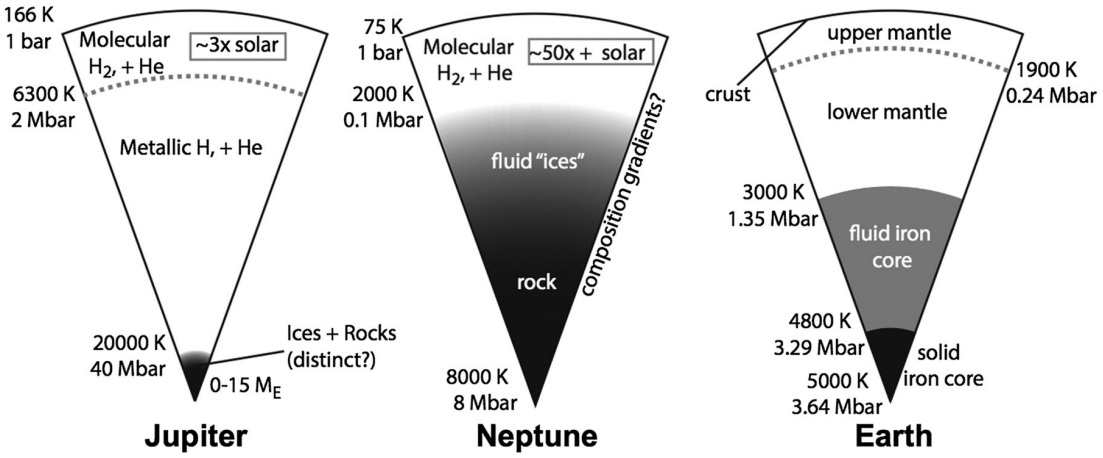

Figure 1.5: Schematic view of the internal structure and composition of gas giants, Neptune-like planets, and terrestrial planets. Figure taken from Spiegel et al. (2014).

- *Terrestrial Planets:* Terrestrial planets are objects as Mercury, Mars, Venus, and the Earth, and they are mainly composed of Si, Mg, Fe, and O. Figure 1.5 shows a representation of Earth's interior. The mantle is made of silicates, whose mineralogy differs between the upper and the lower parts. The core is mostly composed of iron and nickel, with small mixtures of other unknown lighter elements. Despite Venus and the Earth have nearly the same mass and radius, there are major differences between their interior structures. One of the most remarkable ones is the lack of plate tectonics in Venus (Basilevsky & Head 2002). Plate tectonics plays a crucial role on Earth, allowing the Earth's interior to cool and providing an exchange mechanism between the interior, the surface and the atmosphere. Whether this is necessary to form and evolve life on a terrestrial planet is under debate (Walker et al. 1981; Spiegel & Turner 2011).

Initially only the Solar System planets were known, which consist of four terrestrial (Mercury, Mars, Venus, and Earth), two Neptune-like (Uranus and Neptune), and two giant planets (Saturn and Jupiter). Nevertheless, the diversity of exoplanet interiors is much richer than that observed in the Solar System planets. Exoplanet observations have probed that more types and subtypes of planets exist. For example, ocean planets consist of intermediate planets between Super-Earths and Neptune-like, consisting of half-half water ice and rock. GJ-1214b (Cloutier et al. 2021) and the planets of TRAPPIST-1 (Agol et al. 2021) are some of the most likely known candidates for ocean planets. Magma planets, i.e. with a molten silicate mantle, have been

proposed as low-density terrestrial planets (Bower et al. 2019). Another relevant exoplanet type are the evaporated Neptunes, which are Neptune-size planets in which the outer layer has been stripped off due to the intense incoming radiation. The study of these cores provides valuable information to understand planet formation, since they possess the primordial composition before acquiring the envelope.

1.2.1 Interior Models

The study of planetary structures is based on the advances in both experiment and theory of materials at high pressure, better understanding of the Solar System planets thanks to exploratory missions, and data from a wide and diverse variety of exoplanets. In this subsection, I briefly review how planet models are built, with their basic equations and ingredients.

1.2.1.1 The Planetary Structure Equations

There are four equations describing the spherical symmetric static structure of a planet in hydrostatic equilibrium. These planetary structure equations are ordinary differential equations relating the radius, mass, pressure, temperature and luminosity. The equations are the following.

Hydrostatic equilibrium:

The planet is in hydrostatic equilibrium when the outward force due to the pressure-gradient is exactly balanced by the inward force due to gravity. Therefore, the pressure-gradient force prevents gravity the planetary layer from collapsing, whereas the gravity prevents the pressure-gradient from diffusing into outer space. The equation is written as follows:

$$\frac{\partial P}{\partial m} = -\frac{Gm}{r^2}\rho(P,T) \tag{1.13}$$

Mass continuity:

The change in mass of a spherical shell with density $\rho(P,T)$ is simply equal to the density multiplied to the change in volume:

$$\frac{\partial m}{\partial r} = \frac{\partial(\rho \cdot V)}{\partial r} = \rho(P,T) \cdot 4\pi r^2 \tag{1.14}$$

Energy conservation:

All contributions to the energy budget need to be considered in order to make sure that the energy is conserved in the planetary structure model. For a static planet model, the only

contribution to the change of luminosity comes from the release of energy during nuclear reactions, such as the radiogenic decay of unstable isotopes. The energy leaving a spherical shell can then be written as the following:

$$\frac{\partial L}{\partial r} = 4\pi r^2 \rho(P,T) \cdot \epsilon \tag{1.15}$$

where ϵ is the luminosity produced by the radiogenic heat per unit mass per second.

Energy transport:

In a planet, heat can be transported in three different ways. Thermal conduction is the transfer of internal energy through microscopic collisions of particles and movement within the planet, and it can take place in solid, liquid, and plasma phases. Convective heat transfer refers to the transfer of heat due to the movement of fluid, and it is usually the dominant heat transfer mechanism in liquids and gases. And, finally, thermal radiation is the electromagnetic radiation generated by the thermal motion. Each of these heat transfers results in a different temperature gradient, which can be written as:

$$\nabla(P,T) = \frac{d\log(T)}{d\log(P)} = \frac{P}{T}\frac{dT}{dP} \tag{1.16}$$

This equation is also commonly written by multiplying $\frac{\partial P}{\partial M}$ with $\frac{\partial T}{\partial P}$. Then we get the following expression:

$$\frac{\partial T}{\partial m} = -\frac{GmT}{4\pi r^4 P}\nabla(P,T) \tag{1.17}$$

1.2.1.2 The Planetary Layers

In most of the cases, what differentiates the planetary interior models used in the literature is not how they solve the structure equations, but the physical models they use to describe the interior. Equations of state (EOS) are fundamental to describe the properties of the planet interiors, since they relate thermodynamic variables which describe the state of matter under some specific physical conditions. These variables can be extensive (i.e. depend on the size of the sample, such as the entropy or the volume), or intensive (i.e. size independent, such as temperature or pressure). These thermodynamic quantities can be then used to compute state functions, which only depend on the current thermodynamic state and not on the path to reach that state. Some state functions are the internal energy, the enthalpy, Gibbs free energy, and the Helmholtz free energy.

The thermodynamic quantities and the state variables define the thermodynamic state of a system in equilibrium, and any equation relating any of these variables is a EOS. Typically, the EOS in planetary sciences are of the following form:

$$P = f(T, V, X) \tag{1.18}$$

where P is the pressure, T the temperature, V the volume, and X refers to the composition.

We will now revise how the different planetary layers have been typically modelled in the literature.

The Core: Exoplanetary cores are generally assumed to be mostly made of iron. Even if there is no way to test this hypothesis experimentally, there are several theoretical arguments to support it. First, it is a relatively abundant element in the protoplanetary disk. In addition, it is a refractory material, meaning that it condenses at ~1300K like silicon or magnesium (Lodders 2003; Lodders et al. 2009). Consequently, there will be condensed iron throughout all the protoplanetary disk that a growing planet can accrete, except in the most inner part of the disk. Numerical simulations by Ricard et al. (2009) showed that iron's high density and chemical properties make it a very efficient element to form core. They proved that this is also true for very small cores, which suggest that small iron cores may be created early in the planet formation history. Finally, Carter et al. (2015) performed collision simulations and found that the remnant planetesimals are iron-rich in most of the cases.

Nevertheless, the existence of terrestrial planets without iron core has also been predicted under some specific conditions Elkins-Tanton & Seager (2008). In order to constitute this core-less planets, iron must be either fully oxidized before accretion, or it must oxidize later in a well-mixed magma ocean.

Using seismic measurements, Dziewonski & Anderson (1981) probed that the Earth's core contains other elements apart from iron. These lighter elements sink to the core during core formation, diluting the iron with alloys. Recently, Ichikawa & Tsuchiya (2020) performed massive ab initio computations in order to evaluate what light elements can reproduce the seismic data, and identified nickel, silicon, sulfur, oxygen, and hydrogen as the most likely candidates.

Some very widely used EOS for the core are ANEOS (Thompson & Lauson 1972) and SESAME (Johnson & Lyon 1985). These EOS describe the thermodynamic properties of the iron core, rocks, and water, and consists of interpolations between models calibrated on existing

Hugoniot data. Other posterior models in the literature have usually used EOS obtained by fitting material properties measured or computed at pressures of the Earth (e.g. Valencia et al. 2007; Seager et al. 2007; Wagner et al. 2011).

The Mantle: The mantle is usually assumed to be composed of simple oxides of the main refractory elements: SiO_2, MgO, FeO, CaO, Al_2O_3, and NaO_2. Workman et al. (2005) showed that more than 90% of the Earth's mantle mass is made only of SiO_2, MgO, and FeO. Perple_X (Connolly 2009)[c] is a popular collection of Fortran programs designed to calculate and display phase diagrams, phase equilibrium by Gibbs energy minimization, and thermodynamic data. Combined with a thermodynamic model, it can be used to calculate stable minerals at a given temperature, pressure and composition.

A commonly used thermodynamic model for the mantle is the one presented by Stixrude & Lithgow-Bertelloni (2011). This thermodynamic model, as most of the internal structure models in the literature, is based in Earth like conditions and does not include liquid phases. There are several evidences in the literature predicting molten mantle layers. Umemoto & Hsu (2017) concluded that rocky planets more massive than $10M_\oplus$ can reach various TPa at the bottom of the mantle, which is much higher than in the Earth's mantle. On the other hand, according to Lopez & Fortney (2014) planets with equilibrium temperatures above 1000K and significant volatile layers reach temperatures high enough at the top of the mantle to be in liquid phase. Similarly, Vazan et al. (2018) investigated the long-term consequences of the core properties and concluded that most sub-Neptunes with volatile layers more massive than $0.02M_\oplus$ are expected to host a magma ocean.

Despite these evidences, Seager et al. (2007) studied the impact of the temperature of mantle material on its density, and found that the density variation is below 4% for a temperature range between 300K and 6000K.

The Water Layer: Water plays a major role in planetary models, due to its particular chemical properties and its abundance in the universe. It is a key component in the formation and evolution of planets, and it is believed to be a necessary ingredient for the emergence of life (Wiggins 2008). Water is expected to form thick ice layers or oceans in small planets (Sotin et al. 2007), and it is one of the most abundant components in the atmospheres of giants, after hydrogen and helium (van Dishoeck et al. 2014). Indeed, we know that water is present not only in the Earth in our Solar System, but also in Jupiter, Saturn, Uranus, Neptune, some of their moons, and in comets (Grasset et al. 2017). It is crucial, then, to have an accurate description of

[c] www.perplex.ethz.ch

the thermodynamic properties of water to properly model the interior structure of exoplanets.

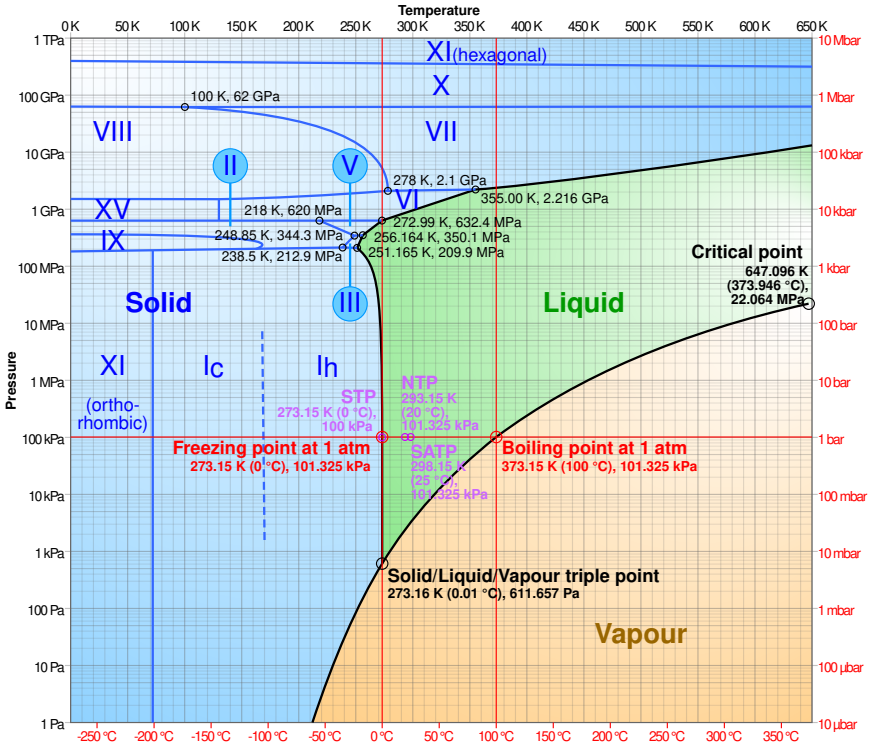

Figure 1.6: Phase diagram of water. The roman numbers indicate the various ice phases. Figure taken from Wikipedia.

Figure 1.6 shows the complex phase diagram of water. We see several triple points in which three different phases can coexist, and the roman numbers correspond to the different ice phases. The physical properties of water at ambient conditions are well understood. The International Association for the Properties of Water and Steam (IAPWS) provided an accurate description of the thermodynamic properties of water up to 1273K and 1GPa in their IAPWS-R6-95 release (Wagner & Pruß 2002), and for the I_h ice in the IAPWS-R10-06 release (Feistel & Wagner 2006).

Some frequently used EOS for water are the quotidian equations of state (QEOS) presented in Vazan et al. (2013). Since according to the Thomas-Fermi model stable molecules do not exist, they treat the H_2O molecules as mixtures of atoms in which the individual Wigner-Seitz cells were fixed by requiring their surface pressures to be equal. An advantage of the approach in Vazan et al. (2013) is that the grain opacity is calculated directly, without using extrapolations as most of the models in the literature before.

Recently, Bollengier et al. (2019) presented accurate water EOS up to 2.3GPa. At higher pressures, experimental data and ab initio calculations are needed. One example is the study

of Mazevet et al. (2019), who presented valid EOS at very high pressures and temperatures, including the ice XI seen in Figure 1.6. Recently a new set of water EOS was presented by Haldemann et al. (2020) combining several EOS in localized regions of the phase diagram. Combining different EOS in each regime of the parameter space allows them to cover a very large range of thermodynamic values.

The Atmosphere: The research on atmospheres is one of the most active fields in planetary sciences. In addition to the Earth, many other astronomical objects in the Solar System have atmospheres. The atmospheres of Jupiter and Saturn are almost entirely made of hydrogen and helium, with traces of methane, water vapor, ammonia, silicon-based compounds, etc. The atmospheres of Uranus and Neptune are rather similar, and are also mainly made of hydrogen and helium, with ~2% of methane. The atmospheres of the terrestrial planets are significantly different. Venus has a hot and dense atmosphere mostly composed of carbon dioxide with a small quantity of nitrogen, and Mars possesses a thin atmosphere of carbon dioxide with some traces of nitrogen and argon.

A common assumption when modeling giant exoplanet atmospheres is that they have similar elemental compositions to the host stars. These compositions are typically close to solar, with some enrichment of heavy elements (Fortney et al. 2008b). Therefore, models frequently use elemental abundance ratios inferred from the star, and then compute the equilibrium compositions at various metallicities (Fegley & Lodders 1994; Lodders & Fegley 2002a). Even if Jupiter and Saturn are known to be enriched with heavy elements by a factor ~2-5 (Atreya et al. 2003; Visscher & Fegley 2005), what processes caused this enrichment and whether they also operate to affect exoplanet atmospheres is still uncertain.

Atmospheric models rely on the study of hydrogen and helium at high pressures. From the point of view of fundamental physics, understanding the metalization of hydrogen has been a major challenge for a long time. One of the most successful equations of state was published by Saumon et al. (1995), and it has been used in numerous publications for planet's interior calculations. During the last decades, the improvements in computing capabilities have allowed a new generation of EOS calculated from ab initio simulations of the properties of hydrogen and helium (e.g. Militzer et al. 2006; Militzer 2008; Nettelmann et al. 2008, 2012; Hubbard 2014; Mazzola et al. 2018; Schöttler & Redmer 2018). Several studies in the literature have combined the EOS from Saumon et al. (1995) in the atomic and molecular domain, with the ab initio calculations in the ionized domain (e.g. Caillabet et al. 2011; Militzer & Hubbard 2013; Becker et al. 2014; Miguel et al. 2016).

In order to calculate the energy transport in the atmosphere, Jin et al. (2014) presented a model in which an irradiated atmosphere is assumed at the top of the gaseous envelope. This irradiated atmosphere is described by the irradiation model of Guillot (2010), which assumes a semi-gray, globally averaged temperature profile. The temperature gradient is then given by the following expression:

$$T^4 = \frac{3T_{int}^4}{4}\left[\frac{2}{3}+\tau\right] + \frac{3T_{eq}^4}{4}\left[\frac{2}{3}+\frac{2}{3\gamma}\left(1+\left(\frac{\gamma\tau}{2}-1\right)e^{\gamma\tau}\right)+\frac{2\gamma}{3}\left(1-\frac{\tau^2}{2}\right)E_2\gamma\tau\right] \tag{1.19}$$

where $\gamma = \kappa_v/\kappa_{th}$ is the ratio between visible and infrared opacity, T_{int} is the intrinsic temperature that characterizes the flux from the planet's interior, and T_{eq} is the equilibrium temperature of the planet. The criterion to distinguish between convective and radiative layers is usually assumed to be the usual Schwarzschild criterion.

1.2.1.3 Other used internal structure models

The discovery of the first extrasolar planet with a mass lower than $10M_\oplus$ by Rivera et al. (2005) triggered the emergence of the first internal structure models of super-Earths and ocean planets. In this subsection, I will go through some of the most relevant models describing the interiors of super-Earths and sub-Neptunes.

- **Valencia et al. (2007):** In this work, Valencia et al. (2007) introduced the first theoretical model to study the structure and properties of the first discovered super-Earth around GJ 876. With a mass of $\sim 7M_\oplus$, GJ 876d can be completely terrestrial or contain an extensive water layer of water and ices. A schematic representation of the interior model presented in Valencia et al. (2007) is shown in Figure 1.7. The terrestrial model is divided into the core, lower mantle, and upper mantle, and the composition of the layers is taken from the mineralogical composition known for the Earth (McDonough & Sun 1995). The solid inner core is made of iron with small amounts of nickel, and the liquid outer core is composed of iron and a light alloy that can contain S, Si, O, C, and H. The lower mantle is composed of perovskite and ferromagnesiowustite, while the upper mantle is made of olivine.

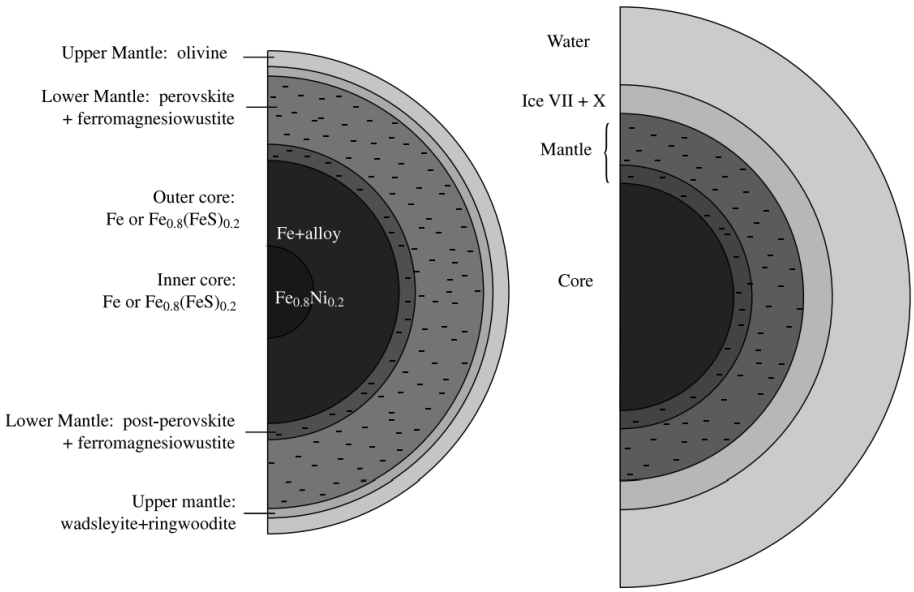

Figure 1.7: Schematic representation of the interior models in Valencia et al. (2007) assuming a similar composition to Earth (left) and an ocean planet (right).

Ocean planets are modeled, adding a water layer on top of the described rocky interior. If the water layer is massive enough, the water in the bottom can solidify in high-pressure phases of ice. The thickness of the layer is determined by the intersection of the pressure-temperature curve of the planet and the melting curve of ice.

- **Sotin et al. (2007):** Another example of structure model for water rich and Earth-like exoplanets was presented by Sotin et al. (2007), which has been widely used due to its simplicity (Grasset et al. 2009; Alibert et al. 2015; Auclair-Desrotour et al. 2019; Bolmont et al. 2020). Similarly to Valencia et al. (2007), they consider three main layers: an iron core, a rocky mantle and an outermost water layer. The core is assumed to be a liquid sphere of Fe-FeS with 87% of pure iron (following terrestrial observations). The mantle is divided into two layers, with similar compositions to those presented by Valencia et al. (2007), and the water layer is composed of a high pressure water ice and liquid water. We note that no vapor is considered.

- **Dorn et al. (2017):** The interior model presented by Dorn et al. (2017) were also relevant because of the inclusion of a H-He atmosphere with heavier elements. In this case, the core is assumed to be pure iron, neglecting other light elements like nickel. Then mantle is modeled with Na_2), CaO, FeO, MgO, Al_2O_3, and SiO_4, and its equilibrium mineralogy and composition are computed by Gibbs energy minimization following Connolly

(2009). The water layer is computed using the QEOS presented in Vazan et al. (2013). On top of this core, they model a gaseous envelope with an elemental composition of H, He, C, and O. An illustration of the interior model and the parametrization used is shown in Figure 1.8. The atmospheric model follows the code presented in Venturini et al. (2015), and the CEA (Chemical Equilibrium with Applications) package (Gordon et al. 1984) is used for the EOS, since more sophisticated EOS are not available for H, He, C, and O mixtures. The presence of an enriched H-He envelope makes this interior model suitable to model sub-Neptunes with relatively low densities.

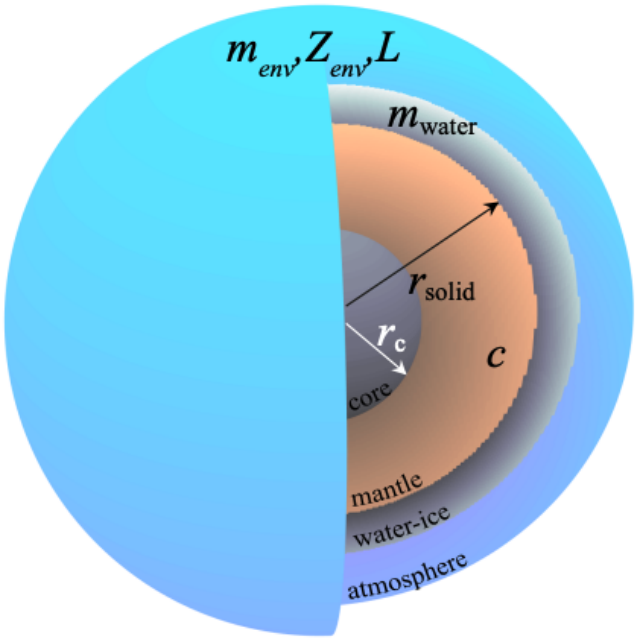

Figure 1.8: Illustration of the interior model parametrization in Dorn et al. (2017). The model parameters are the core radius r_c, mantle radius r_{solid}, mantle composition c, water mass m_{water}, envelope mass m_{env}, envelope luminosity L, and envelope metallicity Z_{env}. Figure taken from Dorn et al. (2017).

- **Brugger et al. (2017) & Mousis et al. (2020):** The interior model in Brugger et al. (2017) is based on the approach described in Sotin et al. (2007), and assumes a core mainly composed of iron with small fractions of nickel and sulfur, a silicate two-layer mantle, and a water layer with high pressure ice and liquid water on top. This model has been recently updated by Mousis et al. (2020) to include steam and a supercritical water layer for highly irradiated planets. This water layer is described using the EOS of Mazevet et al. (2019), and when the pressure and temperature are above the critical point of water, water is in supercritical phase. According to this model, the difference

between super-Earths and sub-Neptunes could simply result from the variations in the water content.

1.2.2 The problem of degeneracy

Inferring the structure and composition of exoplanets located many light years away results challenging. Radial velocity observations combined to transit observations provide the mass and radius of the exoplanet, which can be used as a first indication of an exoplanet's composition. However, there is an intrinsic degeneracy since different interior models can have the same mass and radius. In order to have valid conclusions about the interior of an exoplanet, then, it is necessary to take this degeneracy into account (Rogers & Seager 2010; Carter & Humi 2012; Schmitt et al. 2014; Weiss et al. 2016).

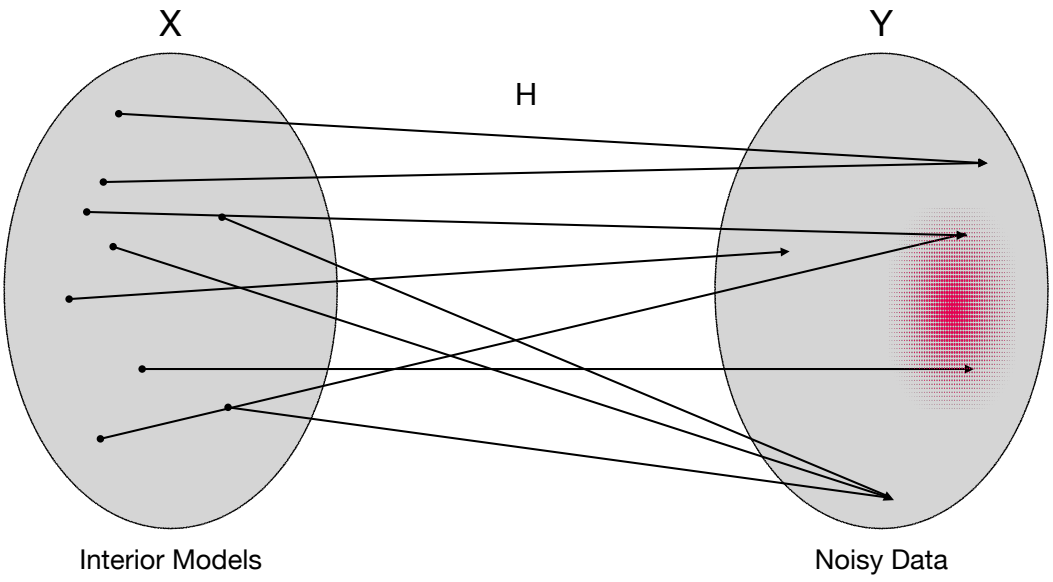

Figure 1.9: Schematic representation of a degenerate inverse problem with noisy data. The red region represents the observed data.

This kind of problems are called "inverse problems", and are common in astrophysics. They are addressed by building a physical model and comparing its predictions with the observations. Then the models that better reproduce the observations are determined. This problem can be formalized as a linear inverse problem:

$$Y = HX \bullet N \tag{1.20}$$

where Y corresponds to the set of noisy measurements, N ins the unknown error term, the $\bullet$ operator translates the way the noise contaminates the data, H is a linear operator, and X is the solution to the problem. Finding X is an inverse problem.

Figure 1.9 shows a schematic illustration of the degenerate inverse problem. The solution space X corresponds to the parameter space of the interiors from our physical model, and the space Y corresponds to the space of the observable parameters, which typically are mass and radius. The operator H generally represents the forward model, which computes the mass and radius of a planet given a set of interior parameters. The error is a stochastic-measurement noise, and the problem can only be solved in a probabilistic way.

1.2.2.1 Bayesian Inference

The most common way to address degenerate inverse problems is to use Bayesian inference. In the so-called 'frequentist' (or classical) methods the probabilities are interpreted as long run frequencies, so that the probability of an event is given by the occurrence rate observed during repeated experiments. In Bayesian inference, probabilities are interpreted as subjective degrees of belief, instead of frequencies. The unknown event is linked to a prior distribution that represents the subjective beliefs about the parameter. Then the posterior distribution of parameters is computed given the data using the Bayes theorem, which states that the probability for an event $\boldsymbol{x}$ given a set of data $\boldsymbol{d}$ is given by the following expression:

$$P(\boldsymbol{x}|\boldsymbol{d}) = \frac{P(\boldsymbol{x})\ P(\boldsymbol{d}|\boldsymbol{x})}{P(\boldsymbol{d})}, \tag{1.21}$$

where $\boldsymbol{x}$ stands for any parameter that may be affected by data. The prior probability $P(\boldsymbol{x})$ is the degree of belief that $\boldsymbol{x}$ is true before the data is observed. The 'Bayesian evidence' $\boldsymbol{d}$ is the new data that were not used when computing the prior probability, and the 'model evidence' P($\boldsymbol{d}$) corresponds to the probability that the observation $\boldsymbol{d}$ is true. The likelihood $P(\boldsymbol{d}|\boldsymbol{x})$ indicates the compatibility of the data with the given hypothesis, i.e. corresponds to the probability that the observation $\boldsymbol{d}$ is true given the model parameter $\boldsymbol{x}$. The likelihood is often given by the following expression:

$$P(\boldsymbol{d}|\boldsymbol{x}) = \frac{1}{(2\pi)^{N/2}(\prod_{i=1}^{N}\sigma_i^2)^{1/2}}\ exp(-\frac{1}{2}\sum_{i=1}^{N}\frac{(g_i(\boldsymbol{x})-\boldsymbol{d}_i)^2}{\sigma_i^2}), \tag{1.22}$$

where N is the number of data points, σ_i the uncertainties of the ith datum, and g($\boldsymbol{x}$) the operator linking the model parameters with the data, that is, $\boldsymbol{d} = g(\boldsymbol{x})$. Finally, the posterior distribution P($\boldsymbol{x}$ | $\boldsymbol{d}$) is what we want to know: the probability of $\boldsymbol{x}$ being true given the observation $\boldsymbol{d}$.

The Bayes theorem provides a formalism to solve the inverse problem. However, in most cases it is not possible to derive it analytically and a sampling method is needed. The most used sampling methods are Markov Chain Monte Carlo and Nested Sampling.

1.2.2.2 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) methods are very popular to solve sampling probability distributions in high-dimensional spaces. 'Monte Carlo' is a technique based on repeated random sampling to obtain numerical results. The underlying principle is to use randomness to solve problems that might be deterministic in principle. On the other hand, a 'Markov chain' is a stochastic model describing a sequence of possible events in which the probability of each event only depends on the previous event. It generates a sequence of random variables in which the current value is probabilistically dependent on the value of the prior variable. MCMC is essentially Monte Carlo integration using Markov chains. Monte Carlo integration draws samples from the required distribution, and then forms sample averages to approximate expectations. MCMC draws these samples by running a cleverly constructed Markov chain for a long time.

A widely used MCMC method to efficiently explore the posterior distribution is the so called Metropolis-Hastings (Metropolis et al. 1953; Hastings 1970). The basic concept can be explained with a simple case. First, an initial starting model is chosen randomly from the prior distribution. We note that the MCMC will always converge to the stationary distribution independently of the initial model. The posterior density of this model is calculated using Eqs. 1.21 and 1.22, and then a new model is created from a distribution centered in the current model. The chain has a probability to move to the new model that depends on the likelihood ratios between the current and new models. The method is repeated iteratively, and the generated samples are distributed, shaping the posterior distribution.

1.2.2.3 Nested Sampling

Nested sampling (Skilling 2004) is another Monte Carlo technique that aims to efficiently evaluate the Bayesian evidence and produce posterior probability distributions. The main strengths of Nested sampling resides on the small amount of problem-specific tuning required and its high efficiency. In parameter estimation problems, like MCMC, the Bayesian evidence $\boldsymbol{d}$ is ignored since it is independent of the parameters $\boldsymbol{x}$. In model selection methods, like Nested sampling, the calculation of the Bayesian evidence takes an important role. For clarity, we rename the likelihood $P(\boldsymbol{d}|\boldsymbol{x})$ as $\mathcal{L}(\boldsymbol{x})$. Since the Bayesian is just the normalization factor,

it is calculated with the following integral:

$$\boldsymbol{d} \equiv \mathcal{Z} = \int \mathcal{L}(\boldsymbol{x}) P(\boldsymbol{x})\, d^N \boldsymbol{x}, \tag{1.23}$$

where N is the dimensionality of the parameter space. Nested sampling use the relation between the likelihood and the prior volume to convert the multidimensional Bayesian evidence into a one-dimensional integral. The 'prior volume' X is defined by the following expression:

$$X(\lambda) = \int_{\mathcal{L}(\boldsymbol{x})>\lambda} P(\boldsymbol{x})\, d^N \boldsymbol{x}, \tag{1.24}$$

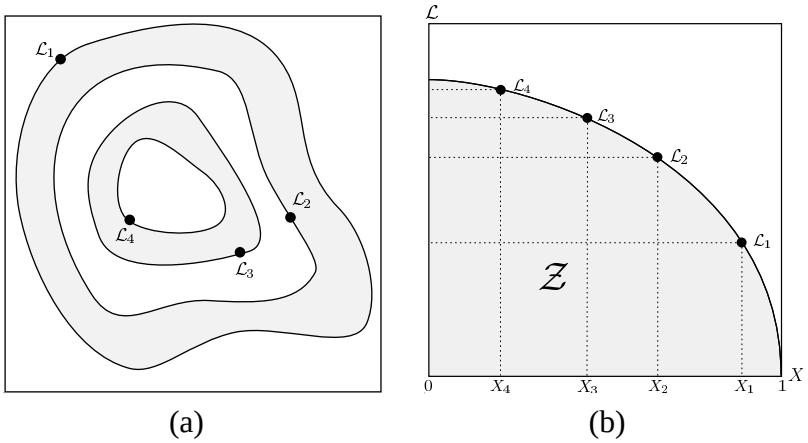

Figure 1.10: Illustration of the posterior of a two-dimensional problem (left), and the transformed $L(X)$ against the prior volumes X_i (right).

where λ is an inverse temperature that is slowly raised from 0 to 1, and the integral extends over the region of the parameter space inside the iso-likelihood contour $P(\boldsymbol{d}|\boldsymbol{x}) > \lambda$. The Bayesian evidence can be re-written as:

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X) dX \tag{1.25}$$

where $\mathcal{L}(X)$ is the inverse of Eq. 1.24. At each iteration, then, the prior volumes X_i shrinks and the likelihoods $\mathcal{L}_i = \mathcal{L}(X_i)$ increase, as shown in the illustration shown in Figure 1.10.

In practice, Nested sample algorithms sample a number of live points randomly from the prior distribution, and the likelihood is evaluated at each point. At each iteration, the point with the lowest likelihood will be replaced by a new point within the prior volume X_i. This process is continued until the Bayesian evidence reaches a specified value.

1.2.3 Link with planet formation and evolution

Determining an exoplanet's structure and composition is crucial to better understand how planets form. An exoplanet's interior structure is a direct result of its formation history, and more knowledge on its composition can lead to better constraints for formation theories. In fact, planet formation theories are continuously evolving due to the large amount of new discovered exoplanets and their properties. An example of how interior models can help to constrain the planet formation history is given in Plotnykov & Valencia (2021), where they compared the inferred derived composition of exoplanets to their host stars, found that the composition of rocky planets spans a wider range than stars, and estimated the maximum iron enrichment of rocky planets during formation.

Stars are formed in regions where the density of the interstellar medium increases. Eventually these dense large molecular clouds gravitationally collapse and flatten out as disks due to initial non-zero rotational inertia. These so-called protoplanetary disks were first observed with infrared space-based telescopes (Smith & Terrile 1984; Strom et al. 1989). Recently, the Atacama Large Millimeter/submillimeter Array (ALMA) has produced images of disks resolving sub-structures such as spiral arms, gaps and rings (Andrews et al. 2018; Andrews 2020).

Planet formation can be roughly divided into two major models: disk instability and core accretion model. In the disk instability model, the planets form from the breakup of the protoplanetary disk due to gravitational instability forming self gravitating clumps of gas, which later evolve into planets. This model requires that a very fast cooling of the protoplanetary disk, and it is still debated this is possible within the disk. On the other hand, in the core accretion model, the formation of a planet begins with a small protoplanetary embryo that accretes gas and solids. The two models are not exclusive. The disk instability model seems to describe better the formation of gas giants at large orbital distances, since the core accretion model requires long formation time scales. However, the core accretion model better reproduces the large diversity of exoplanets, so in this section I will focus on the description of this scenario.

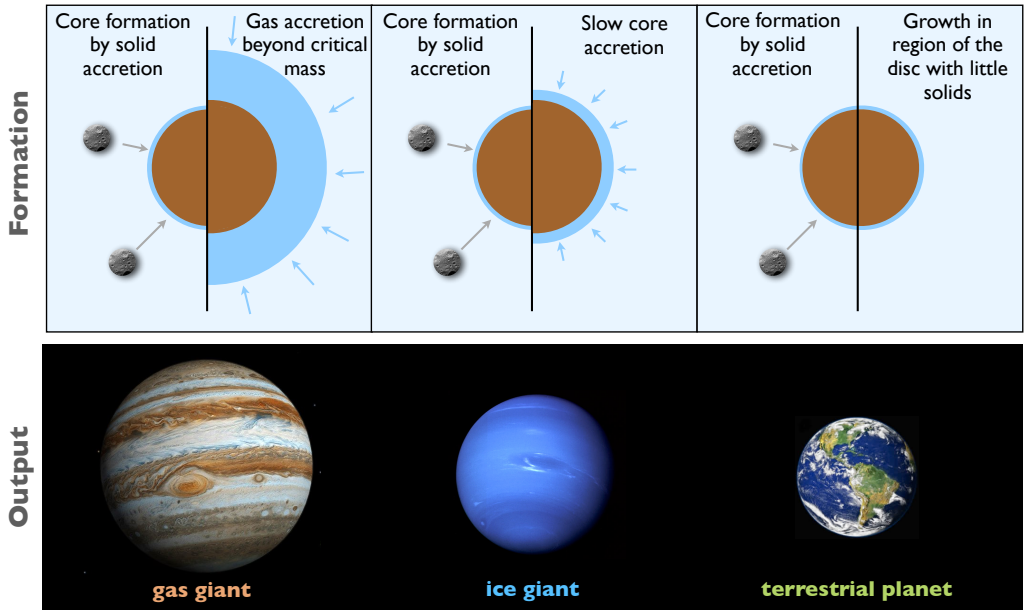

Figure 1.11: Schematic representation of core accretion model and its output in terms of Solar System planets. Illustration taken from Venturini et al. (2020).

At the beginning, the solid accretion rate is higher than the gas accretion rate, so the core grows faster than the volatile envelope. Later, when the mass of the atmosphere becomes comparable to the core mass, the runaway gas accretion starts. At this point, the gaseous envelope starts to be efficiently compressed by its self-gravity, allowing more gas to enter the protoplanet's gravitational sphere of influence and enhancing the envelope's self-gravity even more. The core mass at which this happens is called the critical core mass, and in classical models it is estimated to be between ~5 and 15 $M_\oplus$ (Ikoma et al. 2000).

Figure 1.11 shows a schematic representation of the core accretion model and its output in terms of Solar System planets. Gas giants like Jupiter are formed if there is a lot of gas in the disk when the core reaches the critical mass. This way, runaway gas accretion occurs, and the planets accretes hundreds of $M_\oplus$ of gas. Neptune-like planets are formed when the critical mass is reached and there is not much gas available in the disk. This case will result in a core-dominated planet with a significant H-He envelope. Finally, if the solid accretion timescales are much higher than the disk dispersal ones, or if the protoplanet is in a region where there are not many solids to accrete (e.g. close to the host star), it will remain an embryo during the whole disk's lifetime. After the dissipation of the disk, dynamical instabilities lead to collisions of embryos via giant impacts, and this is how the Earth is thought to be formed (Wetherill 1985, 1990).

Besides planet formation mechanisms, evolution mechanisms play a crucial role shaping the

current structure and composition. Planets inherit a large amount of internal energy from their gravitational contraction, and the progressive loss of this initial energy results in both a cooling and a contraction. The rate at which planets cool and contract is inversely proportional to the planetary age. Therefore, combined measurement of planetary mass, radius and a constraint on the stellar age allows a global determination of the amount of heavy elements present in the planet interior. Another relevant evolution mechanism is planet migration. Planets do not always remain in a fix position with respect to the star, they exchange angular momentum with the disc, leading to orbital migration. This migration usually is inwards (Tanaka et al. 2002), and it explains the presence of most close-in discovered exoplanets which cannot be created at such short orbital periods because the temperatures are too high to allow core accretion. In addition, planets with H-He atmospheres at short orbital distances are likely to suffer mass-loss due to photoevaporation. High-energy irradiation (UV and X-ray) heats up the upper layers of the atmosphere , driving a powerful hydrodynamic outflow that causes the atmosphere to lose mass over time (e.g. Lopez et al. 2012).

1.3 Main objectives of this thesis

This thesis focuses on the study of planetary interiors, combining theoretical models and observational data. The main goals I set four years ago were the following:

- Basing on the Bayesian inversion method of Dorn et al. (2017), to identify what constraints can be placed on the planetary composition and internal structure of super-Earths and sub-Neptunes, and to investigate how the observational quantities mass, radius, and their uncertainties affect the derived internal structure.
- Investigate the demographics and mass-radius relation of transiting planets with robust and reliable masses below $100 M_{\oplus}$, understanding the biases in the measurements and their effects on the data.
- Participate in the observation of exoplanets, focusing on the radial velocity follow-up of TESS transiting planets , and being part of various programs conducted by the Geneva team on CORALIE, HARPS, and ESPRESSO.
- Using theoretical models to identify interesting planetary candidates for future ground- or space-based observations, analyzing what targets and parameter space are be most appropriate for planetary internal structure characterization.

Besides the introduction, this thesis contains three chapters. Chapter 2 is dedicated to the study of the exoplanet demography. We present our work building a revisited exoplanet

catalog which is a reliable as possible, and perform in-depth demographic analyses of the sub-populations found in it. We also present our work investigating the intra-system similarity in multi-planetary systems. At the end of this chapter we discuss different observation programs for CORALIE, HARPS, and ESPRESSO in which I have been strongly involved.

Chapter 3 focuses on the characterization of planetary interiors of individual planets. I initially discuss the updates implemented in the models of Dorn et al. (2017). Then I present our work studying how several aspects as the observational uncertainties, location in the M-R diagram, or model assumptions affect the internal characterization of super-Earths and sub-Neptunes. Finally, I briefly present the contribution to some discovery papers of exoplanets characterized with my interior model and to a project investigating the formation of super-Mercuries by Giant Impacts. In Chapter 4 I summarize the achievements of this work and give an outlook of the possible developments and new questions that can be tackled in the future.

CHAPTER 2

DEMOGRAPHY OF EXOPLANETS

The ultimate goal of exoplanet demography is to determine the distribution of planets as a function of parameters that may affect their formation and evolution. These parameters include the planetary ones as the mass, radius, or orbital properties (semi-mayor axis, orbital period, or eccentricity), and the stellar parameters as the mass, radius, effective temperature, luminosity, metallicity, or age. Imprints left by the physical processes during planet formation and evolution should be reflected on the these planet functions, and therefore a proper understanding of the demography of exoplanets is necessary to set constraints on planet formation and evolution theories. The demographics can then be linked to the physical and chemical processes such as the planetary mass function, primordial atmosphere mass, migration, atmospheric loss, inflation mechanism, etc. providing constraints on the formation models. Therefore, by comparing these planet distributions to the predictions of planet formation theories, we can begin to both test and refine these theories.

During this period, the new-generations telescopes and satellites (particularly TESS) have triggered a rapid expansion of the number of well characterized exoplanets, allowing to perform in-depth analysis of the demography of the different exoplanet populations. In this chapter, we discuss our efforts to further understand exoplanet demography.

2.1 Populating the Mass-Radius diagram

One of the main difficulties when studying the exoplanet demography typically resides in the number of well-characterized planetary systems. In this section, I will show how different techniques and surveys have contributed to increasing the number of discovered exoplanets.

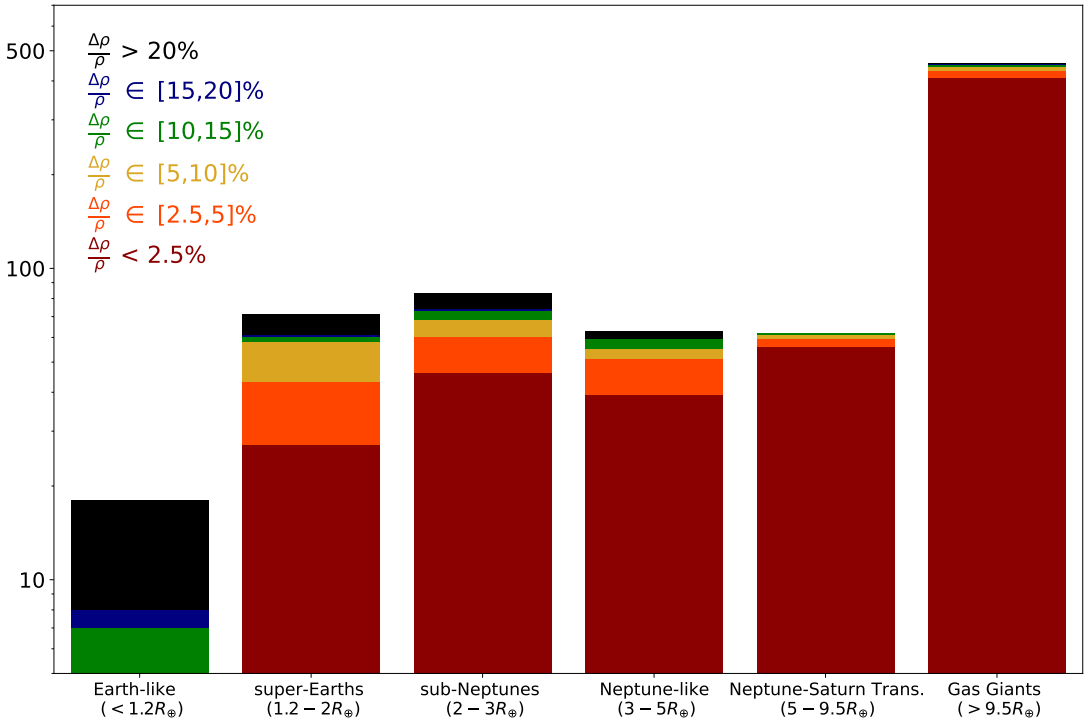

Figure 2.1: Histogram of the different types of discovered planets with measured mass and radius. Different colors indicate the amount of planets with different density uncertainties.

Measuring both the planetary mass and radius allows us to estimate the planetary bulk density and infer the possible compositions and internal structures in the first order. However, out of the more than 4800 exoplanets discovered in January 2022, only 754 have measured mass and radius. Figure 2.1 shows a histogram with discovered planets with mass and radius measurements for different types of planets, from the NASA Exoplanet Archive. More than half of the discovered planets with mass and radius measurements are gas giants, 457 precisely. There are 62 planets with radii between $5R_\oplus$ and 9.5R$_\oplus$, 63 Neptune-like with radii between $3R_\oplus$ and $5R_\oplus$, 83 sub-Neptunes, 71 super-Earths and only 18 Earth-like planets with radii below $1.2R_\oplus$. The colors of the plot indicate the density uncertainties. We can see that a big proportion of the gas giants and planets in the Neptune-Saturn transition that have very precise measured densities with uncertainties below 2.5%. We also note that among the rest

of the planets, most of them have mass and radius uncertainties below 5%, except Earth-like planets that have mostly uncertainties above 20% and are not secured planet detections. This means nowadays there is a relatively large sample of well-characterized planets with a precise estimation of the density that allows performing in-depth studies of the nature of almost all the planetary types.

The investigation of the mass-radius (M-R) relation of exoplanets of different populations can provide important information about exoplanet interiors and their formation and evolution processes. Seager et al. (2007) presented M-R relationships for a range of planet compositions. Other works use parametric laws, power laws in particular, to empirically fit the M-R relation based on the exoplanet data. Weiss & Marcy (2014), for instance, used a sample of 65 exoplanets smaller than $4R_\oplus$ to fit the M-R relation, and found a transition at $1.5R_\oplus$. They found that exoplanets with radii below this value tend to follow the Earth-like composition line, while exoplanet with larger radius tend to have volatile envelopes. Later, Wolfgang et al. (2015) presented the first probabilistic M-R relation for exoplanets smaller than $4R_\oplus$ evaluated in a Bayesian framework, which quantified the intrinsic dispersion and the uncertainties on the M-R relation parameters. Chen & Kipping (2017) also presented a probabilistic M-R relation fitting code, and used it with a sample of 316 exoplanets spanning from terrestrial planets to gas giants. They found a transition from a population of terrestrial planets to a population of less dense planets at $\sim 2M_\oplus$, and a transition between Neptunian to gas giants at $130 \pm 22\ M_\oplus$. Finally, Bashi et al. (2018) used a sample of 274 observed exoplanets to fit the M-R relation, and found a transition between the Neptunian and gas giants planets at a mass of $124 \pm 7\ M_\oplus$ and radius of $12.1 \pm 0.5\ M_\oplus$. Even if the transitions between Jovian and sub-Jovian planets obtained in different works seem to be compatible, the transition between terrestrial-like and Neptune-like planets is still a subject of intense debate.

2.1.1 Revisited Exoplanet Catalog and Mass-Radius relations

One of the most important choices to make when performing any demographic analysis resides on the planetary sample. There are several exoplanet catalogs available, and each one of them uses different information sources and has different criteria when adding information to the database. The "NASA Exoplanet Archive"[a], "Extrasolar Planets Encyclopaedia"[b], and "Open Exoplanet Catalog"[c] currently the most used catalogs by the community. The "Extrasolar Planets Encyclopaedia", for instance, has the largest coverage of planets, but it can include imprecise

[a] https://exoplanetarchive.ipac.caltech.edu/
[b] exoplanet.eu
[c] http://www.openexoplanetcatalogue.com/

estimates (Bashi et al. 2018). On the other hand, The "NASA Exoplanet Archive" is the most up-to-date catalog and the least biased catalog in terms of the interpretation of the mass upper limit, but in many cases, they do not select the most recent paper as a source of the planetary and stellar data. We also realized that some masses and radii present in these catalogs were not very reliable due to the quality or the quantity of the data that they use.

In Otegi et al. (2020a) (next page), we do a careful analysis of the planets with masses below $120M_{\oplus}$ to build an exoplanet catalog that is as robust and reliable as possible. We then use this dataset to obtain a new M-R relation and perform a demographic analysis of the resulting M-R diagram.

Astronomy & Astrophysics manuscript no. output
November 15, 2019

Revisited mass-radius relations for exoplanets below $120 M_\oplus$

J.F. Otegi[1,2], F. Bouchy[2], and R. Helled[1]

[1] Institute for Computational Science, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland
e-mail: jonfr17@gmail.com
[2] Observatoire Astronomique de l'Université de Genève, 51 Ch. des Maillettes, – Sauverny – 1290 Versoix, Switzerland

November 15, 2019

ABSTRACT

The masses and radii of exoplanets are fundamental quantities needed for their characterisation. Studying the different populations of exoplanets is important for understanding the demographics of the different planetary types, which can then be linked to planetary formation and evolution. We present an updated exoplanet catalogueue based on reliable, robust, and, as much as possible accurate mass and radius measurements of transiting planets up to 120 $M_\oplus$. The resulting mass-radius (M-R) diagram shows two distinct populations, corresponding to rocky and volatile-rich exoplanets which overlap in both mass and radius. The rocky exoplanet population shows a relatively small density variability and ends at mass of $\sim 25 M_\oplus$, possibly indicating the maximum core mass that can be formed. We use the composition line of pure water to separate the two populations, and infer two new empirical M-R relations based on this data: $M = (0.9 \pm 0.06)\ R^{(3.45\pm0.12)}$ for the rocky population, and $M = (1.74 \pm 0.38)\ R^{(1.58\pm0.10)}$ for the volatile-rich population. While our results for the two regimes are in agreement with previous studies, the new M-R relations better match the population in the transition region from rocky to volatile-rich exoplanets, which correspond to a mass range of 5-25 $M_\oplus$, and a radius range of 2-3 $R_\oplus$.

1. Introduction

To date, more than 4000 exoplanets have been discovered. The Kepler mission has clearly impacted the field with the detection of more than 2300 exoplanets. For many of the Kepler exoplanets, radial velocity follow-up is restricted to a small fraction corresponding to the brightest host stars. As a result, in order to characterise the exoplanets researchers often rely on a theoretical mass-radius (hereafter M-R) relation. Knowledge of both the planetary mass and radius allows us to estimate the planetary bulk density and infer the possible compositions and internal structures. In addition, the M-R relation is used to explore the demographic of exoplanets in a statistical sense. These demographics can then be linked to the physical and chemical processes driving planet formation and evolution, such as the planetary mass function, primordial atmosphere mass, migration, atmospheric loss, inflation mechanism, etc. providing constraints on the formation models.

Various studies have been dedicated to the investigation of the internal structures of exoplanets (e.g. Sotin et al. 2007; Seager et al. 2007a; Howe et al. 2014; Dorn et al. 2017; Lozovsky et al. 2018) and to the investigation of the M-R relation of exoplanets of different populations. Parametric models, power laws in particular, have been proposed to fit the M-R relation. These are typically empirical relations based on exoplanet data found in the main exoplanet catalogues: Extrasolar Planets Encyclopaedia[1] (Weiss et al. 2013; Weiss & Marcy 2014; Bashi et al. 2017) or NASA Exoplanet Data Archive[2] (Wolfgang et al. 2015). As shown in Bashi et al. (2018), despite the overall good agreement between these catalogues, there are also some differences. On the other hand, the Extrasolar Planets Encyclopaedia has the largest coverage of exoplanets, probably due to its less restrictive selection criteria. On the other, the NASA Exoplanet Database has a 'removed targets' list, providing a more rigorous selection process, and is the most updated catalogue. Recently, Zeng et al. (2016) inferred a semi-empirical M-R relation depending on the core mass fraction, followed by a detailed forecasting model using a probabilistic M-R relation using Markov Chain Monte Carlo (MCMC). (Chen & Kipping 2017).

While such studies are crucial for a more detailed characterisation of exoplanets, it should be noted that the number of discovered exoplanets increases rapidly and the estimates for the masses and radii are continuously being updated. In addition, one has to account for the fact that some of the listed mass and/or radius determinations are not reliable, which can affect the inferred conclusions (see Section 2 for details). In this study, we go through the entire NASA Exoplanet catalogue, in order to create a 'filtered' sample of exoplanets with robust and reliable mass and radius measurements. We consider exoplanets with masses below $120 M_\oplus$ in order to focus on the transition between small-size terrestrial planets and the population of giant gaseous planets. We use our updated catalogue to describe the properties of two distinct populations: rocky and volatile-rich population. We derive updated M-R relations for these two populations and investigate the dependence of the M-R diagram with other external parameters.

[1] exoplanet.eu
[2] exoplanetarchive.ipac.caltech.edu

2. Exoplanet selection with reliable measurements of mass and radius

We use the NASA Exoplanet Archive from June 2019 as a starting point since it is the most up-to-date catalogue, and, in addition, it provides access to all the references relevant for a given observed exoplanet. We build a "reliable and updated" catalogue by applying the following selection criteria:

a) We selected the data from the NASA Exoplanet Archive from July 2019 for planets with masses up to $120M_\oplus$ and filtered the data to consider only exoplanets with measurement uncertainties smaller than $\sigma_M/M = 25\%$ $\sigma_R/R = 8\%$. These thresholds correspond to the median uncertainty and make it possible to have the same impact on the density uncertainty.
b) We added the mass measurements of the exoplanets orbiting around Trappist-1 from Grimm et al. (2018), who used new K2 transit light curves to recompute the masses through TTVs and shrank the mass uncertainties from $30\% - 95\%$ to $5\% - 12\%$.
c) We discarded the mass determinations inferred by Stassun et al. (2017), where the host star masses and radii were replaced by the value derived from GAIA photometry and with uncertainties clearly overestimated. These revised values affect the planetary mass estimation. Therefore, in the cases in where the NASA Exoplanet Archive selects this study as the reference paper, we replace them with the most updated mass estimate (Kepler-78b,Kepler-93b, CoRoT-7b, Kepler-454b, HD 97658b, HIP 116454b, WASP-29b, WASP-69b, WASP-117b, HD 149026b, WASP-63b) [3].
d) In some cases Marcy et al. (2014) give an estimate of the planets masses for single transiting planet with a weak level of validation/confirmation. Several mass estimates are based on very few radial velocity data points with underestimated uncertainties. We therefore did not use mass estimates given by Marcy et al. (2014) for exoplanets with non-robust measurements (Kepler-406c, -97b, -98b, -102b, -48b, -99b, -406b, -100b, -48b, -96b, -102e, -25b, 103b, -106c, -106e, -113b, -103c). These exoplanets do not have any other mass estimates from other studies with measurement uncertainty smaller than $\sigma_M/M = 25\%$ $\sigma_R/R = 8\%$, and therefore were not included in our catalogue.
e) The TTV measurements reported by Xie (2014) differ significantly from the mass measurements reported by other groups (Hadden & Lithwick 2014, 2017). In addition, several of their mass and radius estimates imply that several exoplanets with masses greater than $30M_\oplus$ are denser than pure-iron (e.g., Kepler-128b and Kepler-128c, for which Hadden & Lithwick (2017) estimated masses below $1M_\oplus$). Therefore we also excluded the mass estimates provided by this study.
f) Hadden & Lithwick (2017) provide the planetary masses through TTVs for 150 Kepler exoplanets, which are not used in the NASA Exoplanet Archive. In addition, they introduce a robustness criterion for TTVs, and consider that only 50 out of 150 mass measurements are reliable. We relied on their robustness criterion and discarded the unreliable TTVs mass determination.
g) We updated some mass measurements to the ones presented in more recent publications (Kepler-10b, Kepler-65d, GJ 9827b, 55-Cnc e, K2-55b K2-261b, HAT-P-18b, HAT-P-12b, WASP-20b)[3] and we included several exoplanets that are missing from the NASA Exoplanet Archive (Gl-357b, HD 39091c, HD 3167b, K2-131b, HD 15337c, HD 213885b, EPIC 220674823b, HD 3167c, K2-180b, K2-24c, GJ 143b, HD 21749b, WASP-166b, WASP-107b, HAT-P-48b, HAT-P-47b, Kepler-425b, NGTS-5b, HATS-43b, WASP-160b, Kepler-427b, WASP-181b, K2-295b, EPIC 220501947b, Kepler-426b, Qatar-8b)[3].

Figure 1 shows the M-R diagram after applying the aforementioned selection criteria. It also shows the mass and radius histograms of the exoplanets in our sample. Intermediate steps of the data selection are shown in the appendix in Figure A.1. The planets that are included in our "filtered" catalogue are listed in the appendix (Table A.1), where we also provide the references used by the NASA Exoplanet archive and the ones used in this work. It should be noted that as other catalogues, also ours suffers from observational biases and is incomplete. As a result, it cannot be used to make conclusions about the planetary occurrence rates. Our revised catalogue of transiting planets below $120M_\oplus$ is accessible on the Data & Analysis Center for Exoplanet DACE[4].

3. Analysis of the revisited M-R diagram

3.1. Two distinct exoplanet populations

Our revisited M-R diagram clearly shows two distinct exoplanet populations: one of them closely follows an Earth-like composition, and a second one corresponds to a more volatile-rich composition. It is important to note that even when an exoplanet lies in the M-R diagram of an Earth-like composition, its actual relative composition of iron, silicates and water could be different, given the degeneracy of the problem.

Nevertheless, the amount of water or H/He envelope of these exoplanets is expected to be small in comparison to the refractory materials (e.g. silicates, iron). Therefore, it is reasonable to assume that these exoplanets are mostly rocky[5]. This population follows the Earth-like composition up to a mass of $\sim 25M_\oplus$ including Kepler-411b. However, between 10 and 25 $M_\oplus$, exoplanets appear to be slightly less dense than the ones following the Earth-like composition. These objects might be ice-rich cores, but are unlikely to be volatile-rich. Therefore we include them in the rocky population, which contains 'naked-cores' up 10 $M_\oplus$ and slightly more ice-rich exoplanets from 10 to 25 $M_\oplus$. This suggests that this upper limit corresponds to the maximum core mass that can be formed, and it is important to note that this region of the diagram does not suffer from observational biases since heavier planets are easy to detect in radial velocity. This estimate of the maximum core mass is in fact consistent with theoretical calculations of giant planet formation, and with the estimated core masses of the giant planets in the solar system: for Jupiter structure models typically infer core masses between $7M_\oplus$ and $25M_\oplus$ (e.g. Guillot et al. 2017; Wahl et al. 2017; Helled et al. 2017), and Saturn's core mass is expected to be of the order of $20M_\oplus$ (e.g. Saumon & Guillot 2004; Iess et al. 2019).

In addition, giant planet formation models with pebble accretion estimate the pebble isolation mass to be between $10M_\oplus$ and $20M_\oplus$ (e.g. Johansen & Lambrechts 2017; Bitsch et al. 2019), which is also consistent with our estimated maximum core mass of the order of $25M_\oplus$. This result suggests that our M-R relation can be used to confirm and test theoretical predictions.

It is interesting to note that none of the exoplanets in our sample is found to be consistent with a pure iron composition. A structure model of highest density planet Kepler-107c ($\bar{\rho} = 12.65\ g\,cm^{-3}$) suggests that it has a large iron core and a

[3] References used for these planets are listed in Table A.1

[4] dace.unige.ch

[5] in this study 'rocky exoplanets' refer to exoplanets that mostly consist of metals and rocks

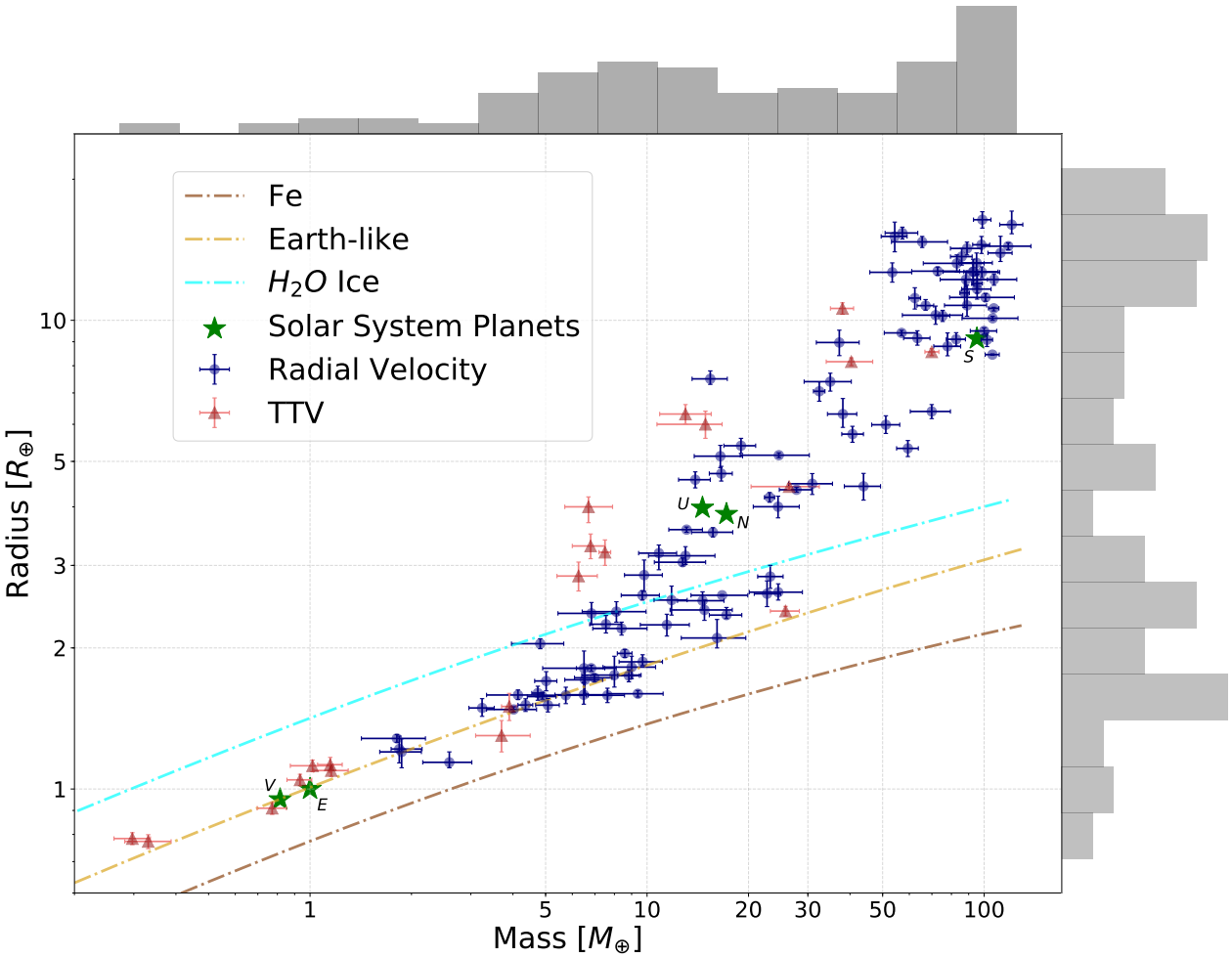

Fig. 1. Revisited M-R diagram after applying our criteria to keep reliable and robust mass measurements with relative uncertainties smaller than 25% for mass and smaller than 8% for radius. The red triangles and blue circles correspond to data with mass determination from TTVs and RVs, respectively. We also display the composition lines of pure-iron (brown), Earth-like planets (light-brown) and water ice (blue) (Dorn et al. 2015). We also plot the contour lines and the distribution of exoplanet mass (top) and radius (right) of our sample.

silicate mantle, corresponding to 70% and 30 % of the planetary mass, respectively (e.g. Bonomo et al. 2019).

The second population shown in our revisited M-R diagram corresponds to less dense planets with a more volatile-rich composition. The density-radius diagram displayed on the right panel of Figure 2 makes it possible to distinguish the two populations separated by the composition line of pure water (see section 3.2 for details). The rocky population presents a nearly flat density up to 2-3 $R_\oplus$, corresponding to behaviour of exoplanets made of refractory materials. The volatile-rich population shows a decreasing density from 2-3 $R_\oplus$ to 12 $R_\oplus$.

Although the M-R is biased toward lower masses, it seems that $5M_\oplus$, like HD39091, is the lower limit from which an exoplanet can accrete and maintain a gaseous envelope. However, it is possible that the data sample is incomplete and suffers from observational biases, since for a given radius it is much easier to detect more massive planets.

The dispersion of the volatile-rich population is significantly larger than the one in the rocky population. It may reflect different core masses accreting gas. Another reason could be that volatile rich exoplanets are very sensitive to insolation, and therefore the dispersion reflects exoplanets with different stellar irradiations (this hypothesis is further discussed in Section 3.2).

Fulton et al. (2017) find that at small radial distances there is a lack of planets with radii between $1.5R_\oplus$ and $2R_\oplus$, known as the Fulton Gap, suggesting a transition between the super-Earth and sub-Neptune populations. Nevertheless, as discussed above, our revisited M-R diagram shows two exoplanet populations with a large overlap in both mass and radius. The overlap in mass ranges between 5 and 25 $M_\oplus$, and in radius from 2 to 3 $R_\oplus$. Although planets smaller than 1.8 $R_\oplus$ are clearly part of the rocky population, planets with a larger radius could belong to both populations. Therefore the planetary mass or radius alone cannot be used to distinguish between the two populations. We therefore use the composition line of pure water to separate the rocky and volatile-rich populations. This provides a more physical criterion to divide the populations as planets that sit above the pure-water line are expected to consist of volatile materials (e.g. H-He).

3.2. The Mass-Radius relations

Published M-R relations to date have divided the different exoplanet populations using mass cutoff (e.g. Chen & Kipping 2017) or radius cutoff (e.g. Weiss & Marcy 2014). In this study, we

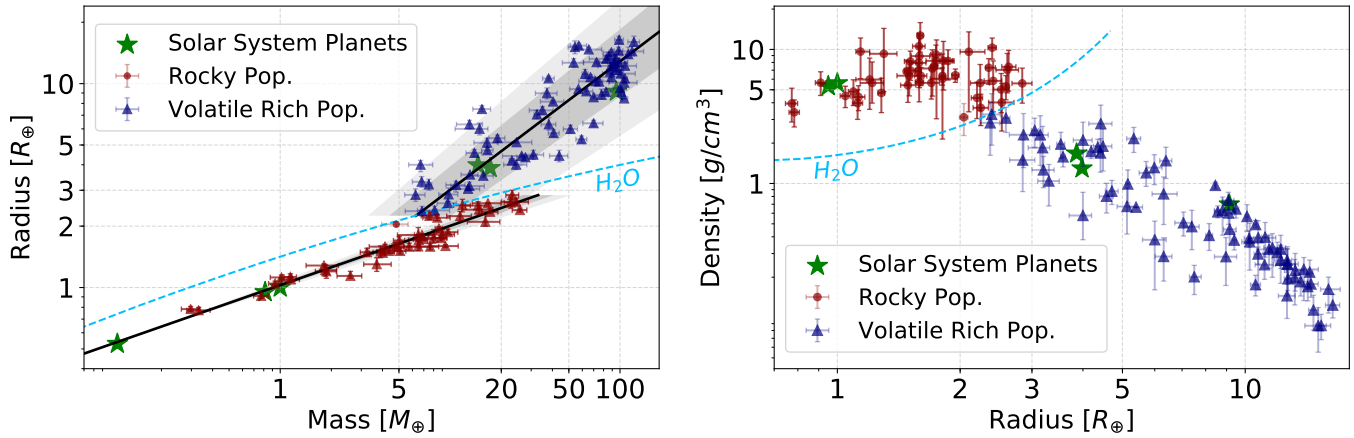

Fig. 2. Left: M-R relations fitting rocky and volatile-rich populations. Dotted line corresponds the composition line of pure water using QEOS for a temperature of 300K (More et al. 1988). The grey and light-grey envelopes represent the $\pm1\sigma$ and $\pm2\sigma$ regions of the fit. Right: Density against radius for our catalogue. Rocky and the volatile-rich populations are separated by the composition line of pure water (Dorn et al. 2015). The grey envelope indicates the region between 2.8 $g\ cm^{-3}$ and 3.3 $g\ cm^{-3}$.

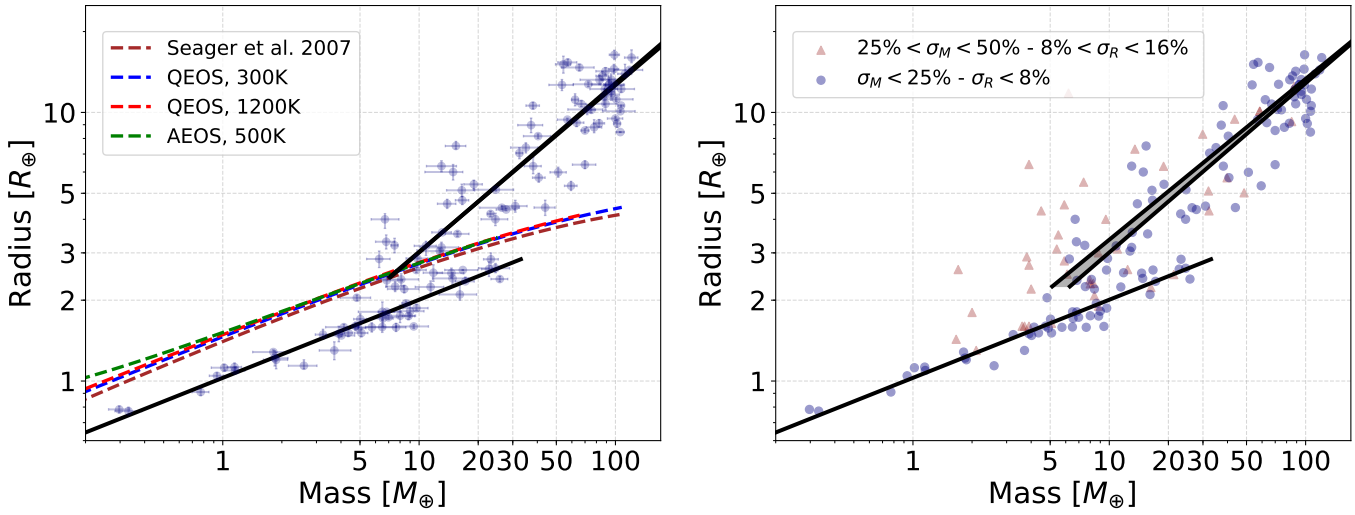

Fig. 3. Right: M-R diagram comparing obtained M-R relations when using different EOS for water: the polytropic EOS of Seager et al. (2007b) (brown), QEOS assuming temperatures of 300K and 1200K (More et al. 1988), and ANEOS (Thompson 1990) with a surface temperature of 500K. Right: M-R diagram comparing obtained M-R relations when using different cuts for the mass and radius uncertainties when building the revisited catalogue: $\sigma_M/M = 25\%$ and $\sigma_R/R = 8\%$, and $\sigma_M/M = 50\%$ and $\sigma_R/R = 16\%$. The grey envelope corresponds to the difference between the two. Blue circles represent the exoplanets with $\sigma_M/M = 25\%$ and $\sigma_R/R = 8\%$, and red triangles represent the planets with uncertainties between $\sigma_M/M = 25\%$ and $\sigma_R/R = 8\%$, and $\sigma_M/M = 50\%$ and $\sigma_R/R = 16\%$.

divide the super-Earth and volatile-rich populations using the composition line of pure-water. We fit the M-R relation of the two populations using a total least squares method, in which observational errors on both dependent and independent variables are considered. Figure 2 (left) shows the inferred M-R relations for the two exoplanet populations, assuming an M-R dependence of $R = AM^B$. The results of the fit are shown in Equations 1 and 2:

$$R = \begin{cases} (1.03 \pm 0.02)\ M^{(0.29\pm0.01)} & ,\ if\ \rho > 3.3\ g\ cm^{-3} \\ (0.70 \pm 0.11)\ M^{(0.63\pm0.04)} & ,\ if\ \rho < 3.3\ g\ cm^{-3} \end{cases} \quad (1)$$

or

$$M = \begin{cases} (0.90 \pm 0.06)\ R^{(3.45\pm0.12)} & ,\ if\ \rho > 3.3\ g\ cm^{-3} \\ (1.74 \pm 0.38)\ R^{(1.58\pm0.10)} & ,\ if\ \rho < 3.3\ g\ cm^{-3} \end{cases} \quad (2)$$

The rocky exoplanet population presents a relatively small dispersion around the M-R relation reflected by the small uncertainties on the fitted parameters. On the other hand, the volatile-rich population presents a larger dispersion around the adjusted relation reflecting a larger diversity in composition. Contrary to previous studies, all the observed masses and radii are found to be at less than 2σ of our M-R relations. In addition, the Solar System planets with masses smaller than $120M_\oplus$ lie on the derived M-R relations (except Mercury, which is anomalously dense).

Figure 2 (right) shows the density-radius diagram of our planetary catalogue, and a density-mass diagram is included in the appendix. Another physically-motivated approach is to divide the rocky and volatile-rich populations by using a density-cutoff. This density cutoff should be between the

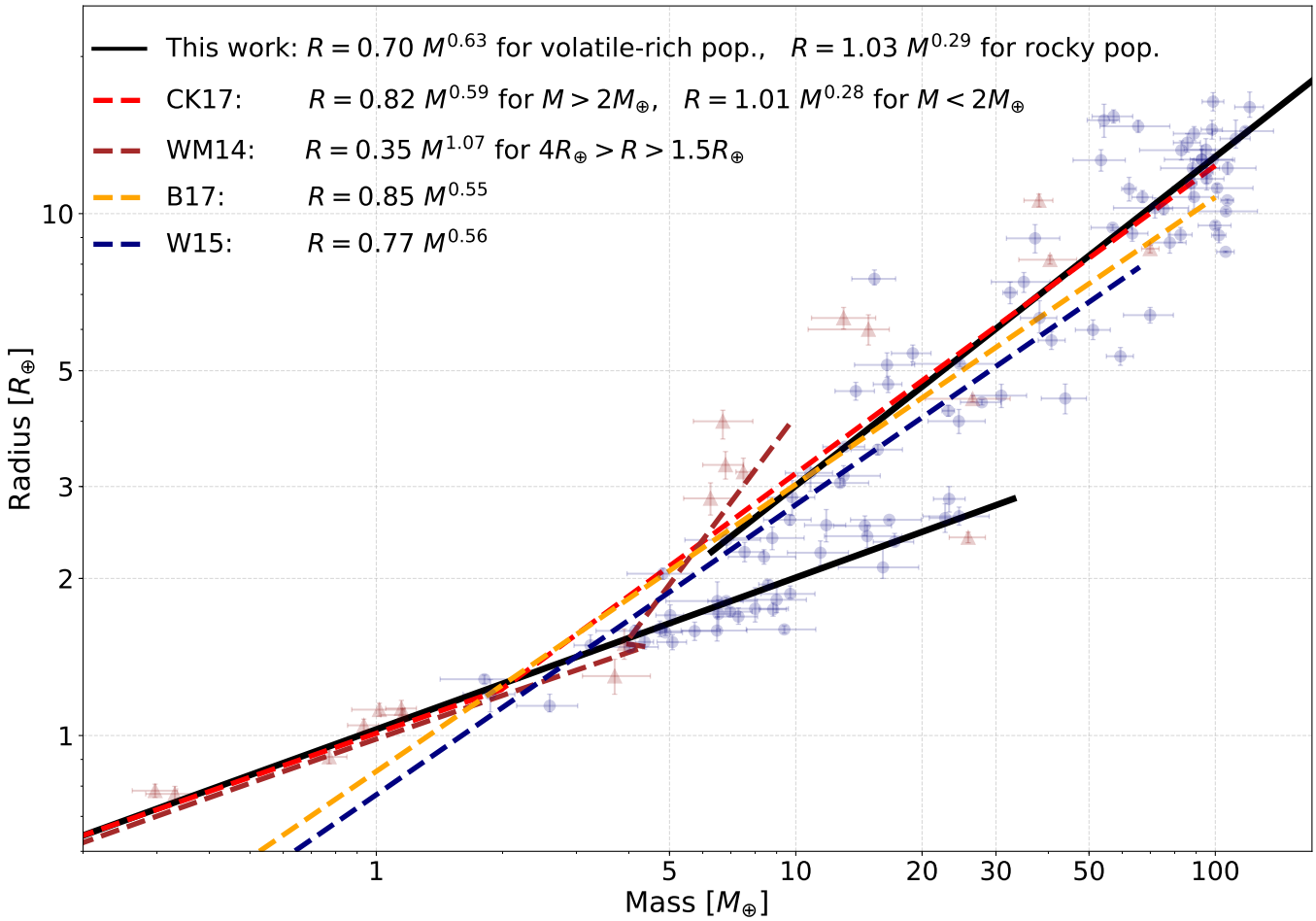

Fig. 4. Comparison of M-R relations in the literature with the one obtained from our revisited catalogue. Red triangles and blue circles correspond to our revised catalogue with mass determination from TTVs and RVs, respectively. The analytic expressions of the M-R relations CK17, WM14, B17 and W16 corresponds to Chen & Kipping (2017), Weiss & Marcy (2014), Bashi et al. (2017) and Wolfgang et al. (2015), respectively.

minimum density of a rocky planet and the maximum density of a volatile-rich planet. The lowest density exoplanets that are expected to be rocky in our sample are the Trappist planets, with $M_{solid}/M_{total} > 84\%$ and $R_{gas}/R_{total} < 2\%$ (e.g. Dorn et al. 2018). This sets a maximum value for the density-cutoff of 3.3 $g\,cm^{-3}$. On the other side, K2-55b is the densest planet in the volatile-rich population and has an estimated envelope mass fraction of 12% (e.g. Dressing et al. 2018). It then sets a minimum limit for the density-cutoff of 2.8 $g\,cm^{-3}$. The grey envelope in Figure 2 (right) shows the region between 2.8 $g\,cm^{-3}$ and 3.3 $g\,cm^{-3}$. Using a physically-motivated density cutoff to divide the two populations, or dividing the two populations by the composition line of pure water, lead to nearly identical results. This suggests that the physically-motivated approach of dividing the two populations using the pure-water curve is essentially equivalent to the division of the populations by a density cutoff of $\sim 3\ g\,cm^{-3}$. Nevertheless, the pure-water composition curve is less arbitrary and is based on physical arguments.

The composition line of pure-water depends on the EOS (Equation of State) used and the planetary temperature. Figure 3 (left) shows how the derived M-R relations depend on these choices, comparing the polytropic EOS of Seager et al. (2007b), QEOS, assuming surfaces temperatures of 300K and 1200K (More et al. 1988), and ANEOS, for a temperature of 500K. We find that the results are insensitive to the water EOS used and the assumed temperature. It is also important to note that the pure-water composition line also depends on the assumed pressure, which here was assumed to be one bar. This assumed pressure corresponds to a water world without a water vapor atmosphere, and, therefore, the M-R relation represents a lower bound for the radii of pure water planets. Nevertheless, we find that the coefficients of the M-R relations are relatively insensitive to the assumed surface pressure.

In order to investigate the impact of mass and radius uncertainties to our M-R relations, we extended our catalogue to planets with uncertainties two times larger in both mass (50%) and radius (16%). The right panel of Figure 3 shows a comparison of the M-R relations obtained using cuts of 25% and 50% for the mass uncertainty, and 8% and 16% for the radius uncertainty. The change of uncertainty has no significant impact on the M-R relations for both the rocky and volatile-rich populations and the adjusted parameters are within error bars. We can therefore conclude that our fit does not depend on the limit in mass and radius uncertainties and is not significantly biased by our selection criteria. For the rocky population, the M-R relation indicates that the bulk density is nearly constant. For volatile-rich populations, the bulk density scale with M^{-1}.

Figure 4 compares the M-R relations in the literature with the one obtained from our revisited catalogue. Our M-R

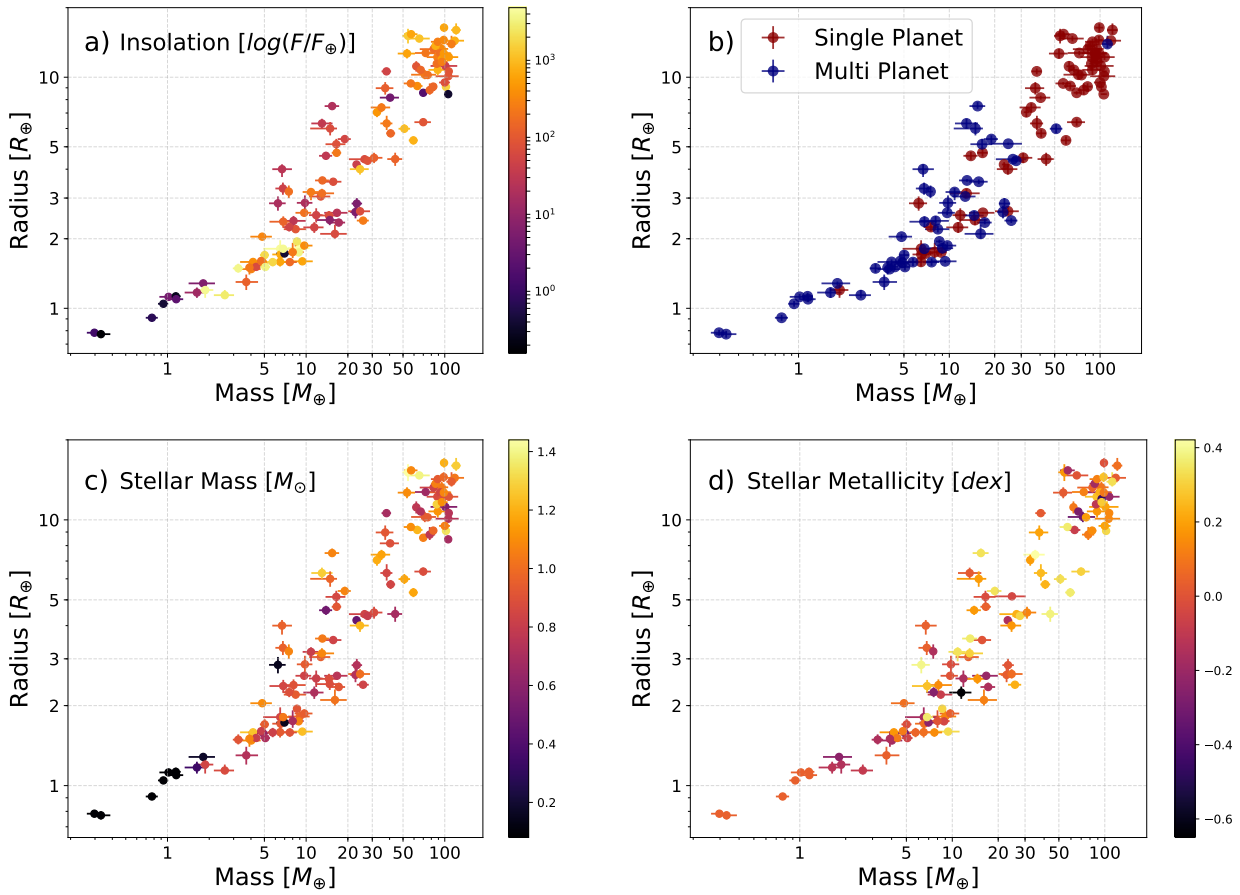

Fig. 5. M-R diagrams from our revisited catalogue showing dependence with insolation (a), multiplicity (b), stellar mass (c) and stellar metallicity (d).

relation is similar to the one inferred by Chen & Kipping (2017), but the transition from the rocky to the volatile-rich regime is defined for a mass of $2M_\oplus$, so they underestimate the masses of most of the rocky exoplanet population. Our inferred transition mass between rocky and volatile-rich planets is found to cover a large range of masses ($10 - 25M_\oplus$) and is significantly higher than the $2M_\oplus$ predicted by Chen & Kipping (2017). Our relation is also very similar to the one derived by Weiss & Marcy (2014) for planets smaller than $R = 1.5R_\oplus$. For larger radii their fit differs from all the relations in the literature. The relations in Bashi et al. (2017) and Wolfgang et al. (2015) are relatively close to our relation for the volatile-rich population, but they use a single and unique relation for all the planets and do not represent the rocky population correctly.

3.3. Dependence on other parameters

The exoplanets in our sample with masses measured via TTV are statistically less massive than the ones measured by RV. As discussed in Steffen (2016), the sensitivity of TTVs and RVs can be expressed by:

$$SNR_{TTV} \sim \frac{M_p R_p^{3/2} P^{5/6}}{\sigma_{TTV}} \quad , \quad SNR_{RV} \sim \frac{M_p}{\sigma_{RV} P^{1/3}} \quad , \qquad (3)$$

where σ is the intrinsic uncertainty of a measurement. This is a clear observational bias since it is easier for the RV technique to derive the mass of a short period planet, while the TTV technique can determine masses more easily for longer periods. Exoplanets orbiting close to their host stars may have smaller atmospheres (lost through evaporation) and therefore, higher densities. In fact, we can see that the discrepancy between TTVs and RVs arises in the volatile-rich population, while in the rocky population, masses measured by TTVs and RVs overlap. Therefore, for volatile-rich exoplanets with densities smaller than 3.3 $g\,cm^{-3}$, the RV method tends to detect more massive exoplanets, while the sensitivity of the TTVs seems to be more uniform. The fact that our exoplanet sample is dominated by RV measurements and that the current TTVs estimates are significantly less accurate and less robust is likely to bias the demography of exoplanets.

It is interesting to investigate whether there is any trend with a third parameter. Figure 5 shows the dependence of M-R diagram with insolation (a), multiplicity (b), stellar mass (c), and stellar metallicity (d). The densest exoplanets of the rocky population (removing the Trappist-1 exoplanets and the ones beyond $M = 10M_\oplus$) are noticeably more irradiated than the population of exoplanets with a volatile envelope although there is no clear difference between these two populations in terms of stellar mass. It suggests that strongly-irradiated exoplanets (with insolation greater than $1000F_\oplus$) are rocky, probably because their H/He envelopes have been photoevaporated by the high-energy radiation from the host star (e.g. Owen & Wu 2013; Lopez & Fortney 2013; Jin et al. 2014; Zeng et al. 2017; Jin & Mordasini 2018).

When comparing the population of single exoplanets and multi-planetary systems, we see that the majority of exoplanets in

the low-mass regime ($< 25M_\oplus$) are multi-planets, while beyond $30M_\oplus$, almost all of them are single, although this could be a result of an observational bias. We do not find clear trends with stellar mass and metallicity. We suggest that more data and a systematic analysis of these results, including the observational biases, are required in order to understand the relation between planetary and stellar properties.

4. Conclusion

We present an updated exoplanet catalogue based on reliable and robust mass and radius measurements up to $120M_\oplus$, which is available in the DACE platform[6]. The resulting M-R diagrams clearly shows two distinct populations, corresponding to rocky exoplanets and volatile-rich exoplanets. The rocky exoplanet population shows a relatively small density variability and ends at a mass of $\sim 25M_\oplus$, possibly indicating the maximum core mass that can be formed.

We present new empirical M-R relations based on this catalogue. Since the two exoplanet populations overlap in mass and radius, we divide the rocky and volatile-rich regimes by the composition line of pure water and fit both populations. We show that the coefficients we get are rather insensitive to the used composition line of pure water and the limits on mass and radius uncertainties chosen for the catalogue. We compare our M-R relations with previous published ones and we identify their limitations to properly describe the two main populations. We also find that for the same mass rocky exoplanets tend to be more irradiated than volatile-rich exoplanets, suggesting that their H/He envelopes may have been photoevaporated by the high-energy stellar radiation.

The ongoing TESS mission, the future missions like CHEOPS and PLATO, and the ground-based radial velocity facilities like ESPRESSO will populate the M-R diagram with precise measurements. This will allow a better understanding of exoplanetary demographics, in particular in the region between 2 and 4 $R_\oplus$, where the transition between the rocky and volatile-rich planets occurs. Finally, it should be noted that the M-R diagram is in fact multi-dimensional, and is affected by other parameters such as the properties of the host stars, the age of the system, etc, and is also affected by the observational biases. Therefore, when more data become available, future studies should investigate the multi-layer nature of the M-R relation, correct for the selection effects, and hopefully, provide a more complete understanding of the characteristics of planets around other stars.

Acknowledgements. We thank the referee for valuable comments which significantly improved our paper. We also thank the Swiss National Science Foundation (SNSF), the Geneva University and the Zurich University for their continuous support to our exoplanet researches. We also thank George Fest for technical support with constructing the catalogue. This work has been in particular carried out in the frame of the National Centre for Competence in Research 'PlanetS' supported by SNSF. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

References

Allart, R., Bourrier, V., Lovis, C., et al. 2018, Science, 362, 1384
Alsubai, K., Tsvetanov, Z. I., Pyrzas, S., et al. 2019, AJ, 157, 224
Anderson, D. R., Collier Cameron, A., Delrez, L., et al. 2017, A&A, 604, A110
Anderson, D. R., Collier Cameron, A., Hellier, C., et al. 2015, A&A, 575, A61
Anderson, D. R., Smith, A. M. S., Madhusudhan, N., et al. 2013, MNRAS, 430, 3422
Awiphan, S., Kerins, E., Pichadee, S., et al. 2016, MNRAS, 463, 2574
Bakos, G. Á., Hartman, J. D., Torres, G., et al. 2016, arXiv e-prints [arXiv:1606.04556]
Bakos, G. Á., Penev, K., Bayliss, D., et al. 2015, ApJ, 813, 111
Barros, S. C. C., Almenara, J. M., Deleuil, M., et al. 2014, A&A, 569, A74
Bashi, D., Helled, R., & Zucker, S. 2018, Geosciences, 8, 325
Bashi, D., Helled, R., Zucker, S., & Mordasini, C. 2017, A&A, 604, A83
Beatty, T. G., Stevens, D. J., Collins, K. A., et al. 2017, AJ, 154, 25
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99
Bitsch, B., Izidoro, A., Johansen, A., et al. 2019, A&A, 623, A88
Bonomo, A. S., Hébrard, G., Raymond, S. N., et al. 2017, A&A, 603, A43
Bonomo, A. S., Sozzetti, A., Lovis, C., et al. 2014, A&A, 572, A2
Bonomo, A. S., Zeng, L., Damasso, M., et al. 2019, Nature Astronomy, 3, 416
Bordé, P., Bouchy, F., Deleuil, M., et al. 2010, A&A, 520, A66
Borucki, W. J., Koch, D. G., Brown, T. M., et al. 2010, ApJ, 713, L126
Bouchy, F., Hebb, L., Skillen, I., et al. 2010, A&A, 519, A98
Brahm, R., Espinoza, N., Rabus, M., et al. 2019, MNRAS, 483, 1970
Buchhave, L. A., Dressing, C. D., Dumusque, X., et al. 2016, AJ, 152, 160
Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Science, 337, 556
Chaplin, W. J., Sanchis-Ojeda, R., Campante, T. L., et al. 2013, ApJ, 766, 101
Chen, J. & Kipping, D. 2017, ApJ, 834, 17
Christiansen, J. L., Vanderburg, A., Burt, J., et al. 2017, AJ, 154, 122
Cochran, W. D., Fabrycky, D. C., Torres, G., et al. 2011, ApJS, 197, 7
Crida, A., Ligi, R., Dorn, C., & Lebreton, Y. 2018, ApJ, 860, 122
Crossfield, I. J. M., Ciardi, D. R., Petigura, E. A., et al. 2016, ApJS, 226, 7
Dai, F., Winn, J. N., Gandolfi, D., et al. 2017, AJ, 154, 226
Demangeon, O. D. S., Faedi, F., Hebrard, G., et al. 2017, VizieR Online Data Catalog, 361
Demory, B.-O., Gillon, M., Madhusudhan, N., & Queloz, D. 2016, MNRAS, 455, 2018
Dorn, C., Khan, A., Heng, K., et al. 2015, A&A, 577, A83
Dorn, C., Mosegaard, K., Grimm, S. L., & Alibert, Y. 2018, ApJ, 865, 20
Dorn, C., Venturini, J., Khan, A., et al. 2017, A&A, 597, A37
Doyle, L. R., Carter, J. A., Fabrycky, D. C., et al. 2011, Science, 333, 1602
Dragomir, D., Teske, J., Günther, M. N., et al. 2019, ApJ, 875, L7
Dressing, C. D., Charbonneau, D., Dumusque, X., et al. 2015, ApJ, 800, 135
Dressing, C. D., Sinukoff, E., Fulton, B. J., et al. 2018, AJ, 156, 70
Dumusque, X., Bonomo, A. S., Haywood, R. D., et al. 2014, ApJ, 789, 154
Dumusque, X., Turner, O., Dorn, C., et al. 2019, A&A, 627, A43
Eigmüller, P., Chaushev, A., Gillen, E., et al. 2019, A&A, 625, A142
Espinoza, N., Brahm, R., Henning, T., et al. 2019, arXiv e-prints [arXiv:1903.07694]
Esposito, M., Covino, E., Mancini, L., et al. 2014, A&A, 564, L13
Esposito, T. M., Duchêne, G., Kalas, P., et al. 2018, AJ, 156, 47
Esteves, L. J., De Mooij, E. J. W., & Jayawardhana, R. 2015, ApJ, 804, 150
Faedi, F., Barros, S. C. C., Anderson, D. R., et al. 2011, A&A, 531, A40
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gandolfi, D., Fossati, L., Livingston, J. H., et al. 2019, ApJ, 876, L24
Gettel, S., Charbonneau, D., Dressing, C. D., et al. 2016, ApJ, 816, 95
Gibson, N. P., Aigrain, S., Barstow, J. K., et al. 2013, MNRAS, 428, 3680
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Nature, 542, 456
Grimm, S. L., Demory, B.-O., Gillon, M., et al. 2018, A&A, 613, A68
Grunblatt, S. K., Howard, A. W., & Haywood, R. D. 2015, ApJ, 808, 127
Guillot, T., Miguel, Y., Hubbard, W. B., et al. 2017, in AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 49, AAS/Division for Planetary Sciences Meeting Abstracts #49, 109.01
Hadden, S. & Lithwick, Y. 2014, ApJ, 787, 80
Hadden, S. & Lithwick, Y. 2017, AJ, 154, 5
Harpsoe, K. B. W., Hardis, S., Hinse, T. C., et al. 2013, A&A, 549, A10
Hartman, J. D., Bakos, G. Á., Sato, B., et al. 2011, ApJ, 726, 52
Hartman, J. D., Bakos, G. Á., Torres, G., et al. 2014, AJ, 147, 128
Hartman, J. D., Bakos, G. Á., Torres, G., et al. 2009, ApJ, 706, 785
Hartman, J. D., Bhatti, W., Bakos, G. Á., et al. 2015, AJ, 150, 168
Helled, R., Lozovsky, M., Vazan, A., et al. 2017, in EGU General Assembly Conference Abstracts, Vol. 19, EGU General Assembly Conference Abstracts, 11876
Hellier, C., Anderson, D. R., Cameron, A. C., et al. 2017, MNRAS, 465, 3693
Hellier, C., Anderson, D. R., Collier Cameron, A., et al. 2015, AJ, 150, 18
Hellier, C., Anderson, D. R., Collier Cameron, A., et al. 2012, MNRAS, 426, 739
Howe, A. R., Burrows, A., & Verne, W. 2014, ApJ, 787, 173
Huang, C. X., Burt, J., Vanderburg, A., et al. 2018, ApJ, 868, L39
Huber, D., Chaplin, W. J., Chontos, A., et al. 2019, AJ, 157, 245
Iess, L., Militzer, B., Kaspi, Y., et al. 2019, Science, 364, aat2965
Jin, S. & Mordasini, C. 2018, ApJ, 853, 163
Jin, S., Mordasini, C., Parmentier, V., et al. 2014, ApJ, 795, 65

[6] dace.unige.ch

Johansen, A. & Lambrechts, M. 2017, Annual Review of Earth and Planetary Sciences, 45, 359
Johnson, M. C., Dai, F., Justesen, A. B., et al. 2018, MNRAS, 481, 596
Jordán, A., Brahm, R., Espinoza, N., et al. 2019, AJ, 157, 100
Korth, J., Csizmadia, S., Gandolfi, D., et al. 2019, MNRAS, 482, 1807
Kosiarek, M. R., Blunt, S., López-Morales, M., et al. 2019, AJ, 157, 116
Lam, K. W. F., Faedi, F., Brown, D. J. A., et al. 2017, A&A, 599, A3
Lam, K. W. F., Santerne, A., Sousa, S. G., et al. 2018, A&A, 620, A77
Lendl, M., Anderson, D. R., Bonfanti, A., et al. 2019, MNRAS, 482, 301
Lendl, M., Triaud, A. H. M. J., Anderson, D. R., et al. 2014, A&A, 568, A81
Lissauer, J. J., Jontof-Hutter, D., Rowe, J. F., et al. 2013, ApJ, 770, 131
Livingston, J. H., Endl, M., Dai, F., et al. 2018, AJ, 156, 78
Lopez, E. D. & Fortney, J. J. 2013, ApJ, 776, 2
Lozovsky, M., Helled, R., Dorn, C., & Venturini, J. 2018, The Astrophysical Journal, 866, 49
Luque, R., Nowak, G., Pallé, E., et al. 2018, A&A, 620, A171
Luque, R., Pallé, E., Kossakowski, D., et al. 2019, arXiv e-prints [arXiv:1904.12818]
MacDonald, M. G., Ragozzine, D., Fabrycky, D. C., et al. 2016, AJ, 152, 105
Malavolta, L., Mayo, A. W., Louden, T., et al. 2018, AJ, 155, 107
Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, ApJS, 210, 20
Masuda, K., Hirano, T., Taruya, A., Nagasawa, M., & Suto, Y. 2013, ApJ, 778, 185
Maxted, P. F. L., Anderson, D. R., Collier Cameron, A., et al. 2016, A&A, 591, A55
Ment, K., Fischer, D. A., Bakos, G., Howard, A. W., & Isaacson, H. 2018, AJ, 156, 213
Mills, S. M., Howard, A. W., Weiss, L. M., et al. 2019, AJ, 157, 145
More, R. M., Warren, K. H., Young, D. A., & Zimmerman, G. B. 1988, Physics of Fluids, 31, 3059
Mortier, A., Bonomo, A. S., Rajpaul, V. M., et al. 2018, MNRAS, 481, 1839
Osborn, H. P., Santerne, A., Barros, S. C. C., et al. 2017, A&A, 604, A19
Owen, J. E. & Wu, Y. 2013, ApJ, 775, 105
Palle, E., Nowak, G., Luque, R., et al. 2019, A&A, 623, A41
Persson, C. M., Fridlund, M., Barragán, O., et al. 2018, A&A, 618, A33
Petigura, E. A., Benneke, B., Batygin, K., et al. 2018, AJ, 156, 89
Petigura, E. A., Sinukoff, E., Lopez, E. D., et al. 2017, AJ, 153, 142
Rajpaul, V., Buchhave, L. A., & Aigrain, S. 2017, MNRAS, 471, L125
Rice, K., Malavolta, L., Mayo, A., et al. 2019, MNRAS, 484, 3731
Rodriguez, J. E., Becker, J. C., Eastman, J. D., et al. 2018, AJ, 156, 245
Santerne, A., Brugger, B., Armstrong, D. J., et al. 2018, Nature Astronomy, 2, 393
Saumon, D. & Guillot, T. 2004, ApJ, 609, 1170
Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007a, ApJ, 669, 1279
Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007b, ApJ, 669, 1279
Sinukoff, E., Howard, A. W., Petigura, E. A., et al. 2017, AJ, 153, 271
Smith, A. M. S., Cabrera, J., Csizmadia, S., et al. 2018a, MNRAS, 474, 5523
Smith, A. M. S., Csizmadia, S., Gandolfi, D., et al. 2018b, arXiv e-prints [arXiv:1807.05865]
Sotin, C., Grasset, O., & Mocquet, A. 2007, Icarus, 191, 337
Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, AJ, 153, 136
Steffen, J. H. 2016, MNRAS, 457, 4384
Sun, L., Ioannidis, P., Gu, S., et al. 2019, A&A, 624, A15
Thompson, S. L. 1990
Turner, J. D., Pearson, K. A., Biddle, L. I., et al. 2016, MNRAS, 459, 789
Turner, O. D., Anderson, D. R., Barkaoui, K., et al. 2019, MNRAS, 485, 5790
Van Grootel, V., Gillon, M., Valencia, D., et al. 2014, ApJ, 786, 2
Vanderburg, A., Becker, J. C., Buchhave, L. A., et al. 2017, AJ, 154, 237
Wahl, S. M., Hubbard, W. B., Militzer, B., et al. 2017, Geophys. Res. Lett., 44, 4649
Weiss, L. M. & Marcy, G. W. 2014, ApJ, 783, L6
Weiss, L. M., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 768, 14
Welsh, W. F., Orosz, J. A., Carter, J. A., et al. 2012, Nature, 481, 475
Wolfgang, A., Rogers, L. A., Ford, E. B., & Laughlin, G. 2015, in AAS/Division for Extreme Solar Systems Abstracts, Vol. 3, AAS/Division for Extreme Solar Systems Abstracts, 501.04
Xie, J.-W. 2014, ApJS, 210, 25
Yee, S. W., Petigura, E. A., Fulton, B. J., et al. 2018, AJ, 155, 255
Yu, L., Rodriguez, J. E., Eastman, J. D., et al. 2018, AJ, 156, 127
Zeng, L., Jacobsen, S. B., & Sasselov, D. D. 2017, Research Notes of the American Astronomical Society, 1, 32
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127
Zhou, G., Bayliss, D., Penev, K., et al. 2014, AJ, 147, 144

Appendix A:

Figure A.1 shows the evolution of the M-R diagram before and after applying our selection criteria as discussed in Section 2.1. Figure A.2 shows density against mass for our catalogue, with the rocky population and the volatile-rich populations being divided by the pure-water line. The grey envelope indicates the region between 2.8 $g\,cm^{-3}$ to 3.3 $g\,cm^{-3}$, showing that the division of the two populations by a density cutoff of $\sim 3\ g\,cm^{-3}$ lead to very similar results. Table A.1 lists all the exoplanets up to a mass of $120 M_{\oplus}$ in our new "filtered" catalogue, and also states the references provided by the NASA Exoplanet Archive with the ones used in this work.

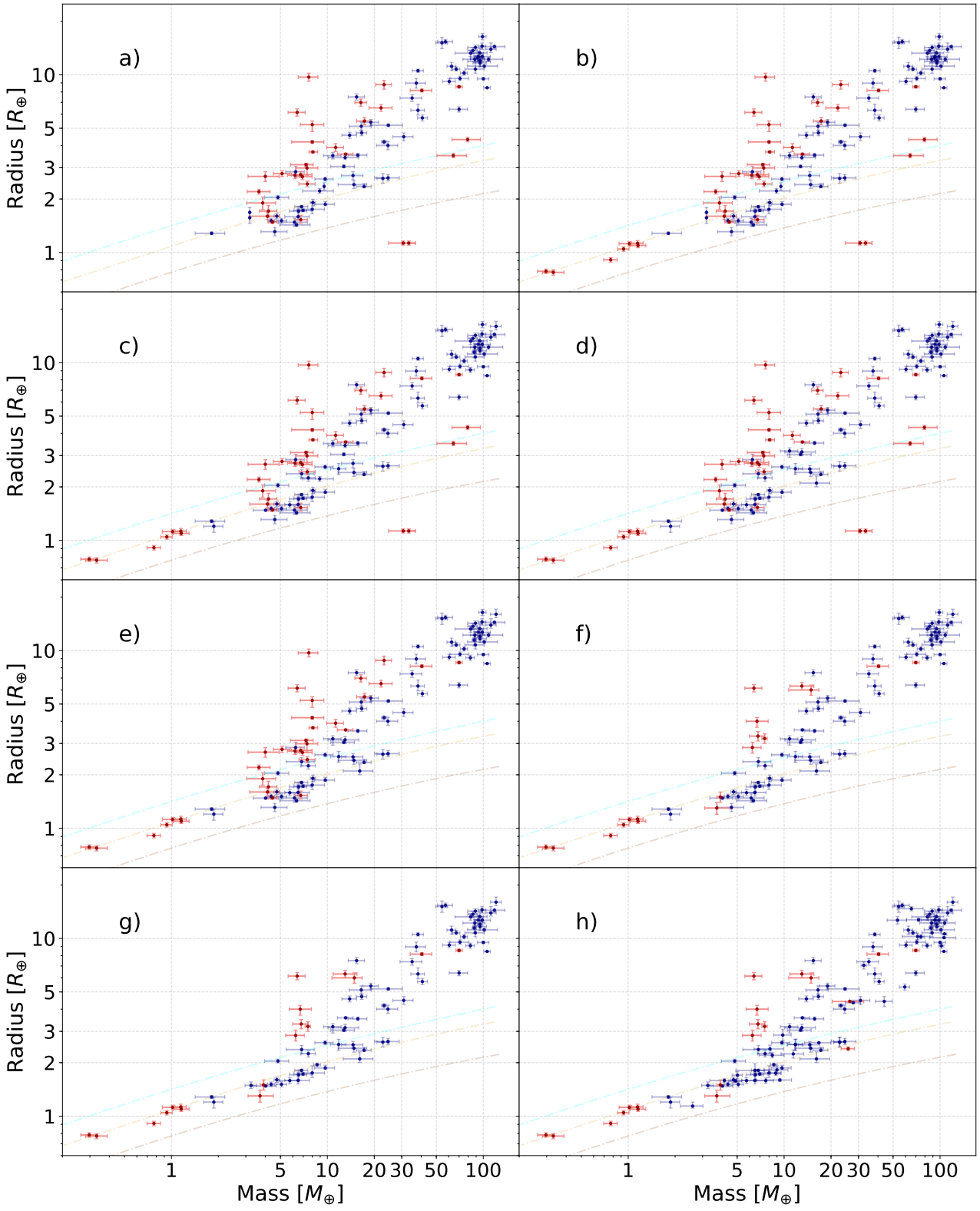

Fig. A.1. The different steps of the data selection starting with initial NASA Exoplanet Archive data and ending with our final "filtered" catalogue.

Table A.1. Filtered catalogue with robust and reliable mass and radius measurements. It compares the reference picked by the NASA Exoplanet Archive (if available) and our selected reference. Where the radius has been obtained from Berger et al. (2018) (due to smaller uncertainty) it is indicated with an asterisk.

Planet	Mass [$M_\oplus$]	Radius [$R_\oplus$]	Reference in NASA Exo. Arch.	Reference in this work
TRAPPIST-1 d	$0.297^{+0.039}_{-0.035}$	$0.784^{+0.023}_{-0.023}$	Gillon et al. (2017)	Grimm et al. (2018)
TRAPPIST-1 h	$0.331^{+0.056}_{-0.049}$	$0.773^{+0.026}_{-0.027}$	-	Grimm et al. (2018)
TRAPPIST-1 e	$0.772^{+0.079}_{-0.075}$	$0.91^{+0.026}_{-0.027}$	Gillon et al. (2017)	Grimm et al. (2018)
TRAPPIST-1 f	$0.934^{+0.078}_{-0.08}$	$1.046^{+0.029}_{-0.03}$	Gillon et al. (2017)	Grimm et al. (2018)
TRAPPIST-1 b	$1.017^{+0.154}_{-0.143}$	$1.121^{+0.032}_{-0.032}$	Gillon et al. (2017)	Grimm et al. (2018)
TRAPPIST-1 g	$1.148^{+0.098}_{-0.095}$	$1.127^{+0.041}_{-0.041}$	Gillon et al. (2017)	Grimm et al. (2018)
TRAPPIST-1 c	$1.156^{+0.142}_{-0.131}$	$1.095^{+0.031}_{-0.031}$	Gillon et al. (2017)	Grimm et al. (2018)
LHS 1140 c	$1.81^{+0.39}_{-0.39}$	$1.282^{+0.024}_{-0.024}$	Ment et al. (2018)	Ment et al. (2018)
GJ 357 b	$1.84^{+0.31}_{-0.31}$	$1.217^{+0.084}_{-0.083}$	-	Luque et al. (2019)
Kepler-78 b	$1.87^{+0.27}_{-0.26}$	$1.2^{+0.08}_{-0.09}$	Stassun et al. (2017)	Grunblatt et al. (2015)
K2-229 b	$2.59^{+0.43}_{-0.43}$	$1.14^{+0.06}_{-0.03}$	Livingston et al. (2018)	Santerne et al. (2018)
Kepler-10 b	$3.24^{+0.28}_{-0.28}$	$1.489^{+0.07}_{-0.06}$	Esteves et al. (2015)	Rajpaul et al. (2017)*
Kepler-80 d	$3.7^{+0.8}_{-0.6}$	$1.3^{+0.1}_{-0.1}$	MacDonald et al. (2016)	Hadden & Lithwick (2017)
Kepler-36 b	$3.9^{+0.2}_{-0.2}$	$1.5^{+0.1}_{-0.1}$	Carter et al. (2012)	Hadden & Lithwick (2017)
Kepler-93 b	$4.02^{+0.68}_{-0.68}$	$1.478^{+0.019}_{-0.019}$	Stassun et al. (2017)	Dressing et al. (2015)
Kepler-65 d	$4.14^{+0.79}_{-0.8}$	$1.587^{+0.04}_{-0.035}$	Chaplin et al. (2013)	Mills et al. (2019)
HD 219134 c	$4.36^{+0.22}_{-0.22}$	$1.511^{+0.047}_{-0.047}$	Gillon et al. (2017)	Gillon et al. (2017)
HD 219134 b	$4.74^{+0.19}_{-0.19}$	$1.602^{+0.055}_{-0.055}$	Gillon et al. (2017)	Gillon et al. (2017)
HD 39091 c	$4.82^{+0.84}_{-0.86}$	$2.042^{+0.05}_{-0.05}$	-	Huang et al. (2018)
GJ 9827 b	$4.89^{+0.477}_{-0.477}$	$1.575^{+0.03}_{-0.02}$	Rodriguez et al. (2018)	Rice et al. (2019)
HD 3167 b	$5.02^{+0.38}_{-0.38}$	$1.7^{+0.08}_{-0.08}$	-	Christiansen et al. (2017)
K2-141 b	$5.08^{+0.41}_{-0.41}$	$1.51^{+0.05}_{-0.05}$	Malavolta et al. (2018)	Malavolta et al. (2018)
CoRoT-7 b	$5.74^{+0.86}_{-0.86}$	$1.585^{+0.064}_{-0.064}$	Stassun et al. (2017)	Barros et al. (2014)
GJ 1214 b	$6.26125^{+0.85814}_{-0.85814}$	$2.847^{+0.202}_{-0.202}$	Harpsoe et al. (2013)	Harpsoe et al. (2013)
K2-291 b	$6.49^{+1.16}_{-1.16}$	$1.589^{+0.095}_{-0.072}$	Kosiarek et al. (2019)	Kosiarek et al. (2019)
K2-131 b	$6.5^{+1.6}_{-1.6}$	$1.81^{+0.16}_{-0.12}$	-	Dai et al. (2017)
K2-265 b	$6.54^{+0.84}_{-0.84}$	$1.71^{+0.11}_{-0.11}$	Lam et al. (2018)	Lam et al. (2018)
Kepler-11 e	$6.7^{+1.2}_{-1.0}$	$4.0^{+0.2}_{-0.3}$	Hadden & Lithwick (2017)	Hadden & Lithwick (2017)
Kepler-11 d	$6.8^{+0.7}_{-0.8}$	$3.3^{+0.2}_{-0.2}$	Lissauer et al. (2013)	Hadden & Lithwick (2017)
WASP-47 e	$6.83^{+0.66}_{-0.66}$	$1.81^{+0.027}_{-0.027}$	Vanderburg et al. (2017)	Vanderburg et al. (2017)
Kepler-454 b	$6.84^{+1.4}_{-1.4}$	$2.37^{+0.13}_{-0.13}$	Stassun et al. (2017)	Gettel et al. (2016)
LHS 1140 b	$6.98^{+0.89}_{-0.89}$	$1.727^{+0.032}_{-0.032}$	Ment et al. (2018)	Ment et al. (2018)
Kepler-36 c	$7.5^{+0.3}_{-0.3}$	$3.2^{+0.2}_{-0.2}$	Carter et al. (2012)	Hadden & Lithwick (2017)
HD 97658 b	$7.55^{+0.83}_{-0.79}$	$2.247^{+0.098}_{-0.095}$	Stassun et al. (2017)	Van Grootel et al. (2014)
HD 15337 b	$7.20^{+0.81}_{-0.81}$	$1.70^{+0.06}_{-0.06}$	Gandolfi et al. (2019)	Dumusque et al. (2019)
K2-216 b	$8.0^{+1.6}_{-1.6}$	$1.75^{+0.17}_{-0.1}$	Persson et al. (2018)	Persson et al. (2018)
Kepler-19 b	$8.4^{+1.6}_{-1.5}$	$2.2^{+0.07}_{-0.07}$	Malavolta et al. (2018)	Malavolta et al. (2018)
55 Cnc e	$8.59^{+0.43}_{-0.43}$	$1.947^{+0.038}_{-0.038}$	Demory et al. (2016)	Crida et al. (2018)
...	...	...	...	...

Planet	Mass [$M_\oplus$]	Radius [$R_\oplus$]	Reference in NASA Exo. Arch.	Reference in this work
HD 15337 c	$8.79^{+1.68}_{-1.68}$	$2.52^{+0.11}_{-0.11}$	-	Dumusque et al. (2019)
HD 213885 b	$8.83^{+0.66}_{-0.65}$	$1.745^{+0.051}_{-0.052}$	-	Espinoza et al. (2019)
EPIC 220674823 b	$9.0^{+1.6}_{-1.6}$	$1.82^{+0.1}_{-0.1}$	-	Sinukoff et al. (2017)
Kepler-107 c	$9.39^{+1.77}_{-1.77}$	$1.597^{+0.026}_{-0.026}$	Bonomo et al. (2019)	Bonomo et al. (2019)
K2-285 b	$9.68^{+1.2}_{-1.3}$	$2.59^{+0.06}_{-0.06}$	Palle et al. (2019)	Palle et al. (2019)
Kepler-20 b	$9.7^{+1.41}_{-1.44}$	$1.868^{+0.066}_{-0.034}$	Buchhave et al. (2016)	Buchhave et al. (2016)
HD 3167 c	$9.8^{+1.3}_{-1.23}$	$2.86^{+0.22}_{-0.22}$	-	Christiansen et al. (2017)
Kepler-94 b	$10.84^{+1.4}_{-1.4}$	$3.186^{+0.13}_{-0.25}$	Marcy et al. (2014)	Marcy et al. (2014)*
K2-180 b	$11.448^{+1.9}_{-1.9}$	$2.24^{+0.12}_{-0.12}$	-	Korth et al. (2019)
HIP 116454 b	$11.82^{+1.33}_{-1.33}$	$2.53^{+0.18}_{-0.18}$	Stassun et al. (2017)	Vanderburg et al. (2017)
Kepler-20 c	$12.75^{+2.17}_{-2.24}$	$3.047^{+0.064}_{-0.056}$	Buchhave et al. (2016)	Buchhave et al. (2016)
Kepler-95 b	$13.0^{+2.9}_{-2.9}$	$3.145^{+0.144}_{-0.132}$	Marcy et al. (2014)	Marcy et al. (2014)*
KOI-94 e	$13.0^{+2.5}_{-2.1}$	$6.31^{+0.3}_{-0.3}$	Weiss et al. (2013)	Masuda et al. (2013)*
WASP-47 d	$13.1^{+1.5}_{-1.5}$	$3.576^{+0.046}_{-0.046}$	Vanderburg et al. (2017)	Vanderburg et al. (2017)
GJ 3470 b	$13.9^{+1.5}_{-1.5}$	$4.57^{+0.18}_{-0.18}$	Awiphan et al. (2016)	Awiphan et al. (2016)
Kepler-48 c	$14.61^{+2.3}_{-2.3}$	$2.522^{+0.113}_{-0.107}$	Marcy et al. (2014)	Marcy et al. (2014)*
K2-263 b	$14.8^{+3.1}_{-3.1}$	$2.41^{+0.12}_{-0.12}$	Mortier et al. (2018)	Mortier et al. (2018)
Kepler-18 d	$14.9^{+1.8}_{-4.2}$	$6.0^{+0.4}_{-0.4}$	Cochran et al. (2011)	Hadden & Lithwick (2017)
K2-24 c	$15.4^{+1.9}_{-1.8}$	$7.5^{+0.3}_{-0.2}$	-	Petigura et al. (2018)
K2-285 c	$15.68^{+2.28}_{-2.13}$	$3.53^{+0.08}_{-0.08}$	Petigura et al. (2018)	Petigura et al. (2018)
Kepler-131 b	$16.13^{+3.5}_{-3.5}$	$2.1^{+0.2}_{-0.1}$	Marcy et al. (2014)	Marcy et al. (2014)*
K2-32 b	$16.5^{+2.7}_{-2.7}$	$5.13^{+0.28}_{-0.28}$	Petigura et al. (2017)	Petigura et al. (2017)
HD 219666 b	$16.6^{+1.3}_{-1.3}$	$4.71^{+0.17}_{-0.17}$	Esposito et al. (2018)	Esposito et al. (2018)
K2-110 b	$16.7^{+3.2}_{-3.2}$	$2.59^{+0.01}_{-0.01}$	Osborn et al. (2017)	Osborn et al. (2017)
Kepler-10 c	$17.2^{+1.9}_{-1.9}$	$2.35^{+0.09}_{-0.04}$	Dumusque et al. (2014)	Dumusque et al. (2014)
K2-24 b	$19.0^{+-2.1}_{--2.0}$	$5.4^{+0.2}_{-0.2}$	Petigura et al. (2018)	Petigura et al. (2018)
GJ 143 b	$22.7^{+2.2}_{-1.9}$	$2.61^{+0.17}_{-0.16}$	-	Dragomir et al. (2019)
GJ 436 b	$23.1^{+0.8}_{-0.8}$	$4.191^{+0.1}_{-0.1}$	Turner et al. (2016)	Turner et al. (2016)
HD 21749 b	$23.2^{+2.13}_{-1.91}$	$2.84^{+0.16}_{-0.16}$	-	Dragomir et al. (2019)
Kepler-4 b	$24.472^{+3.814}_{-3.814}$	$4.002^{+0.213}_{-0.213}$	Borucki et al. (2010)	Borucki et al. (2010)
HD 119130 b	$24.5^{+4.4}_{-4.4}$	$2.63^{+0.11}_{-0.1}$	Luque et al. (2018)	Luque et al. (2018)
Kepler-25 c	$24.6^{+5.7}_{-5.7}$	$5.154^{+0.06}_{-0.06}$	Marcy et al. (2014)	Marcy et al. (2014)*
Kepler-411 b	$25.758^{+2.544}_{-2.544}$	$2.3968^{+0.056}_{-0.056}$	Sun et al. (2019)	Sun et al. (2019)
Kepler-411 c	$26.394^{+6.042}_{-6.042}$	$4.418^{+0.06}_{-0.06}$	Sun et al. (2019)	Sun et al. (2019)
HAT-P-11 b	$27.76^{+3.08}_{-3.08}$	$4.35^{+0.05}_{-0.05}$	Yee et al. (2018)	Allart et al. (2018)
K2-27 b	$30.9^{+4.6}_{-4.6}$	$4.48^{+0.23}_{-0.23}$	Petigura et al. (2017)	Petigura et al. (2017)
WASP-166 b	$32.436^{+1.272}_{-1.272}$	$7.056^{+0.336}_{-0.336}$	-	Hellier et al. (2017)
HD 89345 b	$34.9613^{+5.40311}_{-5.72094}$	$7.398^{+0.314}_{-0.336}$	Yu et al. (2018)	Yu et al. (2018)
WASP-139 b	$37.18611^{+5.40311}_{-5.40311}$	$8.967^{+0.56}_{-0.56}$	Hellier et al. (2017)	Hellier et al. (2017)
WASP-107 b	$38.0^{+3.0}_{-3.0}$	$10.6^{+0.3}_{-0.3}$	-	Anderson et al. (2017)
HATS-7 b	$38.1^{+3.8}_{-3.8}$	$6.31^{+0.5}_{-0.4}$	Bakos et al. (2015)	Bakos et al. (2015)
...	...	...	...	...

Planet	Mass [$M_\oplus$]	Radius [$R_\oplus$]	Reference	
Kepler-35 b	$40.363^{+6.356}_{-6.356}$	$8.16^{+0.157}_{-0.157}$	Welsh et al. (2012)	Welsh et al. (2012)
WASP-156 b	$40.68224^{+3.1783}_{-2.86047}$	$5.717^{+0.224}_{-0.224}$	Demangeon et al. (2017)	Demangeon et al. (2017)
K2-55 b	$43.88^{+5.4}_{-5.4}$	$4.424^{+0.29}_{-0.29}$	Crossfield et al. (2016)	Dressing et al. (2018)
Kepler-101 b	$51.1^{+5.1}_{-4.7}$	$5.986^{+0.27}_{-0.25}$	Bonomo et al. (2014)	Bonomo et al. (2014)*
HAT-P-48 b	$53.4^{+7.6}_{-7.6}$	$12.66^{+0.6}_{-0.6}$	-	Bakos et al. (2016)
KELT-11 b	$54.3^{+4.8}_{-4.8}$	$15.1^{+1.1}_{-1.1}$	Beatty et al. (2017)	Beatty et al. (2017)
K2-261 b	$56.922^{+6.36}_{-6.36}$	$9.4^{+0.12}_{-0.12}$	Johnson et al. (2018)	Brahm et al. (2019)
WASP-127 b	$57.2094^{+6.3566}_{-6.3566}$	$15.356^{+0.448}_{-0.448}$	Lam et al. (2017)	Lam et al. (2017)
K2-108 b	$59.4^{+4.4}_{-4.4}$	$5.33^{+0.21}_{-0.21}$	Petigura et al. (2017)	Petigura et al. (2017)
HAT-P-18 b	$62.3^{+2.5}_{-2.5}$	$11.153^{+0.583}_{-0.583}$	Hartman et al. (2011)	Esposito et al. (2014)
HD 221416 b	$63.4^{+5.7}_{-5.7}$	$9.16^{+0.34}_{-0.31}$	Huber et al. (2019)	Huber et al. (2019)
HAT-P-47 b	$65.508^{+12.4}_{-12.4}$	$14.7^{+0.4}_{-0.4}$	-	Bakos et al. (2016)
HAT-P-12 b	$67.059^{+3.814}_{-3.814}$	$10.749^{+0.325}_{-0.235}$	Hartman et al. (2009)	Huber et al. (2019)
CoRoT-8 b	$69.92^{+9.53}_{-9.53}$	$6.39^{+0.22}_{-0.22}$	Bordé et al. (2010)	Bordé et al. (2010)
Kepler-34 b	$69.92^{+3.496}_{-3.178}$	$8.564^{+0.135}_{-0.157}$	Welsh et al. (2012)	Welsh et al. (2012)
Kepler-425 b	$71.8^{+14.6}_{-14.6}$	$10.255^{+0.47}_{-0.45}$	-	Bonomo et al. (2017)*
NGTS-5 b	$72.8^{+11.8}_{-11.8}$	$12.73^{+0.26}_{-0.26}$	-	Eigmüller et al. (2019)
HATS-5 b	$75.323^{+3.814}_{-3.914}$	$10.223^{+0.28}_{-0.28}$	Zhou et al. (2014)	Zhou et al. (2014)
WASP-29 b	$77.9^{+7.3}_{-7.0}$	$8.8^{+0.6}_{-0.4}$	Stassun et al. (2017)	Gibson et al. (2013)
WASP-69 b	$82.632^{+5.4}_{-5.4}$	$9.11^{+0.3}_{-0.3}$	Stassun et al. (2017)	Anderson et al. (2013)
HATS-43 b	$83.0^{+17.0}_{-17.0}$	$13.23^{+0.56}_{-0.56}$	-	Brahm et al. (2019)
WASP-131 b	$85.8141^{+6.3566}_{-6.3566}$	$13.675^{+0.56}_{-0.56}$	Hellier et al. (2017)	Hellier et al. (2017)
WASP-117 b	$87.55^{+2.86}_{-2.8}$	$11.44^{+0.785}_{-0.785}$	Stassun et al. (2017)	Lendl et al. (2014)
WASP-160 b	$88.35674^{+13.98452}_{-14.30235}$	$12.218^{+0.527}_{-0.46}$	-	Lendl et al. (2019)
WASP-39 b	$88.989^{+9.535}_{-9.535}$	$14.235^{+0.448}_{-0.448}$	Faedi et al. (2011)	Faedi et al. (2011)
WASP-126 b	$88.9924^{+12.7132}_{-12.7132}$	$10.761^{+1.121}_{-0.56}$	Maxted et al. (2016)	Maxted et al. (2016)
Kepler-427 b	$92.8^{+17.2}_{-17.2}$	$12.696^{+0.577}_{-0.547}$	-	Bonomo et al. (2017)*
HAT-P-19 b	$92.802^{+5.721}_{-5.721}$	$12.689^{+0.807}_{-0.807}$	Hartman et al. (2011)	Hartman et al. (2011)
WASP-181 b	$95.0^{+10.8}_{-10.8}$	$13.26^{+0.66}_{-0.79}$	-	Turner et al. (2019)
WASP-21 b	$95.345^{+3.496}_{-3.496}$	$11.99^{+0.56}_{-0.56}$	Bouchy et al. (2010)	Bouchy et al. (2010)
WASP-83 b	$95.349^{+9.5349}_{-9.5349}$	$11.657^{+0.897}_{-0.56}$	Hellier et al. (2015)	Hellier et al. (2015)
HAT-P-51 b	$98.20947^{+5.72094}_{-5.72094}$	$14.493^{+0.605}_{-0.605}$	Hartman et al. (2015)	Hartman et al. (2015)
WASP-151 b	$98.5273^{+12.7132}_{-9.5349}$	$12.666^{+0.336}_{-0.336}$	Demangeon et al. (2017)	Demangeon et al. (2017)
WASP-20 b	$98.8^{+6.0}_{-5.8}$	$16.39^{+0.66}_{-0.66}$	Anderson et al. (2015)	Bonomo et al. (2017)
K2-287 b	$100.0^{+9.0}_{-9.0}$	$9.49^{+0.15}_{-0.15}$	Jordán et al. (2019)	Jordán et al. (2019)
HATS-6 b	$101.0^{+22.0}_{-22.0}$	$11.19^{+0.21}_{-0.21}$	Hartman et al. (2015)	Hartman et al. (2015)
HD 149026 b	$102.0^{+4.0}_{-4.0}$	$9.09^{+0.3}_{-0.3}$	Stassun et al. (2017)	Bonomo et al. (2017)
Kepler-16 b	$105.833^{+5.085}_{-5.085}$	$8.449^{+0.029}_{-0.029}$	Doyle et al. (2011)	Doyle et al. (2011)
K2-295 b	$106.0^{+20.0}_{-20.0}$	$10.1^{+0.1}_{-0.1}$	-	Smith et al. (2018a)
EPIC 220501947 b	$106.848^{+3.816}_{-3.816}$	$10.6064^{+0.056}_{-0.1344}$	-	Smith et al. (2018b)
Kepler-426 b	$107.0^{+18.0}_{-19.0}$	$12.22^{+0.34}_{-0.34}$	-	Bonomo et al. (2017)
HAT-P-44 b	$111.871^{+9.217}_{-9.217}$	$13.922^{+1.188}_{-0.572}$	Hartman et al. (2014)	Hartman et al. (2014)
Qatar-8 b	$117.978^{+19.716}_{-19.716}$	$14.392^{+0.2464}_{-0.2464}$	-	Alsubai et al. (2019)
WASP-63 b	$120.77^{+9.5}_{-9.5}$	$16.0^{+1.1}_{-0.7}$	Stassun et al. (2017)	Hellier et al. (2012)

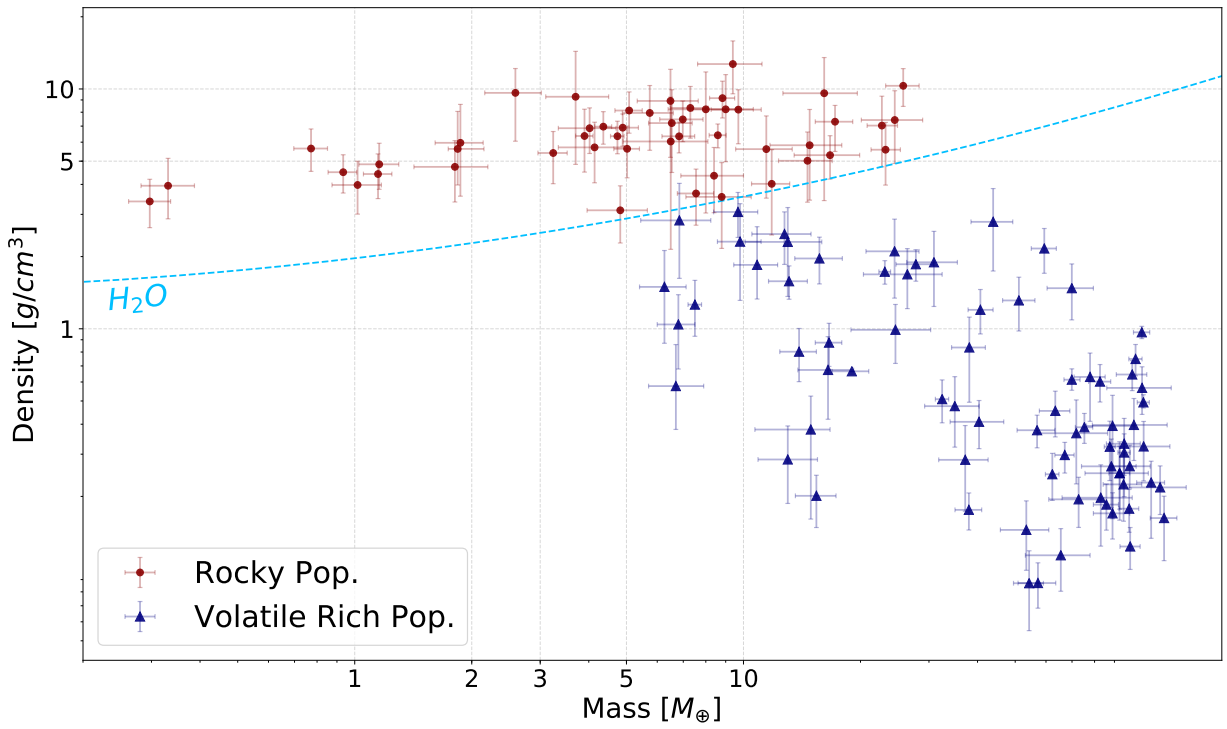

Fig. A.2. Density vs. mass for our revised catalogue. The rocky population and the volatile-rich population are separated by the composition line of pure water (Dorn et al. 2015). The grey envelope indicates the region between 2.8 $g\ cm^{-3}$ and 3.3 $g\ cm^{-3}$.

.

2.1.2 The PlanetS catalog

2.1.2.1 Evolution of the M-R diagram

After the publication of Otegi et al. (2020a), we continued adding the new discovered in our catalog and made it accessible on the Data & Analysis Center for Exoplanet DACE[d]. DACE is a PlanetS web platform located at the University of Geneva and dedicated to exoplanet data visualization, exchange and, analysis. It provides access to exoplanet data with a suite of statistical tools for data analysis. Initially, our revisited catalog was named "Otegi M-R", but later on PlanetS accepted it as their certified catalog, so its name changed to "PlanetS Catalog". DACE allows accessing data from different catalogs and make plots online. During the last years, several useful features to perform simple demographic analyses have been added, such as the addition of composition lines, the possibility to make the color or size of the dot shows an extra parameter, the possibility of removing specific planets, or 3D plots. The data is also available through the DACE API, which allows downloading it and perform more in-depth studies. We also added a tutorial on how to download the data of different catalogs using the DACE API and perform basic demographic analyzes[e].

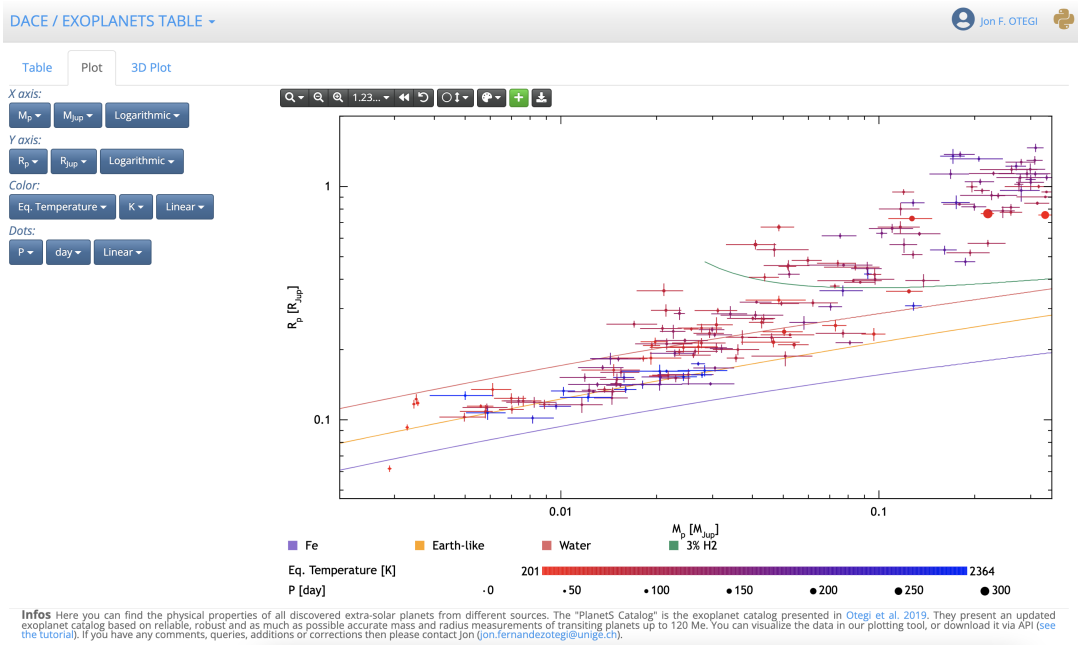

Figure 2.2: Screenshot of M-R diagram interface on the DACE online platform. The exoplanets are taken from th the PlanetS catalog, based on Otegi et al. (2020a). The color and size of the dots indicate the equilibrium temperature and orbital period.

[d] dace.unige.ch
[e] https://dace.unige.ch/tutorials/?tutorialId=30

In Figure 2.3 we can see how the M-R diagram has evolved after Otegi et al. (2020a) was published. One of the main results in the paper consisted of the presence of two distinct exoplanet populations (as seen in the top panel of Figure 2.3): one of them closely following an Earth-like composition, and a second one corresponding to a more volatile-rich composition. It is important to remark that even when an exoplanet lies in the M-R diagram of an Earth-like composition, its actual relative composition of iron, silicates and water could be different, given the degeneracy of the problem. However, the amount of water or H/He envelope of these exoplanets is expected to be small in comparison to the refractory materials (e.g. iron, silicates), and it is reasonable to assume that these exoplanets are mostly rocky[f]. The second population shown in our M-R diagram with exoplanets discovered before 2020 corresponds to less dense planets with a more volatile-rich composition.

However, this bi-modal becomes less evident when we include exoplanets discovered in 2020. When we only consider exoplanets up to 2020 we find a relatively clear gap between the most massive planets of the 'rocky exoplanets' and the less massive of the 'volatile-rich exoplanets'. We find almost no planets with radius between 3-4 $R_\oplus$ and masses between 20-50 $M_\oplus$. Nevertheless, this gap is partially filled with four TESS exoplanets characterized with precise mass and radius measurements in 2020: TOI-824b (Burt et al. 2020a), TOI-849b (Armstrong et al. 2020), TOI-125d (Nielsen et al. 2020a) and TOI-132 (Díaz et al. 2020). Later in 2021 thirty more exoplanets with masses below $120M_\oplus$ were added to the PlanetS catalog, and some of them lied between the two well-defined populations identified in the sample before 2020. Nevertheless, we see that the density of exoplanets in the gap of radius between 3-4 $R_\oplus$ and mass between 20-50 $M_\oplus$ is lower than in the 'rocky' or 'volatile-rich' populations. We conclude that more observational data is needed to have strong conclusions about the gap separating the massive tail of the 'rocky' exoplanets and the less massive 'volatile-rich' exoplanets.

The updated PlanetS catalog also confirms some observed trends in Otegi et al. (2020a). Even if we are considering all planets with radii below the pure water composition line as 'rocky', we note that planets with masses below $10M_\oplus$ tend to follow closely the Earth-like composition line and to be more dense than the 'rocky' planets more massive than $10M_\oplus$. Indeed, it is interesting to see that there are no observed planets below the Earth-like composition line with masses larger than $10M_\oplus$, which cannot be explained by observational biases. We also note that planets with radii larger than $2R_\oplus$ show a very wide diversity in densities. Between $2R_\oplus$ and $3R_\oplus$ we can find either super-Earths or volatile-rich sub-Neptunes with masses from 5 to 30 $M_\oplus$. It is also remarkable the lack of planets with larger radii then $2_\oplus$ and masses below $5_\oplus$.

[f] in this study 'rocky exoplanets' refer to exoplanets that mostly consist of metals and rocks

More observational data and a complete understanding of the observational bias involved are still needed to fully understand the nature of the M-R diagram.

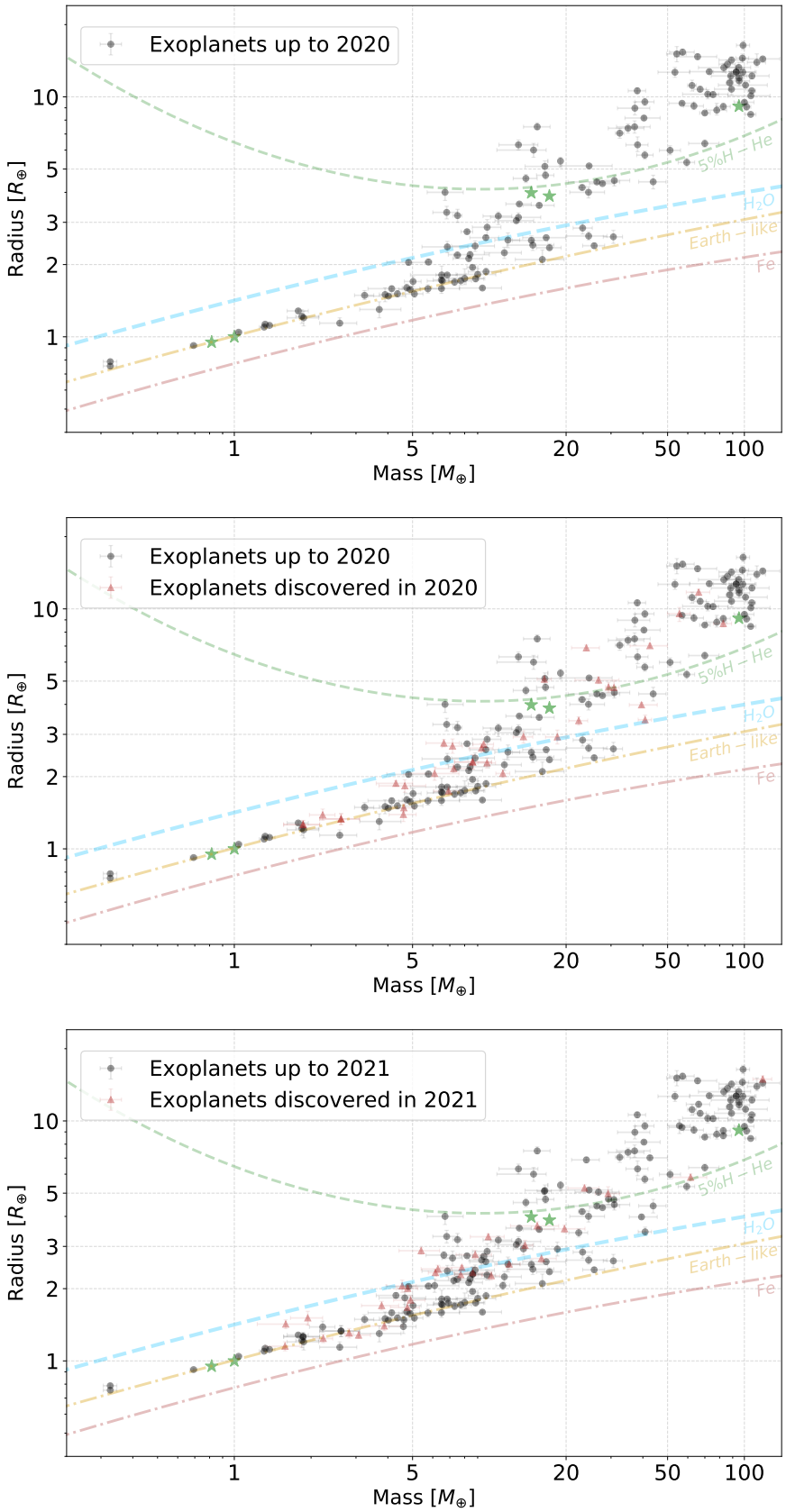

Figure 2.3: Evolution of the Mass-Radius diagram with exoplanets on the PlanetS catalog after the publication of Otegi et al. (2020a).

2.1.2.2 Extension to Gas Giant regime

In Otegi et al. (2020a) we only included exoplanets less massive than $120 M_\oplus$ because our analysis focused on the super-Earth population, sub-Neptune population, and the Neptune-Saturn transition. In the gas giant regime, the electron degeneracy pressure becomes significant and the radius of the planet has a very weak relationship with the mass, and in Otegi et al. (2020a) we wanted to study the regimes with strong M-R dependence. Initially, the PlanetS catalog followed the same criteria as in the paper and added only planets with masses below $120 M_\oplus$. In October 2021 we extended the catalog to planets of all masses using the NASA Exoplanet Catalog. To build a catalog that is as reliable and robust as possible, we followed similar criteria to the ones used in Otegi et al. (2020a):

- We filtered the data to consider only exoplanets with measurement uncertainties smaller than $\sigma_M/M = 25\%$ $\sigma_R/R = 8\%$. These thresholds correspond to the median uncertainties and make it possible to have the same impact on the density uncertainty.
- We selected the most up-to-date paper for as source of the planetary and stellar parameters for all the exoplanets.
- We replaced the mass determinations inferred by Stassun et al. (2017), where the host star masses and radii were replaced by the value derived from GAIA photometry and paralax, and with uncertainties clearly overestimated. Therefore, in the cases in where the NASA Exoplanet Archive selects this study as the reference paper, we replace them with the most updated mass estimate.
- We removed some mass determinations coming from Marcy et al. (2014) which corresponded to single transiting planet with a weak level of validation/confirmation. These exoplanets do not have any other mass estimates from other studies with measurement uncertainty smaller than $\sigma_M/M = 25\%$ $\sigma_R/R = 8\%$, and therefore were not included in our catalog.

The resulting M-R diagram is shown in Figure 2.4. We see a big dispersion in radius in the gas giant regime, which reflects the different compositions and incoming stellar irradiation of the exoplanets. We note that, opposed to the sub-Neptune regime where the PlanetS Catalog shows significantly different demography compared to the other catalogs when applying the same uncertainty limits, the gas giant population from the PlanetS Catalog is very similar to the NASA Exoplanet Archive when applying the same uncertainty limits. This is probably because the selection criteria that we impose for TTVs do not affect the gas giant regime. In the sub-Neptune regime, instead, there are numerous cases in which the other catalogs select mass determinations that are not considered 'robust' by Hadden & Lithwick (2017), and are replaced in the PlanetS Catalog.

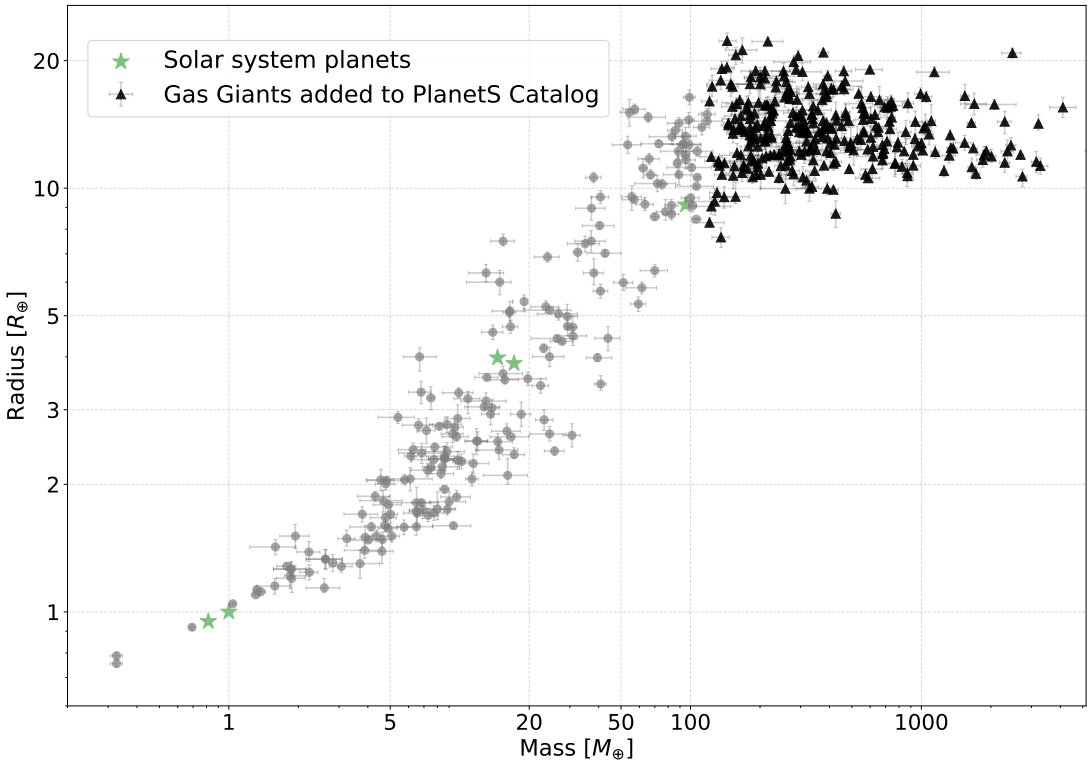

Figure 2.4: Mass-Radius diagram with exoplanets on the PlanetS catalog. The black triangles correspond to the planets more massive than 120$M_\oplus$ that have been added after the publication of Otegi et al. (2020a). The star shaped stars correspond to the Solar System planets.

2.1.2.3 Updated M-R relations

The revisited catalog published in Otegi et al. (2020a) contained 123 exoplanets. Since then, 118 exoplanets with masses below 120$M_\oplus$ have been added to the PlanetS Catalog, meaning that the number of well-characterized planets below this limit has doubled in the last two years. In addition, we added 276 exoplanets to extend the catalog to the Gas Giant regime. We next study what M-R relations can be fitted from this updated catalog.

Previous studies fitting M-R relations have divided the populations of super-Earths and sub-Neptunes using a mass as a threshold (e.g. Chen & Kipping 2017, who used M=2$M_\oplus$) or radius (e.g. Weiss & Marcy 2014, with R=1.5$R_\oplus$). In addition, a study by Fulton et al. (2017) found a lack of planets with radii between 1.5$R_\oplus$ and 2$R_\oplus$, suggesting that it may reflect the transition between these two populations. In the PlanetS catalog we see that even if planets smaller than 1.8$R_\oplus$ are clearly part of the rocky population, planets with a larger radius can be either rocky

Table 2.1: Comparison of the M-R relations derived from the updated PlanetS Catalog and other in the literature. "Small" planets refer to planets with masses below the gas giant regime.

Source	**Small* Planets**	**Gas Giants**	**Transition**
PlanetS Catalog - *Nov 2021*	$R=\begin{cases}1.06M^{0.30} & ,R<R_{H_2O}.\\ 0.58M^{0.67} & ,R>R_{H_2O}.\end{cases}$	R=12.5	M=98$M_\oplus$, R=12.5$R_\oplus$
Otegi et al. (2020a)	$R=\begin{cases}1.03M^{0.29} & ,R<R_{H_2O}.\\ 0.70M^{0.63} & ,R>R_{H_2O}.\end{cases}$	-	-
Bashi et al. (2018)	$R=0.85M^{0.55}$	$R=11.5M^{0.01}$	M=124$M_\oplus$, R=12.1$R_\oplus$
Chen & Kipping (2017)	$R=\begin{cases}1.01M^{0.28} & ,M<2M_\oplus.\\ 0.82M^{0.59} & ,M>2M_\oplus.\end{cases}$	$R=17.6M^{-0.04}$	M=130$M_\oplus$, R=14.4$R_\oplus$
Weiss & Marcy (2014)	ρ=2.43+3.39R ,R<1.5$R_\oplus$. $R=0.35M^{1.07}$,4$R_\oplus$>R>1.5$R_\oplus$.	-	-
Wolfgang et al. (2015)	$R=0.77M^{0.56}$	-	-

or volatile-rich. In Otegi et al. (2020a) we used the pure-water line to separate both populations as it provides a more physical criterion. Nevertheless, some exoplanets discovered in 2020 and 2021 covered the gap between the most massive "rocky" planets and the least massive "volatile-rich" ones, making it less clear what should be the transition between these two populations. It may be possible that the transition between rocky and volatile-rich populations depends on other parameters, as the insolation or stellar parameters. Since so far it is not clear what is the best criterion, we will follow Otegi et al. (2020a) and use the pure-water line to discriminate between "rocky" and "volatile-rich" exoplanets.

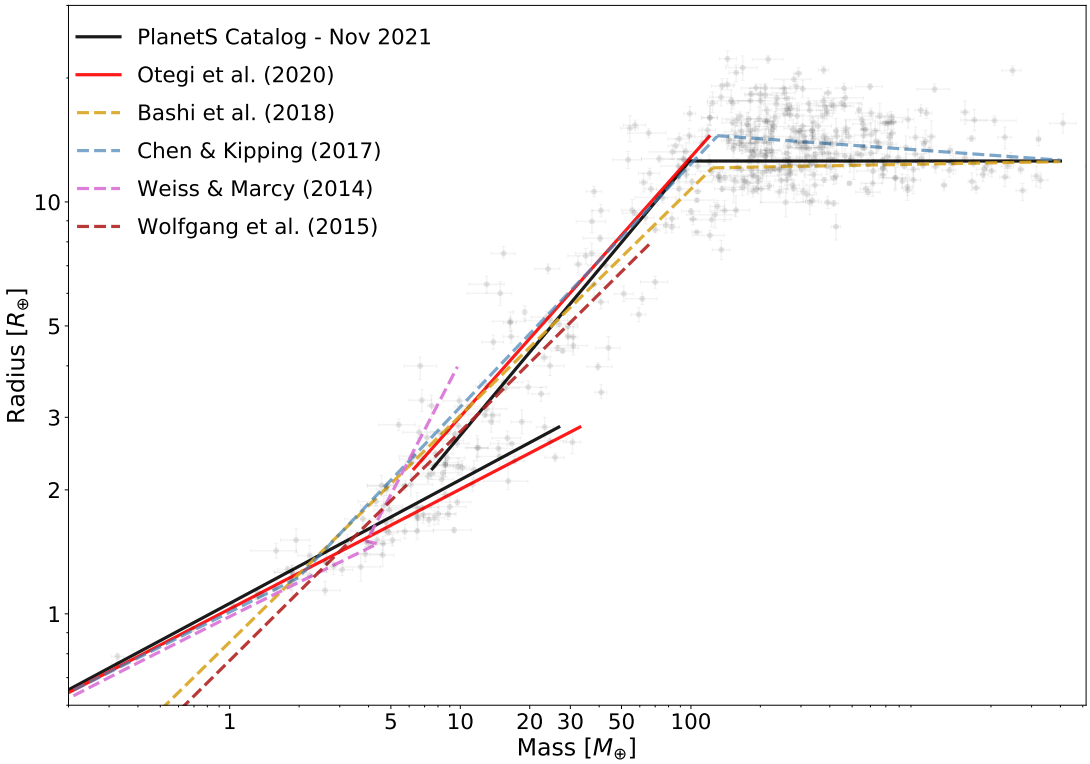

Figure 2.5: Comparison of the M-R relations in the literature with the one obtained from the PlanetS catalog. The observed exoplanets are taken from the PlanetS catalog.

The transition between "small" planets (less massive than gas giants) and giant planets has been extensively studied in the literature. Weiss et al. (2013) found a transition at ~150$M_\oplus$ based on visual estimates of the M-R and mass-density relations. Hatzes & Rauer (2015) analyzed similarly the changes in the slope of the mass-density relation and determined that the transition is at ~95$M_\oplus$. Then Chen & Kipping (2017) used a detailed forecasting model using a probabilistic M-R relation and located the transition at 130±22$M_\oplus$. Finally, Bashi et al. (2018) found two empirical regimes on the M-R relation with a transition at 124±7$M_\oplus$ and 12.1±0.5$R_\oplus$. To identify the transition in our sample, we use the piecewise regression Python package from DataDog[g], which applies piecewise regression with automated breakpoint detection. We find the transition at 98±19$M_\oplus$. Even if this value is under most of the previous results in the literature, it is in agreement within 1σ.

We fit the three populations using a total least squares method, in which observational errors on both dependent and independent variables are considered, and the results are the following:

[g]https://github.com/DataDog/piecewise

$$R = \begin{cases} (1.06 \pm 0.03)\ M^{(0.30\pm0.01)} & , \text{ if } R < R_{H_2O}. \\ (0.58 \pm 0.11)\ M^{(0.67\pm0.05)} & , \text{ if } R > R_{H_2O} \text{ and } M<98M_\oplus. \\ (12.5 \pm 0.60)\ M^{(0.00\pm0.02)} & , \text{ if } M > 98M_\oplus. \end{cases} \quad (2.1)$$

Table 2.1 and Figure 2.5 compare the obtained M-R relations from the updated PlanetS catalog and others found in the literature. We find that the obtained fits for the "rocky" and "volatile-rich" populations are in agreement with the results in Otegi et al. (2020a). Our M-R relations are also similar to the ones inferred in Chen & Kipping (2017), but they underestimate the masses of most rocky exoplanets by defining the transition from rocky to volatile-rich population for a mass of $2M_\oplus$. We see that the M-R relation by Weiss & Marcy (2014) does not match the observed population for planets with larger radii than $1.5R_\oplus$. In the gas giant regime, our relations are relatively close to the ones presented in Bashi et al. (2018). The dispersion of the observed population in this regime is very large, and there is a strong dependence with the coming irradiation, since more irradiated planets have hotter and more expanded atmospheres. Our fitted M-R relation in this regime follows the population of less irradiated gas giants.

2.1.3 Dependence of mass and radius on other parameters

To have a complete understanding of the demography of exoplanets, it is crucial to go beyond mass and radius and study whether there is any trend with other stellar or planetary parameters. It is important to note that the work in the present subsection is preliminary, and should be further developed. In this subsection, we focus on planets with masses below $100M_\oplus$, since we are interested in the transition from rocky to volatile-rich planets and the Neptune-Saturn transition. In addition, for massive planets there is no dependence between the mass and radius due to the fact that they are dominated by their atmosphere.

2.1.3.1 Quantifying the correlation with other parameters

To determine what parameters are more correlated with mass and radius, we use the distance correlation (Székely et al. 2007). The distance correlation is a measure of dependence between random vectors. It goes beyond Pearson's correlation because measures both linear and nonlinear association between the random vectors. Therefore, the distance correlation is zero if and only if the vectors are independent. It is obtained by dividing the distance covariance of the vectors by the product of the distance standard deviations:

$$\mathrm{dCor(X,Y)} = \frac{\mathrm{dCov^2(X,Y)}}{\sqrt{\mathrm{dVar(X)dVar(Y)}}} \tag{2.2}$$

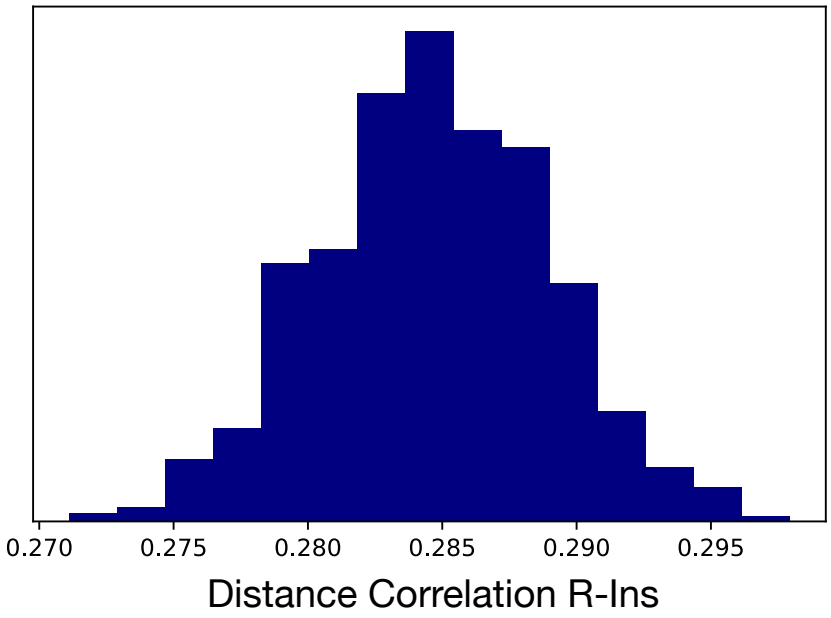

Figure 2.6: Distribition of distance correlation between planetary radius and insolation, with values corresponding to random samples from a Gaussian distribution centered in the measured value.

It ranges from 0 to 1, where 0 implies independence between X and Y and 1 implies that the vectors X and Y are equal. We will use the distance correlation to determine what parameters are more dependent on mass and radius.

We take into account the observational uncertainties using the following approach: instead of using the measured values, we use random samples from a Gaussian distribution centered in the measured value. For each pair of parameters, we calculate 1000 distance correlation from the random samples. An example of the distribution of distance correlations obtained by applying this method to the pair of parameters 'Planet Radius-Insolation' is shown in Fig 2.6. We see that the distributions are relatively narrow, meaning that taking into account the uncertainties does not strongly impact the computed distance correlation. In the example shown in the figure, the standard deviation is ~0.002, and it is similar when we perform the same analysis with other parameters. We also note that this standard deviation is smaller when we make the computation with respect to the planetary radius than with respect to the planetary mass. The distance correlation values shown from this point correspond to the median of the distribution.

Figure 2.7 shows the distance correlations obtained for the studied planetary and stellar parameters with respect to the planetary mass and radius. The parameters we evaluate are the planet insolation, the orbital period, the stellar mass, stellar radius, effective temperature, stellar metallicity, and stellar age. We study four different regimes of the M-R diagram: the 'rocky' and 'volatile-rich' populations as defined in Otegi et al. (2020a) (i.e. separated by the pure-water line), a Neptune-like population of planets with radius between $2R_\oplus$ and $5R_\oplus$, and the whole sample of planets with masses below $100M_\oplus$. We note that there is a overlap between the 'volatile-rich' and 'Neptune-like' population, but we aim to study these two populations separately. First, we realize that all the parameters tend to be more correlated with the planetary radius than with the mass, reflecting the sensitivity of the atmospheric layer. Then we note that the 'rocky' and the Neptune-like populations are the ones with the strongest correlations. In the 'rocky' population, there is a very strong dependence on the orbital period. This is probably explained by an observational bias (planets with very low masses can only be detected when they orbit close to the star). The correlation of stellar mass, radius, and effective temperature with mass and radius in this regime can also be strongly affected by observational bias, since small planets are more easily detectable around small stars. We note that the values obtained for stellar mass and radius are very similar, since most of the sample contains main sequence stars. In the Neptune-like population, the insolation is the parameter with the strongest correlation with the mass, and the stellar metallicity with the radius.

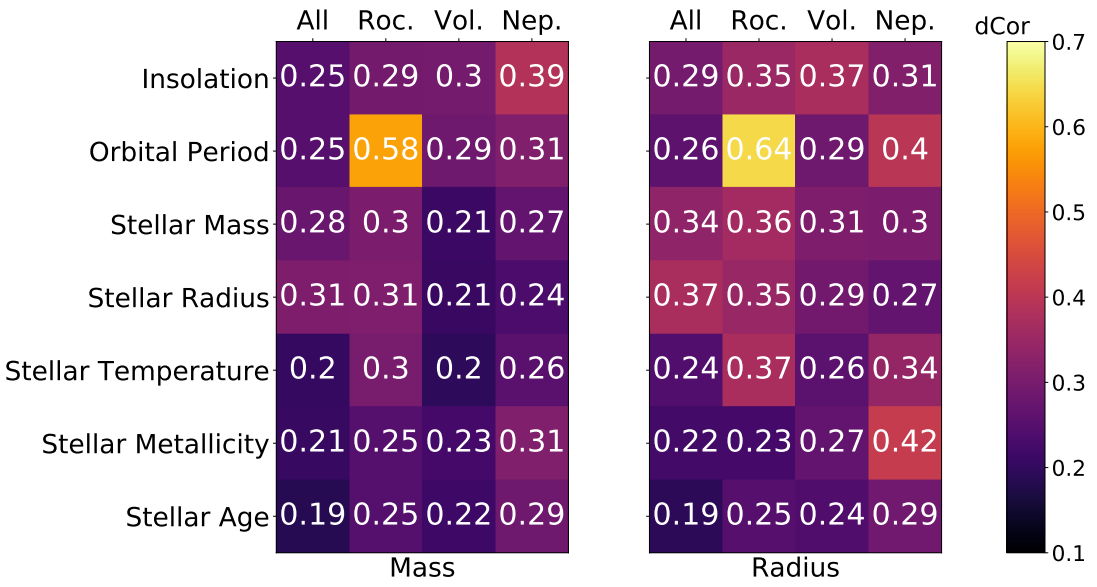

Figure 2.7: Distance correlations for several planetary and stellar parameters with respect to planetary mass (left) and radius (right). Each column corresponds to a given population of planets: 'All' refers to all planets up to $100M_{\oplus}$, 'Roc.' and 'Vol' refer to the 'rocky' and volatile-rich populations as defined in Otegi et al. (2020a), and 'Nep.' corresponds to Neptune-like planets with radii between $2R_{\oplus}$ and $5R_{\oplus}$. The colors reflect the computed distance correlation.

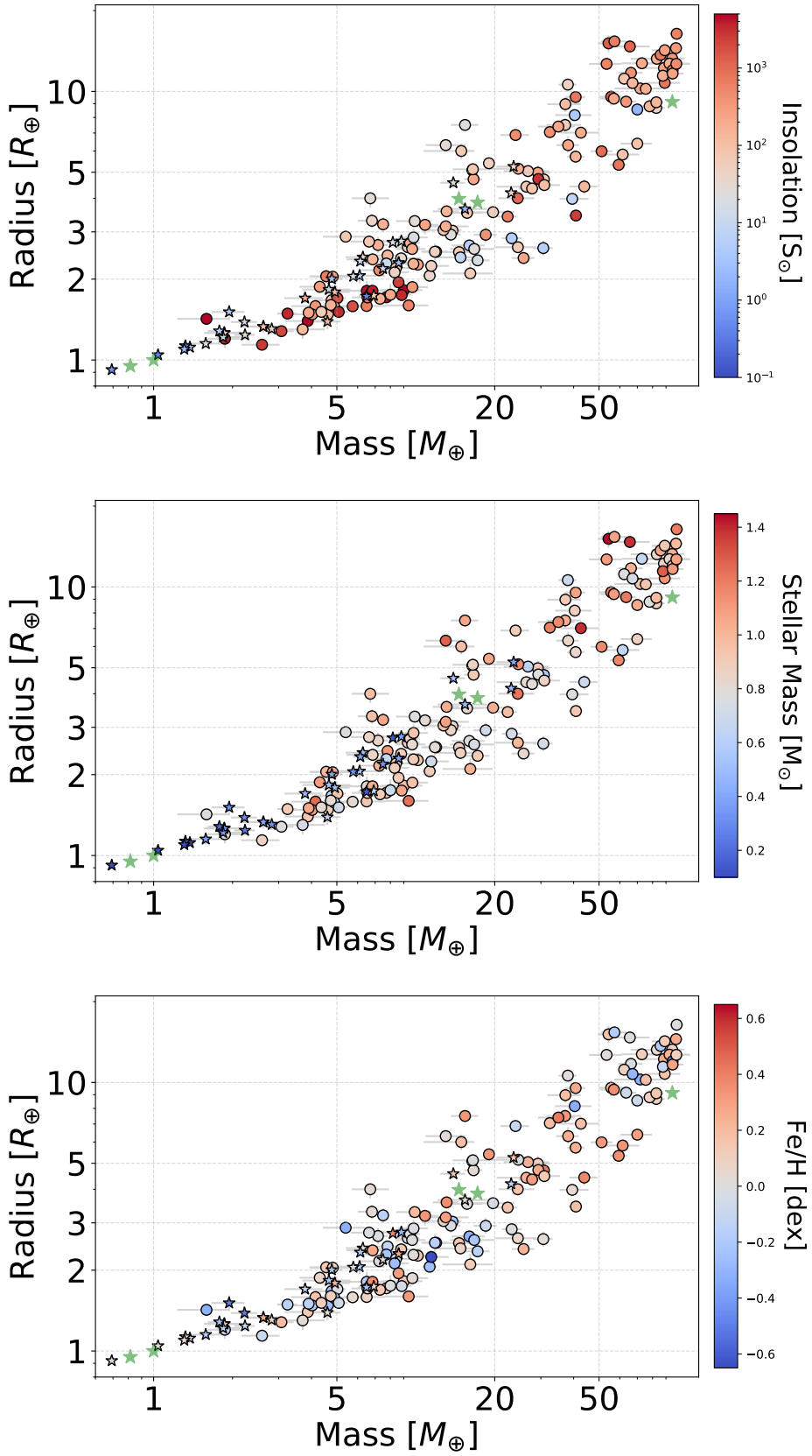

Figure 2.8: M-R diagrams of exoplanets up to 100$M_\oplus$ from the updated PlanetS catalog showing the dependence with insolation (top), stellar mass (middle) and metallicity (bottom). Exoplanets orbiting M-dwarfs are indicated with star shaped dots.

It is interesting to note that the insolation is more correlated with mass than the orbital period in the 'Neptune-like' and 'Volatile-Rich' regimes. In addition, insolation is less affected by observational biases, and it is a more physical parameter. In Figure 2.8 we can see the M-R diagrams up to $100M_\oplus$ from the updated PlanetS catalog showing the dependence with insolation, stellar mass, and metallicity. The presence of M-dwarfs can introduce a bias, since they are not expected to host giant planets. We therefore use a different marker for exoplanets orbiting M-dwarf to see whether the observed features remain when excluding them. A visual inspection of the figure shows that the correlations with insolation, stellar mass, and metallicity remain when not taking into account exoplanets around M-dwarfs. We also find that among planets with masses below $10M_\oplus$ the denser planets are significantly more irradiated than the population of exoplanets with a volatile envelope (excluding the Trappist-1 planets), although there is no clear difference between these two populations in terms of stellar mass. This suggests that these strongly-irradiated exoplanets have probably photoevaporated their H/He by the high-energy radiation from the host star (e.g. Owen & Wu 2013; Lopez & Fortney 2013; Jin et al. 2014; Zeng et al. 2017; Jin & Mordasini 2018). We also see that these dense planets with high incoming irradiation up to $10M_\oplus$ tend to orbit around more metal-rich stars than the planets above in the M-R diagram. It could be explained under the core-accretion formation theory, since planets forming around metal-rich stars have more solid material in the disk to accrete. So it could indicate that these rich protoplanetary disks tend to form denser planets.

We also note that most of the massive planets in the sample with masses ranging from $50M_\oplus$ and $100M_\oplus$ are highly irradiated. Since the formation of such massive planets close to the star is unlikely, this trend could be a hint of migration of planets formed further out in the disk. However, observational biases could also play a role, since it is easier to find massive planets close to the star. Therefore, more data is necessary to know the true ratio of massive planets that are highly irradiated with respect to weakly irradiated. Finally, we also find a positive correlation between planetary mass and radius with stellar metallicity. While it is well-known that the occurrence rate of giant planets and planets with masses between several Earth masses and $40M_\oplus$ increases with stellar metallicity (e.g. Courcol et al. 2016), it was not investigated in the sub-Saturn regime between $40M_\oplus$ and $100M_\oplus$. The fact that the more massive planets of the volatile-rich population tend to orbit around metal-rich stars is consistent with the core accretion model for giant planet formation (see review by Helled et al. (2014) and references therein). However, a more systematic analysis of this result and the corresponding observational biases is requested before robust conclusions are made.

2.1.3.2 Joint analysis of the M-R-I-Z distribution

To go beyond the naked-eye analysis of the M-R diagram, we also perform a joint fit of mass and radius with other parameters. More specifically, we constrain a 4D joint mass-radius-insolation-metallicity distribution of exoplanets using the PlanetS catalog using a probabilistic power law and hierarchical Bayesian modeling similarly as in Chen & Kipping (2017) and Neil & Rogers (2020). We model the relation with the following power-law:

$$\mathrm{R}=\mathrm{C}\mathrm{M}^{\alpha}\mathrm{I}^{\beta}\mathrm{Z}^{\gamma} \tag{2.3}$$

where I and Z correspond to the insolation and metallicity, respectively. The corresponding probabilistic relation can be written as:

$$R \sim \mathcal{N}(\mu = \mathrm{C}\mathrm{M}^{\alpha}\mathrm{I}^{\beta}\mathrm{Z}^{\gamma}, \sigma = \sigma_R) \tag{2.4}$$

We use hierarchical Bayesian modeling to constrain the distribution. The difference with respect to the usual Bayesian method is that in the hierarchical there are two sets of parameters: the hyperparameters and the local parameters. The hyperparameters describe the overall ensemble of properties and sit on top of the local parameters that describe the properties of the individual data. In our case the local parameters are the true M,R,I,Z, and the hyperparameters are those defining the power law. For more details about the methodology, we refer to Chen & Kipping (2017), which uses a similar approach to fit a two-dimensional M-R relation. To fit the model to the data, we use the Python implementation of the Stan statistical software package (Carpenter et al. 2017). It employs the No-U-Turn Sampler MCMC algorithm. The strength of this algorithm resides in its efficiency to handle large dimensional spaces, which makes it ideal to evaluate hierarchical Bayesian models.

Table 2.2: Results of the joint fit of mass, radius, insolation and stellar metallicity using a probabilistic power law with hierarchical Bayesian modelling. Dash lines correspond to not significant correlations.

$R=CM^{\alpha}I^{\beta}Z^{\gamma}$	**All**	**Rocky**	**Vol. Rich**
C	0.97 ±0.02	1.04 ±0.02	0.65 ±0.04
α	0.665 ±0.008	0.28 ±0.01	0.67 ±0.02
β	-0.067 ±0.005	-	-
γ	0.04 ±0.006	-	0.04 ±0.004

We fit the power law in the different regimes of the M-R diagram: the 'rocky' and 'volatile-rich' regimes defined in Otegi et al. (2020a) and the whole exoplanet sample with masses below 100$M_\oplus$. The results of the fit are shown in Table 2.2. When we take into account the entire sample we find significant negative exponents for the insolation (β=-0.067±0.005) and the metallicity (γ=0.04±0.006). In the 'rocky' population, instead, the fit gives non-significant β and γ parameters, indicating a no significant correlation between the insolation and metallicity with the radius. In the volatile-rich population, the obtained β is not significant while γ is (γ=0.04±0.004). This is in disagreement with the results shown in Fig 2.7, where the correlations of insolation and metallicity with radius are found to be stronger for the 'rocky' population. We consider that more data and systematic analysis of these results, including observational biases, are required to understand the relation between planetary and stellar properties.

2.2 Similarity of Multi-Planet Systems

The rich diversity observed in exoplanets also extends to the architecture of multi-planetary systems. The architecture of any planetary system is the result of all the physical processes that lead the system to its present state, and explaining the wide diversity of observed system architecture is still an open problem. One observed trend in the diversity of planetary systems is the so-called 'peas in the pod', i.e. multi-transiting systems tend to have planets with similar sizes and to be regularly spaced. Understanding these trends is crucial since they may provide hints about underlying physical processes and put constraints on formation theories.

The 'peas in the pod' pattern has extensively been studied in terms of radius (Weiss et al. 2018; Zhu 2020; Weiss & Petigura 2020) and period ratios (Weiss et al. 2018; Jiang et al. 2020). However, it has been only explored in terms of mass by Millholland et al. (2017), who used a rather small sample of planets with masses below 50$M_\oplus$ detected by TTVs. After this paper was published, the number of well-characterized systems strongly increased, mainly thanks to TESS. In Otegi et al. (2021b) (see next page), we use the PlanetS catalog to perform an in-depth analysis of the similarity in mass and radius of multi-planetary systems. We also investigate the 'peas in the pod' pattern in terms of density, which was not done by other papers on the topic. Finally, we also explore how different assumptions and setups typically used in the literature when studying the 'peas in the pod' pattern in terms of the period ratio affect the results.

Astronomy & Astrophysics manuscript no. output
December 6, 2021

The similarity of multi-planet systems

J. F. Otegi[1,2], R. Helled[2], and F. Bouchy[1]

[1] Observatoire Astronomique de l'Université de Genève, 51 Ch. des Maillettes, – Sauverny – 1290 Versoix, Switzerland
[2] Institute for Computational Science, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland

December 6, 2021

ABSTRACT

Previous studies using Kepler data suggest that planets orbiting the same star tend to have similar sizes. However, due to the faintness of the stars, only a few of the planets were also detected with radial velocity follow-ups, and therefore the planetary masses were mostly unknown. It therefore yet to be determined whether planetary systems indeed behave as "peas in a pod". Follow-up programs of TESS targets significantly increased the number of confirmed planets with mass measurements, allowing for a more detailed statistical analysis of multi-planet systems. In this work we explore the similarity in radii, masses, densities, and period ratios of planets within planetary systems. We show that planets in the same system that are similar in radii could be rather different in mass and vice versa and that typically the planetary radii of a given planetary system are more similar than the masses. We also find that a transition in the "peas in the pod" pattern for planets more massive than ~ 100 $M_\oplus$ and larger than ~ 10 $R_\oplus$. Planets below these limits are found to be significantly more uniform. We conclude that other quantities like the density may be crucial to fully understand the nature of planetary systems and that, due to the diversity of planets within a planetary system, increasing the number of detected systems is crucial for understanding the exoplanetary demographics.

1. Introduction

The number of discovered exoplanets has increased to over 4000 thanks to various ground-based and space-based surveys, among which NASA's Kepler mission (Borucki et al. 2010) stands out with more than 2300 discovered planets. This large sample of exoplanets has allowed for detailed statistical analyses of hundreds of multi-planetary systems (e.g., Lissauer et al. 2011; Latham et al. 2011; Lissauer et al. 2012; Rowe et al. 2014; Lissauer et al. 2014), which have pushed our knowledge on aspects such as physical compositions (e.g., Carter et al. 2012; Hadden & Lithwick 2014), orbital eccentricities and inclinations (e.g., Fang & Margot 2012; Fabrycky et al. 2014; Xie et al. 2016; Van Eylen et al. 2019). Although these properties in multi-planetary systems have triggered multiple studies of their formation and evolution histories (e.g., Hansen & Murray 2013; Steffen & Hwang 2015; Malhotra 2015; Ballard & Johnson 2016; Mills et al. 2016; Owen & Campos Estrada 2020), our understanding of the diversity of planetary systems is still incomplete.

It was suggested by Weiss et al. (2018) that planetary systems are like "peas in a pod", i.e. multi-transiting systems tend to have planets with similar sizes and to be regularly-spaced. Weiss et al. (2018) used a large sample of Kepler multi-planetary systems whose parameters were refined by the California-Kepler Survey (CKS Petigura et al. 2017), and employed a series of bootstrap tests to quantify the significance of the similarity between sizes and conclude that the observed distribution could not be explained by random resampling. While Zhu (2020) argued that the result by Weiss et al. (2018) is affected by observational biases, a strong evidence that the observed intra-system uniformity has an astrophysical origin has been confirmed by other studies (e.g., Weiss & Petigura 2020; Murchikova & Tremaine 2020; Gilbert & Fabrycky 2020; Jiang et al. 2020; Mishra et al. 2021). A similar statistical approach was presented by Millholland et al. (2017), who used a sample of planets with masses characterized by transit timing variations (TTVs) from Hadden & Lithwick (2017) and found that planets orbiting the same star also tend to have similar masses. The similarity in mass they found was based on a rather restricted sample of planetary systems. They analyzed a sample of 37 systems with masses below $50M_\oplus$ derived by the TTV. Due to the faintness of the stars targeted by Kepler, only a small fraction of the detected planets were suitable for radial velocity (RV) follow up.

After the end of the primary Kepler mission, NASA's K2 mission continued to discover transiting planets orbiting stars near the ecliptic plane (Howell et al. 2014). Compared to Kepler, the K2 mission covered more sky, observed more diverse stellar populations, and focused on brighter targets which are more amenable for radial velocity (RV) follow-up observations. Later, the Transiting Exoplanet Survey Satellite (TESS) was launched in 2018 to survey 85% of the sky for transiting exoplanets around bright stars (Ricker et al. 2015). RV follow-up programs have triggered for a rapid expansion of the number of confirmed planets with a mass measurement, which has allowed to perform an in-depth analysis of the "peas-in-a-pod" pattern in terms of mass with a larger and a more diverse sample of planetary systems.

In this paper we revisit several aspects of the "peas in a pod" pattern (i.e. the radius, mass and period ratio correlation) in detail by accounting for observational biases and using different exoplanet catalogs. The paper is organized as follows: In section 2 we analyze the uniformity of planetary systems in mass and radius. Then, in section 3 we study the uniformity in the period spacing. Our conclusions are summarized in section 4.

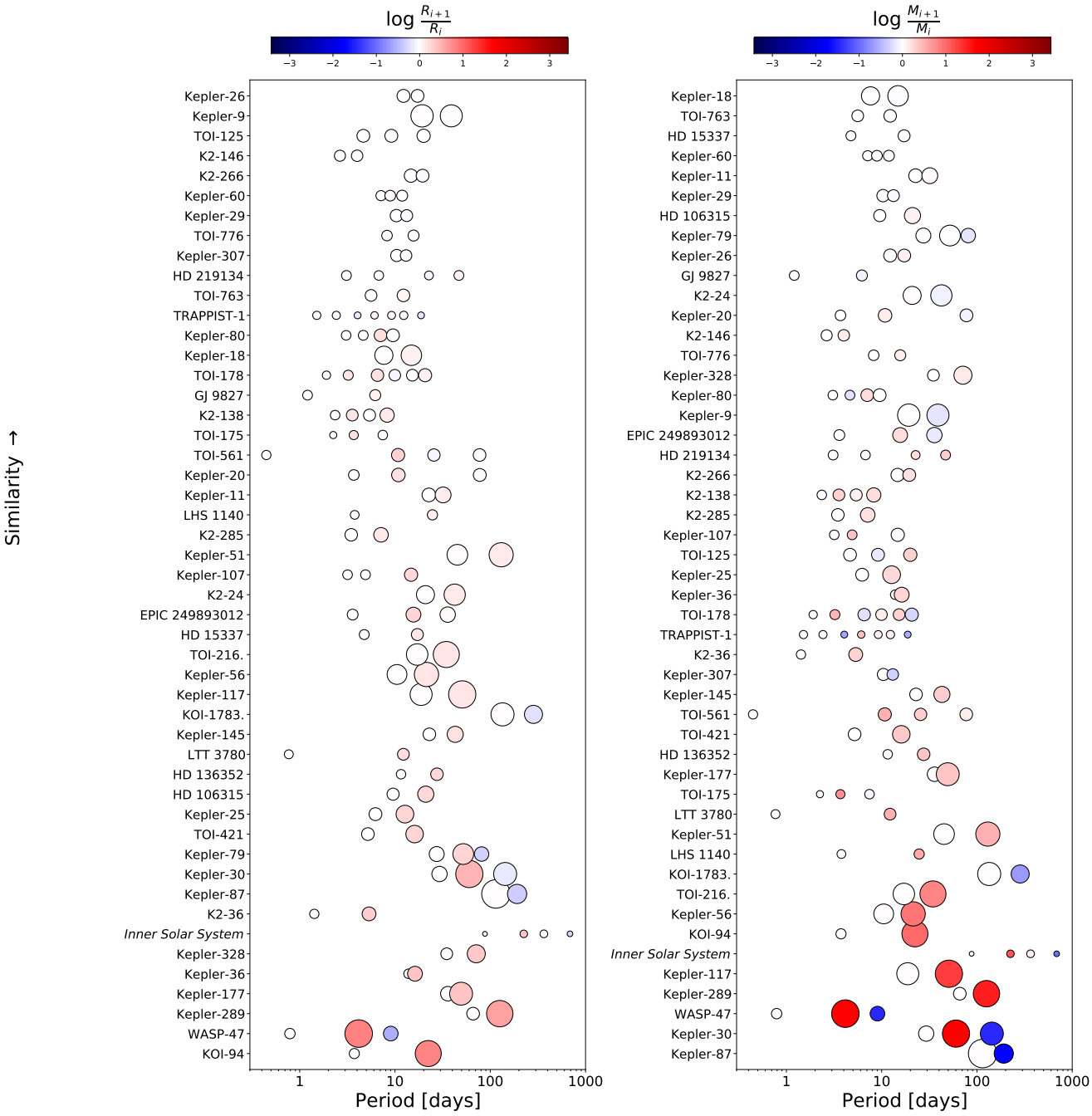

Fig. 1: Orbital architecture of the 48 multi-planet systems in our sample, with measurement uncertainties smaller than $\sigma_M/M = 50\%$ and $\sigma_R/R = 16\%$ in addition with the inner Solar System. In the left panel the color of the points represent the logarithm of the radius of a planet divided by the radius of the previous one, and the systems are ordered by similarity in radius as defined in Section 2.2. The same is represented in the right panel for the mass. The size of the circles are proportional to the radii of the planets.

2. Similarity in mass, radius and density

2.1. Exoplanet sample

We use the NASA Exoplanet Archive[1] (Akeson et al. 2013) on August 2021, since it is the most up-to-date catalog. We exclude the less accurate data by considering only planets with measurement uncertainties smaller than $\sigma_M/M = 50\%$ and $\sigma_R/R = 16\%$, which leaves us with 144 planets part of 48 multi-planet systems. We note that the limits on the mass and radius uncertainties of the planets in our sample are twice of the ones used in Otegi et al. (2020), where we presented an updated exoplanet catalog based on reliable mass and radius measurements of transiting planets with uncertainties smaller than $\sigma_M/M = 25\%$ $\sigma_R/R = 8\%$. This is because in Otegi et al. (2020) we aimed to build a catalog with as much as possible accurate mass and radius measurements to derive a mass-radius relationship, while in this work we prefer to relax the limits in order to include more multi-planet systems. More specifically, the catalog presented in Otegi et al. (2020) contains 23 multi-planet systems, which is half of the sample used here.

[1] exoplanetarchive.ipac.caltech.edu

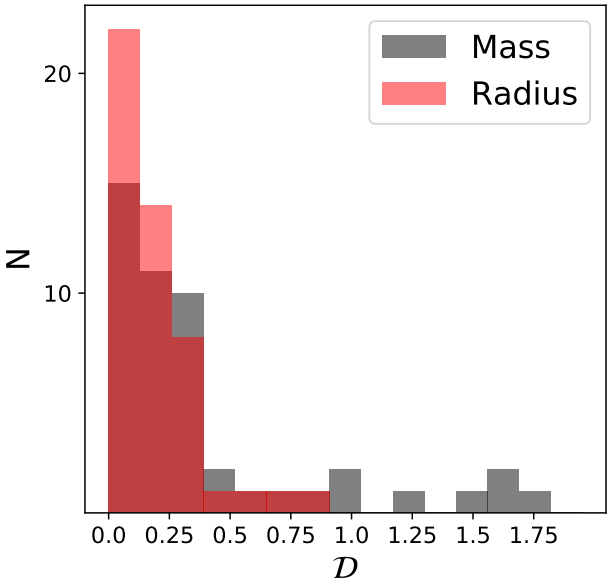

Fig. 2: Histograms of the distances $\mathcal{D}_R$ (red) and $\mathcal{D}_M$ (grey) used to quantify the similarity of a planetary system as defined in Equation 1.

The impact of the choice of the limits on the uncertainties in mass and radius is discussed in Section 2.7. 27 of the systems in our sample have been characterized via TTVs, and 21 by RVs. Out of the 37 systems used in Millholland et al. (2017), 16 have been included in our catalog while 21 systems have been discarded due to large uncertainties or for not being considered "robust" in Hadden & Lithwick (2017). Among the RVs, only one system in our catalog (HD 219134) is also present in the study by Wang (2017), since the rest of their systems are composed by non-transiting planets.

Clearly, the results depend on the available data and could be affected by selection biases from the radial velocity, transit timing variations or transit methods. This effect can be reduced using homogeneous samples of planets with well-determined selection biases. This was done by Weiss et al. (2018) and Millholland et al. (2017), which use only Kepler planets and planets characterized by TTVs, respectively. However, in this work we aim to have a large and diverse sample, which results in a less homogeneous sample.

2.2. Similarity Metric

Several approaches have been taken to analyze the architectures of planetary systems. For example, Kipping (2018) proposed a model to define the entropy of a planetary system's size-ordering; Gilbert & Fabrycky (2020) suggested different descriptive measures to characterize the arrangements of planetary masses, periods, and mutual inclinations. Alibert (2019) defined a new metric to infer the similarity between two planetary systems, which was based on representing the planets of the systems as points on a logarithmic radius-period plane, and spreading the points with a Gaussian kernel whose weights correspond to the planet masses, a similar approach without using a Gaussian kernel estimation of the probability distribution was used by Bashi & Zucker (2021). In this work, we use an approach, similar to the one used in Millholland et al. (2017), where we quantify the similarity of the systems by considering the distance in the logarithmic space, which can be expressed as:

$$\mathcal{D}_M = \sum_{\substack{i=1 \\ P_i<P_{i+1}}}^{N_{pl}-1} \left| \log\frac{M_{i+1}}{M_i} \right| \Big/ \; N_{pl} - 1 \; , \tag{1}$$

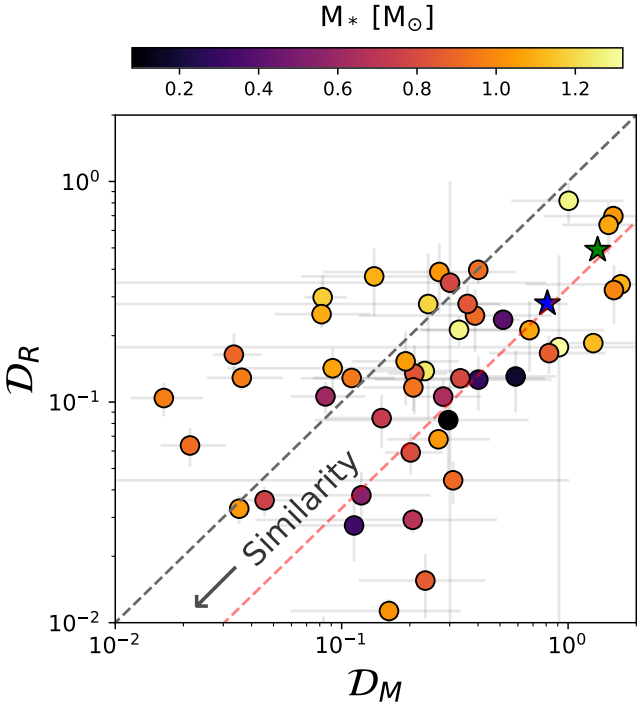

Fig. 3: Similarity in radius $\mathcal{D}_R$ against similarity in mass $\mathcal{D}_M$, as defined in Equation 1. The color of the dots represent the mass of the host star. The black dashed line is the 1:1 line, and the red dashed lines corresponds to $\mathcal{D}_M$= 3x$\mathcal{D}_R$, which would be expected if density tends to be uniform. The green star corresponds to the Solar System and the blue one to the inner Solar System.

with a equivalent expression for $\mathcal{D}_R$. $\mathcal{D}$ is computed by summing the distances in logarithmic space of adjacent planets, normalized by the number of pairs N_p-1 in order to remove the dependency on the number of planets in the system. Note that for this metric lower values correspond to more similarity. We can also consider the global distance in the log M_p-log R_p space as an indicator of the similarity of a planetary system. Similarly to the expression for $\mathcal{D}_M$, the global distance can be given by:

$$\mathcal{D} = \sum_{\substack{i=1 \\ P_i<P_{i+1}}}^{N_{pl}-1} \left[\left(\log\frac{M_{i+1}}{M_i}\right)^2 + \left(\log\frac{R_{i+1}}{R_i}\right)^2 \right]^{1/2} \Big/ \; N_{pl} - 1 \; . \tag{2}$$

Table A.1 lists the radius, mass, and global similarities of all the systems in our sample based on this metric. We also list the results for the planets in the Solar System, as well as for the inner and outer parts of it. In order to ensure that our results do not strongly depend on the metric, we compare the results using the variance in log-space of mass and radius and find that

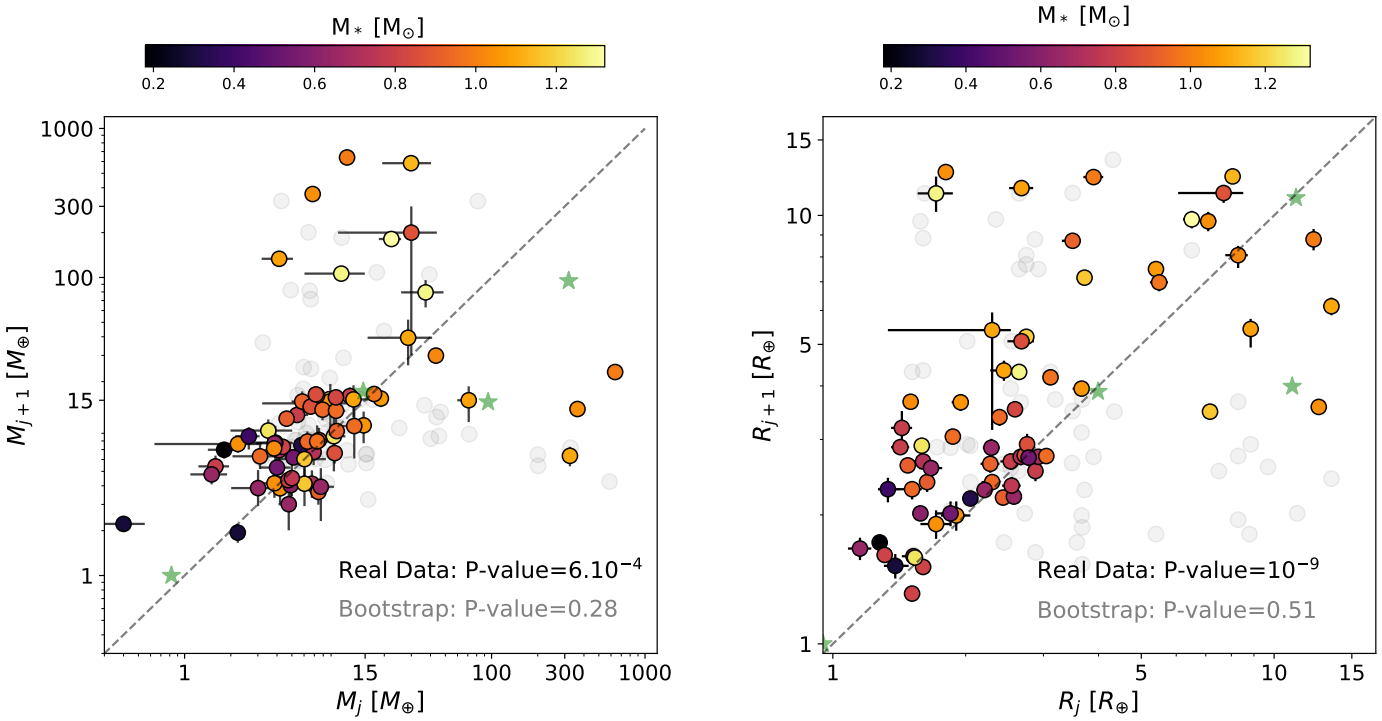

Fig. 4: The mass (left) or radius (right) of a planet against the mass or radius of the next planet (farther from the star). The colors of the dots represent the mass of the host star. The green star shaped markers correspond to the Solar System planets and the gray points to a bootstrap trial.

the inferred ranking is nearly identical. Kepler 60 is found to be the most uniform system, which includes three rocky planets below the radius valley. The next four similar uniform systems include pairs of rocky exoplanets (L 98-56) or sub-Neptunes (Kepler 29, TOI 763, Kepler 26). The less uniform systems are Kepler 87, WASP 47, Kepler 289, Kepler 30, and Kepler 117. However, we may suspect that not all the planets in the systems are detected. The inherent detection limits of the radial velocity, transit timing variation, and transit methods do not allow to have a full picture of the orbital architecture. For each system we may be missing relatively small planets, or planets at large orbital distances. The detection of these missing planets would increase the metric $\mathcal{D}$, making them less similar. The lack of the missing planets may explain the result that the Solar System is the fifth less similar system. Interestingly, even when we consider only the inner Solar System, or only the outer Solar System, it is always found to be least similar planetary system. The weak similarity of the outer Solar System is explained by the difference between Uranus and Neptune with the gas giants, while the weak similarity of the terrestrial planets is mainly due to Mercury.

Figure 1 shows the orbital architecture of the multi-planet systems in our sample. The color of the points represent log $\frac{R_{i+1}}{R_i}$ (left panel) and log $\frac{M_{i+1}}{M_i}$ (right panel), and the planets are ordered by the similarity $\mathcal{D}_R$ in radius (left panel) and by the similarity $\mathcal{D}_M$ (right panel).

Before making a proper quantitative analysis, there are a few architectural features that appear in Figure 1. First, we see that planetary systems in our sample tend to be more similar in radius than in mass. Figure 2 shows the histogram of $\mathcal{D}_R$ and $\mathcal{D}_M$ (the full list of the values can be seen in the appendix in Table A.1), showing that the values $\mathcal{D}_R$ are much lower than those of $\mathcal{D}_M$. We clearly see that there are significantly more systems with $\mathcal{D}_R$<0.25 (35 systems) than with $\mathcal{D}_M$<0.25 (25 systems). This is not caused by the larger range in planetary masses than in planetary radii since the metric we use is insensitive to the size of the range. This could be explained if the density tends to be similar in a planetary system. Since the density is three times more sensitive to radius variations than to mass variations, it is expected to have a stronger uniformity in radius than in mass. More specifically, in this case we would expect the systems to be three times less similar in mass than in radius. We find that the median of $\mathcal{D}_M$ is 1.75 times larger than the one of $\mathcal{D}_R$. Figure 3 shows $\mathcal{D}_M$ against $\mathcal{D}_R$, with a line corresponding to $\mathcal{D}_M$= 3x$\mathcal{D}_R$. It is interesting to note that the $\mathcal{D}_M$= 3x$\mathcal{D}_R$ line does not fit the observed population. However, as we discuss below this could be a result of the the uniformity in density. The uniformity in density within a planetary system is further discussed in Section 2.4. We also find that the distribution of $\mathcal{D}_M$ extends to higher values. The tail seen in the $\mathcal{D}_M$ distribution corresponds to five particular systems that include a giant planet and a sub-Neptune.

Second, we find that generally the planets that are most similar in radius do not correspond to the ones that are most similar in mass, as illustrated in Figure 3. Even if there is a weak correlation between the similarities in mass and radii, the dispersion is very large, especially among the most similar systems. It must also be noted that there is a clear dependence on stellar mass: low-mass stars are more concentrated in the lower part of Figure 3, indicating that planets orbiting low mass stars tend to be more similar in radius than in mass. In addition, M-dwarfs closely follow the $\mathcal{D}_M$= 3x$\mathcal{D}_R$ relation. Since low-mass stars tend to host low-mass planets (Lozovsky et al. 2021), this behavior might hint that the physical processes

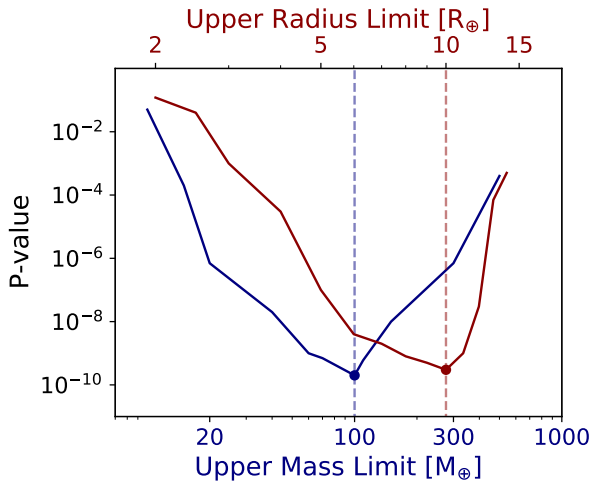

Fig. 5: P-value against the upper mass limit (blue) and upper radius limit (red) of the exoplanet sample. The dashed lined indicate the mass and radius at the minimum.

during planet formation tend to produce planets of similar density in planetary systems around low-mass stars, leading to a similarity in radius that is three times stronger than in mass. The correlations with the host star mass are further discussed in Section 2.6.

2.3. Mass and Radius correlation

Figure 4 shows the mass/radius of a planet against the mass/radius of the next planet. First we note that most of the pairs are above the 1:1 line, especially for the radius. This indicates that larger planets are found at larger orbital periods, which is in agreement to the findings in several other papers (e.g., Ciardi et al. 2013; Millholland et al. 2017; Kipping 2018; Weiss et al. 2018). Even if at long periods it is easier to detect larger/more massive planets, Helled et al. (2016) showed that the observed correlation between planetary radius and orbital period remains even after removing the effect of observational biases and, therefore, it is unlikely to be caused solely by observational biases. Using the Pearson correlation test we find that there is a clear correlation for both mass and radius with P-values of 6.10^{-4} and 10^{-9}, respectively (we note that the calculation of the Pearson coefficients do not include the uncertainties in mass and radius). This indicates that adjacent planets in a multi-planet system are likely to have similar masses and radii. We note that the P-value is significantly smaller for the planetary radii, suggesting that the "peas in the pod" pattern is less strong when it comes to the planetary mass. This is also clear from Figure 1, where the color of dots is more intense on the panel corresponding to the mass than in the panel corresponding to the radius.

It is also interesting to note that there is a clear transition in the uniformity of systems in both mass and radius plots. Figure 4 also shows a transition in the uniformity of planets around 25-100$M_\oplus$ and 5-10$R_\oplus$, i.e. systems below this limits tend to be very uniform and above they are not correlated. We then study how does the 'peas in the pod' pattern depend on the mass range covered by the exoplanet sample. Figure 5 shows the dependence of the P-value on the mass and radius limits applied to the planetary sample. We find that removing the systems containing pairs of planets with masses higher than 100$M_\oplus$ leads to a much stronger correlation. When we impose limits for the mass below 100$M_\oplus$ the P-value increases. However, since samples with less number of points have higher P-values, this could be partly explained by the lower number of pairs in the sample. We therefore explore the impact of the number of pairs when computing the P-value and the R-value. To do so, we randomly remove pairs from the initial sample, and compute the P-value and R-value. We repeat the process 1000 times, and plot the median of the distribution and the 1-sigma error (shown in Figure A.2). We find that randomly removing pairs leads to a rapid increase of the P-value. When changing the upper mass limit of 100$M_\oplus$ to an upper mass limit of 15$M_\oplus$, the number of pairs changes from 58 to 46 and the P-value from 4.10^{-10} to 3.10^{-4}. However, when we randomly remove 12 pairs from an initial sample of 58 we would expect a P-value between 3.10^{-8} and 2.10^{-6}. Consequently, the increase in the P-values when lowering the mass limit to planets below 100$M_\oplus$ cannot be completely explained by the decrease of number of points. We therefore identify a change of tendency in mass at 100$M_\oplus$. This result agrees with the analysis performed by Wang (2017) using a sample of 27 non-transiting multi-planetary systems derived by RVs with minimum masses only.

We repeat the same analysis with the radius, and investigate the dependency of the P-value and the R-value on the radius limit applied to the planetary sample. We find that the results are nearly identical to the ones inferred when applying limits on the mass range. This is because the excluded planets are mostly the same. We find that when we exclude giant planets with radii larger than 10$R_\oplus$ the 'peas in the pod' pattern becomes more significant, but when imposing lower values to the upper limit of the radius there is an increase on the P-value that cannot be explained solely by the lower number of pairs. We conclude that the change in tendency occurs at a radius of ~10$R_\oplus$. We note that both transition points in mass and radius correspond to Saturn-like planets. It is clear that planetary systems tend to consist of planets with similar masses and radii, except for planets more massive or larger than Saturn. Even if current data indicate a minimum of the P-value at around ~100$M_\oplus$ and ~10$R_\oplus$, the precise location of these transitions should be analyzed in detail in future studies.

2.4. Density correlation

We find that planetary systems tend to be more similar in radius than in mass. It is therefore interesting to investigate the similarity in planetary density. If the density tends to be uniform within a planetary system, systems would be more similar in radius than in mass simply because of the stronger dependence of the density on the radius. Figure 5 shows the density of the planets in our sample against the densities of the next planet farther from the star. We find that a very strong correlation between densities of adjacent planets (with P-value of 4.10^{-9}) and, interestingly, we do not find a clear and obvious bi-modality behavior like the ones observed in Figure 4 for mass and radius. **Interestingly, we also find that despite the small number of pairs of planets orbiting around M-dwarfs, they tend to be close the 1:1 line in comparison to pairs around FGK stars. However, this trend is obtained from a sample of seven pairs, and it needs to be confirmed when more data become available.**

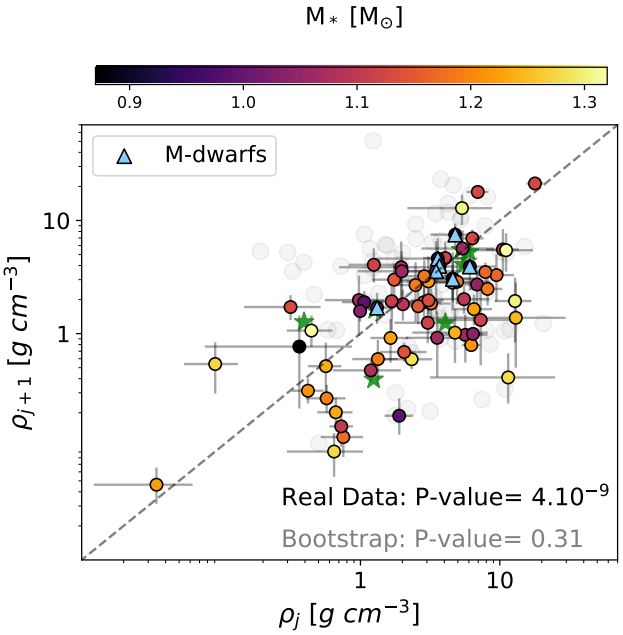

Fig. 6: The planetary density against the densities of the next planet farther from the star. The colors of the dots represent the effective temperature of the host star, and the gray dots correspond to a bootstrap trial. Pairs of planets orbiting around M-dwarfs are represented by light blue triangles. The planet in the Solar System are indicated by the green stars.

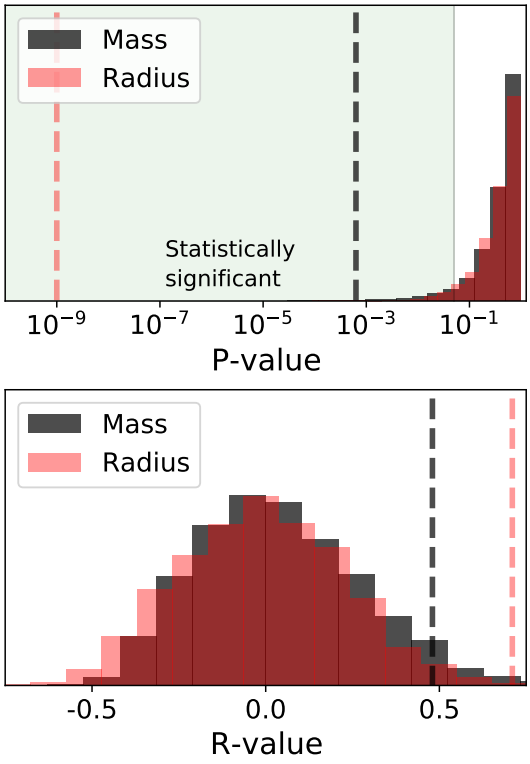

Fig. 7: Calculated P-value (top) and R-value (bottom) corresponding to the correlation between the mass (black) or radius (red) of a planet and the mass or radius of the next (farther) planet, respectively. The dashed lines correspond to the result from the exoplanet sample and the histograms to the results of 3000 bootstrap trials. The green region in the upper panel shows the statistically significant region, where the P-value is smaller than 0.05.

2.5. Significance and biases

Even when the correlation between radii, masses and densities of adjacent planets is clear, it remains unknown whether the correlation is physical or is a consequence of observational biases. In order to investigate the role of detection biases, we perform a series of 'null hypothesis' bootstrap tests. If detection biases are responsible for the observed correlation, then this trend would also be present in a 'mock' exoplanetary population which does not have this trend inherently, but suffers from the same detection biases. The null hypothesis used in the bootstrap tests is that the size and mass of a planet is random and independent of the size and mass of its neighbor. We then subject the resulting sample to the detection biases, and investigate whether a correlation arises. We construct the bootstrap trial drawing random planetary masses from the distribution of the observed masses. Then, to each stellar host we assign a number of planetary masses equal to the number of planets detected and place them at the observed orbital periods. As discussed in Steffen (2016), the sensitivity of TTVs and RVs can be expressed by:

$$SNR_{TTV} \sim \frac{M_p R_p^{3/2} P^{5/6}}{\sigma_{TTV}} \quad , \quad SNR_{RV} \sim \frac{M_p}{\sigma_{RV} P^{1/3} M_*^{2/3}} \quad , \quad (3)$$

where σ is the intrinsic uncertainty of a measurement. In order to consider the detection efficiency on the transit we use the following expression for the SNR:

$$SNR_{transit} = \frac{(R_p/R_*)^2 \sqrt{3.5yr/P}}{CCDP_{6h} \sqrt{6h/T}} \quad (4)$$

with

$$T = 13h \, (P/1yr)^{1/3} \, (\rho_*/\rho_\oplus)^{-1/3} \quad (5)$$

where R_p and R_* are the sizes of the planet and the host star, respectively, P the orbital period, ρ_* the bulk density of the star, and $CCDP_{6h}$ the a 6-hour Combined Differential Photometric Precision (the root mean square of the stellar photometric noise over 6 hours Christiansen et al. 2012). We then follow a similar method as in Weiss & Petigura (2020): we randomly draw planets from a log-normal distribution until its SNR is enough to be detected. We set a limit for SNR_{RV} and SNR_{TTV} of 2, and assume an intrinsic RV uncertainty of 2 m/s. These SNR are rather optimistic, so we also explore higher values and find that the results are similar. We also remove non-detectable transiting planets discarding $SNR_{transit}$ below 7.1, following Mullally et al. (2015). Figure 4 shows an example of a bootstrap trial for mass and radius, resulting in P-values of 0.28 and 0.46, respectively. This indicates that there is no significant correlation. The results for the inferred P-value and R-value when repeating the process 3000 times are shown in Figure 7. It compares the P-value and the R-value of the boostrap trials and the one obtained with the sample of observed planets. We find that only very few of the bootstrap trials are statistically significant for both mass and radius, and that the P-values of the real data is 4.5σ away from the bootstrap trials for the mass distribution and 12σ for the

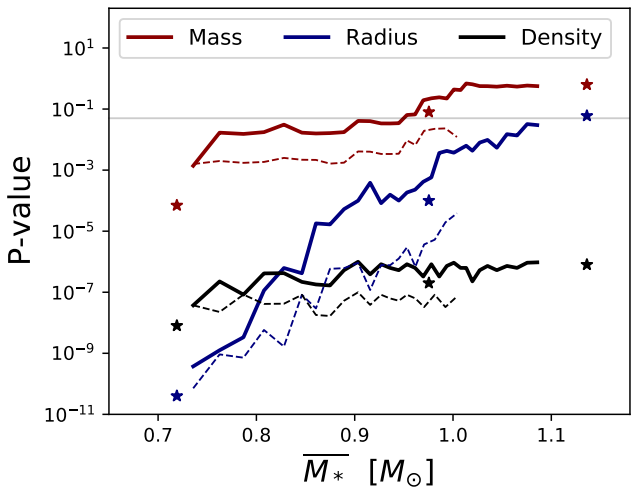

Fig. 8: Results of the moving sample analysis: P-value against the mean value of the stellar mass in the sub-sample. The red, blue, and black lines correspond to the mass, radius, and density distribution, respectively. The green region shows the statistically significant region, with P-value under 0.05. The three star-shaped dots of each color correspond to the P-values of the thirds of the population ordered by stellar mass. The dashed lines correspond to a sample of planets with masses below 100 $M_\oplus$.

radius distribution. We repeat the process using higher SNR limits for the RV and TTV up to 5, but in all cases the bootstrap trials do not show a significant correlation between the masses of adjacent planets. Given that the uniformity is stronger when giant planets are excluded, we perform the same test including only planets less massive than 100 $M_\oplus$ (see Figure A.1). In this case the correlations are more profound, and the P-values of the actual data are 14σ and 17σ away from the bootstrap trials for the mass and radius distributions, respectively. The R-values obtained on the bootstrap trials peak are close to zero. This discrepancy between the observed sample and the synthetic bootstrap samples suggests that a null hypothesis influenced by detection biases cannot produce the observed correlation. As a result, the correlation is likely to be physical.

2.6. Dependence on the stellar mass

Since it is easier to detect small and low-mass planets around small stars and there are only very few giant planets around M-dwarfs, we see a clear dependency of the "peas in the pod" pattern on the stellar mass, as shown in Figures 4 and 6. Since we do not find a bi-modal distribution in Figure 6 it is more difficult to determine whether the stellar effective temperature plays a role in the uniformity of the planetary density. This hypothesis can be tested performing a 'moving sample' technique, which is described as follows:

- First, we sort all the planetary systems in the sample according to the stellar mass,
- Second, we select the 30 first systems and perform the Pearson correlation analysis, obtaining the P-value and the R-value.
- We then move the sub-sample towards larger values of the physical property, removing the system with the lowest value and adding the system with the largest value left outside. We then repeat the Pearson correlation analysis.
- We next repeat the same procedure with a continuously moving sub-sample until the entire sample is covered.

Table 1: Dependence of the Pearson's P and R values on the uncertainty to our exoplanet sample with planets less massive than 100$M_\oplus$. "Np" stands for the number of pairs.

	Np	P-value	R-value
$\frac{\Delta M}{M}$<0.5, $\frac{\Delta R}{R}$<0.16	58	4.10^{-10}	0.71
$\frac{\Delta M}{M}$<0.4, $\frac{\Delta R}{R}$<0.13	53	4.10^{-9}	0.35
$\frac{\Delta M}{M}$<0.3, $\frac{\Delta R}{R}$<0.10	44	4.10^{-6}	0.26
$\frac{\Delta M}{M}$<0.2, $\frac{\Delta R}{R}$<0.07	25	0.001	0.25

Figure 8 shows the result of the moving sample analysis for the stellar mass. We plot the P-value against the median of the moving sample corresponding to the mass, radius and density distributions. The three star-shaped dots of each color correspond to the P-values of the thirds of the population ordered by stellar mass. We find that systems around more massive stars tend to be less "peas in the pod" in mass and radius, but not in density. Despite both lines corresponding to mass and radius distributions have similar shapes, the one corresponding to the mass has significantly higher P-values, suggesting that the uniformity is weaker across all the moving sub-samples. **We also note that the line corresponding to the radius has a higher slope, indicating that the correlation is stronger for planetary radius than mass.** As shown above, planets more massive than 100 $M_\oplus$ do not follow the 'peas in the pod' pattern. Therefore this trend could be easily explained if more massive stars more frequently host more massive planets. **To test this hypothesis, we also plot the dashed blue and red lines corresponding to a sample of planets with masses below 100 $M_\oplus$ for radius and mass respectively. The line corresponding to the radius has a clear positive slope, suggesting that even when gas giants are excluded from the sample the results that more massive stars tend to host less similar planetary systems in radius remains. The line corresponding to the mass also has a positive slope, but rather weak, suggesting that the trend is not as strong as for mass. We also note that given the correlations found between stellar mass and metallicity (Owen & Murray-Clay 2018), and the evidence that metal-rich stars tend to host lightly less uniform planets (Millholland & Winn 2021), it is difficult to interpret whether the dependence is mostly affected by the stellar mass, metallicity, or both.** The nearly flat line corresponding to the planetary density indicates that the uniformity in density is insensitive to the effective temperature of the host star. We perform the same analysis with other parameters as the insolation, however, we do not find a clear dependence with the uniformity in either mass, radius, or density. As with the stellar mass, the line corresponding to the mass is higher than the ones corresponding to the radius and density for all the moving sub-samples. However, in this case all the sub-samples of both mass and radius distributions are well below the P-value of 0.05 and do not have significant slopes, suggesting that the "peas in the pod" pattern is insensitive to the

stellar irradiation.

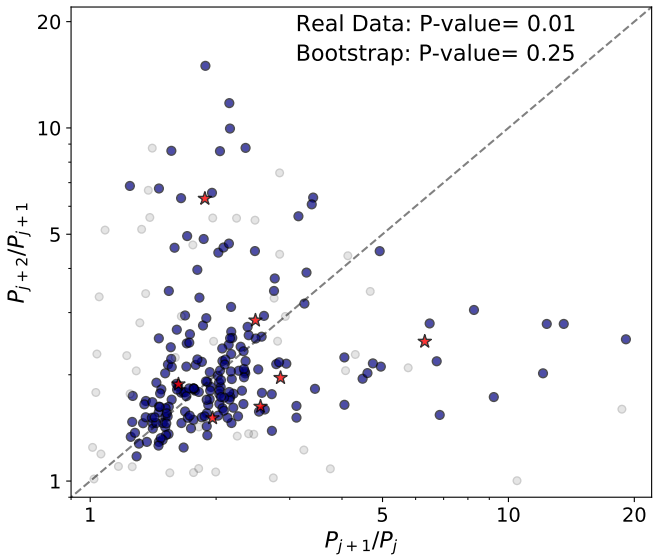

Fig. 9: "Peas in the pod" pattern for the orbital period spacing. The figure shows the period ratio of the outer pairs as function of the period ratio of the inner pairs. The right figure shows the average radius of the pair against its period ratio. The red star-shaped dots correspond to Solar System planets.

2.7. Sample selection effect

Another relevant factor that should be considered when studying the similarity of planetary systems is the effect of the used catalog on the inferred results. It is common to set a limit on the relative uncertainty of the mass and radius measurement in order to discard the least accurate data. We therefore investigate the sensitivity of the results on the limits of the uncertainties on the measured mass and radius. We explore a range of uncertainties ranging from 50% to 16% in mass, since lower values of the upper limits in uncertainties lead to samples with a low number of exoplanets. We set the limits for the radius uncertainties to be a third of the mass uncertainties, in order to have the same impact on the density uncertainty. We find that catalogs with more precise data lead to a weaker mass correlation between adjacent planets. However, this is linked to the previous result on the dependence on the mass limit, since giant planets tend to have lower relative uncertainties. In order to avoid this bias, we discard planets with masses above $100M_\oplus$. Table 2 lists the dependence of the P-value and the R-value on the uncertainties on the this sample of planets less massive than $100M_\oplus$.

We still find that catalogs with more precise data lead to weaker correlation between adjacent planets. As previously argued, this could be because of the lower number of data points in the sample. We note that the P-values obtained when decreasing the limit on uncertainties are higher than the ones inferred when only removing points randomly. For example, when we decrease the uncertainties from $\frac{\Delta M}{M}$<0.5 and $\frac{\Delta R}{R}$<0.16 to $\frac{\Delta M}{M}$<0.3 and $\frac{\Delta R}{R}$<0.1, the number of pairs decreases from 66 to 44 and the P-value increases from 4.10^{-10} to 4.10^{-6}. However, when we remove randomly 22 pairs from the initial sample of 58, we infer a P-value between [9.10^{-9}, 6.10^{-7}]. Therefore, we conclude that the increase on the P-value is not caused by the increase of the number of points, and consequently the systems with more precise mass and radius measurements are less uniform, even after excluding gas giants. This is probably due to a selection effect since very low mass planets tend to have higher uncertainties and tend to be more 'peas in the pod', as seen in Section 2.3.

Table 2: Sensitivity of the P-value when analyzing the period ratio of the outer pairs as function of the period ratio of the inner pairs. We show the dependence of the p-value on the minimum multiplicity and upper limits on planetary radius and period ratio of the sample. "Np" stands for the number of data points in the sample.

	Multiplicity>2		Multiplicity>3	
	Np	P-value	Np	P-value
No limit	272	0.02	168	0.23
PR_{max}:				
PR_{max}<8	261	0.03	162	0.36
PR_{max}<6	254	0.006	154	0.02
PR_{max}<4	237	10^{-4}	150	0.001
PR_{max}<2	111	10^{-4}	84	0.002
R_{max}:				
R_{max}<8$R_\oplus$	238	0.002	142	0.07
R_{max}<6$R_\oplus$	231	0.003	140	0.07
R_{max}<4$R_\oplus$	146	0.007	121	0.16
R_{max}<2$R_\oplus$	49	0.01	25	0.08

3. Uniformity in period spacing

We next focus on the aspect of period spacing. There are various studies which have investigated the orbital spacings of Kepler's multiple system using the Titius-Bode relation (e.g. Lineweaver 2015; Huang & Bakos 2014), their clustering around theoretical stability tresholds (Pu & Wu 2015, e.g.) and the spacing of Kepler planets in terms of the orbital period ratio (e.g. Lissauer et al. 2011; Steffen 2013; Steffen & Hwang 2015). Recently, Weiss et al. (2018) found that planets orbiting the same star tend to have regular orbital spacings, and Jiang et al. (2020) confirmed this pattern and concluded that such a correlation is unlikely to be caused by observational biases.

These results, however, correspond to different exoplanet samples and filtering criteria. In this section we investigate how the inferred uniformity in orbital spacing of planetary systems depends on the used data and the selection criteria. First, Jiang et al. (2020) only consider systems with multiplicities higher than four arguing that systems of lower multiplicities tend not to be dynamically packed and therefore it is more likely to miss non-transiting planets. Second, Weiss et al. (2018) and Jiang et al. (2020) exclude period ratios of adjacent planets higher than four since in these systems the sensitivity to observe larger orbital periods may be incomplete. Finally, they only consider

planets with radii below $6R_\oplus$.

In this section we do not use planetary sample used in Section 2, since we aim to include more systems in our sample. Instead, we also include planets without mass measurements and do not put any constrain on the maximum radius uncertainty, which leaves us with 474 planetary systems and 1220 planets. We perform the same bootstrap test as before, and conclude that the inferred correlation cannot be explained by selection effects from observational biases. We notice that systems with low period ratios of adjacent planets are significantly more uniformly distributed than the ones above. This could be because for systems with high period ratios there is no dynamical interaction between the planets, or simply due to an observational bias.

We next explore the sensitivity of the results on the various choices typically made to filter the planetary samples. The results are listed in Table 3, where we show the dependence of the P-value on the upper limit of the period ratios for exoplanet samples with different upper limit on planetary radii. We show the results corresponding to systems with multiplicities higher than 2 and with multiplicities higher than 3 (as done by Jiang et al. (2020)). We find that the results strongly depend on the the period of adjacent planets. When we exclude the systems with the highest period ratios the P-value decreases very significantly along the whole range of PR_{max} tested. This result is expected since systems with lower period ratios have stronger dynamical interaction.

The effect of setting a limit on the maximum radius is less clear. When we set a lower radius limits the P-value increases, but it could be due to the lower number of data points available in the sample. We analyze dependence of the P-value on the number of data points, as previously done, and conclude that the increase of the P-value when reducing the number of data points could be also explained by the lower number of data points. However, it is remarkable that the small sample with system of radii below $2R_\oplus$ has a P-value of 0.01, which cannot be explained by the number of data points. These systems with rocky planets, therefore, are more uniformly spaced than the systems with larger planets. Finally, the effect of the minimum multiplicity is also not conclusive. Similarly to the radius limit, we find higher P-values that can be explained by the lower number of data points in the sample.

4. Discussion and Conclusions

In this paper we explore the similarity of planets within a multi-planet system. More specifically, we study the 'peas in a pod' pattern of radii, masses, densities and orbital period ratios of adjacent planets, and confirm that their correlation has a physical origin. In addition, we quantify and compare the similarities in radius and mass. Finally, we explore whether different sub-populations of systems show different patterns, and how the obtained results depend on the used exoplanet sample.

Using a similarity metric defined as distance in logarithmic space, we find that planetary systems tend to be more similar in radius than in mass. This could be linked to the fact that the radius has a greater impact on the density and, hence, on the planetary composition than the mass. We also find a strong correlation between densities of adjacent planets. If the density is the main physical quantity that tends to be similar within a planetary system, it would explain that the stronger similarity in radius than in mass. Detection biases are not expected to have a big influence on this result.

We also find that there is a sharp transition in the 'peas in the pod' pattern of planets at $100M_\oplus$ and $10R_\oplus$. Systems with planets below these limits are significantly more uniform. As shown in the bootstrap trials, detection biases could lead to a less obvious 'peas in the pod' pattern among large planets than among smaller ones. However, instead of finding a somewhat smooth transition in the uniformity of systems when transiting from small to large planets, we find two different regimes. It suggests that the physical processes governing the formation of planets with masses below $100M_\oplus$ clearly gives rise to adjacent planets with similar masses and radii, but not for more massive planets. **We find that the P-values of the real data are 4.5σ and 12σ away from the bootstrap trials for the mass and radius distributions, respectively when using the entire sample, and are 14σ and 16σ away when only planets less massive than $100M_\oplus$ are included.** Interestingly, we do not find two regimes when analyzing the 'peas in the pod' pattern in density: there is a rather clear correlation for all types of planets.

The dependence of the 'peas in the pod' pattern with the planetary mass has to be treated with care and be investigated further. First, exoplanet catalogs with more precise mass and radius measurements tend to be have less 'peas in the pod' systems since they contain a lower proportion of low-mass planets. Second, there is a clear dependency of the 'peas in the pod' pattern on the stellar mass. Planetary systems around more massive stars tend to be less uniform in mass and radius, since they tend to host more massive planets. The 'peas in the pod' pattern in density, instead, does not show a clear dependence on the stellar mass, which is expected due to its lack of dependence on the planetary mass.

Finally, the similarity trend orbital period spacing is clearly confirmed, we find that the strength of the correlation strongly depends on the depends on the maximum radius set for the planetary sample, which is often set arbitrarily in other publications. We also find that systems containing planets with small period ratios are more uniformly distributed, which may be an indicative of stronger dynamical interaction.

Ongoing and future space missions like TESS (Ricker et al. 2015), CHEOPS (Broeg et al. 2013) and PLATO (Roxburgh & Catala 2006), as well as ground-based radial velocity facilities like ESPRESSO (Pepe et al. 2014) will rapidly increase the number of characterized exoplanetary systems and will allow to continue monitoring the systems analyzed in this study searching for missing planets. In addition, GAIA will perform high-precision astrometry and characterize missing planet in the outer regions of multi-planet systems. This will lead to a more complete understanding of exoplanetary demographics and of the uniqueness of our own Solar System.

References

Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, Publications of the Astronomical Society of the Pacific, 125, 989–999
Alibert, Y. 2019, A&A, 624, A45
Ballard, S. & Johnson, J. A. 2016, ApJ, 816, 66
Bashi, D. & Zucker, S. 2021, A&A, 651, A61
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
Broeg, C., Fortier, A., Ehrenreich, D., et al. 2013, in European Physical Journal Web of Conferences, Vol. 47, European Physical Journal Web of Conferences, 03005
Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Science, 337, 556
Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. 2012, Publications of the Astronomical Society of the Pacific, 124, 1279–1287
Ciardi, D. R., Fabrycky, D. C., Ford, E. B., et al. 2013, ApJ, 763, 41

Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, ApJ, 790, 146
Fang, J. & Margot, J.-L. 2012, ApJ, 761, 92
Gilbert, G. J. & Fabrycky, D. C. 2020, AJ, 159, 281
Hadden, S. & Lithwick, Y. 2014, ApJ, 787, 80
Hadden, S. & Lithwick, Y. 2017, AJ, 154, 5
Hansen, B. M. S. & Murray, N. 2013, ApJ, 775, 53
Helled, R., Lozovsky, M., & Zucker, S. 2016, MNRAS, 455, L96
Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398
Huang, C. X. & Bakos, G. Á. 2014, MNRAS, 442, 674
Jiang, C.-F., Xie, J.-W., & Zhou, J.-L. 2020, AJ, 160, 180
Kipping, D. 2018, MNRAS, 473, 784
Latham, D. W., Rowe, J. F., Quinn, S. N., et al. 2011, ApJ, 732, L24
Lineweaver, C. H. 2015, in IAU General Assembly, Vol. 29, 2258465
Lissauer, J. J., Marcy, G. W., Bryson, S. T., et al. 2014, ApJ, 784, 44
Lissauer, J. J., Marcy, G. W., Rowe, J. F., et al. 2012, ApJ, 750, 112
Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, ApJS, 197, 8
Lozovsky, M., Helled, R., Pascucci, I., et al. 2021, A&A, 652, A110
Malhotra, R. 2015, ApJ, 808, 71
Millholland, S., Wang, S., & Laughlin, G. 2017, ApJ, 849, L33
Millholland, S. C. & Winn, J. N. 2021, ApJ, 920, L34
Mills, S. M., Fabrycky, D. C., Migaszewski, C., et al. 2016, Nature, 533, 509
Mishra, L., Alibert, Y., Leleu, A., et al. 2021, arXiv e-prints, arXiv:2105.12745
Mullally, F., Coughlin, J. L., Thompson, S. E., et al. 2015, The Astrophysical Journal Supplement Series, 217, 31
Murchikova, L. & Tremaine, S. 2020, AJ, 160, 160
Otegi, J. F., Bouchy, F., & Helled, R. 2020, A&A, 634, A43
Owen, J. E. & Campos Estrada, B. 2020, EvapMass: Minimum mass of planets predictor
Owen, J. E. & Murray-Clay, R. 2018, MNRAS, 480, 2206
Pepe, F., Molaro, P., Cristiani, S., et al. 2014, arXiv e-prints, arXiv:1401.5918
Petigura, E. A., Howard, A. W., Marcy, G. W., et al. 2017, AJ, 154, 107
Pu, B. & Wu, Y. 2015, ApJ, 807, 44
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, ApJ, 784, 45
Roxburgh, I. W. & Catala, C. 2006, in IAU Joint Discussion, Vol. 26, IAU Joint Discussion, 32
Steffen, J. 2013, The Architectures of Near-Resonant Kepler Planets, NASA OSS Proposal
Steffen, J. H. 2016, MNRAS, 457, 4384
Steffen, J. H. & Hwang, J. A. 2015, MNRAS, 448, 1956
Van Eylen, V., Albrecht, S., Huang, X., et al. 2019, AJ, 157, 61
Wang, S. 2017, Research Notes of the American Astronomical Society, 1, 26
Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, AJ, 155, 48
Weiss, L. M. & Petigura, E. A. 2020, ApJ, 893, L1
Xie, J.-W., Dong, S., Zhu, Z., et al. 2016, Proceedings of the National Academy of Science, 113, 11431
Zhu, W. 2020, AJ, 159, 188

Appendix A:

Figure A.1 shows the comparison of the measured P-value and R-value of the observed population of planets excluding planets more massive than $100M_\oplus$ (dashed lines) with the results of 3000 bootstrap trials. We find that when we use a planetary sample with masses below $100M_\oplus$ the calculated P-values of the real data are 14σ and 17σ away from the bootstrap trials for the mass and radius distributions, respectively.

Figure A.2 shows the dependence of the P-value and R-value on the number of pairs on the sample. It shows the evolution of these two valued after randomly removing pairs.

Table A.1 shows the full list of similarities of the multi-planetary systems in our exoplanet sample. We also add the similarities of the Solar System, inner Solar System and outer Solar System for reference.

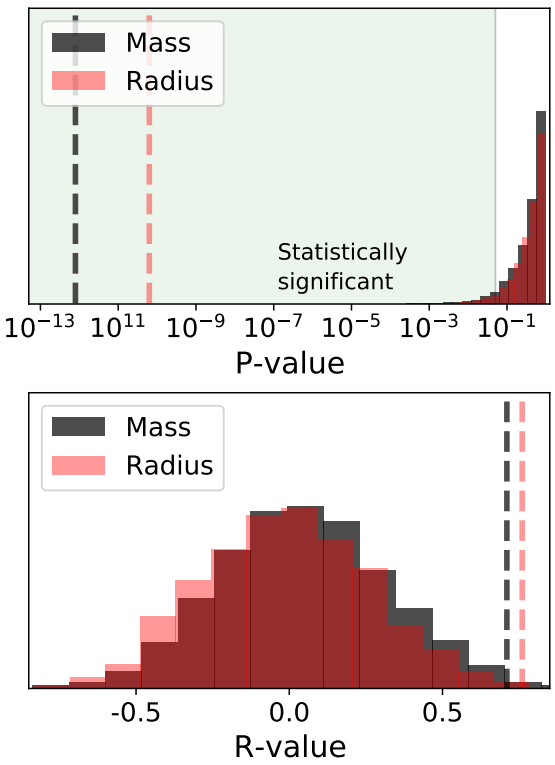

Fig. A.1: Calculated P-value (top) and R-value (bottom) corresponding to the correlation between the mass (black) or radius (red) of a planet and the mass or radius of the next (farther) planet, excluding planets more massive than $100M_\oplus$. The dashed lines correspond to the result from the exoplanet sample and the histograms to the results of 3000 bootstrap trials. The green region in the upper panel shows the statistically significant region, where the P-value is smaller than 0.05.

2.3 Radial Velocity Follow Up of Transiting Planets

While transits provide us the radius and orbital parameters of the planets, performing radial velocity follow-up is essential to confirm or exclude the planetary nature of the transiting companion as well as to determine its accurate mass and bulk density. The planetary density and the position in the M-R diagram with respect to the composition curves provide direct insights into the structure and composition of exoplanets. Validating exoplanet and populating the M-R diagram is, therefore, crucial to understand the exoplanetary nature. During my Ph.D., I have been involved in the target selection, observation, and data analysis of several RV follow-up programs that will be discussed in this section.

2.3.1 NCORES: Searching Naked Cores with HARPS

HARPS (*High-Accuracy Radial velocity Planetary Searcher*) is a high-precision spectrograph, and it is the most successful RV planet finder. The instrument was assembled at the Observatory of Geneva in 2002, and was installed on the ESO 3.6m telescope in La Silla Observatory (Pepe et al. 2002; Mayor et al. 2003). It is a fiber-fed, cross-dispersed echelle spectrograph with a resolution of R=115000. The combination of the instrument stability and the simultaneous ThAr-reference technique allows accurate radial-velocity determination at the level of 1 ms^{-1}, making HARPS suitable for the detection of small-mass planets. After its first detection in 2004, it has discovered 130 to date, making it the most successful planet finder after Kepler and TESS.

The NCORES large program (PI D. Armstrong, 2018-2020) on HARPS was awarded 72 nights to investigate naked cores (i.e. planets in which the outer layer has been stripped off by photoevaporation). Studying the composition of such naked cores is crucial to understanding planets forming via core accretion, since they fold the primordial composition from its formation. These cores contain the material which was present in the protoplanetary disk in the solid accretion phase, before acquiring an envelope. The water-ice mass fraction is key to understanding the formation and migration history of the planet (Unterborn et al. 2018). Cores formed outside the water ice-line in the protoplanetary disk are expected to contain more than 50% of water mass fraction (Lodders 2003; Min et al. 2011), while cores formed within the water-ice line are not expected to contain a significant amount of water. If a close-in planet is found to have a significant amount of ice, it is very likely to have formed further out in the system and migrated inwards to its current position. Jin & Mordasini (2018) showed that the population of potential Kepler cores is consistent with a predominantly rocky population with little ice mass. However, they relied on mass-radius relationships, since most of the Kepler cores transit stars that are too dim to determine the core. Accurately measuring the masses of

cores, then, constitutes a way of testing whether this hypothesis.

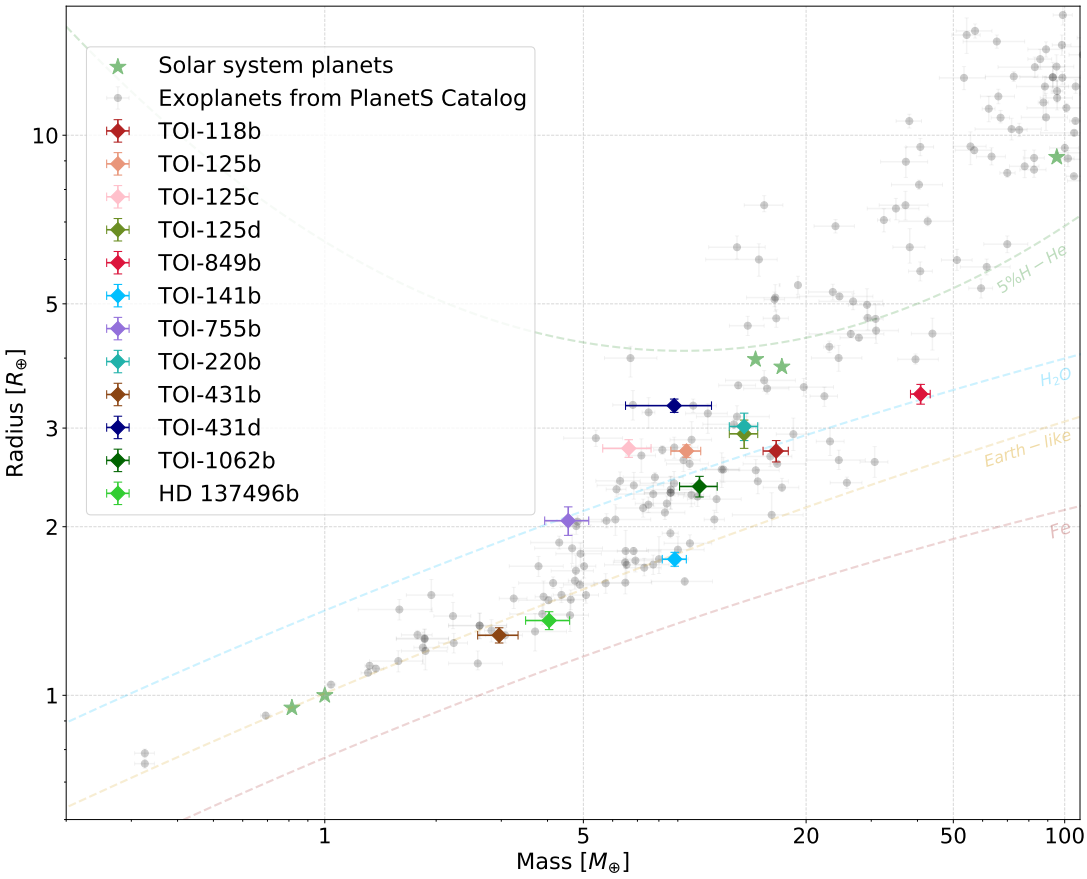

Figure 2.9: Mass-Radius diagram with the exoplanets characterized by the NCORES program. The gray points show the exoplanets in the PlanetS Catalog, and the green star shaped markers correspond to the Solar System planets. We also display composition lines of pure-iron (brown), Earth-like (light-brown), pure-water (blue), and 5% H-He.

NCORES targeted TESS and K2 targets with radii between 1.8$R_\oplus$ and 4$R_\oplus$ orbiting with orbital periods shorter than 5 days. I have contributed to the target selection, the HARPS observations, and the characterization of the internal structure of the discovered planets. In addition, I also analyzed and led the publication of the TOI-1062 system, which is discussed in detail in **??**. Figure 2.9 shows the exoplanets that have been characterized and published under the NCORES program. They are twelve exoplanets distributed in nine planetary systems that with masses going from ~3$M_\oplus$ to ~40$M_\oplus$ and radii from ~1.3$R_\oplus$ to ~3.5$R_\oplus$. The first published discovery was TOI-118b (Esposito et al. 2019), a hot-Neptune from the TESS Sector 1. Nielsen et al. (2020b) discovered the TOI-125 system, which consists of three mini-Neptunes

with similar radii but quite different masses. Another interesting discovery to be announced in 2020 was TOI-141b (Espinoza et al. 2020), a highly irradiated super-Earth in a 1-day orbit. TOI-141b is found to have a mass of 8.83±0.66$M_\oplus$ and a radius of 1.75±0.05$R_\oplus$, making it compatible with an Earth-like composition. Other discoveries published in 2021 under the NCORES program include a hot mini-Neptune in the radius valley was discovered around TOI-755 (Osborn et al. 2021a), a warm mini-Neptune around TOI-220 (Hoyer et al. 2021), and a super-Earth, mini-Neptune and non-transiting planet around TOI-431 (Osborn et al. 2021). Another remarkable discovery is HD 137496b, a hot super Mercury accompanied by a cold Jupiter (Silva et al. 2021). K2 photometry combined with RV from HARPS and CORALIE found a mass of 4.04±0.55$M_\oplus$ and a radius of 1.31±0.05$R_\oplus$, which gives an uncommonly high density of 10.5±2 $\rho_\oplus$. Our interior models found that the planet is composed mainly of iron, with the core representing over 70% of the mass.

Another very remarkable discovery of NCORES is TOI-849b (Armstrong et al. 2020), a dense remnant core of a giant planet with a radius smaller than Neptune and more than twice Neptune's mass (R=3.45±0.14$R_\oplus$, M=39.1±2.7$M_\oplus$). It orbits with a period of just 18.4 hours around a G-type star, leading to an equilibrium temperature of ~1800K. The planet's mass is higher than what is expected for runaway gas accretion, meaning that it may have been a gas giant before undergoing intense mass loss through thermal self-disruption or planet collisions. However, theoretical models of photoevaporation cannot provide the mass loss needed to reduce a Jupiter-like giant planet to a core like TOI-849b in a timescale of several Gyr. Therefore, the most likely scenarios are that either it suffered extreme mass loss through thermal self-disruption or giant planet collisions, or it avoided substantial gas accretion, maybe because of late formation or due to a gap opening. In any case, photoevaporation is expected to have removed a few $M_\oplus$ of H-He, so the remaining atmosphere is likely to be enriched, making it an interesting target for future atmospheric studies.

In addition, I leaded the discovery paper of the two sub-Neptunes orbiting around TOI-1062 (also presented in Section 3.2.2.1). We found that the host star is rotating with a period of 21.9 days, and also noted the existence of a stellar spot whose lifetime is similar to the stellar rotational period. The spot causes a significant stellar activity, which is handled using Gaussian processes (GP) on the RV data. The RV fit is shown in Figure 2.10. We find no hint of transit for TOI 1062c, and that TOI 1062b has an orbital period of 4.114 days, a radius of $2.265^{+0.096}_{-0.091}$ $R_\oplus$, and a mass of $10.15^{+0.81}_{-0.84}$ $M_\oplus$. We find an eccentricity of $0.18^{+0.07}_{-0.06}$ for TOI-1062b and 0.14±0.07 for TOI-1062c. The eccentricity for TOI-1062b is slightly higher than usual for its orbital period, it may not be significant due to the bias towards higher eccentricities for nearly circular planets (Lucy & Sweeney 1971). TOI 1062 c instead is found to have an orbital period

of 7.988 ± 0.04 days and a minimum mass of $9.78^{+1.26}_{-1.18}$ $M_{\oplus}$. TOI 1062 c is in nearly 2:1 motion resonance with its inner companion. This publication entirely presented and further discussed in Section 3.2.2.1.

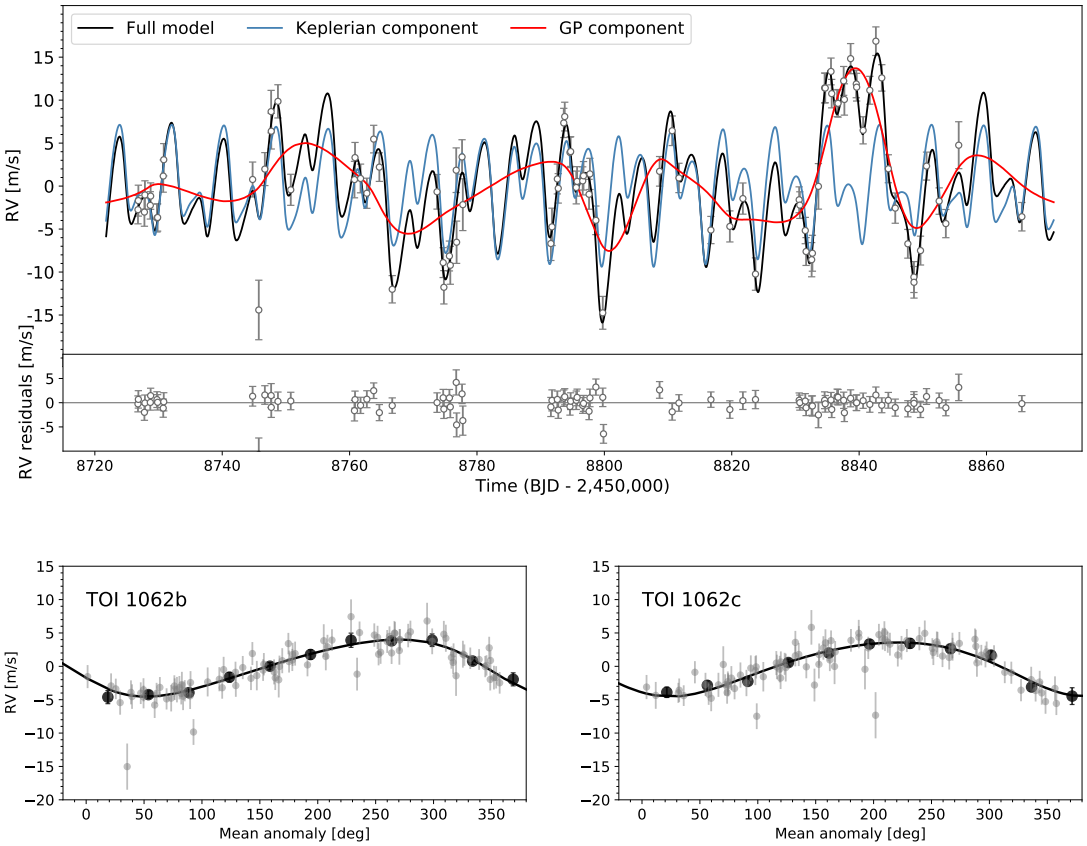

Figure 2.10: HARPS RVs for TOI 1062 with a two planet model including eccentric orbits and Gaussian processes with Matern kernel.Top panel: RVs along with the residuals to the best fit right below the time series. Bottom panel: Phase-folded data for each planet. The gray dots represent the observed RVs and the black dots the binned data. Figure taken from Otegi et al. (2021a).

We note that eight out of the twelve characterized exoplanets lie close to the pure-water line, so they are close to the transition between 'rocky' and 'volatile-rich' populations as defined in Otegi et al. (2020a). The position in the M-R diagram of TOI-118b, TOI-125b,c,d, TOI-755b, TOI-220b, TOI-431d and, TOI-1062 suggests a significant presence of water-ice, implying that they are likely to have formed from outside the water ice-line and migrated inwards.

2.3.2 ESPRESSO Programs

ESPRESSO (*Echelle Spectrograph for Rocky Exoplanet and Stable Spectroscopic Observations*) is a new-generation spectrograph mounted on the ESO's Very Large Telescope (VLT) on Paranal Observatory in Chile. It is the successor of a line of echelle spectrometers including CORALIE and HARPS. Indeed, it builds on the foundations laid by HARPS, but it benefits not only from the larger capacity to collect light but also from the improvements in the stability and calibration accuracy thanks to the laser frequency comb technology. It can observe with any of the four Unit Telescopes of the VLT at a resolving power up to 190000 on the 378.2 to 788.7 nm wavelength range. It can reach a precision level of 10 cm/s, which is approximately the radial-velocity variation that the Earth induces on the Sun. One of the main goals of the Guaranteed-Time Observations (GTO) of ESPRESSO is to detect Earth-like planets in the habitable zone. The high RV precision enables ESPRESSO to find Earth-size planets inside solar-type stars' habitable zone if they are bright and 'quite' enough. In addition, the ESPRESSO GTO is also dedicated to the characterization of planetary atmospheres, asteroseismology of K-dwarf stars, and stellar metallicity studies.

During my Ph.D., I have been involved in the target selection and development of the scientific case of two observing programs designed for ESPRESSO. In this subsection, we will discuss them.

2.3.2.1 Warm mini-Neptunes

The objective of this program is to accurately measure masses of long-period sub-Neptune candidates that are not expected to suffer intense irradiation causing atmospheric erosion. More specifically, it targets planets with a radius between $2R_\oplus$ and $4R_\oplus$ and stellar irradiation lower than $30F_\oplus$. This program is complementary to the ESPRESSO GTO, which aims to characterize the rocky population by targeting planets with radii below $2R_\oplus$, and to study the transition from gas-rich to gas-poor planets in order to constrain evaporation processes by observing sub-Neptune-size objects with irradiations between $50F_\oplus$ and $200F_\oplus$.

To date, three mechanisms have been proposed to explain the transition from super-Earths to sub-Neptunes: photoevaporation of the H-He by high-energy radiation from the host star (Owen & Wu 2017; Jin & Mordasini 2018), atmospheric mass-loss powered by the cooling luminosity of the planet's core (Gupta & Schlichting 2019) and gas-poor formation in which the gas accretion is delayed by dynamical friction (Lopez & Rice 2018). These three mechanisms predict different locations of the transition between the two populations in the period-radius-mass space. Figure 2.11 shows the parameter space covered in a Radius-Insolation diagram.

We see that more than 4500 candidates released by TESS clearly show a cloud of sub-Neptune-size objects in agreement with the Fulton distribution (Fulton et al. 2017). This program aims to study whether these sub-Neptunes with radii larger than $2R_\oplus$ are mainly rocky or possess volatile envelopes that contribute significantly to the planetary size, and how such envelope is affected by the stellar irradiation.

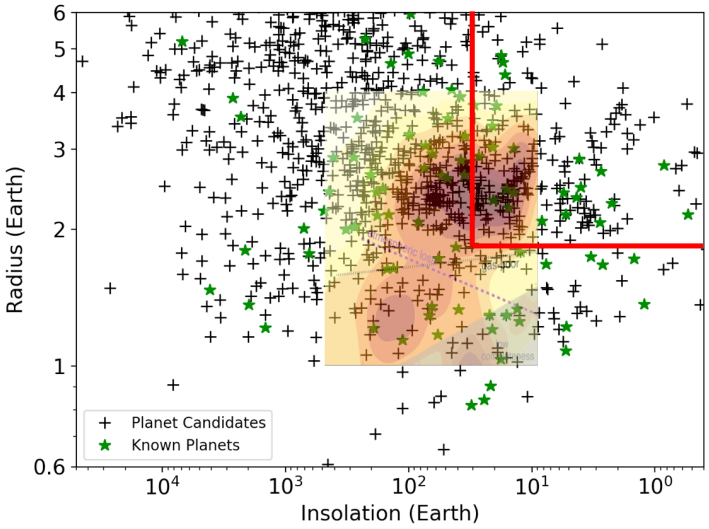

Figure 2.11: Planetary radius against incident stellar flux of TESS planetary candidates. The color code indicates the status of the candidates (Green: known and validated planets; Black: planet candidates). The relative occurrence of Kepler exoplanets from Fulton et al. (2017) is superimposed in transparency. The red box represent the domain covered in the Warm mini-Neptune ESPRESSO program.

In Otegi et al. (2020a) we showed that sub-Neptunes are spread over a wide range of masses (5-30$M_\oplus$) and densities (1-10$\rho_\oplus$) and that there is an overlap in both mass and radius between volatile-rich and "rocky" planets. In addition, the M-R diagram is multidimensional, and it is affected by other parameters as stellar irradiation, host star properties, planetary architecture, and age of the system. Populating the M-R diagram with precise measurements is crucial to explore the multi-layer nature of the M-R relationship and understand the transition between the two populations. Figure 2.12 shows the M-R diagram with exoplanets orbiting around FGK stars from the PlanetS Catalog. We see that sub-Neptunes orbiting FGK stars with low stellar irradiation seem to have a slightly higher density than those subjected to higher irradiation. It could be due to their atmospheres that are more compressed and denser, but accurate mass measurements of similar exoplanets with precise radial velocities will allow confirmation of this trend.

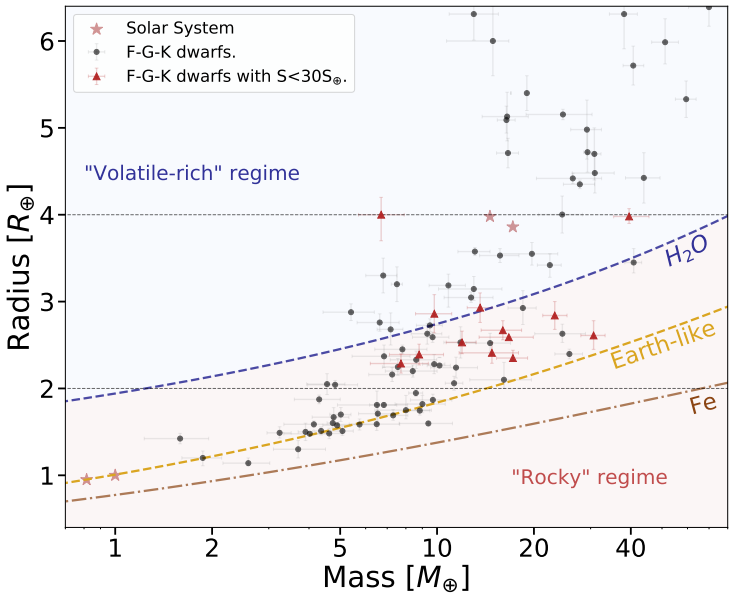

Figure 2.12: Mass-Radius diagram with exoplanets orbiting around FGK stars from the PlanetS Catalog. The red triangles correspond to exoplanets with stellar irradiation lower than $30S_{\oplus}$. The composition lines of pure-iron, Earth-like planets, and pure water are displayed in brown, light-brown, and blue, respectively.

The proposal was accepted in rank B in P105 and rescheduled in P107. We observed the following six targets from the 1st April to 30th September 2021: TOI-815 (18 points), TOI-1064 (22 points), TOI-1231 (14 points), TOI-133 (10 points), TOI-757 (22 points), and TOI-913 (30 points). TOI-815 and TOI-133 show strong variability, and more RVs are needed to find a solution. TOI-1231 has two sub-Neptunes that have already been characterized using HARPS data (Burt et al. 2021), but we have improved the RV fit. Similarly, TOI-1064b has been resolved with HARPS data and there is a publication in prep., but we have an improved mass estimation. For TOI-757 and TOI-913 require further analysis to determine whether more observations are needed since a significant proportion of the observations were taken in bad conditions.

The proposal has been accepted for period 109 (April-September 2022) for a total of 76 hours. Our target list is composed fourteen stars: K2-56, K2-186, TOI-208, TOI-213, TOI-215, TOI-282, TOI-444, TOI-458, TOI-533, TOI-687, TOI-707, TOI-1057, TOI-1221, and TOI-4302. Considering that the range of expected masses for planets with radii between $2R_{\oplus}$ and $4R_{\oplus}$ is between $5M_{\oplus}$ and $30_{\oplus}$, we need a mas determination better than 20%. The expected RV amplitudes of such objects range between $1m/s^{-1}$ and $2.5m/s^{-1}$, which are out of reach of HARPS and require the use of ESPRESSO. ESPRESSO provides a 2-magnitude gain and

improved RV precision with respect to HARPS-like instruments, enabling the exploration of small-size exoplanets at relatively long orbital period. The domain of warm mini-Neptune-size planets with relatively low stellar irradiation and relatively long orbital periods (larger than 20 days) is not well covered. SO far only ten exoplanet in this parameter space are orbiting FGK stars brighter than V=12 and have their mass and radius measured with uncertainties smaller than 20% and 8% respectively. This program will continue exploring a planetary population that is still poorly understood, and yet plays a very important role in constraining planetary formation and evolution models.

2.3.2.2 Mini-Neptunes around M-dwarfs

Planets orbiting M-dwarfs result particularly interesting since their formation environment is different from Solar-type stars. They have a long hot protostellar phase, lower protoplanetary disk mass, and intense activity at young stellar ages. All these different conditions may cause differences in the formation and composition on the planets around them. However, it is not still clear how these differences affect the observed planet population.

Cloutier & Menou (2020) found an increase in the frequency of close-in rocky planets orbiting around increasingly lower-mass stars. They also found that the occurrence rate of rocky to non-rocky planets is ~6-30 times around M-dwarfs compared to K-dwarfs, but they are not able to firmly identify the physical cause of this trend. A key question then is whether planets around M-dwarfs tend to form rocky planets. Several planet formation and evolution studies have pointed out that low mass stars tend to host richer planets in refractory materials (Alibert & Benz 2017; Schoonenberg et al. 2019; Burn et al. 2021). In addition, Kubyshkina & Vidotto (2021) modeled the evolution of a wide range of sub-Neptunes orbiting stars of different masses and evolutionary histories and found that atmospheric escape occurs more efficiently around lower-mass stars.

However, despite this theoretical evidence suggesting that M-dwarfs should host more rocky planets, when we look at Figure 2.13 with planets from the PlanetS Catalog we see sub-Neptunes around M-dwarfs actually seem to be less dense compared to those around other stars. It is particularly intriguing that exoplanets around FGK stars subjected to low irradiation appear to behave higher densities than those with higher irradiation, while planets orbiting M-dwarfs appear to be less dense than those orbiting FGK stars. This clearly indicates that the difference between planets orbiting M-dwarfs goes beyond the low irradiation suffered by the planets. This program aims for precise mass measurements of these planets that will provide essential data to confirm these differences between planets around M-dwarfs as opposed to those around

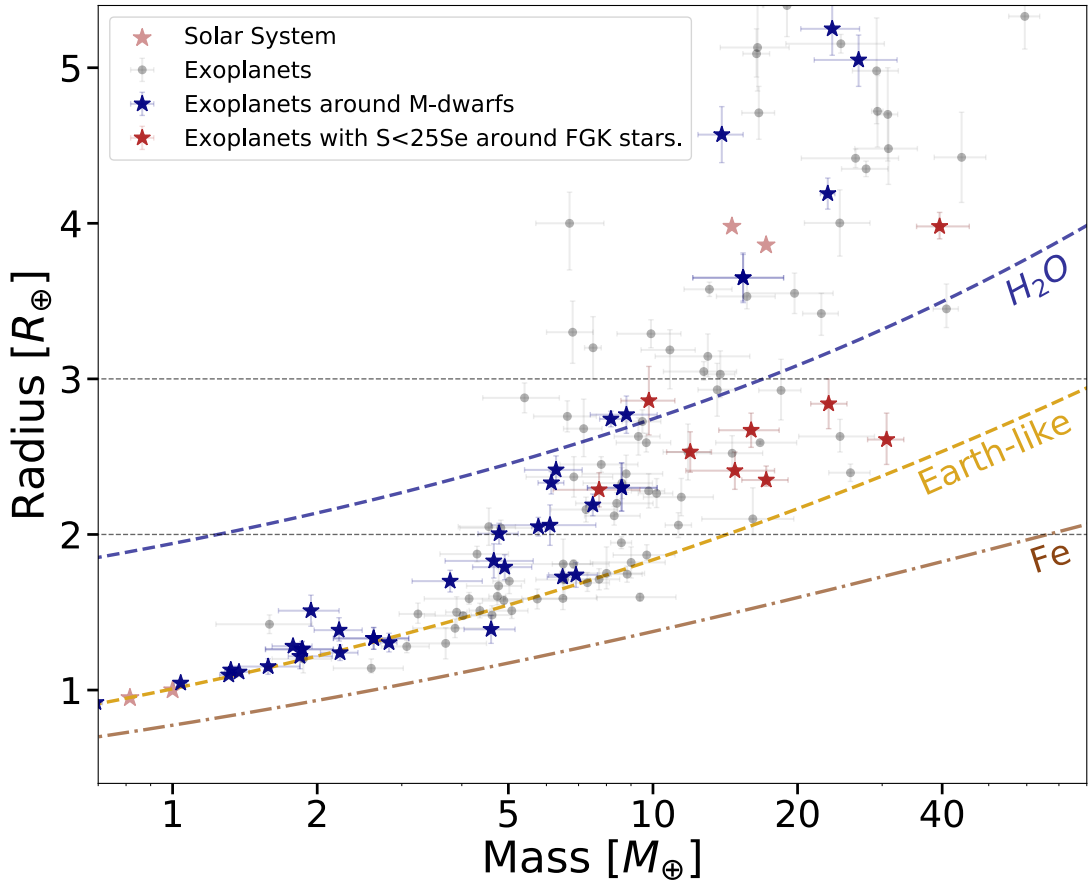

Figure 2.13: Mass-Radius diagram with exoplanets orbiting around FGK stars from the PlanetS Catalog. Blue stars correspond to exoplanets around M-dwarfs, and red stars to exoplanets subjected to stellar irradiation of less than $25S_{\oplus}$. The green shaded region between $2R_{\oplus}$ and $3R_{\oplus}$ corresponds to the range covered by the ESPRESSO program. The composition lines of pure-iron, Earth-like planets, and pure water are displayed in brown, light-brown, and blue, respectively.

FGK stars.

The proposal has been accepted for the period 109 (April-September 2022) for a total of 63 hours. The target list is composed of TOI-4342, TOI-4336, TOI-870, and TOI-4364. We aim to obtain accurate ($\lesssim$20%) mass measurements and constrain their orbital parameters. The expected RV semi-amplitude of the targets are ~1.5-3 m/s and given the faintness of our targets (V~12-13.6) these signals are only obtainable with the high precision of ESPRESSO and large collecting area of the 8.2 m VLT. All our targets have been vetted for asteroid crossing, contamination by nearby eclipsing binaries, and systematic errors from light curves processing. Additionally, all but one (which will be) of our targets have been spectroscopically vetted with

1-2 m class telescopes to further screen out false positives. Our targets have radii of 2-3 $R_\oplus$ and orbital periods of 5 - 81 days including a multi-planet system holding two sub-Neptunes with periods of 5.5 and 10.7 days. We will model the planet compositions and their implications on formation models to further understand the unique planet formation around M-dwarfs.

The scientific goals of this ESPRESSO program are partly overlap with the objectives of Near Infra-Red Planet Searcher (NIRPS, Bouchy et al. 2017), which currently is under commissioning. NIRPS is an Echelle spectrograph equipped with adaptive optics with a resolution of at least 80000 that will deliver RV precision below 1 m/s. It will operate in the infrared, as this is the range of wavelengths emitted by M-dwarfs, and will focus on the characterization of low-mass planets, including rocky planets in the habitable zone. The Front-End has already been installed and tested, and the fiber link was recently integrated in November 2021. The shipment of the Back-End is expected for January, and the first light is scheduled for June 2022.

2.3.3 Long Term Follow Up of planets in the Neptune desert with CORALIE

CORALIE (Queloz et al. 2000) is a high-resolution echelle spectrograph on the Swiss 1.2m Euler telescope, located at La Silla Observatory in Chile. It is fed by a 2" fiber and has a spectral resolution of R~60000. The instrument can reach a precision of 3 ms^{-1} for bright stars.

This program focuses on the study of planets in the Neptune desert. The hot Neptunian desert refers to a deficiency of Neptune-like planets close to their host stars (see Figure 2.14), and it is one of the key signatures of the planet population. The existence of the desert was first proposed by Szabó & Kiss (2011), who found a lack of intermediate mass planets ($0.02<M_p<0.8M_J$) with orbital period below 2.5 days. The physical mechanisms that result in the Neptunian desert are still unknown, and have been the subject of intensive research (e.g. Beaugé & Nesvorný 2013; Mazeh et al. 2016; Owen & Lai 2018; McDonald et al. 2019). Since short-period planets are easier to detect, and we can observe planets at smaller masses and longer periods, the desert cannot be an observational bias. Photoevaporation driven by high-energy radiation is expected to have dramatic consequences on the evolution and properties of the Neptune population, and it has been suggested as a mechanism to strip the H/He envelopes from Neptune-like planets (Baraffe et al. 2006; Jackson et al. 2010; Lopez & Fortney 2013; Chen & Rogers 2016). This mechanism involves the photoionization of hydrogen from the absorption of X-rays and extreme ultraviolet radiation (XUV) and occurs at atmospheric pressures of around a nano bar where a H/He-rich atmosphere becomes optically thin to the XUV. This process can produce a hydrodynamic escape of the H/He in the atmosphere. However,

a study by Owen & Lai (2018) finds that photoevaporation alone cannot explain the Neptune desert, as planets more massive than $0.5M_J$ can resist photoevaporation even at extremely short orbital periods. The latest consensus is that the desert arises due to a combination of tidal disruption and photoevaporation, with photoevaporation being primarily responsible for the lower boundary and tidal disruption for the upper boundary of the desert.

In addition, disk driven migration may also play a significant role in the formation of all close-in planets (Ida & Lin 2008; Mordasini et al. 2009). High eccentricity migration by the Lidov-Kozai mechanism may explain the existence of hot Jupiters, pumping the planet into a very eccentric orbit as a result of gravitational interactions with other planets or a distant stellar companion, followed by tidal dissipation that makes the planetary orbit circular (Petrovich 2015; Anderson et al. 2016; Muñoz et al. 2016).

The goal of this project is to perform a long-term follow-up of planets in the Neptunian desert to find an outer companion and/or improve the RV fit and refine the planetary parameters. The NCORES program for HARPS led by D. Armstrong (further discussed in Section 2.3.1) focuses on the same parameter space, so eventually, we also may use this program to vet targets that will later be characterized by HARPS. The target list was built in August 2021 following the following criteria:

- We prioritize strongly irradiated planets or planets orbiting very close to the host star. Even if the insolation is more directly linked to the physical state of the planet, planets orbiting close to the star are also interesting in the framework of planetary migration.
- We select planets with minimum RV semi-amplitude of ~8m/s. Semi-amplitude is higher than ~15m/s are preferable, since it may be possible to refine the parameters of the inner planet with CORALIE alone.
- We prioritize targets with already a large time span of RV data with CORALIE and/or HARPS.
- We exclude targets with V-magnitude larger than 12.

We select 19 targets, which are shown in Figure 2.14 and Table 2.3. Figure 2.14 shows the radius against insolation and orbital period for our selected targets and the population of known exoplanets from the NASA Exoplanet Archive. We note that most of our targets are inside or on the border of the Neptune desert in the radius-insolation diagram. Eight of them (TOI-849, LTT-9779, K2-39, TOI696, TOI-954, TOI-132, TOI-824, and K2-290) are well inside the Neptune desert. We note that some targets are not inside or in the border of the Neptune desert in the radius-insolation space (e.g., LP 714-47, TOI-181, WASP-29), but they are located in

the border of the desert the in radius-orbital period. We note that all selected planets have higher radii than $\sim 3R_{\oplus}$ since CORALIE does not have the precision to characterize smaller planets. This means that our planets are mainly located in the transition from Neptune-like to Saturn-like planets.

This program started running in September 2021, with an allocated time of six hours per month. Among the 19 targets we made a selection of 11 high priority targets based on the amount of available data, time span, luminosity, semi-amplitude and position in the Neptune desert. The high priority targets are in the top part of the Table 2.3. Among the high priority targets, we selected 6 that were observable, and observed each of them every two weeks with thirty minute exposures. The observed targets have changed due to their observability. Ten targets have been observed to date: WASP 166, WASP 29, TOI-954, TOI-618, TOI-969, TOI-118, TOI-181, K2-39, and TOI-824.

Among the observable targets, a selection of the six most promising targets was made among the observable ones, and each of them was observed every two weeks with thirty minutes exposures. The observed targets have changed due to their observability. Ten targets have been observed to date: WASP 166, WASP 29, TOI-954, TOI-618, TOI-969, TOI-118, TOI-181, K2-39, and TOI-824. TOI-618 and TOI-969 have a clear additional signal that requires long term monitoring. The RV time-series of TOI-618 is shown in Figure 2.15. The period of the outer companion is not well constrained because the available data do not cover a full period. We note that the CORALIE data is in agreement with HARPS, and the mass of the outer companion is in the giant planet regime ($6\pm1M_J$). On the other hand, the RV time-series of TOI-969 is shown in Figure 2.16. The outer companion has an estimated mass of $10\pm2M_J$ and its period is also not constrained yet. We find an additional significant signal at 12 days. This could be due to stellar activity since we find some correlation of the RV data with the bissector. These two targets are promising and we will continue monitoring them.

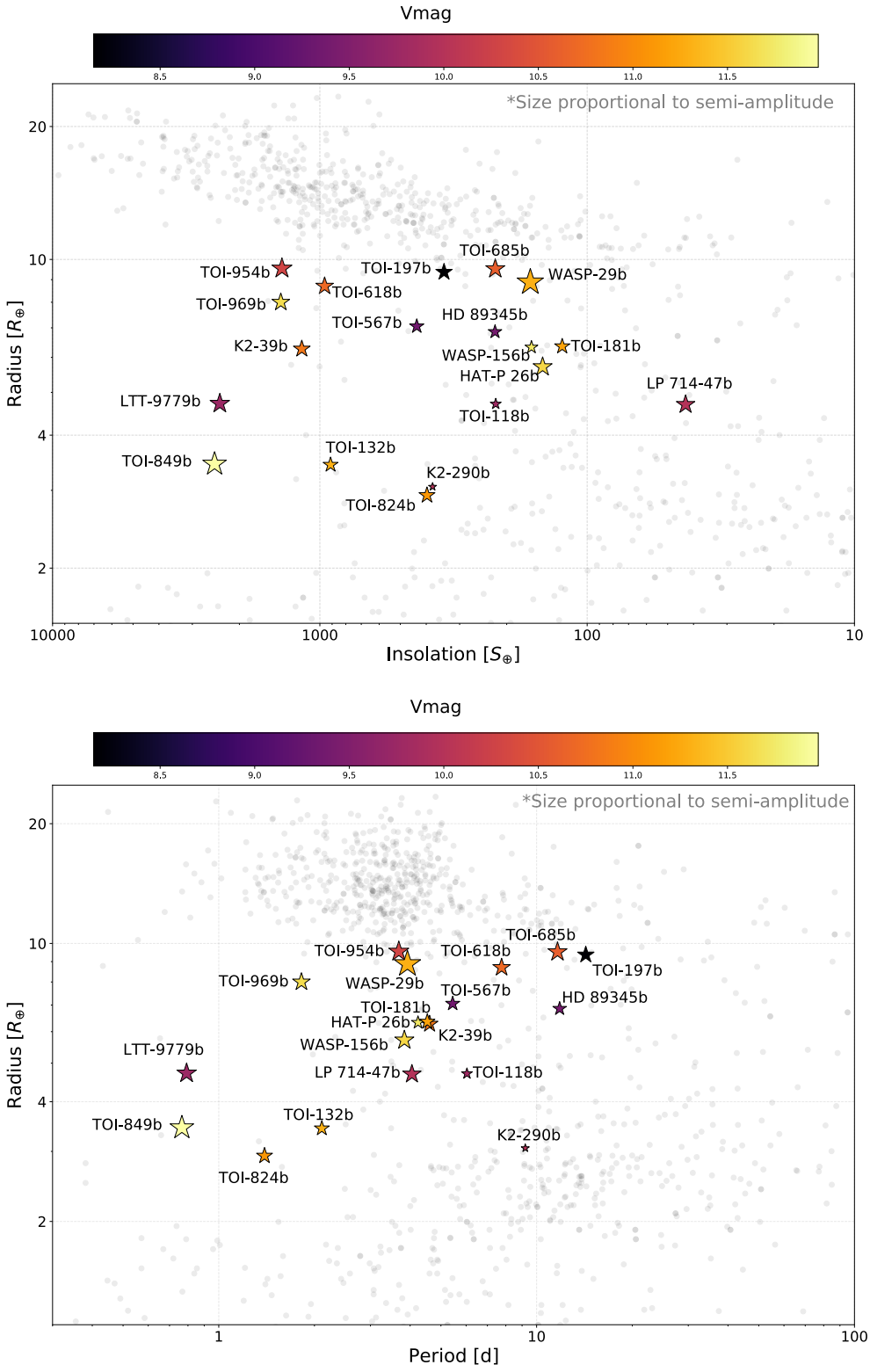

Figure 2.14: Radius against insolation (top) and against period (bottom) with the targets selected for the Long-Term Follow-Up of Planets in the Neptune Desert for CORALIE. The size of the stars is proportional to the semi-amplitude of the planets, and the color indicates the V-magnitude. In light-gray we show the population of detected exoplanets from the NASA Exoplanet ARCHIVE.

Table 2.3: Target selection for the Long Term Follow-Up of Planets in the Neptune Desert for CORALIE. The three planets without RV amplitude correspond to not published planets. The top part of the table shows the high priority targets.

Star	Ins [$S_\oplus$]	P [d]	R [$R_\oplus$]	V [mag]	K [m/s]	Sp. Type	Time Span [yr]	Nb. Obser.
K2-39	1168	4.61	6.27	10.83	14.40	K	5.4	8
TOI-954	1386	3.68	9.54	10.30	20.70	G	2.4	8
TOI-824	397	1.39	2.92	11.13	13.20	K	2.5	5
TOI-118	219	6.04	4.71	9.81	6.14	G	3.4	8
WASP-156	147	3.83	5.71	11.60	18.61	K	7.3	8
WASP-29	164	3.92	8.88	11.30	35.54	K	12.4	9
HAT-P 26	162	4.23	6.32	11.74	8.46	K	12.1	2
TOI-181	124	4.53	6.35	11.19	-	K	3.2	9
TOI-969	1403	1.82	8.00	11.65	-	K	1.2	6
TOI-618	961	7.76	8.71	10.70	-	F	2.2	3
TOI-567	434	5.44	7.05	9.36	10.47	G	G	0
TOI-849	2477	0.76	3.44	11.98	28.64	G	2.5	0
TOI-132	912	2.11	3.43	11.30	11.38	G	2.8	0
LTT 9779	2368	0.79	4.72	9.76	19.67	G	1.4	0
K2-290	379	9.21	3.06	10.02	<3.3	M	3.4	0
TOI-197	343	14.27	9.36	8.15	14.07	K	3.2	0
TOI-685	220	11.63	9.52	10.61	18.80	G	2.5	0
HD 89345	221	11.81	6.85	9.40	9.47	G	4.1	0
LP 714-47	43	4.05	4.69	10.07	17.53	K	2.4	0

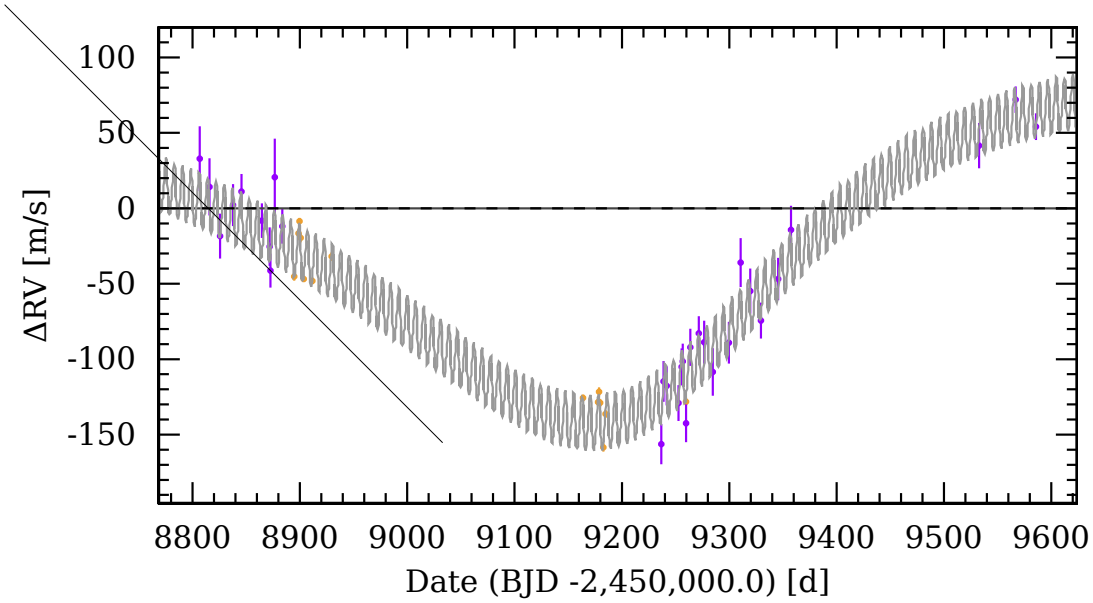

Figure 2.15: RV time series of TOI-618. The purple dots correspond to CORALIE data, and the yellow to HARPS data.

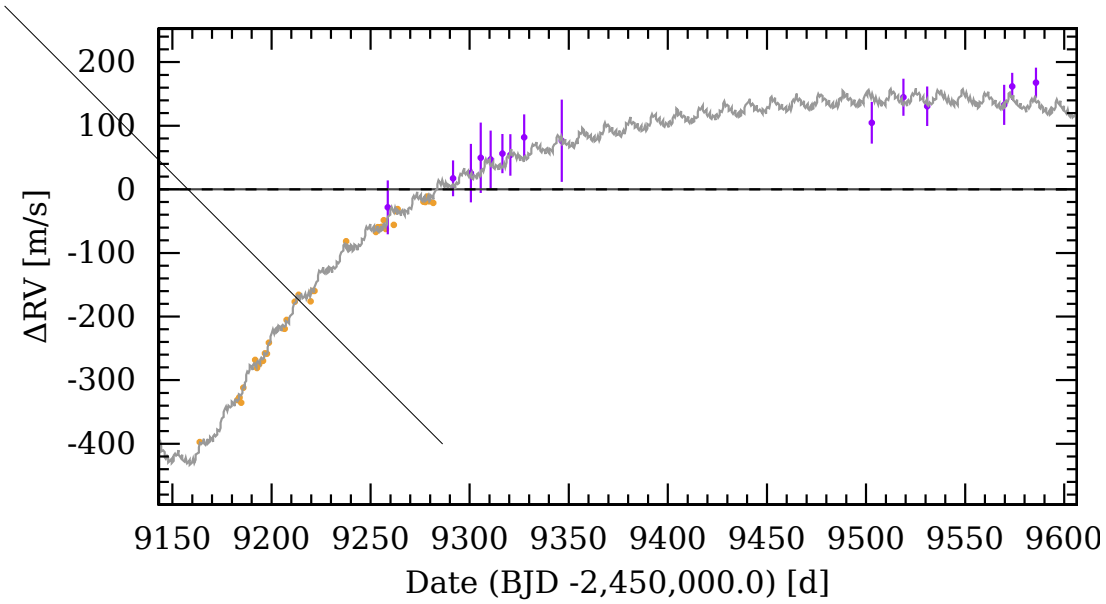

Figure 2.16: RV time series of TOI-969. The purple dots correspond to CORALIE data, and the yellow to HARPS data.

Chapter 3

Linking Theory and Observations

In the previous chapter we have seen that over the last years the number of planets with precise mass and radius measurements has substantially increased, allowing us to identify multiple planet populations. Characterizing the detected exoplanets (i.e. determining their structure and composition) is one of the main foci of current exoplanetary science, but it is not an easy task. Unlike in most scientific domains where samples can be studied in situ in a laboratory, the structure and composition of an exoplanet is not directly accesible to us. The only planetary layer that can be probed is the atmosphere, using transmission spectroscopy. However, direct measurements of atmospheres are still limited to a small number of planets.

In most cases there are only four main observables that allow us inferring the planetary structure and composition: the planetary mass, the planetary radius, its distance to the host star, and the luminosity of the host star. The mass and radius provide the bulk density of the planet, while the its semi-mayor axis and stellar luminosity determines its incident stellar irradiation. The bulk density can be used as a first indication of the planetary composition. Dense planets are more likely to be rocky and rich in refractory materials, while less dense planets are likely to host a layer of volatile materials. However, determining the internal structure and composition is extremely challenging as a consequence of intrinsic degeneracy because several compositions can lead to identical bulk density (e.g. Rogers & Seager 2010; Lopez & Fortney 2014; Dorn et al. 2015, 2017; Lozovsky et al. 2018; Otegi et al. 2020b). In this chapter we discuss the planetary model and statistical approach used to infer the planetary structure and composition from observable quantities, and present several published applications.

3.1 Inferring the interior structure from the observed mass and radius

3.1.1 Our planet model

In this first section we will describe the physical models used to solve the planetary structure. These physical models describe the density and compositional properties of the materials in the planetary layers. In the literature there are numerous physical models describing the internal structure of exoplanets, and some of them have been presented in Section 1.2.1.3. These models usually do not differ in the way in which the internal structure equations are solved, but in the physical models they use to describe the interior of the exoplanets. Figure 3.1 shows a schematic view of our interior structure model and the different concentric layers that compose it: an iron core, silicate mantle, water layer, and an atmosphere.

I started using the internal structure model presented in Dorn et al. (2015) and Dorn et al. (2017), but during the Ph.D. I introduced several updates, together with Caroline Dorn and Jonas Haldemann. In this section I will describe the most relevant modifications implemented to the initial internal structure model.

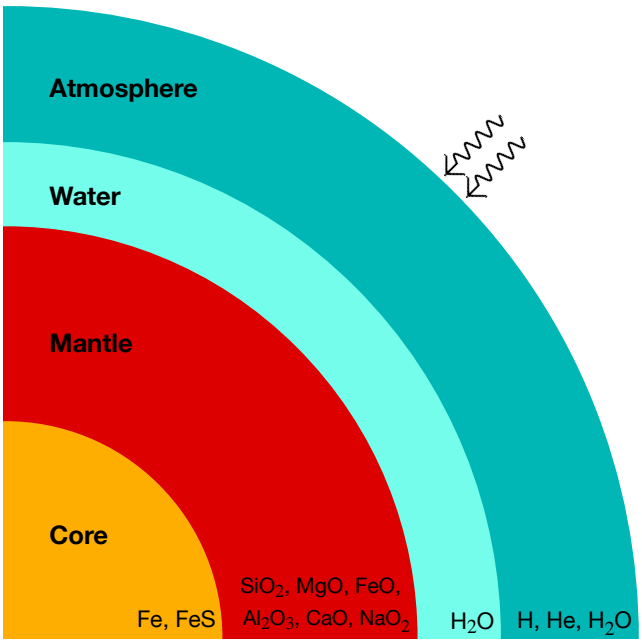

Figure 3.1: Schematic representation of the interior structure model with its four concentric layers: iron core, silicate mantle, water layer, and the atmosphere.

3.1.1.1 Core

In Dorn et al. (2017) the core is assumed to be completely made of iron, and its structure is calculated using the EoS of Bouchet et al. (2013). They used ab initio molecular dynamics simulations to calculate the EoS of iron in the solid phase for both face-centered cubic (fcc) and body-centered cubic (bcc) forms up to ~1.5TPa. Later, Hakim et al. (2018) developed a new EoS of hexagonal close packed (hcp) iron based on density functional theory results. These new EoS are adapted for super-Earth conditions, and they are valid for pressures up to ~137TPa. We therefore implemented these new EoS in our structure model.

We also implemented the possibility of adding FeS in the core. Since it is still unknown which lighter elements are present in exoplanets' cores, in our model sulphur accounts for other elements with may reduce the density of the core. We use the EoS of FeS from Hakim et al. (2018), and assume that it is uniformly distributed in the core (Valencia et al. 2007). The mixture between Fe and FeS is based on the so called additive volume law, and therefore it does not take into account the interactions between them.

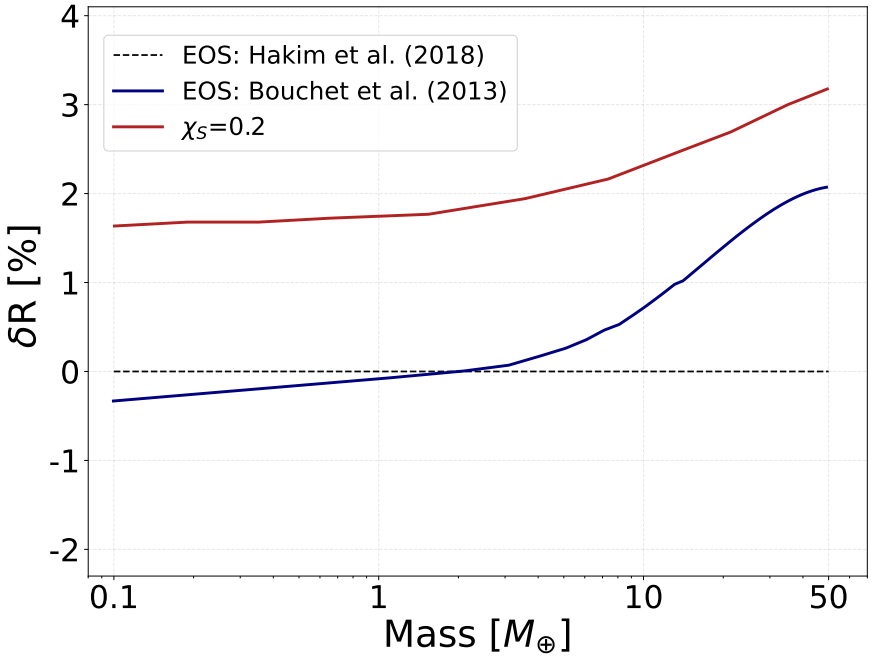

Figure 3.2: Relative radius difference between cores calculated with different models. The reference model is a pure iron core calculated with the EoS of Hakim et al. (2018). The blue line corresponds to the difference with respect to a pure iron core calculated with the EoS from Bouchet et al. (2013), and the red line with respect to an iron core with a molar fraction of sulfur of X_S=20%.

Figure 3.2 shows the relative difference in radius between cores calculated with different

models. The reference model is a pure iron core calculated with the EoS of Hakim et al. (2018). The relative difference is calculated as follows:

$$\delta R = 100 \left(1 - \frac{R}{R_{\text{Hakim et al. (2018)}}} \right) \tag{3.1}$$

We see that the difference of the iron core radius calculated with Bouchet et al. (2013) with respect to Hakim et al. (2018) is nearly 2% for cores of $50M_{\oplus}$. We also note that the difference is below 0.3% for cores less massive than $5M_{\oplus}$. The figure also shows the difference of the iron core radius with a molar fraction of sulphur of 20% with respect to a pure iron core. We use a sulphur molar fraction of 20% because it is close to the concentration of 20% FeS used in Valencia et al. (2007) and the maximum concentration of 19% that can be modelled by the EoS in Ichikawa & Tsuchiya (2020). We see that variation of the radius of the core when introducing a 20% of sulphur is of ∼1.7% at low masses and ∼3% at $50M_{\oplus}$.

3.1.1.2 Mantle

The silicate mantle is computed following Dorn et al. (2015), and no major changes were implemented. The constituents of the mantle layer are assumed to be SiO_2, MgO, FeO, CaO, Al_2O_3, or NaO_2. The stable minerals at a given pressure, temperature and composition are calculated by Gibbs energy minimization using the Perple_X code presented in Connolly (2009) and the thermodynamical model of Stixrude & Lithgow-Bertelloni (2011). The thermal gradient in the mantle is fixed by the adiabatic gradients on the Earth's mantle. For more details on the model we refer to Dorn et al. (2015).

3.1.1.3 Water layer

The phase diagram of water is highly complex, and contains multiple ice phases and a variety of transition points depending on temperature and pressure. For this reason, water EoS usually only cover certain parts of the pressure-temperature space. Dorn et al. (2017) computes the structure of the water layer using the quotidian equation of state (QEOS Vazan et al. 2013), which combines the Cowan ion EoS with the Thomas-Fermi model for electrons and treats H_2O as a mixture of atoms. Above 44.3GPa, the tabulated EoS of Seager et al. (2007) are used, which are derived from density functional theory simulations. This approach has proven to be in agreement with the widely used analytical equation of state (ANEOS Thompson & Lauson 1972).

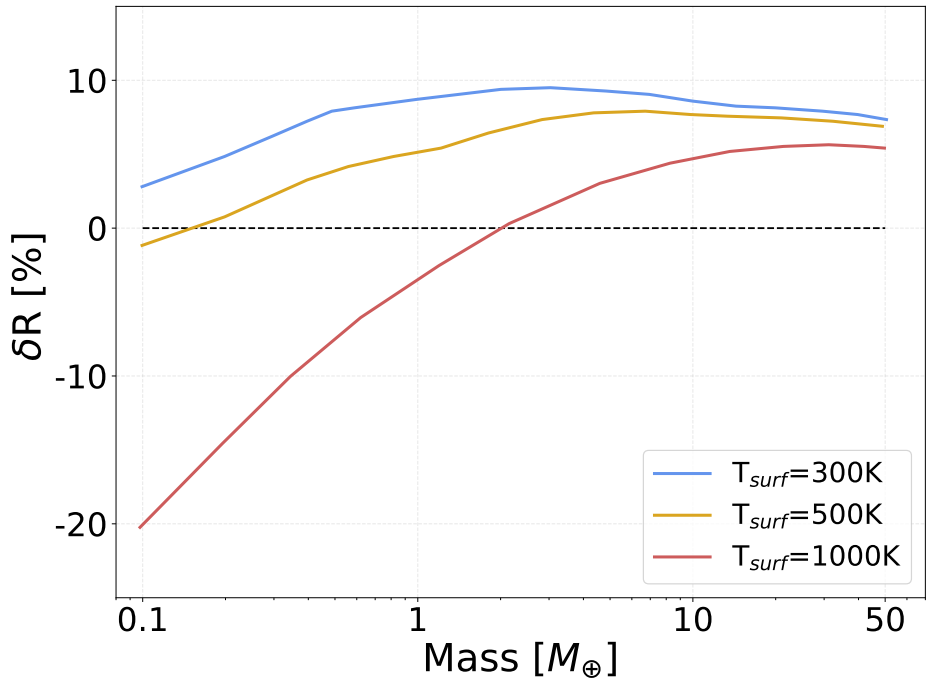

Figure 3.3: Relative radius difference between water dominated planets (50% H_2O, 25% mantle and 25% core) using QEOS from Vazan et al. (2013) compared to using AQUA EoS from Haldemann et al. (2020). Different colors represent different surface temperatures.

However, recently Haldemann et al. (2020) combined several EoS valid in localised regions of the parameter space to form a continuous EoS spanning a very wide range of pressures (0.1Pa to 400TPa) and temperatures (150K to 10^5K). The description of the ice phases at low, intermediate and high pressures is obtained combining the water EoS of Mazevet et al. (2019), Feistel & Wagner (2006), Journaux et al. (2020), and French & Redmer (2015). The liquid and gas phases at low pressures are modeled with the EoS from Wagner & Pruß (2002) and the high temperature low pressure regime with the CEA package (Gordon et al. 1984). The resulting EoS are called AQUA and are accesible online[a]. The strength of the AQUA EoS resides on large range of thermodynamical values at which they are valid. We therefore implemented these EoS on our internal structure model.

The impact of choosing AQUA EoS for the water layer instead of the QEOS from Vazan et al. (2013) is shown in Figure 3.3. It shows the relative radius difference between water dominated planets (also called ocean planets), with 50% H_2O, 25% mantle and 25% core using QEOS from Vazan et al. (2013) compared to using AQUA EoS. We see that the choice of the

[a]https://github.com/mnijh/aqua

water EoS can have a very significant impact, with a relative difference of nearly 10% for warm planets with masses between ~1$M_\oplus$ and ~10$M_\oplus$. We also note that AQUA tends to provide larger radii than QEOS.

3.1.1.4 Atmosphere

The atmospheric model in Dorn et al. (2017) assumes and envelope with an elemental composition of H, He, C, and O, which are basic for the formation of fundamental molecules in the atmosphere as H_2, CO, CO_2, or CH_4 (Lodders & Fegley 2002b; Visscher & Moses 2011; Madhusudhan 2012; Heng & Lyons 2016). It uses the Chemical Equilibrium with Applications (CEA) package (e.g. Gordon et al. 1984) for the EoS, with calculates the chemical equilibrium for an arbitrary gas mixture. For the energy transport, the model presented in Jin et al. (2014) is used, which assumes an irradiated atmosphere at the top of the atmosphere. We also incorporated the possibility of modeling a volatile layer with H, He and H_2O using the analytical equation of state (ANEOS Melosh 2007), which is one of the most widely EoS in planetary science.

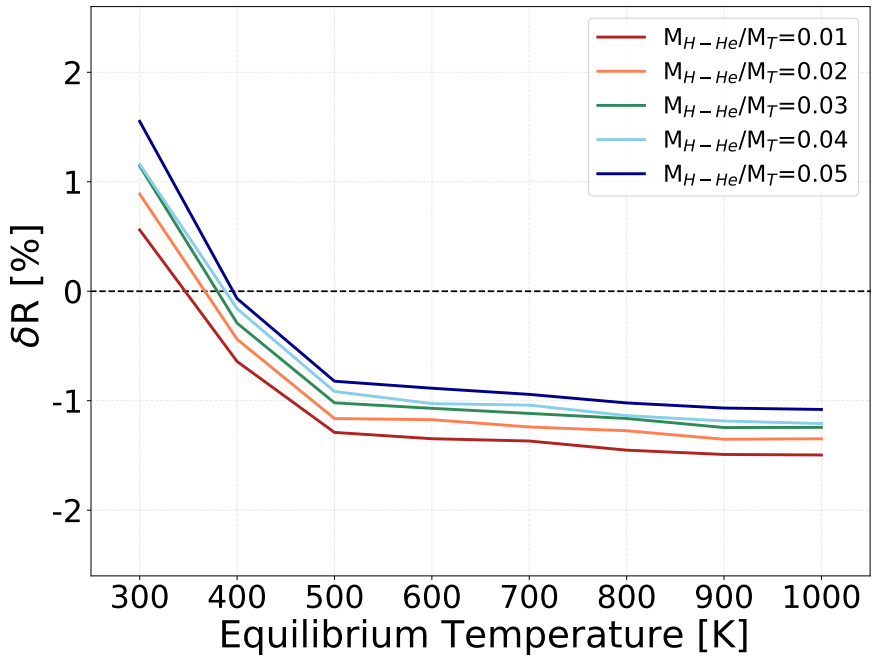

Figure 3.4: Relative radius difference between 5$M_\oplus$ planets with H-He envelopes calculated with EoS from Chabrier et al. (2019) (as reference) and Saumon et al. (1995) against the equilibrium temperature of the planet. The different colors correspond to different H-He mass fractions. The core of the planet is made of 50% iron core and 50% silicate mantle.

When restricted to a pure H-He gas layer (Z_{env}=0), Dorn et al. (2017) uses the EoS of

Saumon et al. (1995). However, recently Chabrier et al. (2019) presented a new EoS for dense H-He mixtures covering a large range of densities, pressures and temperatures. These EoS combine of Saumon et al. (1995) in the low density and temperature regime with EoS from Chabrier & Potekhin (1998) in the high density and temperature domain. In the intermediate density and temperature regime it uses ab initio quantum molecular dynamics (QMD) calculations. The tables are publicly available, and we implemented them in our atmospheric model.

Figure 3.4 compares the relative radius difference between $5M_\oplus$ planets with H-He envelopes calculated with the EoS from Saumon et al. (1995) and from Chabrier et al. (2019). The comparison is made for planets with different H-He mass fractions. We see that the difference for a planet of $5M_\oplus$ and 1% of H-He in mass is ~1.3% for equilibrium temperatures higher than 500K. We also note that at low temperatures (<400K) the difference is larger for planets with large volatile envelopes, while at higher temperatures (>400K) the difference is larger for planets with smaller envelopes. In any case, over the range of H-He fractions and equilibrium temperatures tested the difference between the two EoS is not very significant (<1.5%) and not measurable.

An important assumption when calculating the planetary structure is the luminosity of the planet. A recent study by Mordasini (2020) presented planetary evolution models that can be used to model the cooling of exoplanets after their formation. Indeed, after their formation, exoplanets retain a certain amount of heat and, in addition, the continuous decay of radiogenic elements heat the planet. Therefore, the internal luminosity is function of time and the planet composition. We use the relation given in Mordasini (2020) and calculate the internal luminosity as follows:

$$\frac{L}{L_J} = a_0 + b_1\left(\frac{M_T - M_{H-He}}{M_\oplus}\right) + b_2\left(\frac{M_T - M_{H-He}}{M_\oplus}\right)^2 + c_1\left(\frac{M_{H-He}}{M_\oplus}\right) + c_2\left(\frac{M_{H-He}}{M_\oplus}\right)^2 \tag{3.2}$$

where M_T and M_{H-He} are the total mass of the planet and the mass of the H-He envelope, respectively. the parameters a_0, b_1, b_2, c_1, and c_2 depend on the age of the planet and are given in Table A.1 of Mordasini (2020). The relation was obtained by modeling exoplanets with masses between $1M_\oplus$ and $40M_\oplus$, taking into account the cooling and contraction of the core and envelope together with the radiogenic luminosity (Mordasini et al. 2012).

3.1.1.5 Atmospheric structure with MESA:

As mentioned before, the atmospheric radiative model in Dorn et al. (2017) is based in Jin et al. (2014), for which the analytic irradiation model of Guillot (2010) is adopted. The opacities in Guillot (2010) are free parameters, and Jin et al. (2014) calibrates them for different equilibrium temperatures. This way they reproduce results from more sophisticated atmospheric models in which a wavelength-dependent opacity function is used while solving the radiative equilibrium (e.g. Fortney et al. 2008a; Parmentier et al. 2013). Dorn et al. (2017) follows this approach, and interpolate the opacities for a given equilibrium temperature from Table 2 in Jin et al. (2014). By doing this they aim to mimic the fundamental physics underlying atmospheric absorption and re-irradiation and a more simple and numerically inexpensive way.

Even if our atmospheric model for a pure H-He envelope is reasonably reliable with the use of state-of-the-art EOS from Chabrier et al. (2019), when we model heavy elements we use either ANEOS (Melosh 2007) or EoS from Saumon et al. (1995) with fitted opacities from Jin et al. (2014). Recently Müller et al. (2020) (M20) developed a EoS model in Module for Experiments in Stellar Astrophysics (MESA Paxton et al. 2010; Paxton et al. 2013, 2018) code to properly model the mixture of H-He and water. MESA is an open stellar evolution software that can construct a model for the interior of a star and evolve it over time. One of the main advantages of MESA resides in its versatility, and it has recently been adapted to model the evolution of planets (Müller et al. 2020; Valletta & Helled 2020). The EoS implemented by M20 combines the EoS from Chabrier et al. (2019) for H-He and QEOS (More et al. 1988) for water. In addition, MESA provides the option of using the Rosseland mean opacity values from Freedman et al. (2014), which are more rigorous and reliable than the extrapolated values in Jin et al. (2014).

We therefore implemented MESA's atmospheric model to our interior models. In addition to providing a better treatment of the gas opacities, it also gives access to a large database of modules and functionalities that can be easily added, and it allows to simulate the evolution of the planet. In Figure 3.5 we see a comparison between the results using ANEOS and MESA for a planet of M=15$M_\oplus$, T_{eq} = 300, and Z=0.5. We model three planets, with H-He mass fractions of 5%, 10%, and 15%, while the rest of the planets is half core and half mantle in mass. The Outer Boundary Condition (OBC) is set to where the optical depth becomes 2/3. We see on the top panel that the radius of the planet in MESA stabilizes at ~300Myr, and that the radius obtained are ~10% higher. In the middle and bottom plot we see internal profiles of the planets computed with ANEOS and with MESA after 2000Myr. Even if both approaches use the usual Schwarzschild criterion to distinguish between convective and radiative layers, we can see in the bottom plot that they do not provide consistent outputs. Even if the outer

layers of the atmosphere with MESA is more dense, its inner layers are hotter and, therefore, it results in a more expanded planet.

This is preliminary work, and so far we have only tested our forward model with MESA's atmospheric model. More testing is needed before implementing it to the Bayesian inference scheme with Nested Sampling, to fully understand the differences with our current model.

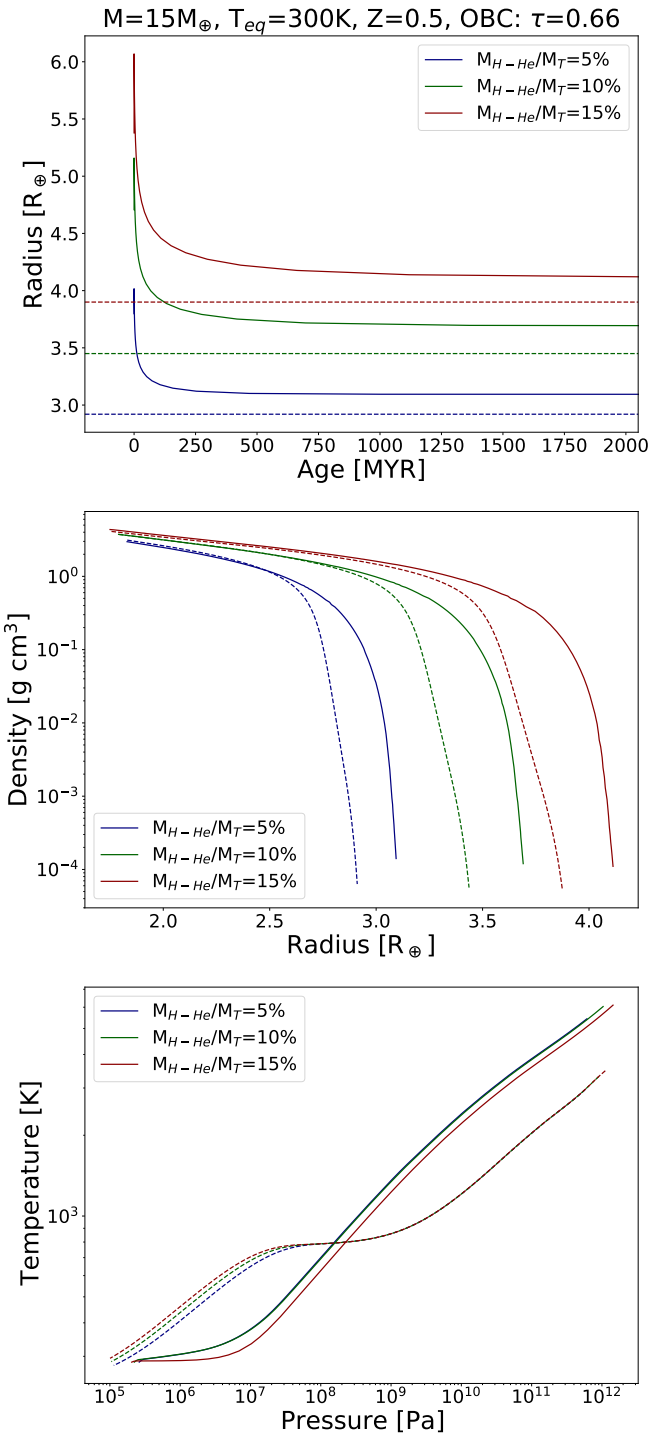

Figure 3.5: Comparison between atmospheric models using ANEOS (dashed lines) and MESA's (solid lines) atmospheric models for a planet of M=15$M_\oplus$, T_{eq} = 300, and Z=0.5. We model three planets, with H-He mass fractions of 5%, 10%, and 15%, while the rest of the planets is half core and half mantle in mass. The OBC is set to where the optical depth becomes 2/3. The top panel shows the evolution of planetary radius over time, while the middle and bottom panels shown the interior density and pressure-temperature profiles, respectively.

3.1.2 Solving the structure

The planet's internal structure is calculated by solving the equations of mass conservation, hydrostatic equilibrium, thermal transport, and energy conservation (see Section 1.2.1.1 more more details). In Dorn et al. (2017), these equations are integrated from the outer boundary to the inside. An illustration of the planet model and how the structure equations are solved is shown in Figure 3.6. First the structure of the volatile layer is calculated, followed by the water layer, silicate mantle, and metallic core. However, to start integrating from the outside some boundary conditions are needed. In our chosen setup he masses of the different layers is given as an input, but the pressures in the center (P_{cent}) and between layers (P_{CMB}, P_{MWB}, P^{surf}_{solid}) are needed for the integration (CMB and MWB refer to Core-Mantle Boundary and Mantle-Water Boundary, respectively). Previously these pressure estimations were fixed, but now we use the approximate analytic solutions of the dimensionless equations of planetary structure presented in Seager et al. (2007) to approximate them. We therefore have the following:

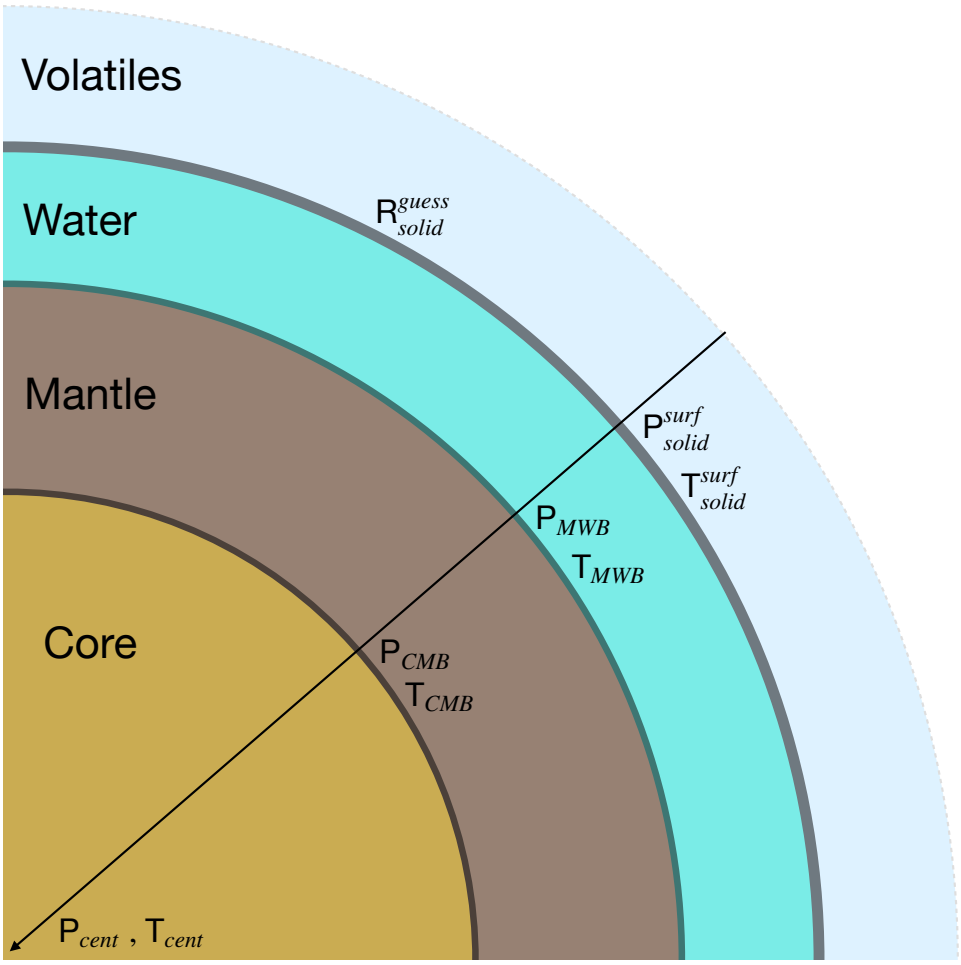

Figure 3.6: Illustration of the planet structure in our model. The arrow indicates the direction in which structure equations are solved.

$$P_{\rm cent} = \frac{3G}{8\pi}\frac{M_p^2}{(R_{\rm P}^{\rm guess})^4} \tag{3.3}$$

$$P_{\rm CMB} = P_{\rm cent} - \frac{2}{3}\pi G\frac{M_{\rm solid}^2}{(R_{\rm solid}^{\rm guess})^4} \tag{3.4}$$

$$P_{\rm MWB} = P_{\rm cent} - \frac{2}{3}\pi G\frac{M_{\rm core}^2}{R_{\rm core}^4\ P_{\rm cent}} \tag{3.5}$$

$$P_{\rm solid}^{\rm surf} = P_{\rm cent} - \frac{2}{3}\pi G\frac{(M_{\rm core}+M_{\rm core})^2}{(R_{\rm core}+R_{\rm mantle})^4\ P_{\rm cent}} \tag{3.6}$$

The radii of the core, mantle and water layer are estimated using M-R relations. We draw composition lines corresponding to a pure iron core using EoS of Hakim et al. (2018), silicate

mantle with Earth-like abundances, and pure water line with the AQUA EoS from Haldemann et al. (2020). We them fit the M-R relations using a total least squares method and assuming a dependence of R=AM^B. We get the following results:

$$R_{\rm core} = 0.808\ M_{\rm core}^{0.247} \tag{3.7}$$

$$R_{\rm mantle} = 1.061\ M_{\rm mantle}^{0.266} \tag{3.8}$$

$$R_{\rm water} = 1.403\ M_{\rm water}^{0.261} \tag{3.9}$$

We therefore use the following guesses for the solid and total radii:

$$R_{\rm solid}^{\rm guess} = \frac{M_{\rm core}}{M_{\rm P}}\ 0.808\ M_{\rm solid}^{0.247} + \frac{M_{\rm mantle}}{M_{\rm P}}\ 1.061\ M_{\rm solid}^{0.0.266} + \frac{M_{\rm water}}{M_{\rm P}}\ 1.403\ M_{\rm solid}^{0.261} \tag{3.10}$$

$$R_{\rm P}^{\rm guess} = \frac{M_{\rm core}}{M_{\rm P}}\ 0.808\ M_{\rm P}^{0.247} + \frac{M_{\rm mantle}}{M_{\rm P}}\ 1.061\ M_{\rm P}^{0.0.266} + \frac{M_{\rm water}}{M_{\rm P}}\ 1.403\ M_{\rm P}^{0.261} \tag{3.11}$$

After integrating the internal structure we obtain as an output the solid radius ($R'_{\rm solid}$). We use a shooting method (Keller 1992) to address this problem, which consist in guessing the missing values at one boundary, then integrate to the other boundary, and check if the values obtained through integration are within a given tolerance. So after the first integration we repeat the same process using $R_{\rm P}^{\rm guess} = R'_{\rm solid}$. We then repeat the process until the difference between $R_{\rm P}^{\rm guess}$ and the resulting $R'_{\rm solid}$ is below the tolerance.

3.1.3 Bayesian Inference with Nested Sampling

For the interior characterization we use a Bayesian inference method, which computes the posterior probability distribution function for each model parameter from the observational data. Since the posterior distribution cannot be derived analytically, a Monte Carlo sampling method is needed to sample the prior parameter space and evaluate how well each candidate model fits the data. In Dorn et al. (2017) a MCMC algorithm based on a Metropolis-Hastings method is used to exoplore the posterior distribution. We replaced this approach by a nested smapling scheme (Skilling 2004). This Monte Carlo technique can efficiently evaluate the Bayesian evidence and produce posterior probability distributions. Its main strengths with respect to the previous MCMC technique are the higher efficiency and the small amount of

problem-specific tuning required. We integrated the nested sampling scheme to the interior model using the PyMultiNest package (Buchner et al. 2014).

This sampling scheme is controlled by two main parameters: the number of live points and the maximum efficiency parameter. The number of live points needs to be large enough to properly sample the parameter space, for which we typically use a recommended value of 1000. The maximum efficiency controls the sampling volume at each iteration, which is equal to the sum of the volumes enclosing the active point set. We set it to 1 to obtain the maximum efficiency.

As already mentioned, the composition of the planets are given by the masses of the iron core, mantle layer, water layer, and H-He envelope. The composition of the core is determined by the FeS fraction, and the mantle composition is given by the $\mathrm{Mg/Si_{mantle}}$, $\mathrm{Fe/Si_{mantle}}$, $\mathrm{Al/Si_{mantle}}$, $\mathrm{Ca/Si_{mantle}}$, and $\mathrm{Na/Si_{mantle}}$ fractions. The water layer is considered to be pure H_2O and for the atmosphere the metallicity Z, and the irradiation temperature $\mathrm{T_{irr}}$ are used to model the atmospheric structure. We generally fix the $\mathrm{Mg/Si_{mantle}}$, $\mathrm{Fe/Si_{mantle}}$, $\mathrm{Al/Si_{mantle}}$, $\mathrm{Ca/Si_{mantle}}$, and $\mathrm{Na/Si_{mantle}}$ fractions, Z, and T_{irr}, and only let the layer masses vary. For the nested sampling algorithm we assume uniform prior distributions on the layer masses from 0 to the total planet mass. However, we would like to emphasize that for a given target planet, depending on the scientific question, a different choice of priors might be necessary. The set of prior presented here is one possible way of setting up the analysis but not the only one.

We use the total mass and total radius of the planet as data variables. In some cases we also use the bulk Mg/Si and bulk Fe/Si fractions as data, and in this case the $\mathrm{Mg/Si_{mantle}}$ and $\mathrm{Fe/Si_{mantle}}$ fractions are also varied, assuming a uniform prior. The total radius and the bulk Mg/Si and Fe/Si fractions are an outcome of the structure model, while the total mass is calculated as the sum of the layer masses. In Table 3.1 we show the model parameters, their priors, and the data variables used in the Bayesian inference. In some cases, when the stellar abundances are known, they can be used as a proxy for the bulk abundances of the planet (Dorn et al. 2015; Adibekyan et al. 2021). In this cases, we set the Fe/Si and Mg/Si ratios of the mantle as free parameters.

Table 3.1: Summary of model parameters, priors, and data.

Model Param.	**Priors**	**Data**
M_{Core}	$\mathcal{U}(0, M)$	Total Mass
M_{Mantle}	$\mathcal{U}(0, M)$	Total Radius
M_{Water}	$\mathcal{U}(0, M)$	Bulk Fe/Si
M_{H-He}	$\mathcal{U}(0, M)$	Bulk Mg/Si
χ_S	const.	
Fe/Si_{Mantle}	const. or $\mathcal{U}(0, 2Fe/Si_{\odot})$	
Mg/Si_{Mantle}	const. or $\mathcal{U}(0, 2Mg/Si_{\odot})$	
Al/Si_{Mantle}	const.	
Ca/Si_{Mantle}	const.	
Na/Si_{Mantle}	const.	
Z	const.	
T_{irr}	const.	

3.2 Applications

In this section I will present the projects in which I have used our internal structure model. These projects include a first author paper analyzing the influence of measured parameters on the inferred internal structure of exoplanets, eight published discovery papers as co-author, and two paper in preparation as co-author.

3.2.1 Impact of measured parameters of exoplanets on the inferred internal structure

In this work we aimed to study several aspects that affect the internal characterization of super-Earths and sub-Neptunes: observational uncertainties, location on the M-R diagram, impact of additional constraints such as bulk abundances or irradiation, and model assumptions. One of our main focuses in this work is to explore how the internal structure determination depends on the observational uncertainties for exoplanets of different masses and radii. Figure 3.7 shows the relative mass and radius error for the observed exoplanet population. They peak at ~8% in radius and ~35% in mass, but they will decrease rapidly with new generation instruments. We therefore how these observational improvements will help to constrain the internal structure of exoplanets in different M-R regimes.

We find that a decrease in observational uncertainties for planets below the Earth-like composition line leads to a better determination of the core mass. In this density regime, planets do not have a significant gaseous envelope and they require a large amount of iron. Then, variation in the mass or radius distributions is reflected in the core mass, since there is no other interior layer that can account for such high planet bulk densities. Similarly, we also find that a

decrease in observational uncertainties for planets above the pure-water composition line leads to a better determination of the atmospheric mass. The explanation is similar to the previous one for dense planets: for a given mass, low-density planets require thick gaseous envelopes to match the observed radius, and this significantly reduces the degeneracy. Unfortunately, the transition between rocky planets and volatile-rich planets is strongly degenerated.

We also explore how uncertainties in the temperature gradient of the planet can lead to radius uncertainties. Observational uncertainties are expected to decrease significantly in the near future, and we consider that the importance of the theoretical uncertainty related to model assumptions (e.g., envelope structure, and composition of the planetary layers) may soon overcome the observational uncertainties.

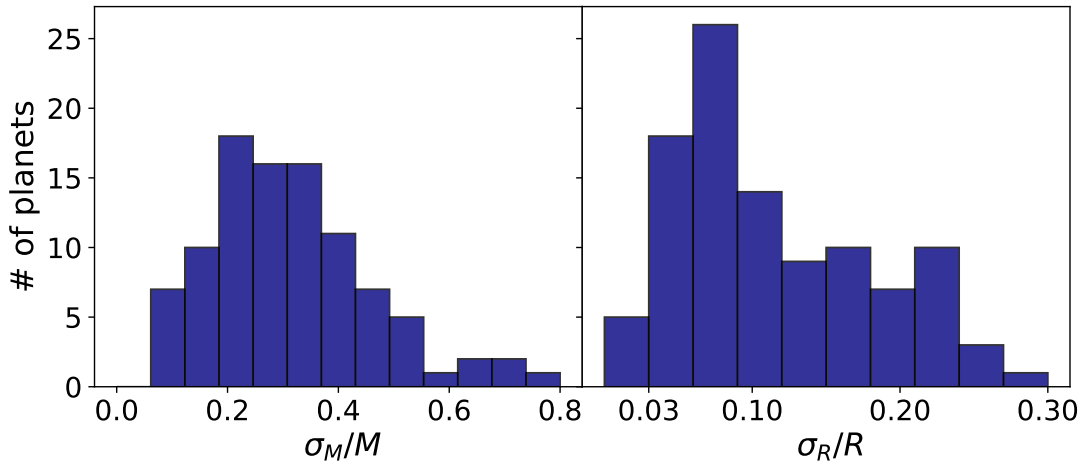

Figure 3.7: Distributions of the relative mass (left) and radius (right) uncertainties for the 115 observed exoplanets less massive than $25M_\oplus$ from the NASA Exoplanet Archive. Figure taken from Otegi et al. (2020b).

Astronomy & Astrophysics manuscript no. main
June 11, 2021

Impact of the measured parameters of exoplanets on the inferred internal structure

J.F. Otegi[1,2], C. Dorn[1], R. Helled[1], F. Bouchy[2], J. Haldemann[3], and Y. Alibert[3]

[1] Institute for Computational Science, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland
e-mail: jon.fernandezotegi@unige.ch
[2] Observatoire Astronomique de l'Université de Genève, 51 Ch. des Maillettes, – Sauverny – 1290 Versoix, Switzerland

[3] Departement of Space Research & Planetary Sciences, University of Bern, Gesellschaftsstrasse 6, CH-3012 Bern, Switzerland

June 11, 2021

ABSTRACT

Context. Exoplanet characterization is one of the main foci of current exoplanetary science. For super-Earths and sub-Neptunes, we mostly rely on mass and radius measurements, which allow us to derive the mean density of the body and give a rough estimate of the bulk composition of the planet. However, the determination of planetary interiors is a very challenging task. In addition to the uncertainty in the observed fundamental parameters, theoretical models are limited owing to the degeneracy in determining the planetary composition.
Aims. We aim to study several aspects that affect the internal characterization of super-Earths and sub-Neptunes: observational uncertainties, location on the M-R diagram, impact of additional constraints such as bulk abundances or irradiation, and model assumptions.
Methods. We used a full probabilistic Bayesian inference analysis that accounts for observational and model uncertainties. We employed a nested sampling scheme to efficiently produce the posterior probability distributions for all the planetary structural parameter of interest. We included a structural model based on self-consistent thermodynamics of core, mantle, high-pressure ice, liquid water, and H-He envelope.
Results. Regarding the effect of mass and radius uncertainties on the determination of the internal structure, we find three different regimes: below the Earth-like composition line and above the pure-water composition line smaller observational uncertainties lead to better determination of the core and atmosphere mass, respectively; and between these regimes internal structure characterization only weakly depends on the observational uncertainties. We also find that using the stellar Fe/Si and Mg/Si abundances as a proxy for the bulk planetary abundances does not always provide additional constraints on the internal structure. Finally we show that small variations in the temperature or entropy profiles lead to radius variations that are comparable to the observational uncertainty. This suggests that uncertainties linked to model assumptions can eventually become more relevant to determine the internal structure than observational uncertainties.
Conclusions.

1. Introduction

Over the past few years, the characterization of planet interiors has been the subject of extensive research. The large number and diversity of discovered exoplanets has allowed us to identify multiple planet populations. Among these, there is an increasing interest in super-Earths and sub-Neptunes, which cover the transition from terrestrial planets to gas giants and have no analog in our solar system. Major improvements in observational techniques allow for relatively precise measurements of mass and radius. The precision of the planetary radius is limited by the uncertainty of the stellar size, since the transit depth scales as R_P^2/R_*^2. Recently Berger et al. (2018) presented revised radii of more than 180 000 Kepler stars, leading to a remarkable improvement of the median radius precision. In some cases, space missions can perform high precision photometry and asteroseismology and can reach relative radius uncertainties of about 3% (e.g., Hatzes 2016). In addition, the current most advanced spectrographs have radial velocity precision of 1m/s. This precision was recently improved with instruments such as ESPRESSO (Echelle Spectrograph for Rocky Exoplanets and Stable Spectroscopic Observations; e.g., Pepe et al. (2018) and references therein), which is expected to have an accuracy close to 10 cm/s. Therefore, a significant improvement in the mass determination is also expected, allowing us to reach a relative uncertainty better than 10%.

The masses and radii can be used to estimate the interior structure and composition of a planet. However, determining the internal structure is extremely challenging as a consequence of intrinsic degeneracy because several compositions can lead to identical mass and radius (e.g., Rogers & Seager 2010; Lopez & Fortney 2014; Dorn et al. 2015, 2017; Lozovsky et al. 2018). Furthermore, for a planet of given mass and composition, the radius depends on several aspects such as the choice of equation of state (EOS), the envelope structure (differentiated, fully mixed, or with a compositional gradient), or the temperature. This degeneracy is critical owing to the large number of free parameters needed to model the interior of an exoplanet and the few observational constraints. In order to determine how well one interior model compares with the other possible models that also fit the data and which structural parameters can be constrained, Dorn et al. (2017) presented a generalized Bayesian inference method to quantify the degeneracy and correlation of the planetary structural parameters.

In this work we explore the limitations of constraining the internal structure of super-Earths and sub-Neptunes (focusing on planets with masses up to $25M_\oplus$ and radii up to $3.5M_\oplus$). We use

a Bayesian inference analysis together with a nested dampling technique (e.g., Skilling 2004) to discuss several aspects that affect interior characterization: observational uncertainties, location in the mass-radius (M-R) diagram, additional observational constraints as bulk abundances, masses derived from published M-R relationships for fixed radii, and the uncertainty related to model assumptions. First, we study the influence of the data uncertainty on the determination of the internal structure. In Dorn et al. (2015), they quantified the information gained by higher data precision. This study aims to show systematically the effect of data uncertainty for a wide range of masses and bulk densities ($0.5 - 14.7\ g.cm^{-3}$). Additional constraints are crucial to reduce the degeneracy, and the assumption that relative abundances of refractory elements (e.g., Fe/Si, Mg/Si) of a planet are similar to that of its host star has been proposed to reduce the existing degeneracy (e.g., Grasset et al. 2009; Dorn et al. 2015). Several solar system and planet formation studies have stated that there is a direct correlation between stellar and planetary relative bulk abundances (e.g., Carter et al. 2012; Lodders 2003; Drake & Righter 2002; McDonough & Sun 1995; Bond et al. 2010; Elser et al. 2012; Johnson et al. 2012; Thiabaud et al. 2015; Wang et al. 2018). We explore under what conditions interior estimates can be improved by constraints on planetary bulk abundances taken from stellar proxies.

Most of the discovered exoplanets do not have measured masses and radii, and published M-R relationships allow us to estimate the mass for a given radius and vice versa. In addition, M-R relationships describe the main properties of various classes of exoplanets. Several studies have been dedicated to the investigation of the M-R relationship of observed exoplanets. The M-R relationships are power laws of the type $M = AR^B$ and are based on exoplanet data (e.g., Weiss et al. 2013; Weiss & Marcy 2014; Wolfgang et al. 2015; Bashi et al. 2017). Recently, Zeng et al. (2016) inferred a semiempirical M-R relationship depending on the core mass fraction, followed by a detailed forecasting model using a probabilistic M-R relationship via Markov chain Monte Carlo (MCMC) (e.g., Chen & Kipping 2017). In addition, in Otegi et al. (2019) we presented an updated exoplanet catalog based on reliable, robust, and as much as possible accurate mass and radius measurements of transiting planets up to 120 $M_\oplus$, and we inferred two new empirical M-R relationships corresponding to rocky and volatile-rich populations. We study inferred interior parameters using the mass calculated from above-mentioned M-R relationships.

Finally, we study how variations in the temperature profiles lead to radius uncertainties comparable to observational uncertainties. In this work, we assess the importance of the various sources of uncertainty related to model assumptions and study the effect of variations of the temperature profile on the planetary radius, which has not been explored yet.

2. Method

2.1. Synthetic sample

We studied the internal structure of a sample of 20 synthetic planets with different masses and radii up to $25M_\oplus$ and $3.5R_\oplus$ (listed in Table 1). We aim to better understand the transition between rocky and volatile-rich exoplanets. Fulton et al. (2017) found that there is a lack of planets with radii between $1.5R_\oplus$ and $2R_\oplus$, known as "evaporation valley" (e.g., Owen & Wu 2013; Jin et al. 2014; Lopez & Fortney 2014), suggesting a transition between the super-Earth and sub-Neptune populations. In addition, in Otegi et al. (2019) we did a careful analysis to build an exoplanet catalog as reliable as possible and we found a transition region from rocky to volatile-rich exoplanets, which corresponds to a mass in the range 5-25 $M_\oplus$, and a radius in the range 2-3 $R_\oplus$. The M-R range covered by the synthetic planets includes the transition between these two populations. Furthermore, more massive planets may have massive atmospheres in which the effects of electron degeneracy pressure are significant, and these are not accounted for in our atmospheric model.

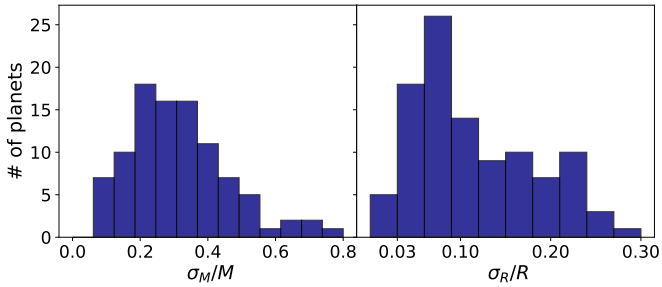

Fig. 1. Distributions of the relative mass (left) and radius (right) uncertainties for the 115 observed exoplanets less massive than $25M_\oplus$ from the NASA Exoplanet Archive.

The 20 synthetic planets for our sample were chosen to represent the properties of the exoplanet catalog presented in Otegi et al. (2019), in which we presented a revisited exoplanet catalog based on robust, reliable, and precise mass and radius measurements for transiting exoplanets (with $\sigma_M/M = 25\%$ and $\sigma_R/R = 8\%$). This exoplanet catalog is dominated by exoplanets for which the masses have been measured through radial velocity, so it is dominated by relatively short-period exoplanets. Exoplanets orbiting close to their host stars are expected to have smaller atmospheres (lost through evaporation) and therefore the sample used in this work may be biased toward higher densities. Figure 1 shows histograms of the relative mass and radius error for the observed exoplanets from the NASA Exoplanet Archive [1]. Currently, the distributions are peaked at $\sim 8\%$ in radius and $\sim 35\%$ in mass, but upcoming space and ground-based missions are expected to improve these values to a few percent. We used uncertainties of $3\%, 5\%$, and 10% in radius and $5\%, 15\%$, and 30% in mass for our synthetic planets to cover the range from the current most common uncertainties to the smallest uncertainties. Figure 2 shows the synthetic planets, the observed population exoplanets (from our revisited catalog in Otegi et al. (2019)), and M-R curves for idealized compositions of iron, Earth-like, and water ice. The bulk densities of the synthetic planets cover the range from the physically motivated high limit (pure iron) to the minimum density of an observed exoplanet with a mass up to $20M_\oplus$.

The irradiation has a significant effect on the interior structure determination of exoplanets with thick volatile envelopes. The exoplanets belonging to the volatile-rich population in the revisited exoplanet catalog of Otegi et al. (2019) are typically irradiated with fluxes of few hundreds times the Earth flux. We therefore used a default irradiation of $100F_\oplus$ for the synthetic

[1] exoplanetarchive.ipac.caltech.edu

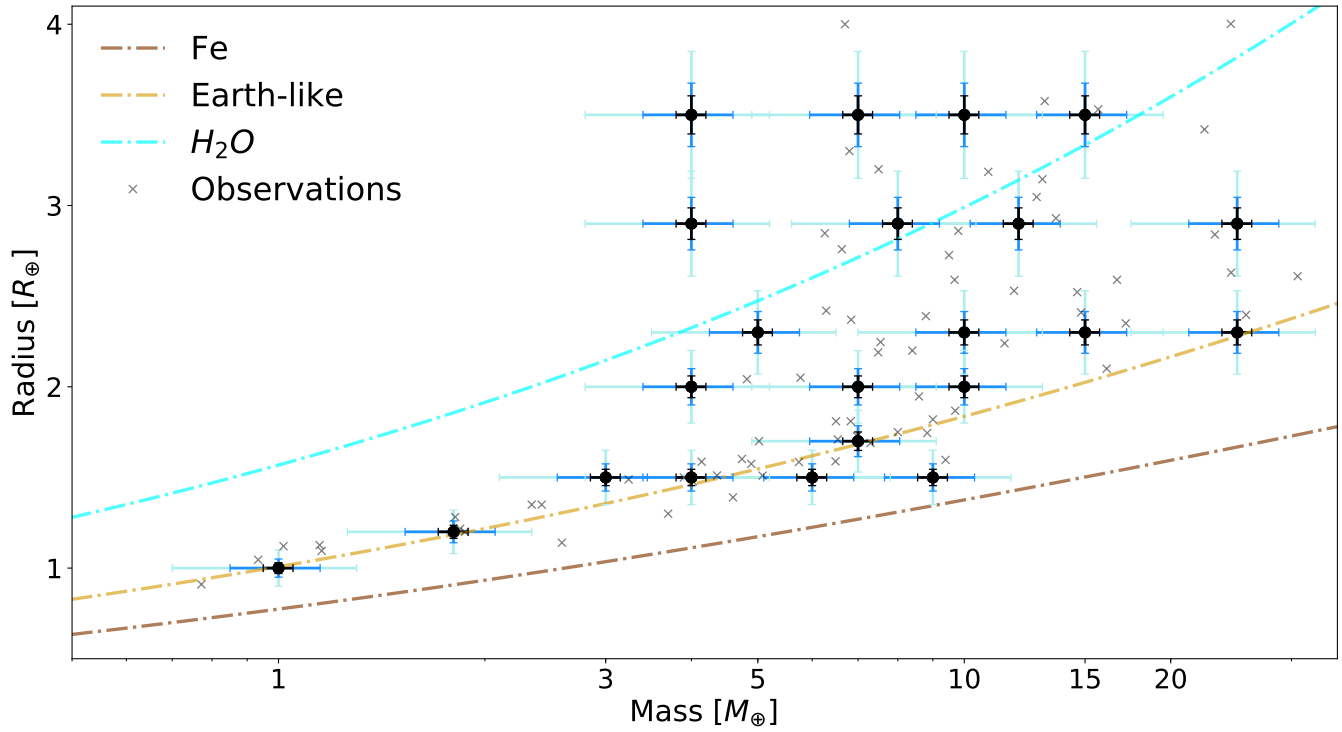

Fig. 2. M-R diagram with the synthetic planets used in the study (black dots). The black, blue, and light blue bars correspond to uncertainties of 5%, 10%, and 20% in mass and 3%, 5%, and 10% in radius, respectively. The crosses represent observed planets from the revisited catalog in Otegi et al. (2019). The synthetic planets are plotted against the composition lines of pure-iron (brown), Earth-like (light brown), and water ice (blue).

Table 1. Mass, radius, and bulk density of the synthetic planets.

Case	Mass [$M_\oplus$]	Radius [$R_\oplus$]	Density [g/cm^3]
A	1	1	5.51
B	1.2	1.8	5.73
C	9	1.5	14.7
D	6	1.5	9.8
E	3	1.5	4.9
F	10	2	6.9
G	7	2	4.8
H	4	2	2.8
I	25	2.3	11.4
J	15	2.3	6.8
K	10	2.3	4.5
L	5	2.3	2.3
M	25	2.9	5.7
N	12	2.9	2.7
O	8	2.9	1.8
P	4	2.9	0.9
Q	15	3.5	1.93
R	10	3.5	1.3
S	7	3.5	0.9
T	4	3.5	0.5

planets. The impact of the irradiation on the internal structure determination is studied further in section 3.5.

2.2. Structure model

We used the structure model presented in Dorn et al. (2017), which assumes a pure iron core, silicate mantle, pure water layer, and H-He atmosphere. This model uses self-consistent thermodynamics in the core, mantle, high-pressure ice, and water ocean. However in this study we used EOS for hexagonal close packed iron for super-Earth conditions presented in Hakim et al. (2018) for the core density profile. These are based on density functional theory results up to 137 TPa. Unlike Earth's core (e.g., Badro et al. 2007), we did not consider the presence of light elements in the core.

The silicate mantle is assumed to be made of oxides $Na_2O - CaO - FeO - MgO - Al_2O_3 - SiO_2$. Equilibrium mineralogy and density were computed as a function of pressure, temperature, and bulk composition by minimizing Gibbs free energy (e.g., Connolly 2009). For the water layers, we followed the approach presented in Sotin et al. (2010), which uses a temperate Birch-Murnaghan EOS including thermal corrections. The water can be in the solid, liquid, or supercritical phase depending on the pressure and temperature. The surface temperature of the water layer is set to be equal to the temperature of the bottom of the gas layer.

For the gas layer, the equations of hydrostatic equilibrium, mass conservation, and energy transport are solved. We assumed an envelope with an elemental composition of H, He, C, and O, which are fundamental for the formation of key atmospheric molecules such as H_2, CO, CO_2, and CH_4 (e.g., Madhusudhan 2012; Lodders & Fegley 2002; Visscher & Moses 2011; Heng & Lyons 2016). We used the Chemical Equilibrium with Applications (CEA) package (e.g., Gordon et al. 1984) for the EOS, which performs chemical equilibrium calculations for an arbitrary gaseous mixture. For the energy transport, we used the model presented in Jin et al. (2014), where an irradiated atmosphere is assumed at the top of the gaseous envelope.

Within the envelope, the usual Schwarzschild criterion is used to distinguish between convective and radiative layers. More details on the structural model are found in Dorn et al. (2017).

2.3. Bayesian inference based on nested sampling scheme

For the interior characterization we used the Bayesian inference analysis based on a nested sampling scheme using the PyMultiNest package (e.g., Buchner et al. 2014). Bayesian inference computes the posterior probability according to Bayes' theorem, which states that the probability for a fixed model parameter $\boldsymbol{x}$ given a set of data $\boldsymbol{d}$ is given by

$$P(\boldsymbol{x}|\boldsymbol{d}) = \frac{P(\boldsymbol{x})\, P(\boldsymbol{d}|\boldsymbol{x})}{P(\boldsymbol{d})}, \quad (1)$$

where $P(\boldsymbol{x})$ is the prior probability of $\boldsymbol{x}$ before the data is observed, $P(\boldsymbol{d})$ the Bayesian evidence, and $P(\boldsymbol{d}|\boldsymbol{x})$ the likelihood function. The likelihood function represents the probability of observing the data given the model parameter $\boldsymbol{x}$, and is given by

$$P(\boldsymbol{d}|\boldsymbol{x}) = \frac{1}{(2\pi)^{N/2}(\prod_{i=1}^{N}\sigma_i^2)^{1/2}}\, exp(-\frac{1}{2}\sum_{i=1}^{N}\frac{(g_i(\boldsymbol{x})-\boldsymbol{d}_i)^2}{\sigma_i^2}), \quad (2)$$

where N is the number of data points, σ_i the uncertainties of the i*th* datum, and g($\boldsymbol{x}$) the operator linking the model parameters with the data, that is, $\boldsymbol{d} = g(\boldsymbol{x})$. Our posterior probability distribution cannot be derived analytically, so we used a nested sampling scheme (e.g., Skilling 2004). The aim of Monte Carlo technique is to efficiently evaluate of the Bayesian evidence, but also to produce posterior probability distributions. The main strengths of nested sampling with respect to other sampling methods are the small amount of problem-specific tuning required and high efficiency.

In short, the nested sampling scheme works as follows. The algorithm samples some number of live points randomly from the prior $P(\boldsymbol{x})$. The likelihood $P(\boldsymbol{d}|\boldsymbol{x})$ is evaluated at each of these points. At each iteration the point with the lowest likelihood L_{min} is replaced by a new point sampled from the region of prior with likelihood $P(\boldsymbol{d}|\boldsymbol{x}) > L_{min}$, keeping the number of live points constant. This process is continued until Bayesian evidence reaches some specified value (typically 0.5 in log-evidence). It produces a list of samples that can be used to produce marginalized posterior distributions. The sampling scheme is controlled by two main parameters: the number of live points and the maximum efficiency parameter. The number of live points has to be large enough to adequately sample the parameter space, for which we used a recommended value of 1000. The maximum efficiency controls the sampling volume at each iteration, which is equal to the sum of the volumes enclosing the active point set. We set it to 1 to obtain the maximum efficiency.

2.4. Parameters and priors

The composition of the planets are given by the masses of the iron core, mantle layer, water layer, and H-He envelope. As already mentioned we assume a pure iron core, the mantle composition is given by the Mg/Si_{mantle}, Fe/Si_{mantle}, Al/Si_{mantle}, Ca/Si_{mantle}, and Na/Si_{mantle} fractions. The water layer is considered to be pure H_2O and for the atmosphere the metallicity Z, luminosity L, and the irradiation temperature T_{irr} are used to model the atmospheric structure. If not stated otherwise we fix the Mg/Si_{mantle}, Fe/Si_{mantle}, Al/Si_{mantle}, Ca/Si_{mantle}, and Na/Si_{mantle} fractions, Z, and T_{irr} while only the layer masses are varied. The luminosity is scaled as $L \propto M^{2.76}$, which fits the Jovian planets (e.g., LaViolette 2006).

For the nested sampling scheme we assumed a uniform prior distributions on the layer masses from 0 to the target planets mass. We would like to emphasize that for a given target planet, depending on the scientific question, a different choice of priors might be necessary. The set of prior presented here is one possible way of setting up the analysis but not the only one[2]. The impact of different priors should be subject of further studies.

As data variables we chose the total mass and total radius of the planet, except in section 3.4 where we also use the bulk Mg/Si and bulk Fe/Si fractions as data; in that case the Mg/Si_{mantle} and Fe/Si_{mantle} fractions are also varied, assuming a uniform prior. The total mass is calculated as the sum of all layer masses. The total radius and the bulk Mg/Si and Fe/Si fractions are an outcome of the structure model. In Table 2 we summarize the model parameters, their priors, and the data variables.

Table 2. Summary of model parameters, priors, and data.

Model Param.	Priors	Data
M_{Core}	$\mathcal{U}(0, M)$	Total Mass
M_{Mantle}	$\mathcal{U}(0, M)$	Total Radius
M_{Water}	$\mathcal{U}(0, M)$	Bulk Fe/Si
M_{H-He}	$\mathcal{U}(0, M)$	Bulk Mg/Si
Fe/Si_{Mantle}	const. or $\mathcal{U}(0, 2Fe/Si_{\odot})$	
Mg/Si_{Mantle}	const. or $\mathcal{U}(0, 2Mg/Si_{\odot})$	
Al/Si_{Mantle}	const.	
Ca/Si_{Mantle}	const.	
Na/Si_{Mantle}	const.	
Z	const.	
L	const.	
T_{irr}	const.	

3. Dependence of internal structure determination on observational uncertainties

In this section we explore how the internal structure determination depends on the observational uncertainties for exoplanets of different masses and radii fixing the irradiation to $100F/F_{\oplus}$ and solar Fe/Si and Mg/Si abundances. In particular, we explore how the posterior distribution of the internal parameters depend on the observational uncertainty for the synthetic planets in Table 1. An example is shown in Figure 3, in which the posterior distributions of the internal parameters corresponding to a planet of observed mass and radius of $6M_{\oplus}$ and $1.5R_{\oplus}$ with different observational uncertainties are shown. The first four columns show the posterior distributions of the iron core, mantle, water, and atmospheric masses (output) renormalized by the sum, and the last column shows the mass distribution of the planet (input). The rows correspond to different observed radius uncertainties and the colors to different observed mass uncertainties; black, red, and green for 30%, 15%, and 5%, respectively. We find that for a planet of $6M_{\oplus}$ and $1.5R_{\oplus}$ a decrease in the uncertainties improves the determination of the core mass, but has a negli-

[2] For example, we could also choose to sample the layer mass fractions, instead of the layer masses, from the 3D probability simplex and use a uniform prior on the total mass of the planet.

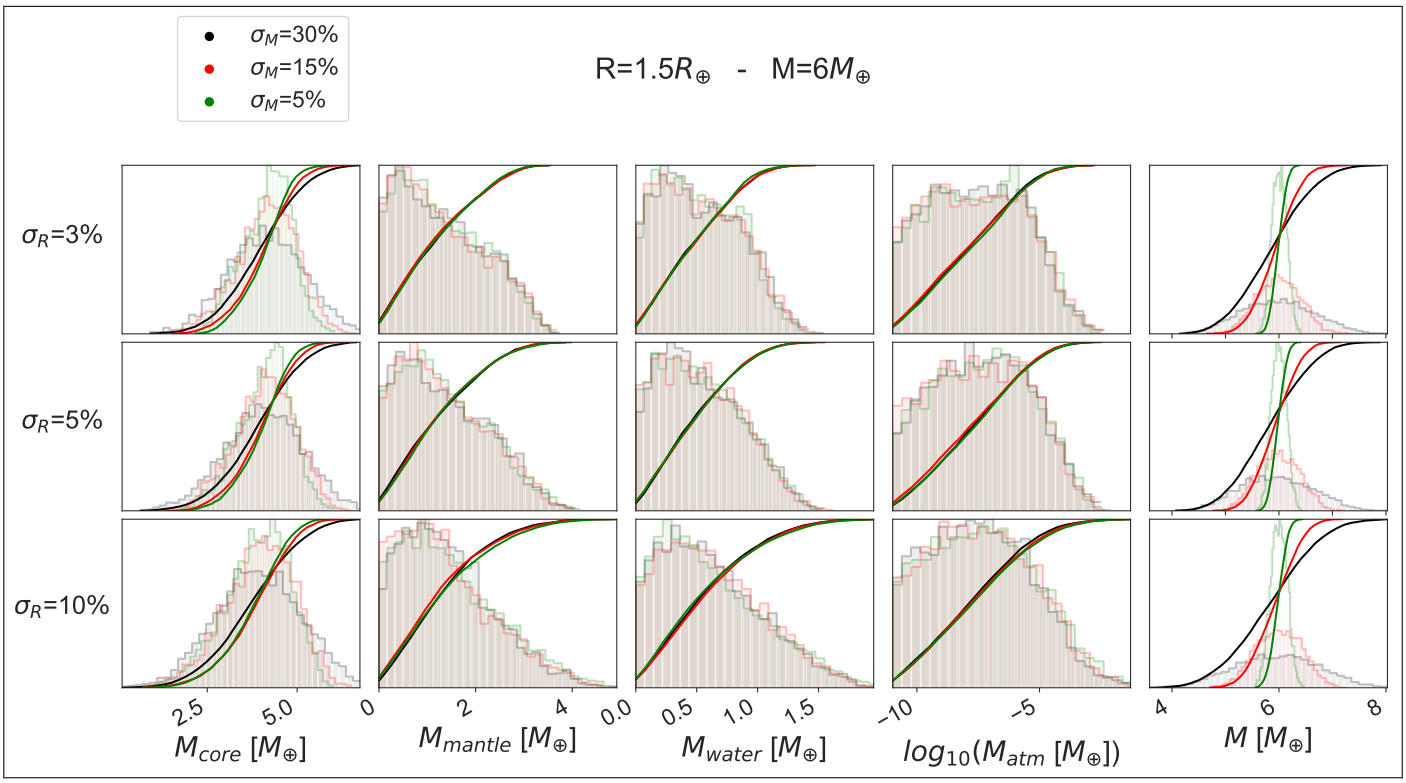

Fig. 3. Posterior distributions of core, mantle, water, atmospheric, and total masses of a planet with mass of $6M_\oplus$ and radius of $1.5R_\oplus$. The rows correspond to different radius uncertainties and colors in the subplots correspond to different mass uncertainties.

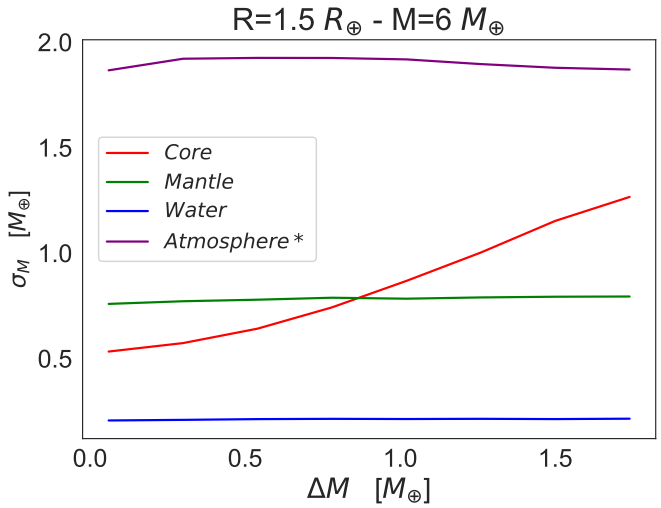

Fig. 4. Standard deviation of the posterior distributions of core, mantle, water, and atmospheric masses vs. width of the total mass distribution for a planet of $R = 1.5R_\oplus$ and $M = 6M_\oplus$. For the atmosphere, the standard deviation of the posterior distribution in logarithmic scale is used.

gible effect on the determination of the other internal parameters.

In Figure 4 the standard deviation (σ) of the posterior distributions of the internal structure parameters are shown versus the mass uncertainty (ΔM), for a radius of $1.5R_\oplus$ and a mass of $6M_\oplus$. We aim to use this approach to explore what interior parameters are better constrained when the mass precision is higher, keeping the radius fixed with zero uncertainty. In this work we assume a flat uniform distributed uncertainty for the planetary mass since it simplifies the analysis and allows us to consider cases with uncertainties close to 0. The equation $\Delta M = 0$ shows the intrinsic degeneracy of the interior parameters when data uncertainty is zero for a planet of $1.5R_\oplus$ and $6M_\oplus$. When ΔM increases, $\boldsymbol{\sigma}$ of the core mass increases, meaning that this is the only tested interior parameter that is sensitive to a change in observational uncertainties. It is in agreement with what is shown in Figure 3 where we use a more realistic Gaussian distribution for the mass and radius. It is interesting to note that the evolution of σ_M with ΔM is nearly linear. The slope $\boldsymbol{\delta\sigma_M/\delta\Delta M}$ contains very valuable information: it indicates how much interior estimates (i.e., layer mass fractions) can be improved by increasing data precision (i.e., mass). Figure 5 shows these slopes $\delta\sigma_M/\delta\Delta M$ of the core, mantle, water, and atmospheric masses for all synthetic planets. The results allow us to differentiate three different regimes, described in the following subsection.

3.1. Planets below Earth-like composition line

We refer to planets C and D (Table 1). For these planets, a decrease of the observational uncertainties leads to a better determination of the core mass. In this density regime, planets do not have a significant gaseous envelope and they require a large amount of iron. Then, variation in the mass or radius distributions is reflected in the core mass, since there is no other interior layer that can account for such high planet bulk densities.

As mentioned above, Figure 5 is constructed assuming flat uniform error distributions for mass and radius. To check whether the conclusions are also valid for more realistic distributions we compare the results using the Gaussian and flat distributions, where σ of the Gaussian is equal to ΔM of the flat

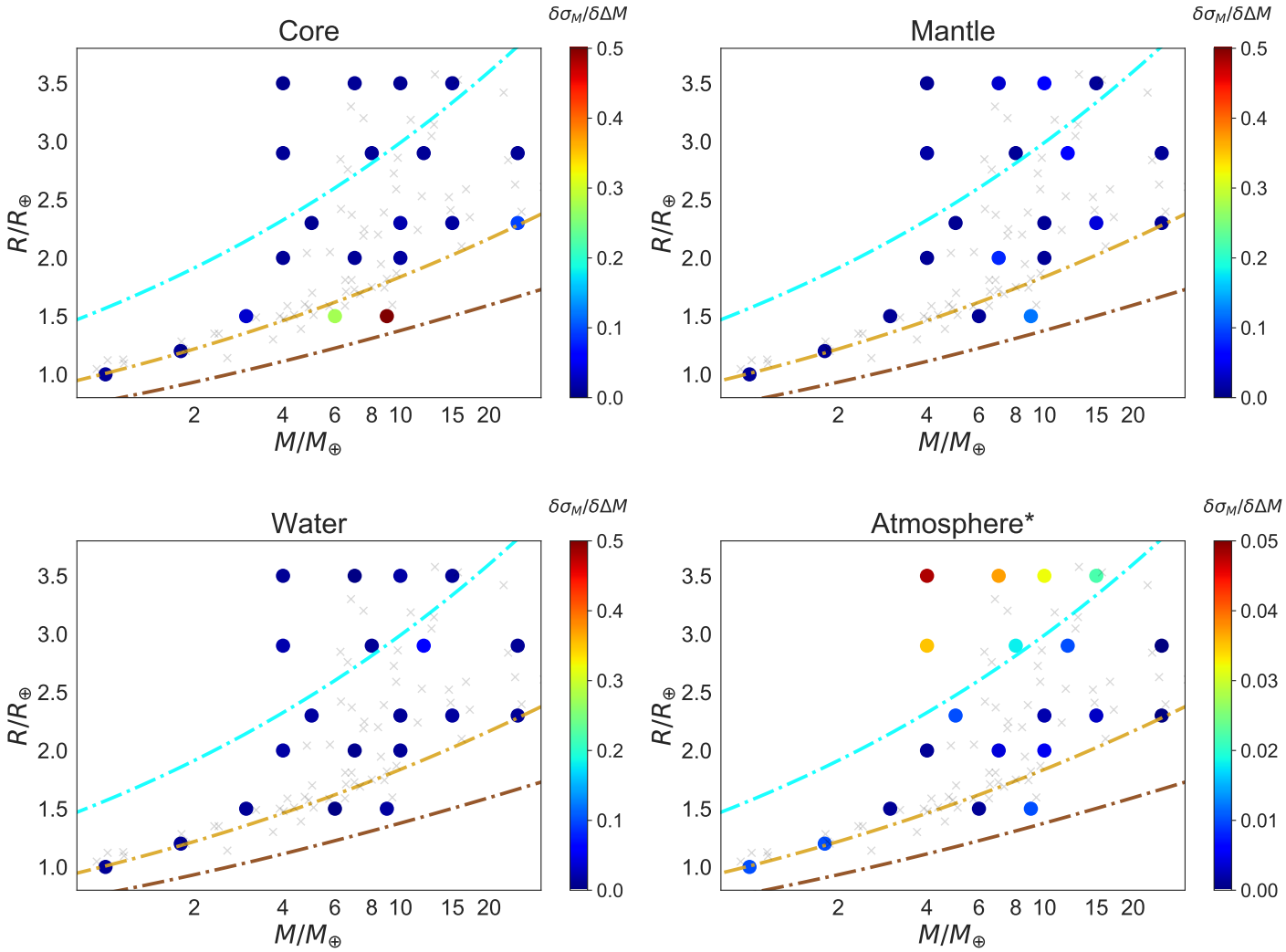

Fig. 5. Model parameters and synthetic planets $\Delta\sigma/\Delta M$ (see text for further details). A higher value of $\Delta\sigma/\Delta M$ indicates that a decrease of the observational uncertainty provides additional information of the internal parameter. The crosses represent the observed planets from Otegi et al. 2019. Composition lines of pure-iron (brown), Earth-like planet (light brown), and water ice (blue) are also plotted.

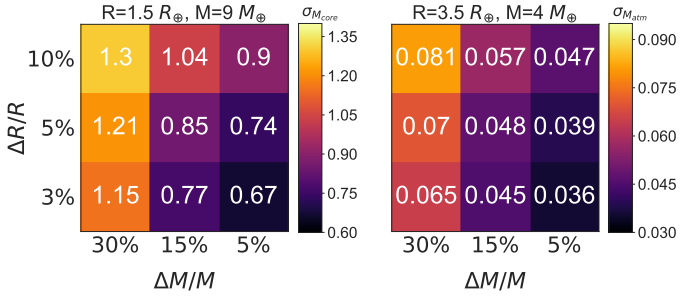

Fig. 6. Observational uncertainties in mass and radius versus σ_{core} (left) and σ_{atm} (right) for the two synthetic planets with the highest and lowest densities.

distribution. The obtained results with the flat distributions are consistently slightly higher than with the Gaussian results, but the general shape is nearly equal. Figure 6 shows σ_{core} versus observational uncertainties in mass and radius using Gaussian distributions. It shows that σ does not decrease uniformly with uncertainty and that the internal structure determination is more sensitive to an improvement of data uncertainties when data uncertainties are large. Furthermore, we find that a decrease of the observed mass uncertainty is much more effective to better determine the core mass than a decrease in the radius uncertainty. This is because the core mass directly affects the planetary mass, while its effect on radius is smaller.

3.2. Planets above pure-water composition line

We find that when a planet lies above the pure-water composition line a decrease in the mass and radius uncertainties leads to a better determination of the H-He envelope mass. The explanation is similar to that discussed above for high-density planets: for a given mass, low-density planets require thick gaseous envelopes to match the observed radius, and this significantly reduces the degeneracy. Therefore a change in the total mass/radius distributions has a direct impact on the properties of the gaseous envelope (i.e., mass fraction, thickness).

The right panel of Figure 6 shows the σ_{atm} versus the observational uncertainties in mass and radius for the low-density synthetic planets with realistic Gaussian mass and radius distributions. Interestingly, in this regime decreasing the

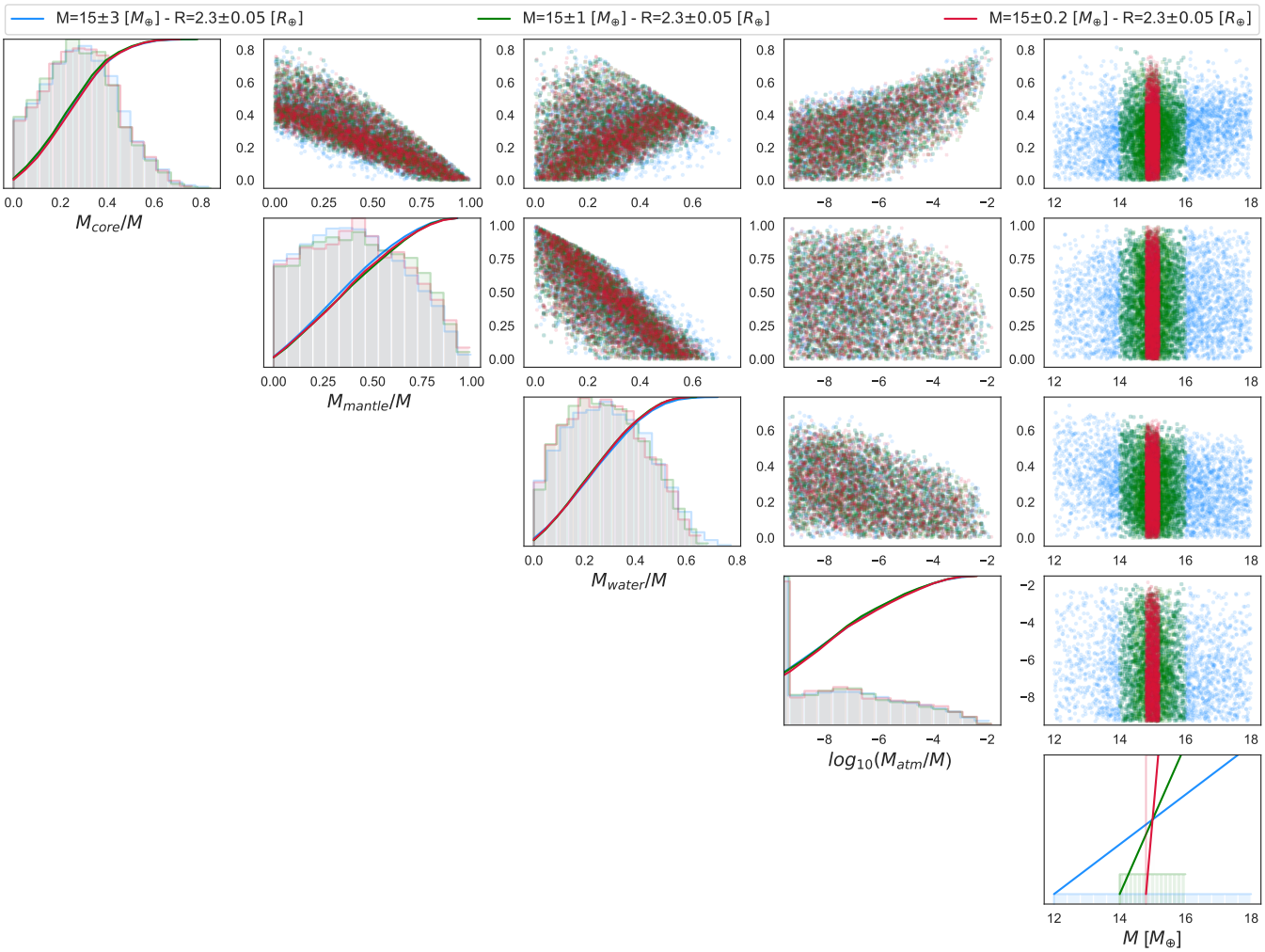

Fig. 7. Illustration of the degeneracy in the intermediate density regime previously defined. The posterior distributions the internal parameters of a planet with fixed radius of $2.3R_\oplus$ and mass of $15M_\oplus$ are shown The different colors represent different widths of the mass distribution.

radius uncertainty is more effective than decreasing the mass uncertainty to constrain the atmospheric mass. This can be understood by the fact that the atmospheric mass fraction of the low-density planets in the synthetic sample is nearly negligible, but the radius fraction is significant.

3.3. Planets between Earth-like and pure-water composition lines

Finally, we find that planets with densities between the Earth-like and pure-water composition lines are the most degenerate. For these planets, decreasing the mass and radius uncertainties does not improve the determination of any internal structure parameter. Even when removing one compositional layer, there is a substantial degeneracy among the other three. Figure 7 shows the posterior distributions the interior parameters of a planet with fixed radius of $2.3R_\oplus$ and mass of $15M_\oplus$. The different colors represent different widths of the total mass distribution. The posterior distributions of the internal parameters do not cover the same region in the two-dimensional parameter space. The posterior distributions only overlap when projecting them into a one-dimensional histogram. Therefore, for these planets it would be crucial to have additional information that can further constrain the internal structure, such as as atmospheric metallicity measurements from space missions like ARIEL (Atmospheric Remote-sensing Infrared Exoplanet Large-survey; e.g., Tinetti et al. (2016)) and JWST (James Webb Space Telescope; e.g., Beichman & Greene (2018)).

3.4. Using stellar abundances as an additional constraint

As shown in Dorn et al. (2017), abundance constraints of Fe/Si and Mg/Si from the host star can serve as a proxy for the planet bulk abundance and can reduce model degeneracy. This assumption reproduces the composition of the Earth (e.g., Javoy et al. 2010) and Mars (e.g., Khan & Connolly 2008), but not the one of Mercury, for which a post-formation giant impact scenario is often considered (e.g., Benz et al. 2007; Chau et al. 2018). Thiabaud et al. (2015) used a chemical model to link the composition of 18 synthetic stars with solar mass and luminosity with the composition of the hosted planets, and they found a close relation. Nevertheless, Wang et al. (2018) showed that there are some differences for the Earth compared to the Sun, and that volatilization trends in the bulk composition of exoplanets should be considered (e.g., Dorn et al. 2019).

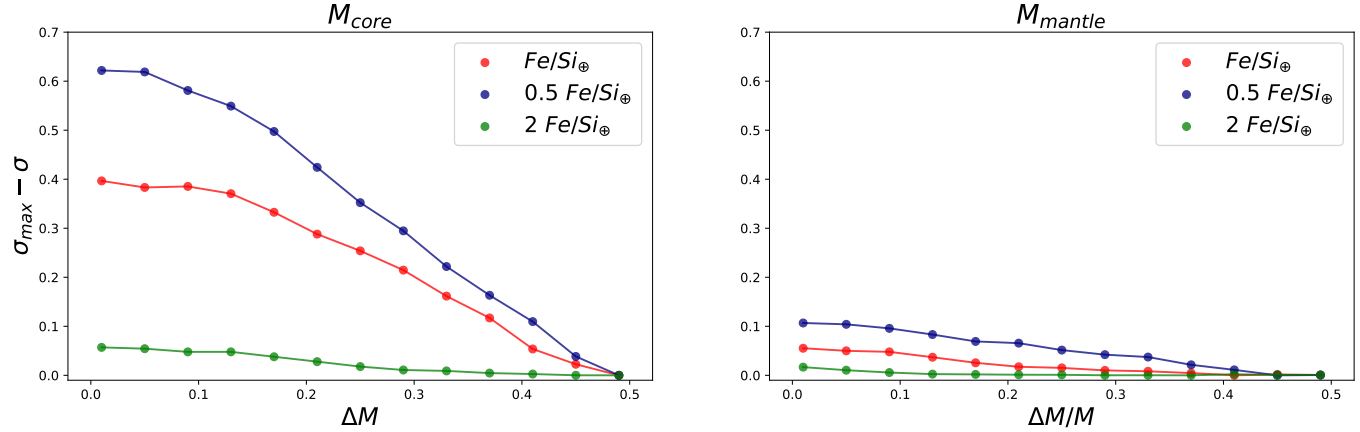

Fig. 8. Standard deviation of the core mass (left) and mantle mass (right) for a planet of mass $15M_\oplus$ and radius $2.3R_\oplus$ vs. the width of the mass distribution. The red, green, and blue lines represent different Fe/Si in the planet. A higher value of $\sigma_{max} - \sigma$ indicates that the internal parameter is better constrained with respect to the reference, at which $\sigma = \sigma_{max}$.

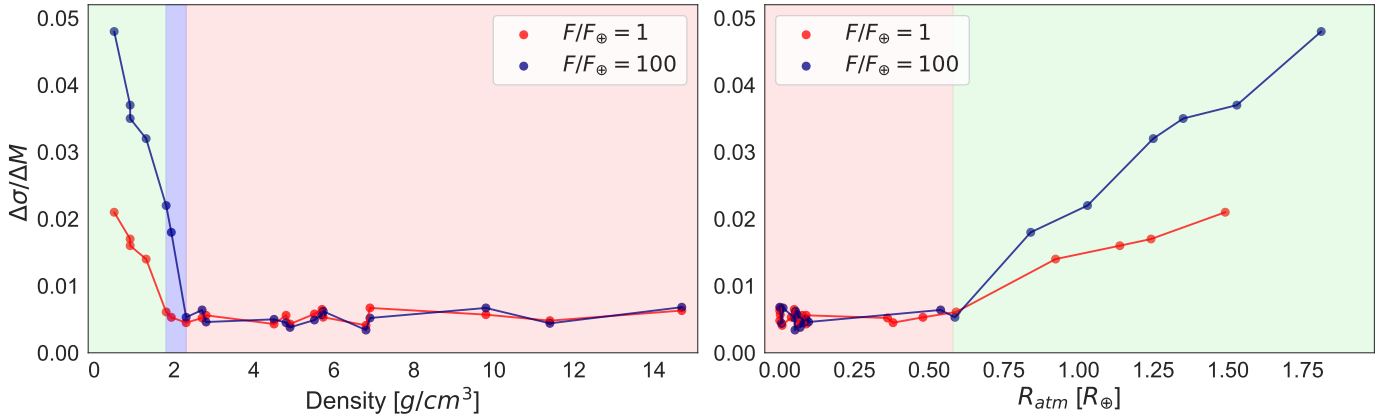

Fig. 9. Atmospheric mass $\Delta\sigma/\Delta M$ for the synthetic planets vs. the bulk density (left) and median atmospheric thickness (right) for different irradiation. The light green and red regions represent the ranges in which decreasing the observational uncertainties lead and do not lead to better constraints on the atmospheric mass, respectively. The blue region represents the density range in which planets with an irradiation of 100 Earth fluxes have better constrained atmospheric mass with decreasing uncertainties and planets with an irradiation of one Earth flux do not.

We next explore how the results are affected when assuming that the stellar relative abundances reflect the planetary relative abundances. We then use the bulk Fe/Si and Mg/Si as additional constraints, and introduce two additional model parameters: the Fe/Si and Mg/Si ratios in the mantle. The goal of this section is to evaluate whether using the bulk abundances allows us to improve the estimation of the internal structure using a more detailed interior model. We find that using the bulk Fe/Si and Mg/Si does not necessarily further constrain the planetary internal structure when adding Fe/Si and Mg/Si ratios in the mantle as free parameters. We assume a small uncertainty on the bulk Fe/Si and Mg/Si of 3%. Figure 8 shows the standard deviation of the core mass and mantle mass for a planet of the intermediate-density regime in terms of the width of the total mass distribution. In this case $\sigma_{max} - \sigma$ is displayed, so that a higher value means that the interior parameter is better constrained with respect to the reference, at which $\sigma = \sigma_{max}$. In the case of the chosen synthetic planets, when we introduce the constraint of $\mathrm{Fe/Si_{planet}} = \mathrm{Fe/Si_{star}}$ with $\mathrm{Fe/Si_{star}} < \mathrm{Fe/Si_\odot}$, their core mass fraction is better constrained. The lower $\mathrm{Fe/Si_{star}}$, the better the core mass gets constrained. Regarding their mantle mass fraction, it is only better constrained when $\mathrm{Fe/Si_{star}} < 1/2\ \mathrm{Fe/Si_\odot}$. Considering the bulk stellar abundances mostly affects the core mass fraction determination. For $\mathrm{Fe/Si_{star}} > 2\ \mathrm{Fe/Si_\odot}$ the abundance constraints do not provide additional information on the internal parameters for the tested planets.

3.5. Sensitivity of atmosphere-rich planets to irradiation

The calculations presented above correspond to an irradiation of $100F_\oplus$. In this subsection we explore the sensitivity of the results to the assumed stellar irradiation. It should be noted that our model does not include atmospheric evaporation and the effect of the insolation is only reflected on the equilibrium temperature of the planet. Figure 9 shows $\Delta\sigma/\Delta M$ of the envelope mass for all the simulated planets versus the bulk density and atmosphere radius for different assumed insolations. Clearly, there is a difference when changing the insolation for planets with relatively large atmospheres. A higher insolation increases the temperature of the H-He atmosphere, which leads to an expansion and a decrease in H-He layer density. In general, the atmospheric mass can be better constrained for strongly irradiated planets.

Figure 9 shows whether an improvement in observational uncertainties contributes to constrain the envelope mass depend-

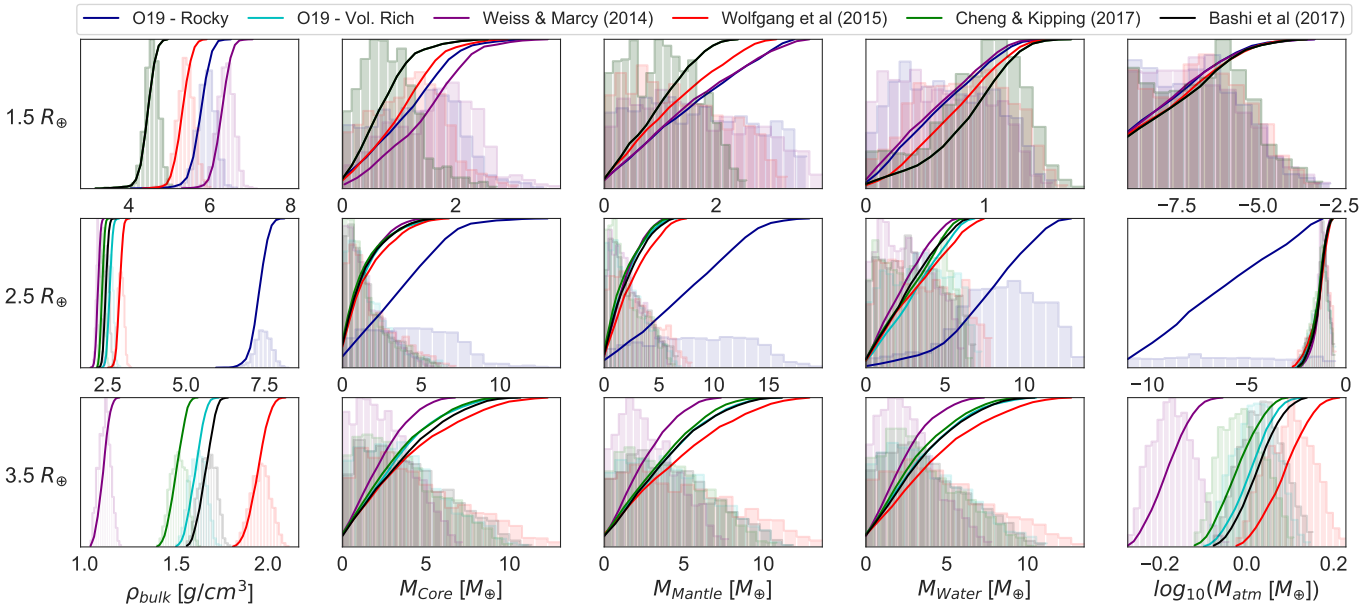

Fig. 10. Posterior distributions of the bulk density and internal parameters of the planet for different radii and masses computed with various published M-R relationships.

Table 3. Masses from published M-R relationships for different radii. O19-R: Otegi et al. (2019)–rocky population; O19-V: Otegi et al. (2019)–volatile-rich population; W14: Weiss & Marcy (2014); W15: Wolfgang et al. (2015); C17: Chen & Kipping (2017); B17: Bashi et al. (2017).

	O19-R	**O19-V**	**W14**	**W16**	**C17**	**B17**
1.5 $R_\oplus$	3.64 $M_\oplus$	-	3.92 $M_\oplus$	3.32 $M_\oplus$	2.80 $M_\oplus$	2.79 $M_\oplus$
2.5 $R_\oplus$	21.24 $M_\oplus$	7.4 $M_\oplus$	6.30 $M_\oplus$	8.32 $M_\oplus$	6.66 $M_\oplus$	7.05 $M_\oplus$
3.5 $R_\oplus$	-	12.59 $M_\oplus$	8.62 $M_\oplus$	15.25 $M_\oplus$	11.77 $M_\oplus$	13.00 $M_\oplus$

ing on the planetary density and irradiation. Planet O ($M = 8M_\oplus, R = 2.9R_\oplus$) has a null $\Delta\sigma/\Delta M$ for Earth-like irradiation, but it is positive for high irradiation. If we look at this depending on the atmosphere radius (median of the posterior distribution) instead of the density, we see that the breakpoint coincides for the two insolations. The determination of atmospheric mass of planets holding atmospheres up to $0.5R_{atm}$ is insensitive to an improvement of observational uncertainties. When the radius of the H-He envelope is lower than $0.5R_{atm}$ there is a strong degeneracy with the other layers, but otherwise a high amount H-He is needed to fit the observed mass and radius, and therefore, variations on the observational uncertainties affect the determination of the H-He envelope. It is important to note that a more realistic model would include water in the atmosphere. However, this would be another free parameter that would increase the degeneracy and would be difficult to determine.

4. Dependence of internal structure determination on the mass derived from M-R relationships.

Since not all planets have measured masses and radii, (e.g., most Kepler planets in the past), published M-R relationships are often used to estimate the mass of a planet for a given radius and vice versa. Figure 4 of Otegi et al. (2019) compares some of the published mass radius relations. This figure shows significant disagreement in mass of approximately 25% for a given radius. Despite these differences, given the degeneracy when determining the internal structure, different masses inferred from various M-R relations could lead to very similar internal structures. In such a case, even if only the radius is measured, it is still possible to infer information on the planetary structure and bulk composition. In this section we explore whether these differences in mass are reflected in the determination of the internal structure parameters, or whether they get diluted because of the degeneracy.

Figure 10 shows the posterior distribution of the internal structure parameters for planets with the same radii and different masses as computed from the published M-R relationships (the masses and radii are listed in Table 3). We do not include other published M-R relationships as those presented in Weiss et al. (2013) or Zeng et al. (2016) because they also depend on an additional parameter (irradiation and core mass fraction, respectively). Since for a radius of $1.5R_\oplus$ the masses corresponding to Chen & Kipping (2017) and Bashi et al. (2017) are almost identical we only show one of them (Chen & Kipping (2017)). We assume a very small uncertainty of 1% for mass and radius.

For a radius of $1.5R_\oplus$ the masses from different M-R relationships lead to significantly different inferred core mass distributions. The mass from Weiss & Marcy (2014) leads to an inferred core mass of $2.1 \pm 0.7M_\oplus$ and Chen & Kipping (2017) to $0.8 \pm 0.5M_\oplus$. At this radius the bulk densities lie between 4g/cm^3 and 7g/cm^3, so these planets are expected to have a negligible gaseous envelope. In addition, the water mass is small compared to the mass of refractory materials. Consequently, the degeneracy is reduced and the internal structure is more sensitive to a change in bulk density. As previously discussed, a change in the bulk density of planets dominated by refractory materials is mainly reflected in the core mass, and for this reason

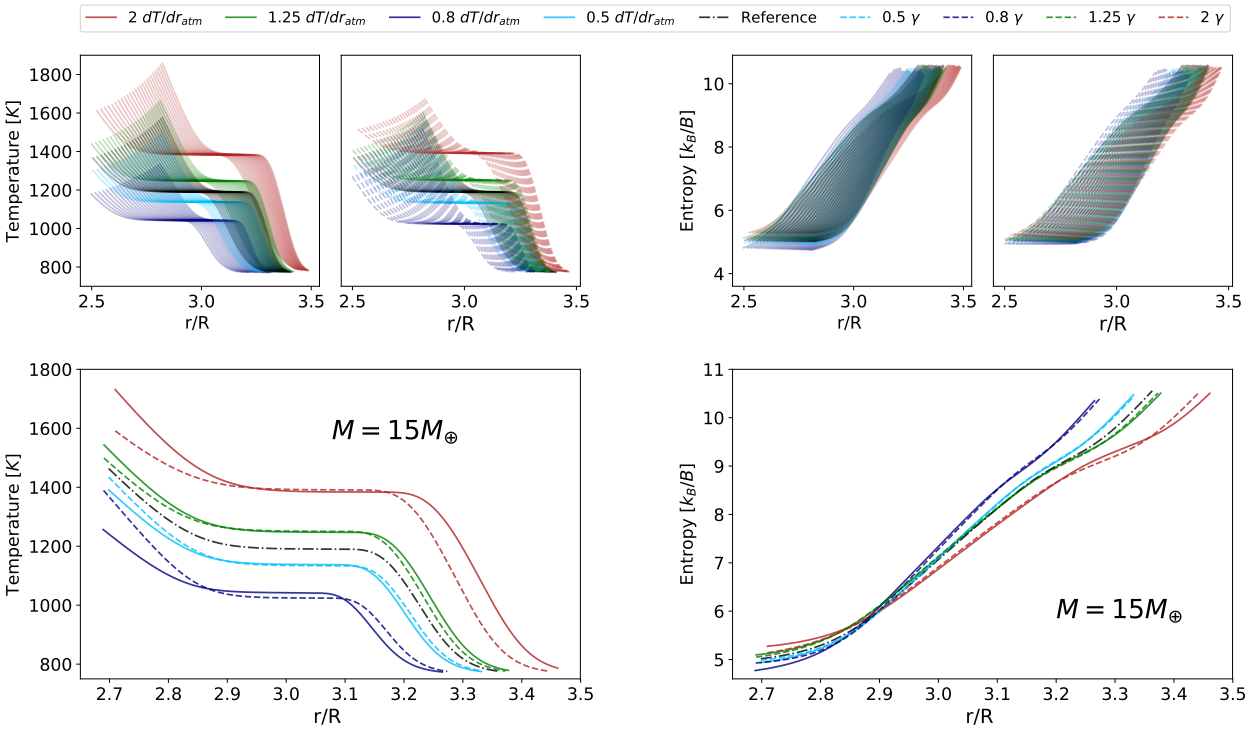

Fig. 11. Atmospheric temperature and entropy profiles when artificially shifting the temperature gradient (solid lines) and γ (dashed lines). The temperature and entropy profiles of 25 planets with masses from $3M_\oplus$ to $20M_\oplus$ (from bottom to top) with a fixed composition of 0.5% H-He mass fraction, 40% water mass fraction, and 59.5% of a rocky core. On the bottom, the profiles corresponding to a mass of $15M_\oplus$.

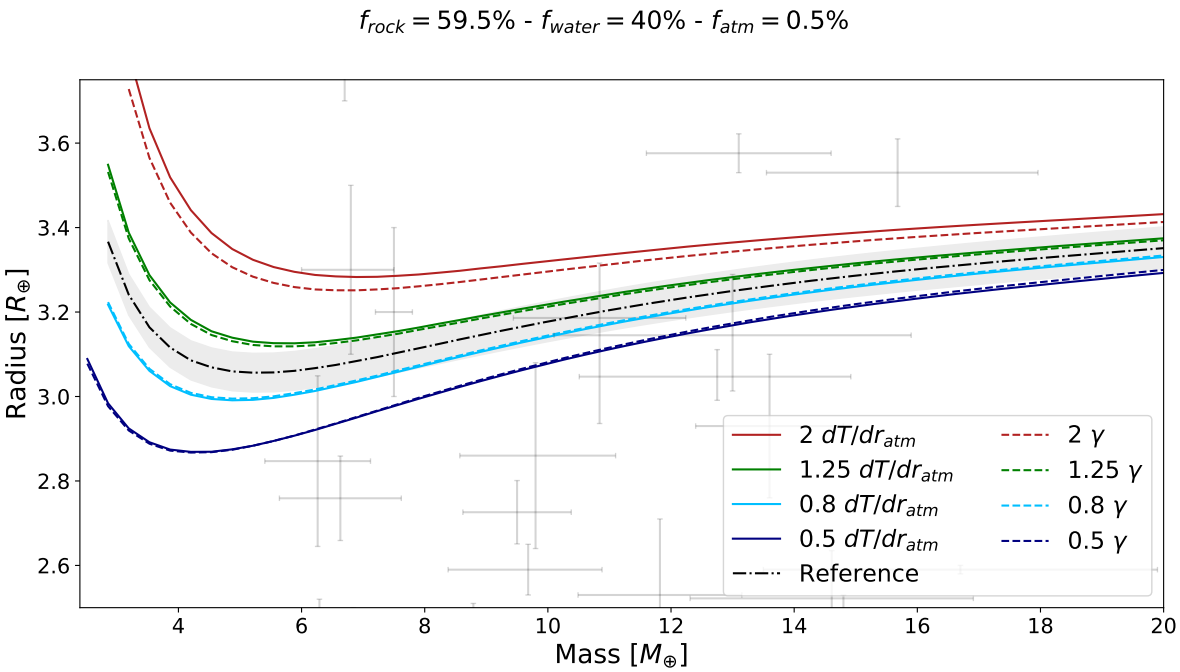

Fig. 12. M-R relationships for a planet with fixed composition and the perturbations applied to the temperature profile as presented in Figure 11. The gray envelope corresponds to a radius uncertainty of 3%, which will soon be reached with future missions.

the core mass distributions for planets with $1.5R_\oplus$ are sensitive to slight mass variations. There are some minor differences between the mantle and water mass distributions, especially between Chen & Kipping (2017) and the other considered studies.

For a radius of $2.5R_\oplus$ the mass computed from the M-R relationship for rocky exoplanets in Otegi et al. (2019) is much higher compared to the masses from other published M-R relationships, and consequently the inferred internal structure is drastically different. Nevertheless, the difference between the internal structures inferred with the masses from other M-R

relationships are very similar to each other. They lie in the intermediate density regime previously discussed, which is mostly degenerate. Consequently, there are not significant differences in the internal structure independently of the M-R relationship used, except for that corresponding to rocky exoplanets as in Otegi et al. (2019); this work estimates a mass for $2.5R_\oplus$ which is approximately three times larger than the other M-R relationships.

Finally, for radius of $3.5R_\oplus$ Weiss & Marcy (2014) leads to much lower masses than the other M-R relationships, and therefore it leads to a significantly different estimated internal structure. The main difference at this radius arises in the posterior distribution of the atmospheric mass. We find that the choice of the M-R relationship for internal structure determination only in the solid-dominated and volatile-rich radius regimes is relevant.

We conclude that planetary internal structure should not be inferred from the radius alone. Even if for radii smaller than $2R_\oplus$ most of the observed exoplanets closely follow the M-R relationships, relatively small variations in mass have a significant impact on the inferred core mass fraction. For radii in the range 2-3$R_\oplus$ where the degeneracy is the strongest, instead, small variations on the mass lead to almost identical inferred internal structure. However, this radius regime corresponds to the transition from rocky to volatile-rich exoplanets, and the large diversity of masses for a given radius does not allow us to infer the internal structure by radius alone.

5. Dependence of the radius on the temperature profile.

Internal structure models depend on the assumed EOS and on several theoretical assumptions such as the internal structure (mixed, differentiated), the envelope structure and its luminosity, and the composition of the planetary layers. Since observational uncertainties are expected to decrease significantly in the upcoming years, it is desirable to explore whether the theoretical uncertainties related to the model assumptions could dominate compared to the observational uncertainties. In this section we investigate how the assumed temperature profiles affect the inferred planetary radius and study what uncertainty in the temperature profile of a planet is required to match the observational uncertainty.

Relative changes in material density have the largest effects on the radius for low-density materials, for example, variations in the metallicity distribution in the envelope significantly affect the inferred planetary radius (e.g., Lozovsky et al. 2018). We therefore explore how variations in the density, temperature, and entropy profiles in the H-He envelope affect the radius of a planet. A common assumption in the envelope structure models is to consider a fully adiabatic temperature profile. However, more realistic models should account for compositional gradients along the envelope, which leads to non-adiabatic temperature profiles. **We do not know a priori what composition gradients might exist in most planets**. In this section, we artificially perturb the temperature profile from the adiabat and study how it affects the inferred planetary radius.

We perturb the temperature profiles in two different ways: shifting the temperature gradient dT/dR in the atmosphere and the γ as defined in Hansen (2008), which is the ratio of visible to infrared opacities. The perturbations in dT/dR and γ are selected to produce a change in the radius comparable to the observational uncertainties. The effect of these perturbations on the temperature and entropy profiles is shown in Figure 11. The top panels show the temperature and entropy profiles of 25 planets with masses from $3M_\oplus$ to $20M_\oplus$ and different assumed atmospheric temperature gradients and γ. The bottom panels illustrate the effect of these variations in the internal temperature and entropy profiles of a planet with a mass of $15M_\oplus$. Figure 12 shows the composition lines corresponding to a fixed composition and the perturbations applied to the temperature profile as presented in Figure 11. The gray envelope corresponds to a radius uncertainty of 3%, which is expected to be reached with future missions. We find that small variations in the temperature profile, such as those corresponding to uncertainties of 20% in the temperature gradient or the ratio of visible and infrared opacities lead to radius uncertainties of 3%. This implies that theoretical uncertainties that are associated with the model assumptions could be larger than the observational uncertainties and therefore dominate the uncertainties in internal structure models. It would be necessary to do a systematic and detailed analysis of the uncertainties introduced by the model assumptions (e.g., EOS, composition of the layers, layer boundaries, intrinsic luminosity of the planet, and atmospheric opacities).

6. Summary and conclusions

We present new internal structure models based on the work of Dorn et al. (2017) with a Bayesian inference analysis using a nested sampling scheme. We explore several aspects that affect the characterization of exoplanets with masses up to $25M_\oplus$ and radii up to $3.5R_\oplus$, such as how variations in the observational uncertainties or location in the M-R affect the inferred internal structure, how the choice of the mass using different M-R relationships for a fixed radius influences the inferred internal structure, and how the atmospheric temperature profile affects the inferred radius. We should keep in mind that these results were computed for a given set of priors. The sensitivity of the results given various priors should be a topic for future investigation. Regarding the sensitivity of internal characterization to observed mass and radii, our main findings are summarized as follows:

- A decrease in observational uncertainties for planets below the Earth-like composition line leads to a better determination of the core mass.
- A decrease in observational uncertainties for planets above the pure-water composition line leads to a better determination of the atmospheric mass.
- A decrease in observational uncertainties for planets between Earth-like and pure-water composition lines does not significantly improve the determination of any internal structure parameter.
- The density boundaries listed above slightly depend on the used interior model (e.g., luminosity, insolation).
- The atmospheric mass of strongly irradiated atmosphere-rich planets can be better determined than for weakly irradiated planets.
- Using the stellar Fe/Si and Mg/Si abundances as a proxy for the bulk planetary abundances does not always help to constrain the planetary internal parameters when adding two extra model parameters related to the mantle composition. This depends on the actual value of the measured stellar abundances and their uncertainties and on the data of planetary mass and radius. Low stellar Fe/Si ratios improve the determination of the core mass for planets with densities between

2.3g/cm^3 and 6.9g/cm^3, where the planetary internal parameters are most degenerate.

We also find that internal structure must not be estimated using radius alone. For a fixed radius, the inferred planetary internal structure using different masses from various published M-R relationships can vary significantly depending on the measured radius. For radii of nearly $1.5R_\oplus$ the choice of the M-R relationship significantly affects the inferred core mass, and for radii above $3R_\oplus$ this choice significantly affects the inferred atmospheric mass. For planets with radii of nearly $2.5R_\oplus$ small differences in mass lead to similar inferred internal structures, but the wide range of possible masses does not allow us to infer the internal structure with radius alone.

We find that uncertainties of 20% in the temperature gradient or the ratio of visible and infrared opacities γ can lead to radius uncertainties of 3%. Observational uncertainties are expected to decrease significantly in the near future, thereby decreasing the uncertainties in both mass (with advanced spectrographs like ESPRESSO) and radius (with next generation space missions like TESS, CHEOPS or PLATO). While these improved data will certainly help us to better understand planetary populations around other stars, a detailed characterization of individual planets is expected to remain somewhat limited. While the characterization of volatile-rich and dense exoplanets is expected to improve with decreasing observational uncertainties, a significant degeneracy in the internal structure of most of the super-Earth population is expected to remain. We also emphasize the importance of the theoretical uncertainty related to model assumptions (e.g., envelope structure, and composition of the planetary layers), which may overcome the observational uncertainties soon. We therefore suggest that, along with the great efforts to improve the data, similar efforts should be made on the theoretical front.

Acknowledgements. We thank the referee for valuable comments which significantly improved our paper. C.D. acknowledges support from the Swiss National Science Foundation under grant PZ00P2_174028. J.H. acknowledges the support from the Swiss National Science Foundation (SNSF) under grant 200020_172746. This work has been carried out within the frame of the National Center for Competence in Research PlanetS supported by the SNSF.

References

Badro, J., Fiquet, G., Guyot, F., et al. 2007, Earth and Planetary Science Letters, 254, 233
Bashi, D., Helled, R., Zucker, S., & Mordasini, C. 2017, A&A, 604, A83
Beichman, C. A. & Greene, T. P. 2018, arXiv e-prints, arXiv:1803.03730
Benz, W., Anic, A., Horner, J., & Whitby, J. A. 2007, Space Sci. Rev., 132, 189
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99
Bond, J. C., O'Brien, D. P., & Lauretta, D. S. 2010, ApJ, 715, 1050
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Carter, J.-M., Shankar, V. V., Zeb, M. A., & Kee, H.-Y. 2012, Phys. Rev. B, 85, 115105
Chau, A., Reinhardt, C., Helled, R., & Stadel, J. 2018, ApJ, 865, 35
Chen, J. & Kipping, D. 2017, ApJ, 834, 17
Connolly, J. A. D. 2009, Geochemistry, Geophysics, Geosystems, 10, Q10014
Dorn, C., Harrison, J. H. D., Bonsor, A., & Hands, T. O. 2019, MNRAS, 484, 712
Dorn, C., Khan, A., Heng, K., et al. 2015, A&A, 577, A83
Dorn, C., Venturini, J., Khan, A., et al. 2017, A&A, 597, A37
Drake, M. J. & Righter, K. 2002, Nature, 416, 39
Elser, S., Meyer, M. R., & Moore, B. 2012, Icarus, 221, 859
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gordon, S., McBride, B., & Zeleznik, F. J. 1984, Computer program for calculation of complex chemical equilibrium compositions and applications. Supplement 1: Transport properties
Grasset, O., Schneider, J., & Sotin, C. 2009, ApJ, 693, 722
Hakim, K., Rivoldini, A., Van Hoolst, T., et al. 2018, Icarus, 313, 61
Hansen, B. M. S. 2008, ApJS, 179, 484
Hatzes, A. P. 2016, in Astrophysics and Space Science Library, Vol. 428, Methods of Detecting Exoplanets: 1st Advanced School on Exoplanetary Science, ed. V. Bozza, L. Mancini, & A. Sozzetti, 3
Heng, K. & Lyons, J. R. 2016, ApJ, 817, 149
Javoy, M., Kaminski, E., Guyot, F., et al. 2010, Earth and Planetary Science Letters, 293, 259
Jin, S., Mordasini, C., Parmentier, V., et al. 2014, ApJ, 795, 65
Johnson, T. V., Mousis, O., Lunine, J. I., & Madhusudhan, N. 2012, in EGU General Assembly Conference Abstracts, Vol. 14, EGU General Assembly Conference Abstracts, ed. A. Abbasi & N. Giesen, 11532
Khan, A. & Connolly, J. A. D. 2008, in Lunar and Planetary Science Conference, Vol. 39, Lunar and Planetary Science Conference, 1245
LaViolette, P. 2006, Physics Essays, 18
Lodders, K. 2003, ApJ, 591, 1220
Lodders, K. & Fegley, B. 2002, Icarus, 155, 393
Lopez, E. D. & Fortney, J. J. 2014, ApJ, 792, 1
Lozovsky, M., Helled, R., Dorn, C., & Venturini, J. 2018, ApJ, 866, 49
Madhusudhan, N. 2012, ApJ, 758, 36
McDonough, W. F. & Sun, S.-s. 1995, Chemical Geology, 120, 223
Otegi, J. F., Bouchy, F., & Helled, R. 2019, arXiv e-prints, arXiv:1911.04745
Owen, J. E. & Wu, Y. 2013, ApJ, 775, 105
Pepe, F., Bouchy, F., Mayor, M., & Udry, S. 2018, High-Precision Spectrographs for Exoplanet Research: CORAVEL, ELODIE, CORALIE, SOPHIE and HARPS, ed. H. J. Deeg & J. A. Belmonte (Cham: Springer International Publishing), 1–28
Rogers, L. A. & Seager, S. 2010, ApJ, 716, 1208
Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, American Institute of Physics Conference Series, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405
Sotin, C., Jackson, J. M., & Seager, S. 2010, Terrestrial Planet Interiors, ed. S. Seager, 375–395
Thiabaud, A., Marboeuf, U., Alibert, Y., Leya, I., & Mezger, K. 2015, A&A, 580, A30
Tinetti, G., Drossart, P., Eccleston, P., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9904, Proc. SPIE, 99041X
Visscher, C. & Moses, J. I. 2011, ApJ, 738, 72
Wang, H. S., Lineweaver, C. H., & Ireland, T. R. 2018, Icarus, 299, 460
Weiss, L. M. & Marcy, G. W. 2014, ApJ, 783, L6
Weiss, L. M., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 768, 14
Wolfgang, A., Rogers, L. A., & Ford, E. B. 2015, IAU General Assembly, 22, 2258521
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127

3.2.2 Characterization of discovered exoplanets:

During my Ph.D. I have been co-author of several discovery papers in which I was in charge of the internal structure modeling. In this subsection I make a brief summary of these papers and my contribution on them.

3.2.2.1 TESS and HARPS reveal two sub-Neptunes around TOI 1062: *Otegi et al. 2021, A&A, A105, 653*

In Otegi et al. (2021a) we report the discovery of two sub-Neptunes orbiting around TOI-1062, using TESS photometry and HARPS data under the NCORES program. TOI-1062b is found to have a radius of $2.27\pm0.09 R_\oplus$ and a mass of $10.15\pm0.8 M_\oplus$, while the second planet is not transiting and has a minimum mass of $9.78\pm1.2 M_\oplus$. The outer planet is near 2:1 mean motion resonance with the innermost planets, so we perform a dynamical analysis to explore the proximity of the system to the resonance and to attempt to further constrain the orbital parameters. The inner planet is found to have a density of $4.85^{+0.84}_{-0.74}$ g cm^{-3}, slightly lower than the Earth. Internal structure models indicate that TOI 1062b is expected to be composed of significant iron core and silicate mantle, accounting for nearly 40% of the planetary mass each, and that it is expected to have a small volatile envelope of 0.35% of the mass at most.

The position of TOI 1062b with respect to the radius valley and its high equilibrium temperature of ~1000K make it an interesting candidate for atmospheric characterization. The strong stellar irradiation may result in a significant loss of the H-He envelope leaving a high-metallicity atmosphere that can be an excellent laboratory to study carbon chemistry at high temperature.

Astronomy & Astrophysics manuscript no. output
June 4, 2021

TESS and HARPS reveal two sub-Neptunes around TOI 1062

J. F. Otegi[1,2], F. Bouchy[1], R. Helled [2], D.J. Armstrong[19,20], M. Stalport[1], A. Psaridi[1], J.-B. Delisle[1], K.G. Stassun[13], E. Delgado-Mena[17], N.C. Santos[17,18], N.C. Hara[1], K. Collins[7], S. Gandhi [19,20], C. Dorn[2], M. Brogi[19,20], M. Fridlund[24,25], H.P. Osborn[23], S. Hoyer[26], S. Udry[1], S. Hojjatpanah[17,18], L.D. Nielsen [1], X. Dumusque[1], V. Adibekyan[17,18], D. Conti[8], R. Schwarz[9], G. Wang[10], P. Figueira [17,27], J. Lillo-Box[3], A. Hadjigeorghiou[19,20], D. Bayliss[19,20], P.A. Strøm[19,20], S.G. Sousa[17], D. Barrado[3], A. Osborn[19,20], S.C.C. Barros[17,18], D.J.A. Brown[19,20], J.D. Eastman [7], D.R. Ciardi[16], A. Vanderburg[15], R.F. Goeke[4], N.M. Guerrero [4], P.T. Boyd[14], D.A. Caldwell [12,21], C.E. Henze[12], B. McLean [22], G. Ricker[4], R. Vanderspek[4], D.W. Latham[7], S. Seager[4,5,6], J. Winn[11], and J.M. Jenkins [12]

[1] Observatoire Astronomique de l'Université de Genève, 51 Ch. des Maillettes, – Sauverny – 1290 Versoix, Switzerland
[2] Institute for Computational Science, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland
[3] Center for Astrobiology (CAB, INTA-CSIC), Dept. de Astrofísica, ESAC campus 28692 Villanueva de la Cañada (Madrid), Spain
[4] Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[5] Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[6] Department of Aeronautics and Astronautics, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[7] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[8] American Association of Variable Star Observers, 49 Bay State Road, Cambridge, MA 02138, USA
[9] Patashnick Voorheesville Observatory, Voorheesville, NY 12186, USA
[10] Tsinghua International School, Beijing 100084, China
[11] Department of Astrophysical Sciences, Princeton University, NJ 08544, USA
[12] NASA Ames Research Center, Moffett Field, CA 94035, USA
[13] Vanderbilt University, Department of Physics & Astronomy, Nashville, TN 37235, USA
[14] Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[15] Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA
[16] Caltech/IPAC-NExScI, M/S 100-22, 1200 E California Blvd, Pasadena, CA 91125, USA
[17] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal
[18] Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, Porto, Portugal
[19] Department of Physics, University of Warwick, Coventry CV4 7AL, UK
[20] Centre for Exoplanets and Habitability, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
[21] SETI Institute, Mountain View, CA 94043, USA
[22] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[23] NCCR/PlanetS, Centre for Space & Habitability, University of Bern, Bern, Switzerland AND Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[24] Leiden Observatory, University of Leiden, PO Box 9513, 2300 RA, Leiden, The Netherlands
[25] Department of Earth and Space Sciences, Chalmers University of Technology, Onsala Space Observatory, 439 92, Onsala, Sweden
[26] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[27] European Southern Observatory, Alonso de Córdova 3107, Vitacura, Santiago, Chile
[28] International Center for Advanced Studies (ICAS) and ICIFI (CONICET), ECyT-UNSAM, Campus Miguelete, 25 de Mayo y Francia, (1650) Buenos Aires, Argentina.

June 4, 2021

ABSTRACT

The Transiting Exoplanet Survey Satellite (*TESS*) mission was designed to perform an all-sky search of planets around bright and nearby stars. Here we report the discovery of two sub-Neptunes orbiting around TOI 1062 (TIC 299799658), a V=10.25 G9V star observed in the *TESS* Sectors 1, 13, 27, and 28. We use precise radial velocity observations from HARPS to confirm and characterize these two planets. TOI 1062b has a radius of $2.265^{+0.096}_{-0.091}$ $R_\oplus$, a mass of 10.15 ±0.8 $M_\oplus$, and an orbital period of 4.1130±0.0015 days. The second planet is not transiting, has a minimum mass of $9.78^{+1.26}_{-1.18}$ $M_\oplus$ and is near the 2:1 mean motion resonance with the innermost planet with an orbital period of $7.972^{+0.018}_{-0.024}$ days. We performed a dynamical analysis to explore the proximity of the system to this resonance, and to attempt further constraining the orbital parameters. The transiting planet has a mean density of $4.85^{+0.84}_{-0.74}$ g cm^{-3} and an analysis of its internal structure reveals that it is expected to have a small volatile envelope accounting for 0.35% of the mass at most. The star's brightness and the proximity of the inner planet to what is know as the radius gap make it an interesting candidate for transmission spectroscopy, which could further constrain the composition and internal structure of TOI 1062b.

1. Introduction

The Kepler mission was the first exoplanet mission to perform a large statistical survey of transiting exoplanets (Borucki et al. 2010; Howell et al. 2014); it impacted the field significantly with the detection of over 2300 exoplanets (see NASA Exoplanet Archive Akeson et al. 2013). Nevertheless, due to the faintness of the stars targeted by Kepler, only a small fraction are suitable for radial velocity (RV) follow-up. The Transiting Exoplanet Survey Satellite (*TESS*) has been designed to survey 85% of the sky for transiting exoplanets around bright, nearby stars (Ricker et al. 2015). It spent the first year of its mission searching for planets in the Ecliptic Southern Hemisphere. More than 2000 candidates have been detected and over 80 have been confirmed, expanding the number of small planets around cool stars (Guerrero 2020). RV follow-up programs have allowed us to quickly validate transiting planet candidates and to constrain their masses. Accurate measurements of the planetary mass and radius provide information on the planet's bulk density. However, a full characterization on the planetary composition and internal structure is extremely challenging due to the degenerate nature of this problem, where various compositions and interiors can lead to the same average density. Nevertheless, by studying the population of planets in this mass and size range, it is possible to better understand the dominating processes of planet formation and evolution in a statistical manner (Alibert et al. 2010, 2015).

A key signature in the planet population is the hot Neptune desert, which corresponds to a lack of Neptune-mass planets close to their host stars (Szabó & Kiss 2011). The desert is likely to arise from a combination of photoevaporation and tidal disruption (Beaugé & Nesvorný 2013; Mazeh et al. 2016). At small orbital distances planets are subject to intense radiation from the their host stars, and Neptune-like planets are probably not massive enough to retain their gaseous envelopes. An important advantage of the all-sky nature of *TESS* resides in its ability to find the brightest examples of stars hosting rare sub-populations of planets, such as the exoplanets lying inside the desert. *TESS* has been able to find the first planets deep inside the desert, for example TOI-849b (Armstrong et al. 2020), a recently discovered remnant core of a giant planet, and LTT9779b (Jenkins et al. 2020), an ultra-hot Neptune. The NCORES HARPS large program is designed to further study planets that may have undergone substantial envelope loss. Some of the recently discovered planets under this program include the mentioned TOI-849b and three mini Neptunes TOI-125b, c, d (Nielsen et al. 2020).

Another relevant feature in the exoplanet population is a lack of planets with radii between $1.5R_\oplus$ and $2R_\oplus$ (Fulton et al. 2017) known as the radius valley. This suggests that there is a transition between the super-Earth and sub-Neptune populations (Owen & Jackson 2012). This gap could be a result of stellar irradiation (Lopez & Fortney 2013) or core-powered mass-loss (e.g., Ginzburg et al. 2018), which leads to atmospheric evaporation that strips the planet down to its core or, alternatively, the consequence of a gas-poor formation (e.g., Gupta & Schlichting 2019). The observed features of the exoplanet populations are not well understood, and more exoplanet discoveries are crucial to reveal the overall picture of planets in the super-Earth and Neptune size and mass regimes.

In this paper we present the discovery and confirmation of two highly irradiated sub-Neptunes hosted by the *TESS* Object of Interest (TOI) 1062, a nearby (d= 82 pc) and bright (V_{mag} ~10.2) G9V star (see Table 1 for a full summary of the stellar

Table 1: Stellar parameters of TOI-1062.

Parameter	Value	Source
Identifying Information		
TOI	TOI-1062	*TESS*
TIC ID	299799658	*TESS*
2MASS ID	J02322848-7801256	2MASS
Gaia ID	4632865331094140928	Gaia DR3
Astrometric Parameters		
R.A. (J2015.5, h:m:s)	02:32:29	Gaia DR3
Dec (J2015.5, h:m:s)	-78:01:25	Gaia DR3
Parallax (mas)	12.14 ±0.01	Gaia DR3
Distance (pc)	82.2±0.2	Gaia DR3
Photometric Parameters		
B	11.02 ± 0.07	Tycho
V	10.25 ± 0.01	Tycho
T	9.48 ± 0.01	*TESS*
G	10.00 ± 0.01	Gaia
J	8.78 ± 0.02	2MASS
H	8.41 ± 0.03	2MASS
K	8.30± 0.03	2MASS
W1	8.257 ± 0.02	WISE
W2	8.322 ± 0.02	WISE
W3	8.256 ± 0.02	WISE
W4	8.35 ± 0.19	WISE
A_V	0.13 ± 0.04	Sec. 3
Abundances		
[Fe/H] (dex)	0.14 ± 0.04	Sec. 3
[O/H] (dex)	-0.14 ± 0.16	Sec. 3
[C/H] (dex)	0.13 ± 0.02	Sec. 3
[Cu/H] (dex)	0.25 ± 0.04	Sec. 3
[Zn/H] (dex)	0.08 ± 0.04	Sec. 3
[Sr/H] (dex)	0.17 ± 0.08	Sec. 3
[Y/H] (dex)	0.04 ± 0.08	Sec. 3
[Zr/H] (dex)	0.20 ± 0.06	Sec. 3
[Ba/H] (dex)	-0.02 ± 0.04	Sec. 3
[Ce/H] (dex)	0.16 ± 0.10	Sec. 3
[Nd/H] (dex)	0.10 ± 0.06	Sec. 3
[MgI/H] (dex)	0.16 ± 0.07	Sec. 3
[AlI/H] (dex)	0.25 ± 0.04	Sec. 3
[SiI/H] (dex)	0.20 ± 0.07	Sec. 3
[CaI/H] (dex)	0.10 ± 0.08	Sec. 3
[TiI/H] (dex)	0.23 ± 0.07	Sec. 3
[CrI/H] (dex)	0.17 ± 0.07	Sec. 3
[NiI/H] (dex)	0.17 ± 0.05	Sec. 3
Bulk Parameters		
Mass ($M_\odot$)	0.94 ± 0.02	Sec. 3
Radius ($R_\odot$)	0.84 ± 0.09	Sec. 3
$T_{\rm eff}$ (K)	5328 ± 56	Sec. 3
log g (cm s^{-2})	4.55 ± 0.09	Sec. 3
Spectral type	G9V	Sec. 3
ρ (g cm^{-3})	2.22 ± 0.10	Sec. 3
$v \sin i$ (km s^{-1})	2.13 ± 0.5	Sec. 3
P_{rot} (d)	21.8 ± 2.4	Sec. 3
Age (Gyrs)	2.5 ± 0.3	Sec. 3

2MASS (Skrutskie et al. 2006);Tycho (Høg et al. 2000); WISE (Wright et al. 2010); and Gaia (Gaia Collaboration et al. 2018)

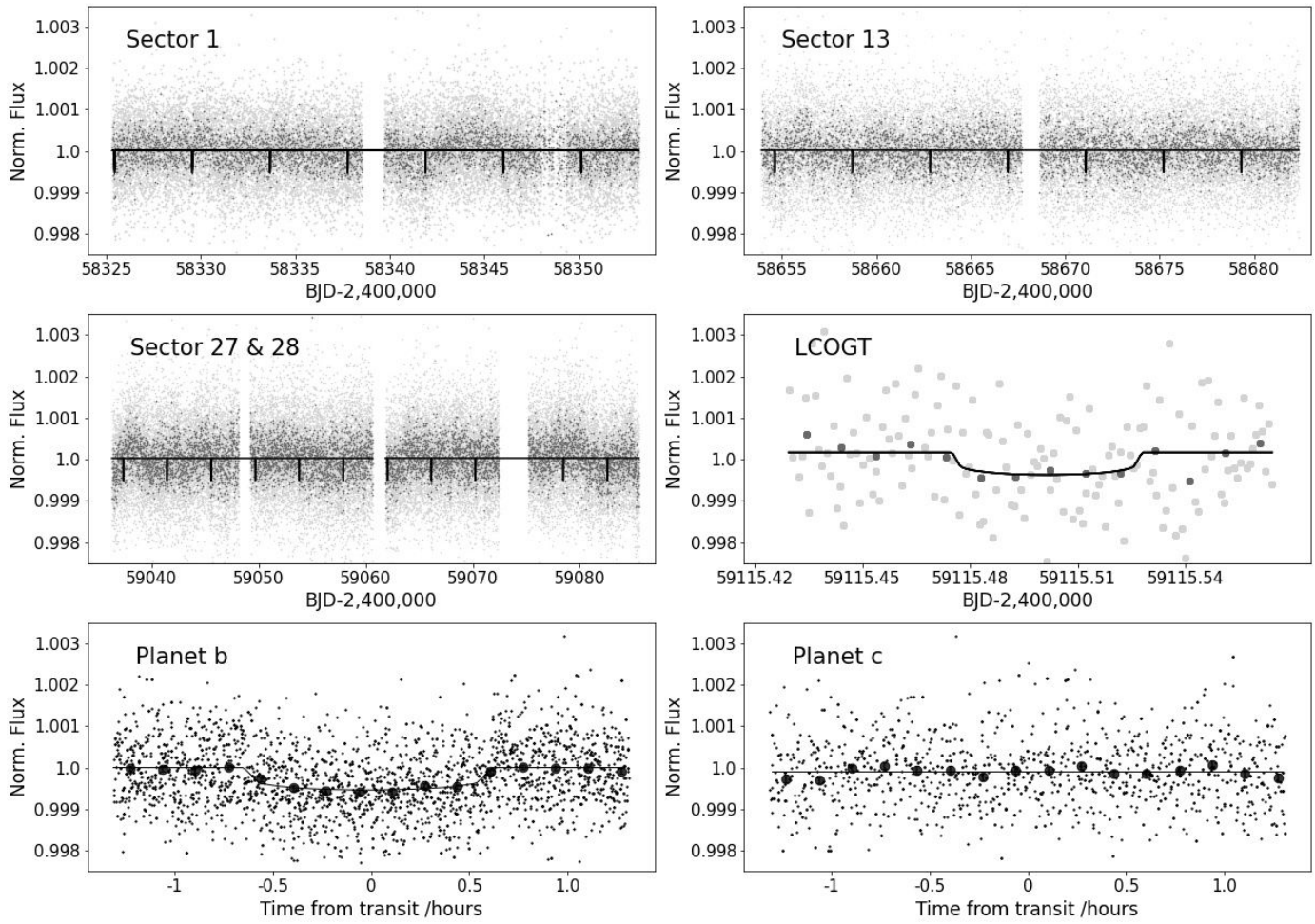

Fig. 1: *TESS* photometry from the of TOI 1062 from Sectors 1 (upper left), 13 (upper right), and Sectors 27 and 28 (middle left), and LCOGT photometry (middle right). The *TESS* light curves correspond to the PDC-SAP flux time series provided by SPOC. The *TESS* full light curve with the 2 min cadence data is shown in light gray, and the same data binned to 10 min is in dark gray. The gaps in the data coverage are due to observation interruptions of the *TESS* spacecraft that occur after each *TESS* orbit of 13.7 days. The dots of the LCOGT are bigger for better visualization. The lower panels shows the phase-folded *TESS* light curves for TOI 1062b (left) and TOI 1062c (right) with the 2 min cadence data in light gray and binned to 10 min in black. We find that TOI 1062c does not transit.

properties). We also used intensive radial velocity follow-up observations with HARPS to confirm the planetary nature of the transit detected in the *TESS* data and precisely determine the properties of the planetary system. The paper is organized as follows: in Section 2 we present the data collected on the system to discover and validate the planets, in Section 3 we analyze the host star fundamental parameters, and in Section 4 we characterize and discuss the planets. Our conclusions are presented in Section 5.

2. Observations

2.1. TESS photometry

TESS observed TOI 1062 (TIC 299799658) in Sectors 1 and 13 during the first year, and during Sectors 27 and 28 of the extended mission, obtaining data from 25 July to 22 August 2018, from 19 June to 17 July 2018, and from 4 July to 26 August 2020. The two-minute cadence data were reduced with the Science Processing Operations Center (SPOC) pipeline (Jenkins et al. 2016) adapted from the pipeline for the Kepler mission at the NASA Ames Research Center in order to produce calibrated pixels and light curves. On 29 September 2018 a transit candidate was announced around TOI 1062. TOI 1062.01 is a planet candidate with a period of 4.11506±0.00002 days, a transit depth of 487.90 ± 45 ppm, and an estimated planet radius of 2.32 ± 0.46 $R_\oplus$. The candidate passed all the tests from the Threshold Crossing Event (TCE) Data Validation Report (DVR; Twicken et al. 2018; Li et al. 2019).

We used the publicly available Presearch Data Conditioning (PDC-SAP) flux time series (e.g., Twicken et al. 2010; Smith et al. 2012) provided by SPOC for the transit modeling, which is corrected for common trends and artifacts, for crowding, and for the finite flux fraction in the photometric aperture. Figure 1 shows the 2 min cadence *TESS* light curve, along with the phase-folded curve for TOI 1062b.

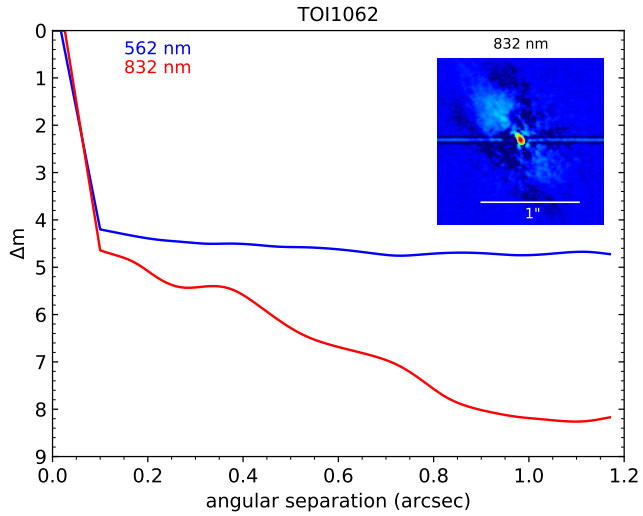

Fig. 2: Gemini/Zorro contrast curves and 1.2" x 1.2" reconstructed image of the 832 nm band.

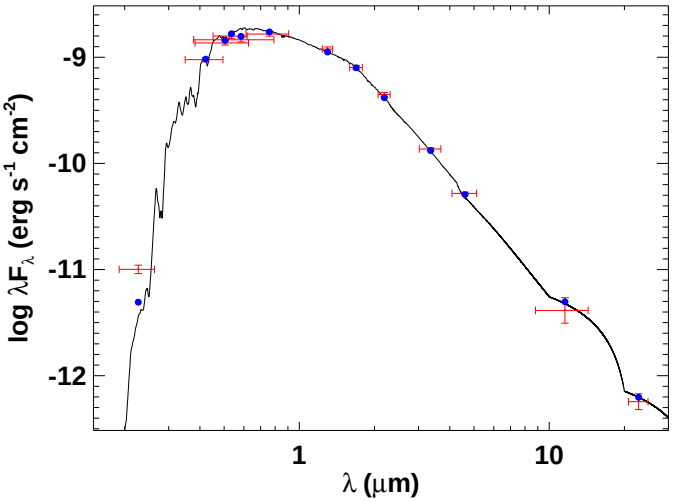

Fig. 3: Spectral energy distribution of TOI 1062. Red symbols represent the observed photometric measurements; the horizontal bars represent the effective width of the passband. Blue symbols are the model fluxes from the best-fit Kurucz atmosphere model (black).

2.2. Ground-based photometry with LCOGT

We obtained ground-based time series follow-up photometry of full transits of TOI 1062.01 on 24 August 2019 in i' band and on 22 September 2020 in Pan-STARSS z-shortband from the Las Cumbres Observatory Global Telescope (LCOGT; Brown et al. 2013) 1.0 m network node at the South Africa Astronomical Observatory. We used the *TESS* Transit Finder, which is a customized version of the Tapir software package (Jensen 2013), to schedule the observations. The 4096x4096 LCO SINISTRO cameras have an image scale of 0.389" per pixel, providing a field of view of 26'x26'. The standard LCOGT BANZAI pipeline was used to calibrate the images, and the photometric data were extracted with the AstroImageJ (AIJ) software package (Collins et al. 2017).

The initial i'-band observation intentionally saturated the target star to check for a faint nearby eclipsing binary (NEB) that could be contaminating the *TESS* photometric aperture. To account for possible contamination from the wings of neighboring star PSFs, we searched for NEBs out to 2.5' from the target star. If fully blended in the SPOC aperture, a neighboring star that is fainter than the target star by 8.8 magnitudes in *TESS* band could produce the SPOC-reported flux deficit at mid-transit (assuming a 100% eclipse). To account for possible delta-magnitude differences between *TESS* band and i' band, we included an extra 0.5 magnitudes fainter (down to *TESS*-band magnitude 18.3). Our search ruled out NEBs in all 24 neighboring stars that meet our search criteria. The z-shortband observation was defocused and exposed appropriately to measure the TOI 1062.01 light curve at high precision. For the data reduction process we used a 10.5" radius aperture and selected four comparison stars with brightness comparable to that of the target star. The inner and outer radii of the background annulus were then set to the corresponding AIJ calculated values of 18.7" and 28", respectively. This combination was chosen as it minimized the amount of noise in the light curve. A 0.5 ppt transit, with model residuals of 0.3 ppt in 10 min bins, was detected in the uncontaminated target aperture.

2.3. HARPS follow-up

We collected 90 high-resolution spectra of TOI 1062 using the High Accuracy Radial Velocity Searcher (HARPS) spectrograph mounted at the ESO 3.6m telescope of La Silla Observatory, Chile (Mayor et al. 2003a), with the goal of precisely determining the mass of the planet candidate and searching for additional planets in the system. The observations were carried out as part of the NCORES large program (ID 1102.C-0249, PI: Armstrong) between 31 August 2019 and 17 January 2020. HARPS is a stabilized high-resolution (R 115000) echelle spectrograph which can reach sub-$m\,s^{-1}$ RV precision (Mayor et al. 2003b). The instrument was used in high-accuracy mode (HAM), with a 1" fiber on the star and another one to monitor the sky background. We used exposure times of 30 minutes.

The standard HARPS Data Reduction Software (DRS) was used to reduce the data, using a K0 mask for both the cross-correlation function (CCF) (Pepe et al. 2002; Baranne et al. 1996) and the color correction, and reaching a typical signal-to-noise ratio per pixel of 60 and a photon-noise uncertainty of 1.4 m s^{-1} . For each spectrum the usual activity indicators (S-index, Hα-index, Na-index, Ca-index, $\log R'_{\rm HK}$), the full width half maximum (FWHM), the line bisector, and the contrast of the CCF were measured.

2.4. High-resolution imaging

To check for the presence of contaminating stars in the *TESS* photometric aperture, TOI 1062 was observed on 14 January 2020 using the Zorro speckle imager (Scott 2019), mounted on the 8.1m Gemini South telescope in Cerro Pachon, Chile. Zorro uses high-speed electron-multiplying CCDs (EMCCDs) to simultaneously acquire data in two bands centered at 562 nm and 832 nm. The data is reduced following the procedures described in Howell et al. (2011), and provides output data including a reconstructed image and robust limits on companion detections. The contrast achieved in the resulting reconstructed image is shown in Figure 2. We see

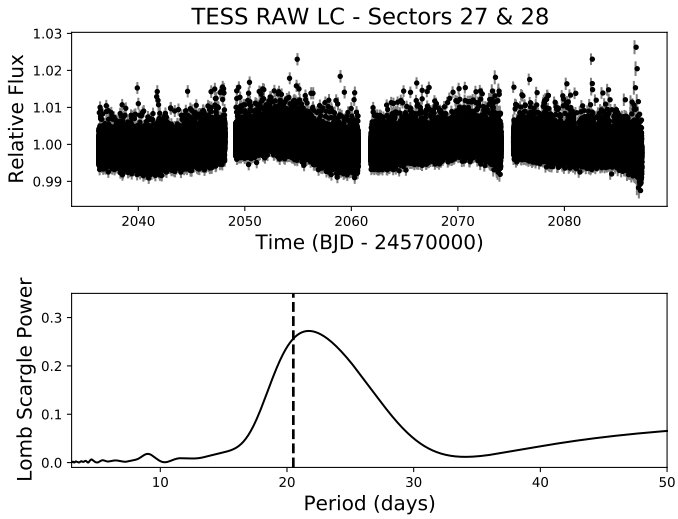

Fig. 4: *TESS* RAW photometric data of sectors 27 and 28 (top) and the corresponding periodogram. The peak of the periodogram shows a possible rotational modulation with a period of 21.8 days, which is in agreement with the rotational period obtained from HARPS calibration.

that the Zorro speckle images reach a contrast of Δmag=8.25 at a separation of 1.1" in the 832 nm band, and cover a spatial range of 9.9 to 98.6 au around the star with contrasts between 5 and 8 mag in both bands, showing no close companions to TOI 1062b.

3. Host star fundamental parameters

3.1. Analysis of the spectrum

The stellar atmospheric parameters ($T_{\rm eff}$, $\log g$, and [Fe/H]) and respective error bars were derived using the methodology described in Sousa (2014); Santos et al. (2013). Briefly, we make use of the equivalent widths (EWs) of iron lines, as measured in the combined HARPS spectrum of TOI 1062 using the ARES v2 code[1] (Sousa et al. 2015), and we assume ionization and excitation equilibrium. This method makes use of a grid of Kurucz model atmospheres (Kurucz 1993) and the radiative transfer code MOOG (Sneden 1973). This approach leads to the following stellar parameters: $T_{\rm eff}$= 5328 ± 56K, [Fe/H]=0.14 ± 0.04 dex, and $\log g$= 4.55 ± 0.09. This is in agreement with the $\log g$ derived making use of the *Gaia* parallax and luminosity (Santos et al. 2004) which gives a value of 4.53. Using the Torres et al. (2010) calibration leads to a stellar mass of 0.94±0.02 $M_\odot$ and radius of 0.84±0.09 $R_\odot$. These values are used as priors in the global fit with `juliet` in Section 4.

3.2. Analysis of the spectral energy distribution

As an independent determination of the stellar parameters, we also performed an analysis of the broadband spectral energy distribution (SED) of the star together with the *Gaia* DR2 parallax (adjusted by +0.08 mas to account for the systematic offset reported by Stassun & Torres 2018), in order to determine an empirical measurement of the stellar radius, following the procedures described in Stassun & Torres (2016), Stassun et al. (2017), and Stassun et al. (2018). We pulled the $B_T V_T$ magnitudes from *Tycho-2*; the *BVi* magnitudes from *APASS*; the JHK_S magnitudes from *2MASS*; the W1–W4 magnitudes from *WISE*; the G, $G_{\rm BP}$, $G_{\rm RP}$ magnitudes from *Gaia*; and the near-ultraviolet (NUV) magnitude from *GALEX*. Together, the available photometry spans the full stellar SED over the wavelength range 0.2–22 μm (see Figure 3).

We performed a fit using Kurucz stellar atmosphere models, using the effective temperature ($T_{\rm eff}$), metallicity ([Fe/H]), and surface gravity ($\log g$) adopted from the spectroscopic analysis. The only additional free parameter is the extinction (A_V), which we fixed to be zero due to the star's proximity. The resulting fit is very good (Figure 3) with a reduced χ^2 of 1.8. The reduced χ^2 is improved to 1.2 by excluding the GALEX NUV flux, which exhibits a modest excess; using the empirical relations of Findeisen et al. (2011), the GALEX NUV excess implies an activity level $\log R'_{\rm HK} = -4.9 \pm 0.1$, consistent with the spectroscopically determined value.

Integrating the (unreddened) model SED gives the bolometric flux at Earth, $F_{\rm bol} = 2.429 \pm 0.057 \times 10^{-9}$ erg s^{-1} cm^{-2}. Taking the $F_{\rm bol}$ and $T_{\rm eff}$ together with the *Gaia* DR2 parallax gives the stellar radius $R_\star = 0.838 \pm 0.020\,R_\odot$. In addition, we can use the $R_\star$ together with the spectroscopic $\log g$ to obtain an empirical mass estimate of $M_\star = 0.91 \pm 0.19 M_\odot$, which is consistent with that obtained via empirical relations of Torres et al. (2010) and a 6% error from the empirical relation itself, $M_\star = 0.97 \pm 0.06 M_\odot$. These values are in agreement with those derived from the spectral analysis in Section 3.1.

3.3. Rotational period and age

The stellar age can be derived from empirical rotation-age relations of Mamajek & Hillenbrand (2008). The HARPS calibration from the CCF-FWHM and the color index B-V gives $v \sin i$= 2.13 ± 0.5 km/s; then together with $R_\star$, we estimate the stellar rotation period $P_{\rm rot}/\sin i$ = 20.5 ± 2.9 d. In addition, Fig. 4 shows that the peak of the periodogram of the *TESS* light curves of Sectors 27 and 28 is at 21.82.4 days. The measured rotational modulation is nearly identical in the other *TESS* sectors. Another estimate of the expected rotation can be inferred from the average value of the Ca II H & K chromospheric activity indicator for TOI 1062, which is $\log R'_{\rm HK} = -4.80 \pm 0.05$. Using the empirical relations in Mamajek & Hillenbrand (2008) we get a rotation period of 21.9^{+3}_{-2} days, in agreement with the result derived from the HARPS calibration and the modulation of the RAW light curves. Another way of constraining the rotational period of the star consists of fitting the RVs with time-dependent Gaussian processes (GPs). We use the `celerite` package (e.g., Foreman-Mackey et al. 2017) with a quasi-periodic kernel, and obtain a rotational period of $22.4^{+4.6}_{-6.7}$ days, which is in agreement with the previous estimations.

Using the rotational period obtained from the *TESS* data and the relations of Mamajek & Hillenbrand (2008), the age estimate is $\tau_\star$=2.5±0.3 Gyr. Because the $v \sin i$ is a projected rotational velocity, we note that it is a lower limit on the true rotational velocity, so $P_{\rm rot}/\sin i$ is an upper limit on the true $P_{\rm rot}$, which implies that the age estimate above provides an upper limit to the stellar age.

[1] The most recent version of the ARES code (ARES v2) can be downloaded at http://www.astro.up.pt/~sousasag/ares

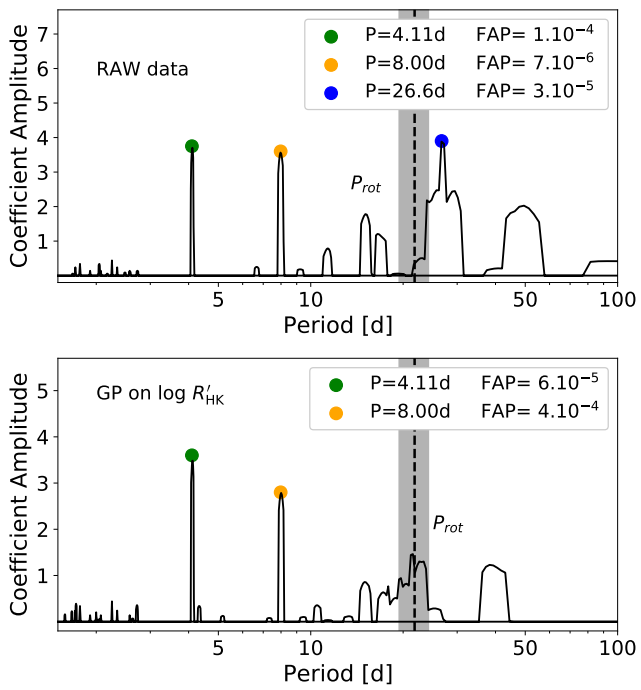

Fig. 5: Undetrended RVs (top) and detrended RVs (bottom) obtained via the l_1 periodogram using Gaussian processes on the $\log R'_{\rm HK}$ indicator with `SPLEAF`. The periods at which the significant peaks occur (FAP<10^{-3}) are color-coded (see inset). The stellar rotational period determined through the *TESS* data is highlighted in gray.

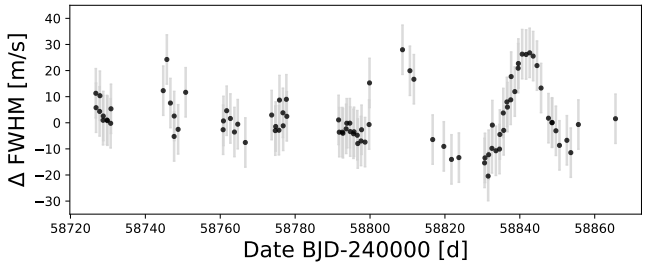

Fig. 6: Time series of the CCF-FWHM.

We also attempted to derive the stellar age by making use of chemical clocks (e.g., Delgado Mena et al. 2019); however, we obtain a value much higher than with other methods described below. The reason is that the low $T_{\rm eff}$ of TOI 1062 lies at the limit of applicability of such empirical relations, and thus we did not use the age provided by this method.

3.4. Stellar abundances

Stellar abundances of refractory elements were derived using the classical curve-of-growth analysis method assuming local thermodynamic equilibrium (e.g., Adibekyan et al. 2012; Delgado Mena et al. 2017). Abundances of the volatile elements C and O were derived following the method of Delgado Mena et al. (2010); Bertran de Lis et al. (2015). Since the two spectral lines of oxygen are usually weak and the 6300.3Å line is blended with Ni and CN lines, the EWs of these lines were manually measured with the task `splot` in IRAF. All the [X/H] ratios are obtained by doing a differential analysis with respect to a high S/N solar (Vesta) spectrum from HARPS. The obtained values for this star are normal considering its metallicity except for the lower-than-expected value of oxygen. We find a high dispersion between both oxygen indicators that has to be taken with caution due to the difficulty of measuring reliable oxygen abundances for cool metal-rich stars. The stellar abundances of the elements are presented in Table 1.

Table 2: Prior parameter distribution of the global fit with `juliet`.

Parameter		Prior distribution
Instrumental parameters:		
$q_{1,TESS}$		$\mathcal{U}(0,1)$
$q_{2,TESS}$		$\mathcal{U}(0,1)$
$m_{flux,TESS}$		$\mathcal{N}(0,0.1)$
σ_{TESS}	[ppm]	$\mathcal{J}(0.1,1000)$
σ_{HARPS}	[m/s]	$\mathcal{J}(0.1,1000)$
GP parameters:		
GP_σ	[m/s]	$\mathcal{J}(10^{-3},10^5)$
GP_ρ	[d]	$\mathcal{J}(19.4,10^5)$
Planetary parameters:		
P_b	[d]	$\mathcal{N}(4.11,0.1)$
P_c	[d]	$\mathcal{N}(8.00,0.1)$
$T_{0,b}$	$[BJD_{TBD}]$	$\mathcal{N}(2459082.59485,0.1)$
$T^*_{0,c}$	$[BJD_{TBD}]$	$\mathcal{U}(2459082,2459091)$
K_b	[m/s]	$\mathcal{U}(1,100)$
K_c	[m/s]	$\mathcal{U}(1,100)$
$\sqrt{e}\sin(\omega)$		$\mathcal{U}(-1,1)$
$\sqrt{e}\cos(\omega)$		$\mathcal{U}(-1,1)$
r_1		$\mathcal{U}(0,1)$
r_2		$\mathcal{U}(0,1)$

* $\mathcal{U}(a,b)$ indicates a uniform distribution between a and b; $\mathcal{J}(a,b)$ a Jeffrey or log-uniform distribution between a and b; and $\mathcal{N}(a,b)$ a normal distribution with mean a and standard deviation b.

3.5. Signal identification

In order to search for periodicities in the RV data we computed the l_1 periodogram (e.g., Hara et al. 2017, 2020). This tool is designed to find periodicities in time series data corresponding to exoplanets detected using radial velocity data. It can be used similarly to a typical Lomb-Scargle periodogram or its variants (e.g., Baluev 2008), but it is based on a different principle. The l_1 periodogram can analyze the radial velocity without the need to estimate the frequency iteratively, using the theory of compressed sensing adapted for handling correlated noise. Instead of fitting a sinusoidal function at each frequency as the Lomb–Scargle-type periodograms, the l_1 periodogram searches directly for a representation of the input signal as a sum of sinusoids. As a result, all the frequencies of the signal are searched simultaneously. This relies on the so-called sparse recovery tools, which are designed to find a representation of an input signal as a linear combination of sine functions.

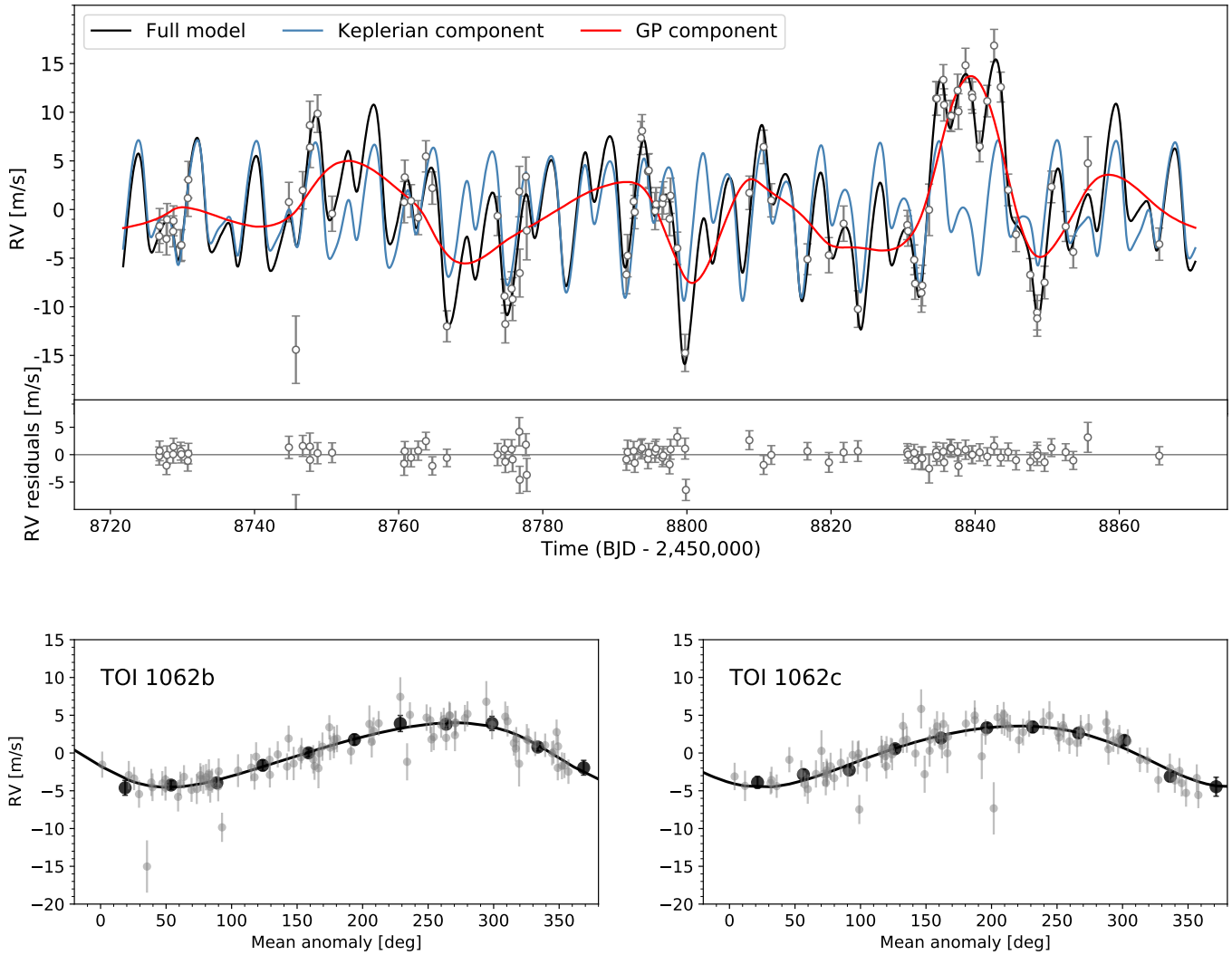

Fig. 7: HARPS RVs for TOI 1062 with a two planet model including eccentric orbits and Gaussian processes with Matern kernel.Top panel: RVs along with the residuals to the best fit right below the time series. Bottom panel: Phase-folded data for each planet. The gray dots represent the observed RVs and the black dots the binned data.

The l_1 periodogram outputs a figure that has a similar aspect to that of a standard Lomb–Scargle-type periodogram, but with fewer peaks due to aliasing. We use the l_1 periodogram of the HARPS data on a grid of frequencies between 0.01 and 0.3 cycles per day. In order to measure the significance of a possible signal the false alarm probability (FAP) is calculated, which encodes the probability of measuring a peak of a given height conditioned on the assumption that the data consists of Gaussian noise with no periodic component. Typically, a signal at the 1% FAP level is considered suggestive, and a 0.1 % FAP level statistically significant. Figure 5 shows two l_1 periodograms of the HARPS RVs. The top panel corresponds to the undetrended RVs. In this case we detect three significant signals at 4.11d (FAP 1.10^{-4}%), 7.98 days (FAP 7.10^{-6}%), and 26.6 days (FAP 3.10^{-5}%). Although the coefficient amplitude of the second and third peaks are relatively close, their FAP values can differ significantly since it is estimated in a sequential manner: the highest peak is computed against zero, the second peak against the first, and the third peak against the second. Figure 6 clearly shows the effects of activity in the FWHM at the end of the time series. It suggests that the star was more active at that moment, probably due to one spot crossing event during one single rotational period, standing for only 20 days (from BJD 58830 to 58850). However, we find that the signal at 26.6 days disappears when we detrend the RVs using GP modeling with *SPLEAF* (e.g., Delisle et al. 2020) (see bottom panel of Figure 5). The GP is trained on the $logR'_{\rm HK}$ indicator, and then it is used to detrend the RVs using a linear fit. This method gives two significant peaks at at 4.11d with FAP 6.10^{-5}% and 8.00 days with FAP 4.10^{-4}%.

Another way to detrend the stellar activity is to fit the correlations between activity indicators as the full width at half maximum (FWHM), the bisector or the S-index, and subtract them in order to correct for any long-term effects resulting from a magnetic cycle. We do so dividing the data into different time regimes selected in order to maximize the correlation of the RV data with the activity indicators in each bin. This approach also identifies two peaks at 4.11d with FAP 7.10^{-8}% and 8.01 days with FAP 3.10^{-4}%. Finally, we also find the same two significant signals when we do not use the data TBJD 58730 to TBJD 58750, which corresponds to periods when the star was more active. The signal at 26.6 days only appears when using the full undetrended RV sample: it is clearly affected by stellar

activity and it is within 1.6σ of the rotation period derived from the chromospheric activity indicator. We therefore conclude that only the peaks at ~4.11 days and ~8 days are caused by planets. In addition, even though the ~8-day signal is nearly double the ~4.11-day signal, in all the cases presented we find the 8-day signal after removing the ~4.11-day signal. We also evaluate the evidence of the second planet by comparing the difference in the Bayesian information criterion (ΔBIC) between the one-eccentric-planet and the two-eccentric-planets models. We obtain a value of ΔBIC=11, which suggests strong evidence.

3.6. Global fit with `juliet`

We determine the planetary parameters using the publicly available software `juliet` (e.g., Espinoza et al. 2019). `juliet` allows us to jointly fit the *TESS* and LCOGT photometry and the HARPS radial velocities with GP. The software is built on several publicly available tools used to model the photometric data (the `batman` package; Kreidberg 2015), the radial velocities (the `radvel` package; Fulton et al. 2018), and also to incorporate GP (the `celerite` package, Foreman-Mackey et al. 2017). The parameter space is explored using nested sampling, with the `MultiNest` algorithm (e.g., Feroz & Hobson 2008) in its Python implementation, `PyMultiNest` (e.g., Buchner et al. 2014). In addition, `juliet` computes the Bayesian evidence using `dynesty` (e.g., Speagle 2020), a Python package that estimates Bayesian posteriors and evidence using dynamic nested sampling. In short, nested sampling algorithms work as follows. The algorithm samples a number of live points randomly from the prior distribution, and the likelihood is evaluated at each of these points. At each iteration the point with the lowest likelihood is replaced by a new sampled point, keeping the number of live points constant. This process is continued until Bayesian evidence reaches a specified value. The number of live points used has to be large enough to adequately sample the parameter space.

The transit model fits the stellar density $\rho_\star$ together with the planetary and jitter parameters. For the stellar density we use the value obtained in Section 3, and the priors of the orbital parameters of the inner planet are taken from ExoFOP. We use the quadratic limb darkening coefficients (q_1, q_2) introduced by Kipping (2013) for the photometric data, since it was shown to be appropriate for space-based missions (Espinoza & Jordán 2015). In addition, instead of fitting the planet-to-star radius ratio and the impact parameter of the orbit, we use the parameterization introduced in Espinoza (2018) and fit the parameters r_1 and r_2 to make sure the full exploration of physically plausible values in the (p,b) space. We parameterize the eccentricity and the argument of periastron with $\sqrt{e}\sin\omega$ and $\sqrt{e}\cos\omega$, always ensuring that $e \leq 1$. We use the `celerite` aproximate Matern multiplied by exponential kernel to account for the stellar activity. The fit obtained with this kernel is nearly identical to the one obtained with other kernels as the quasi-periodic or the exp-sine-squared kernel.

We fit both planets with `juliet` and find no hint of transit for TOI 1062c, as shown in Figure 1. The final median planetary parameters determined by the `juliet` fit are listed in Table 2. We find that TOI 1062b has an orbital period of 4.114 days, a radius of $2.265^{+0.096}_{-0.091}$ $R_\oplus$, a mass of $10.15^{+0.81}_{-0.84}$ $M_\oplus$, and a density of $4.85^{+0.84}_{-0.74}$ g cm^{-3}, slightly below the Earth's mean density. Nevertheless, as discussed in detail in Section 5, its composition and internal structure are unlikely to be similar to that of the Earth. We find an eccentricity of $0.18^{+0.07}_{-0.06}$ for TOI-1062b and 0.14±0.07 for TOI-1062c. Figure 8 shows the eccentricity against the orbital period for the population of observed exoplanets from the NASA Exoplanet Archive, and we see that the TOI-1062b value is slightly higher than usual for its orbital period. However, it may not be significant due to the bias towards higher eccentricities for nearly circular planets (Lucy & Sweeney 1971). TOI 1062 c instead is found to have an orbital period of 7.988 ± 0.04 days and a minimum mass of $9.78^{+1.26}_{-1.18}$ $M_\oplus$. TOI 1062 c is in nearly 2:1 motion resonance with its inner companion.

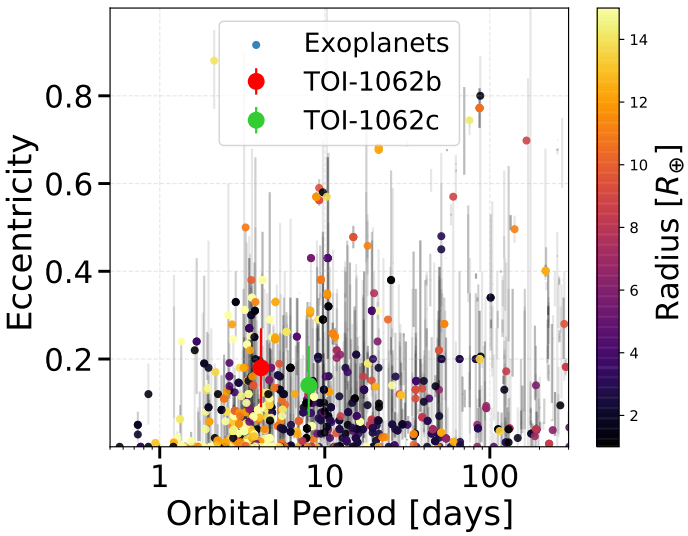

Fig. 8: Orbital eccentricity against orbital period for observed exoplanets from the NASA Exoplanet Archive. The color represents the planetary radius (see color scale at right).

4. Discussion

4.1. Internal structure

In order to characterize the internal structure of TOI 1062 b, we modeled its interior considering a pure-iron core, a silicate mantle, a pure-water layer, and a H-He atmosphere. The equations of state (EOSs) used for the iron core are from Hakim et al. (2018), the EOS of the silicate-mantle was calculated with PERPLE_X from Connolly (2009) using the thermodynamic data of Stixrude & Lithgow-Bertelloni (2011) and assuming Earth-like abundances, and the EOS for the H-He envelope are from Chabrier et al. (2019) assuming a proto-solar composition. For the pure-water layer we used the AQUA EOS from Haldemann et al. (2020). We assumed an envelope luminosity of L=$10^{22.52}$ erg s^{-1} (equal to Neptune's luminosity). The thickness values of the planetary layers were set by defining their masses and solving the structure equations. To obtain the transit radius, we followed Guillot (2010) and evaluated the location where the chord optical depth τ_{ch} is 2/3. We did not use stellar abundances as an additional constraint since it is not clear whether they are a good proxy for the planetary bulk abundances (Wang et al. 2019; Plotnykov & Valencia 2020), and it has been shown that the stellar abundances are not always useful for constraining the internal composition (Otegi et al. 2020b).

Table 3: TOI-1062 parameters from `juliet`: Median and 68% confidence interval.

Instrumental Parameters:			
$q_{1,TESS}$ ...	Quadratic limb-darkening parametrization	$0.845^{+0.035}_{-0.033}$	
$q_{2,TESS}$ ...	Quadratic limb-darkening parametrization	$0.113^{+0.114}_{-0.067}$	
$m_{flux,TESS}$	Offset relative flux	$-0.0000217^{+0.0000016}_{-0.0000017}$	
σ_{TESS}	Jitter (ppm)	$118.42^{+5.27}_{-5.03}$	
σ_{HARPS} ...	Jitter (m/s)	$1.141^{+0.157}_{-0.101}$	
Parameters of the GP with quasi-periodic kernel:			
GP_σ	Amplitude (m/s)	$21.19^{+3.1}_{-2.6}$	
GP_ρ	Time-scale (days)	$21.80^{+1.91}_{-1.33}$	
Planetary Parameters:		b	c
P	Period (days)	$4.11296^{+0.0015}_{-0.0015}$	$7.972^{+0.018}_{-0.024}$
T_0	Time of transit center* (BJD_{TBD})	$2459082.59^{+0.09}_{-0.09}$	$2459087.61^{+1.84}_{-1.63}$
K	Radial velocity semi-amplitude (m/s) ...	$4.27^{+0.31}_{-0.33}$	$4.05^{+0.46}_{-0.41}$
$\sqrt{e}sin(\omega)$.	Parametrization for e and ω	$0.14^{+0.08}_{-0.08}$	$-0.02^{+0.06}_{-0.06}$
$\sqrt{e}cos(\omega)$	Parametrization for e and ω	$-0.07^{+0.05}_{-0.06}$	$-0.09^{+0.12}_{-0.81}$
r_1	Parametrization for p and b	$0.908^{+0.006}_{-0.007}$	-
r_2	Parametrization for p and b	$0.0235^{+0.0003}_{-0.0004}$	-
Derived transit and RV parameters:		b	c
e	Eccentricity or the orbit	$0.177^{+0.067}_{-0.064}$	$0.140^{+0.073}_{-0.069}$
ω	Argument of periastron (deg)	117^{+27}_{-19}	159^{+16}_{-35}
i	Inclination (deg)	$85.913^{+0.062}_{-0.041}$	-
$p = R_p/R_\star$	Planet-to-star radius ratio	$0.0234^{+0.0004}_{-0.0004}$	-
b	Impact parameter of the orbit	$0.861^{+0.007}_{-0.009}$	-
Derived physical parameters:		b	c
M_p	Planetary mass ($M_\oplus$)	$10.15^{+0.81}_{-0.84}$	-
R_p	Planetary radius ($R_\oplus$)	$2.265^{+0.096}_{-0.091}$	-
ρ_p	Planetary density (g/cm^3)	$4.85^{+0.84}_{-0.74}$	-
a_p	Semi-mayor axis (AU)	$0.052^{+0.024}_{-0.025}$	$0.080^{+0.013}_{-0.012}$
S	Insolation ($S_\oplus$)	232^{+11}_{-10}	95^{+5}_{-4}
T_{eq}	Equilibrium Temperature (K)	1077^{+10}_{-9}	859^{+9}_{-8}
$M_P \sin i$ ..	Minimum planetary mass ($M_\oplus$)	$10.15^{+0.82}_{-0.85}$	$9.78^{+1.26}_{-1.18}$

* For TOI 1062c T_0 corresponds to the time when the planet would have transited.

Figure 9 shows M-R curves tracing the compositions of Earth-like planets (with a CMF=0.33) and pure water subjected to a stellar radiation of $F/F_\oplus$= 230 (similar to that of TOI 1062 b). For reference, we also show exoplanets with accurate and reliable mass and radius determinations (Otegi et al. 2020a, accessible on the Data Analysis Center for Exoplanets, DACE[2]). TOI 1062 b sits above the Earth-like curve and below the pure-water curve, suggesting that it contains a small amount of volatile materials of less than 1% in mass. Figure 9 also displays the insolation flux relative to Earth against radii for the known exoplanets extracted from the NASA Exoplanet Archive, which shows the separate populations of super-Earths and mini-Neptunes. We see that TOI 1062 *b* sits in the mini-Neptune regime, close to the radius valley. In Otegi et al. (2020a) we identified two distinct populations, volatile-rich and "rocky" exoplanets (those expected to have small amounts of volatiles) separated by the water line. Using these results, we see that TOI 1062b lies in the rocky exoplanet regime defined in Otegi et al. (2020a), even if its radius is above the radius valley, suggesting that it is mostly composed of refractory materials by mass.

We used a generalized Bayesian inference analysis using a nested sampling scheme (Buchner et al. 2014) to quantify the degeneracy between various interior parameters and produce posterior probability distributions. Figure 10 shows ternary diagrams of the inferred composition of TOI 1062 b. The ternary diagram shows the degeneracy associated with the determination of the composition of exoplanets with measured mass and radius. We find a median H-He mass fraction of 0.1%, which corresponds to a lower bound since enriched H-He atmospheres are more com-

[2] https://dace.unige.ch/exoplanets/

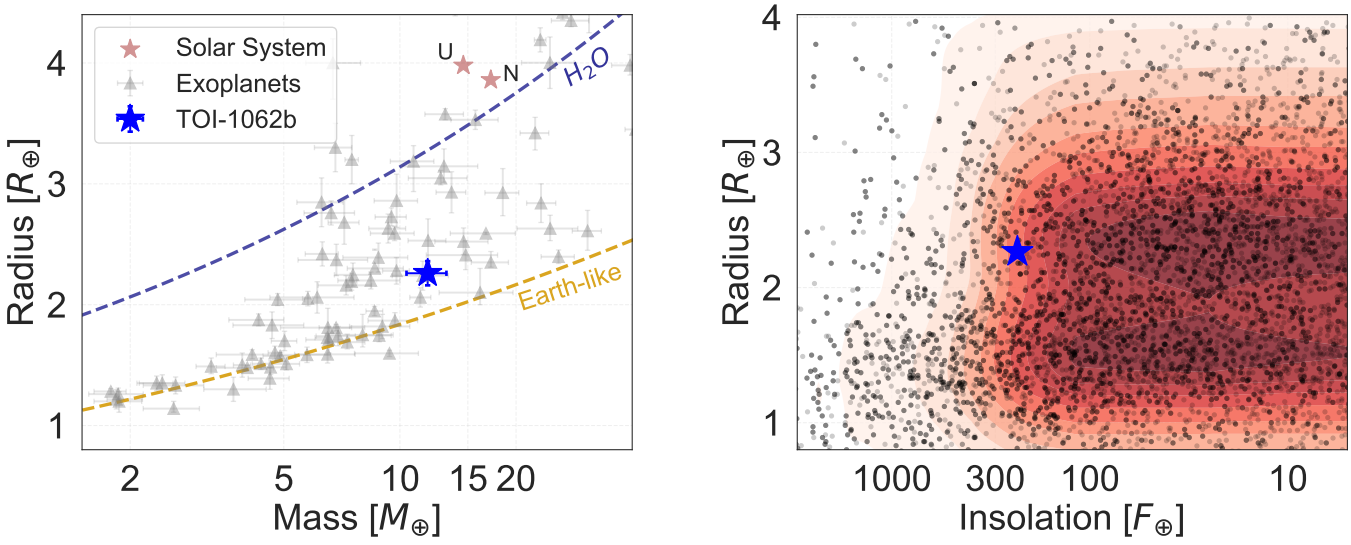

Fig. 9: Mass–radius and insolation–radius diagrams of observed exoplanets. Left: Mass–radius diagram of exoplanets with accurate mass and radius determination from Otegi et al. (2020a). Also shown are the composition lines for Earth-like planets and pure water subjected to a stellar radiation of $F/F_{\oplus}$= 230 (similar to that of TOI 1062 b). Right: Radius against insolation flux for known exoplanets from the NASA Exoplanet Archive. The colormap indicates the point density. TOI 1062 *b* is represented by a blue star.

Table 4: Inferred interior structure properties of TOI 1062b.

Constituent	**4-layer [%]**	**Without H_2O [%]**
$M_{\rm core}/M_{\rm total}$	44^{+26}_{-24}	26^{+26}_{-11}
$M_{\rm mantle}/M_{\rm total}$	40^{+27}_{-16}	73^{+11}_{-14}
$M_{\rm water}/M_{\rm total}$	15^{+9}_{-6}	-
$M_{\rm H-He}/M_{\rm total}$	$0.11^{+0.06}_{-0.04}$	$0.35^{+0.12}_{-0.08}$

pressed, and therefore increase the planetary H-He mass fraction. Formation models suggest that sub-Neptunes are likely formed by envelope enrichment (Venturini & Helled 2017). We also find that TOI 1062 b is expected to have a very significant iron core and silicate mantle, accounting for nearly 40% of the planetary mass and thicknesses of $1R_{\oplus}$ and $0.5R_{\oplus}$, respectively. The water layer has an estimated relative mass fraction of 15%. Nevertheless, the degeneracy between the core, silicate mantle, and water layer in this M-R regime is particularly high (Otegi et al. 2020b), and it does not allow accurate estimates of the masses of these constituents. Interior models cannot distinguish between water and H-He as the source of low-density material, so we also ran a three-layer model without the H_2O envelope. Table 4 lists the inferred mass fractions of the core, mantle, water layer, and H-He envelope for the four-layer model and for the water-free model. In this case we find that the planet is 0.35% H-He, 26% iron, and 73% rock by mass, setting maximum limits for the atmospheric and rock mass since any water added would decrease these mass fractions.

4.2. Dynamical analysis

As a general note for this section, the studies presented here were done under the hypothesis of a co-planar planetary system (i.e., the orbits of planets b and c evolve in the same plane). This is consistent with the observations, since TOI 1062c would not transit if it had the same orbital inclination as TOI 1062b.

4.2.1. Resonance

As can be seen from Table 2, the period ratio of the planets is $P_c/P_b = 1.941$. The planet pair thus lies close to the 2:1 mean-motion resonance (MMR). How close is the system to the MMR? We explore the structure of the parameter space of TOI-1062 near that resonance in order to investigate its dynamical state.

We therefore designed a two-dimensional section of the parameter space defined by the period ratio on one axis and the eccentricity of the outer planet e_c on the other. Our section has a resolution of 101x101, meaning that we explore the dynamics of 10201 initial configurations defined by a unique set of (P_c/P_b,e_c). All the other orbital parameters were initially fixed at the best values reported in Table 2. We numerically computed the future evolution of each configuration over 20 kyr. This was performed with the adaptive time-step high-order N-body integrator IAS15, which is available from the python package REBOUND (Rein & Liu 2012; Rein & Spiegel 2015). The perturbative effect of general relativity described in Anderson et al. (1975) was included via the library REBOUNDx (Tamayo et al. 2019). From these numerical simulations, the level of chaos of each configuration was evaluated with the Numerical Analysis of Fundamental Frequencies (NAFF) fast chaos indicator (Laskar et al. 1992; Laskar 1993). The result is presented in Fig. 11.

The NAFF computes precisely, for each planet, the average mean motion $n = \frac{2\pi}{P}$ over the two halves of the integration and compares these two estimations. Due to the secular constancy of the semi-major axis of regular orbits, this difference should be small in non-chaotic orbits. The higher the drift in the average mean-motion, the more chaotic the orbit. In this work we took as the NAFF of the system the maximum value of this drift over the planetary orbits in logarithmic scale, $NAFF = \max_i \left[log_{10} \frac{\Delta n_i}{n_{0,i}}\right]$, where the subscript i refers to the planet b or c, Δn_i is the drift of the average mean motion over the two halves of the integration, and $n_{0,i}$ is the initial mean motion of planet i. The color-coding in

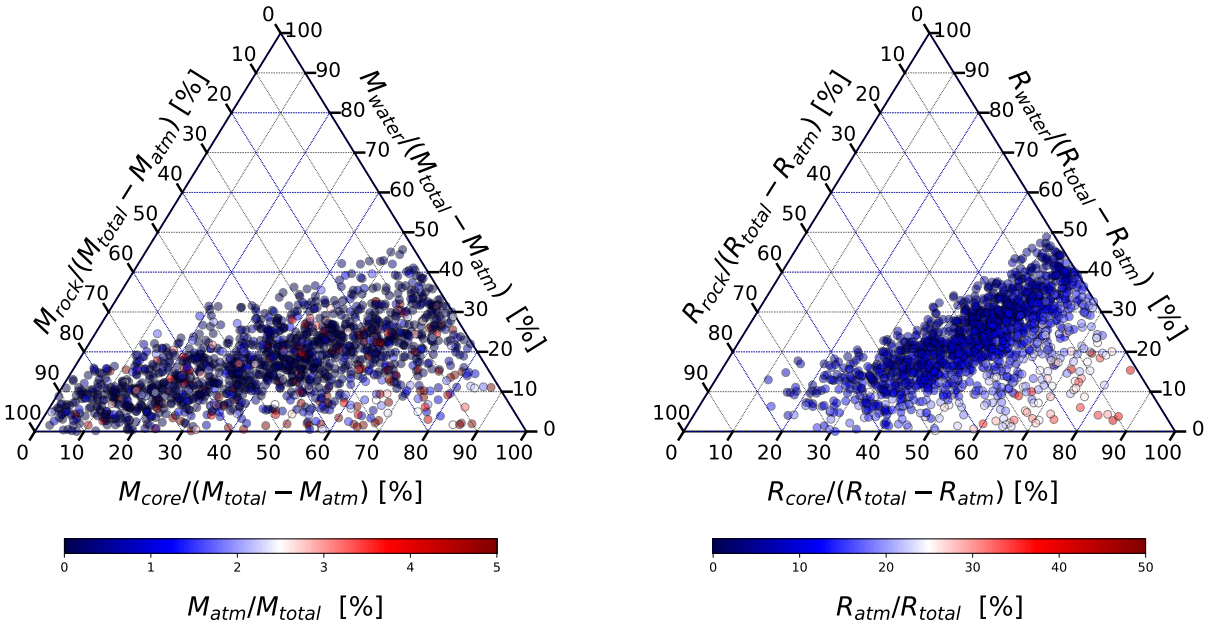

Fig. 10: Ternary diagram of the inferred internal composition of TOI 1062 b. We show the parameter space covered by the posterior distributions in the four-layer model in mass (left) and radius (right).

Fig. 11 depicts the NAFF defined above: the bluer colors indicate more regular systems. We also distinguish the strongly chaotic configurations from the ones that did not finish the integration because of either a close-encounter or an escape of a body (white boxes). We finally explain our choice of 20 kyr for the total integration time. Over this time span, the planetary orbits in the TOI-1062 system are expected to cover several secular cycles during which the semi-major axes oscillate. Covering several of these cycles allows to properly average the secular variations and isolate the chaotic diffusion. We verified numerically that several secular cycles are made up over an integration.

In this map, the 2:1 mean motion resonance appears clearly as the orange band in the middle of the plot. For the current system's parameters and the resolution of our chaoticity map, this resonance therefore seems chaotic. It is important to note that this picture is highly dependent on the system's parameters. For instance, in this case the arguments of periastron ω_b and ω_c are close to the alignment. The opposite configuration where the orbits are anti-aligned would show a drastically different picture, with a different strength and apparent stability of the 2:1 MMR. The two vertical lines depict the 1σ window of P_c. With the currently estimated orbital parameters and planet masses, the TOI 1062 system certainly lies outside of the 2:1 MMR. Despite the influence that a revision of the parameters may have on the resonance, it seems very unlikely that this conclusion will change.

Finally, we note that the uncertainty on the eccentricity of the inner planet e_b is quite large. Modifying this parameter will directly impact the strength of the 2:1 MMR as the resonance width is expected to increase with the eccentricity. The same also applies for the planetary masses.

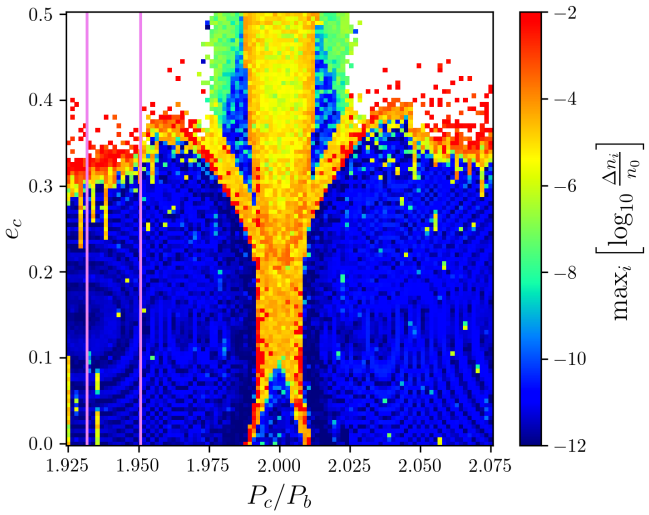

Fig. 11: Chaoticity map around the solution for the two-planet model (Table 2). The eccentricity of the outer planet e_c and the period ratio P_c/P_b are explored on a grid of 101x101 different configurations. A color is assigned to each box after a short numerical integration, according to its level of chaos using the NAFF indicator (see text).

4.2.2. Stability constraints

While the chaoticity map presented above (Fig. 11) is informative about the maximum eccentricity of the outer planet allowed by orbital stability, this estimate should be considered with caution. This map explores the parameter space in two dimensions, without taking into account potential correlations with the other parameters. As already discussed, if the best-fit estimate of the other parameters changes, the entire map might change as well. In order to give proper stability constraints on the orbital parameters

and masses of the planets, we also performed a global stability analysis from the posterior distribution of the joint photometry + RV analysis. The only fixed parameter is the inclination of the outer planet, taken equal to the inclination of the inner planet (i.e., we explored the co-planar case, which is compatible with the outer planet non transiting). We selected a sample of 10000 solutions from that posterior. Each is numerically integrated over 20 kyr using the same integrator set-up described in the previous section. The level of chaos is again computed with the NAFF indicator. Over the 10000 configurations, 8455 survived the entire 20 kyr simulation; the others were unstable (escape or close encounter of two bodies). Imposing a more strict stability threshold by further removing all solutions with $NAFF > -5$ leaves us with a stable posterior of 7224 configurations.

In any case, no stringent constraint can be added on the orbital parameters and masses of the planets. With the sorting in NAFF as defined above, only a slight cut in the eccentricities is observed. The new median estimates and 1σ confidence intervals are $e_b = 0.162$ [0.098,0.228] and $e_c = 0.129$ [0.065,0.205]. In particular, the global stability analysis does not allow us to further constrain the planetary masses (again assuming co-planar orbits).

4.3. Atmospheric characterization

TOI 1062b poses an interesting target for atmospheric characterization given its equilibrium temperature of ∼1000 K. We expect a high-metallicity atmosphere for a planet with such a small radius and mass, as a result of the strong stellar irradiation received which would result in a significant loss of the H_2/He envelope. However, such a target is an excellent laboratory to study carbon chemistry at high temperature given its high expected metallicity. Beyond metallicities of $\gtrsim 300\times$ solar, the primary carbon bearing species in the atmosphere transitions from being CH_4/CO to CO_2 in thermochemical equilibrium (e.g., Moses et al. 2013). This planet is thus an ideal target to explore the CO, CH_4, and CO_2 chemical stability boundary which occurs at ∼800 K in photospheric conditions at such high metallicities. The mean molecular weight for such an atmosphere at $\sim 300\times$ metallicity is expected to be ∼4 g/mol, and hence significant features are still present in transmission spectra compared to a solar metallicity atmosphere (at $\sim$ 2.35 g/mol). Using the transmission spectroscopy metric (TSM) defined in Kempton et al. (2018), we determine that TOI 1062b has a TSM value of 31.8 assuming a mean molecular mass of 2.3 g/mol and 18.3 assuming 4 g/mol. The lower TSM value for the higher mean molecular weight is due to the reduced atmospheric scale height of a heavier atmosphere. This value is comparable to targets such as Trappist-1f and several simulated *TESS* targets from Sullivan et al. (2015).

On the other hand, TOI 1062b may have completely lost all of its H_2/He envelope which would result in an ultra-high-metallicity secondary atmosphere. At metallicities $\gtrsim 3000\times$ solar, the mean molecular weight of the heavier species such as CO_2 and H_2O now completely dominate due to the lack of significant H_2 or He. Therefore, the mean molecular weight is likely to be > 18 g/mol, and thus the scale height of spectral features in the transmission spectra are significantly reduced by a factor of $\gtrsim 7\times$ over a solar metallicity atmosphere. This reduces the TSM value for such an atmosphere to $\lesssim$5, making constraints on the abundances difficult. However, observations of the terminator may still be able to place a lower limit on the molecular weight, and thus the metallicity. They would also provide a direct contrast to planets such as 55 Cancri e which also indicate a high mean molecular weight atmosphere (Jindal et al. 2020). In addition, secondary eclipse spectroscopy for such a cool target is challenging, but has been achieved for GJ436b (Stevenson et al. 2010). However, for the case of TOI 1062b, the host star is a G-type and thus will result in a very weak planet–star flux contrast in emission spectroscopy.

5. Conclusions

We present the discovery of two new planets from the *TESS* mission in the TOI 1062 system. The analysis is based on 2 min cadence *TESS* observations from four sectors, ground-based follow-up from LCOGT, and RV data from the HARPS spectrograph. High-resolution imaging from Zorro speckle imager rules out the presence of nearby companions and potential nearby eclipsing binaries. We find that the host star is rotating with a period of 21.9 days, and also note the existence of a stellar spot whose lifetime is similar to the stellar rotational period.

TOI 1062b is expected to be a mini-Neptune with a period of 4.11 days, radius of 2.27±0.1 $R_\oplus$, and mass of 10.2±0.8 $M_\oplus$. Internal structure models indicate that TOI 1062b is expected to be composed of significant iron core and silicate mantle, accounting for nearly 40% of the planetary mass each, and that it is expected to have a small volatile envelope of 0.35% of the mass at most. TOI 1062c, which is not transiting, is found in the HARPS RV data. We find that its minimum mass is inferred to be 9.8±1.2 $M_\oplus$, and that it is close to a 2:1 motion resonance with its inner companion, with a period of 7.97 days. The position of TOI 1062b with respect to the radius valley and its high equilibrium temperature of ∼1000K make it an interesting candidate for atmospheric characterization. The strong stellar irradiation may result in a significant loss of the H-He envelope leaving a high-metallicity atmosphere that can be an excellent laboratory to study carbon chemistry at high temperature.

Acknowledgements. This work has been in particular carried out in the frame of the National Centre for Competence in Research 'PlanetS' supported by SNSF. D.J.A. acknowledges support from the STFC via an Ernest Rutherford Fellowship (ST/R00384X/1). This work was supported by FCT - Fundação para a Ciência e a Tecnologia through national funds and by FEDER through COMPETE2020 - Programa Operacional Competitividade e Internacionalização by these grants: UID/FIS/04434/2019; UIDB/04434/2020; UIDP/04434/2020; PTDC/FIS-AST/32113/2017 & POCI-01-0145-FEDER-032113; PTDC/FIS-AST/28953/2017 & POCI-01-0145-FEDER-028953. V.A. and E.D.M acknowledge the support from FCT through Investigador FCT contracts nr. IF/00650/2015/CP1273/CT0001, IF/00849/2015/CP1273/CT0003, respectively. C.D. acknowledges support from the Swiss National Science Foundation under grant PZ00P2_174028. Siddharth Gandhi acknowledges support from the UK Science and Technology Facilities Council (STFC) research grant ST/S000631/1. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center for the production of the SPOC data products. We acknowledge the use of public *TESS* Alert data from pipelines at the *TESS* Science Office and at the *TESS* Science Processing Operations Center. This research has been partly funded by the Spanish State Research Agency (AEI) Projects No.ESP2017-87676-C5-1-R and No. MDM-2017-0737 Unidad de Excelencia "María de Maeztu"- Centro de Astrobiología (INTA-CSIC). S.G.S acknowledges the support from FCT through Investigador FCT contract nr. CEECIND/00826/2018 and POPH/FSE (EC). H.P.O. acknowledges that this work has been carried out within the framework of the NCCR PlanetS supported by the Swiss National Science Foundation. S.H. acknowledge support by the fellowships PD/BD/128119/2016 funded by FCT (Portugal). X.D is grateful to The Branco Weiss Fellowship–Society in Science for its financial support. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 851555/SCORE. A.O acknowledges support from an STFC studentship. S.H acknowledges CNES funding through the grant 837319.This work was supported by FCT through national funds (PTDC/FIS-AST/28953/2017) and by FEDER - Fundo Europeu de Desenvolvimento Regional through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI-01-0145-FEDER-028953) and through national funds (PIDDAC) by the grant UID/FIS/04434/2019. Funding for the *TESS* mission is

provided by NASA's Science Mission directorate. This work makes use of observations from the LCOGT network.

References

Adibekyan, V. Z., Sousa, S. G., Santos, N. C., et al. 2012, A&A, 545, A32
Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, PASP, 125, 989
Alibert, Y., Mordasini, C., Benz, W., & Naef, D. 2010, in EAS Publications Series, Vol. 42, EAS Publications Series, ed. K. Gożdziewski, A. Niedzielski, & J. Schneider, 209–225
Alibert, Y., thiabaud, a., marboeuf, u., et al. 2015, in AAS/Division for Extreme Solar Systems Abstracts, Vol. 47, AAS/Division for Extreme Solar Systems Abstracts, 115.02
Anderson, J. D., Esposito, P. B., Martin, W., Thornton, C. L., & Muhleman, D. O. 1975, ApJ, 200, 221
Armstrong, D. J., Lopez, T. A., Adibekyan, V., et al. 2020, Nature, 583, 39
Baluev, R. V. 2008, MNRAS, 385, 1279
Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
Beaugé, C. & Nesvorný, D. 2013, ApJ, 763, 12
Bertran de Lis, S., Delgado Mena, E., Adibekyan, V. Z., Santos, N. C., & Sousa, S. G. 2015, A&A, 576, A89
Borucki, W. J., Koch, D., & Kepler Science Team. 2010, in AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 42, AAS/Division for Planetary Sciences Meeting Abstracts #42, 47.03
Brown, T. M., Baliber, N., Bianco, F. B., et al. 2013, PASP, 125, 1031
Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125
Chabrier, G., Mazevet, S., & Soubiran, F. 2019, ApJ, 872, 51
Collins, K. A., Kielkopf, J. F., Stassun, K. G., & Hessman, F. V. 2017, AJ, 153, 77
Connolly, J. A. D. 2009, Geochemistry, Geophysics, Geosystems, 10, Q10014
Delgado Mena, E., Israelian, G., González Hernández, J. I., et al. 2010, ApJ, 725, 2349
Delgado Mena, E., Moya, A., Adibekyan, V., et al. 2019, A&A, 624, A78
Delgado Mena, E., Tsantaki, M., Adibekyan, V. Z., et al. 2017, A&A, 606, A94
Delisle, J. B., Hara, N., & Ségransan, D. 2020, A&A, 638, A95
Espinoza, N. 2018, Research Notes of the American Astronomical Society, 2, 209
Espinoza, N. & Jordán, A. 2015, MNRAS, 450, 1879
Espinoza, N., Kossakowski, D., & Brahm, R. 2019, MNRAS, 490, 2262
Feroz, F. & Hobson, M. P. 2008, MNRAS, 384, 449
Findeisen, K., Hillenbrand, L., & Soderblom, D. 2011, AJ, 142, 23
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220
Fulton, B. J., Petigura, E. A., Blunt, S., & Sinukoff, E. 2018, PASP, 130, 044504
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, ArXiv e-prints [arXiv:1804.09365]
Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759
Guerrero, N. 2020, in American Astronomical Society Meeting Abstracts, Vol. 235, American Astronomical Society Meeting Abstracts #235, 327.03
Guillot, T. 2010, A&A, 520, A27
Gupta, A. & Schlichting, H. E. 2019, MNRAS, 487, 24
Hakim, K., Rivoldini, A., Van Hoolst, T., et al. 2018, Icarus, 313, 61
Haldemann, J., Alibert, Y., Mordasini, C., & Benz, W. 2020, arXiv e-prints, arXiv:2009.10098
Hara, N. C., Bouchy, F., Stalport, M., et al. 2020, A&A, 636, L6
Hara, N. C., Boué, G., Laskar, J., & Correia, A. C. M. 2017, MNRAS, 464, 1220
Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27
Howell, S. B., Everett, M. E., Sherry, W., Horch, E., & Ciardi, D. R. 2011, AJ, 142, 19
Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398
Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, in Proc. SPIE, Vol. 9913, Software and Cyberinfrastructure for Astronomy IV, 99133E
Jenkins, J. S., Díaz, M. R., Kurtovic, N. T., et al. 2020, arXiv e-prints, arXiv:2009.12832
Jensen, E. 2013, Tapir: A web interface for transit/eclipse observability
Jindal, A., de Mooij, E. J. W., Jayawardhana, R., et al. 2020, AJ, 160, 101
Kempton, E. M. R., Bean, J. L., Louie, D. R., et al. 2018, PASP, 130, 114401
Kipping, D. M. 2013, MNRAS, 435, 2152
Kreidberg, L. 2015, PASP, 127, 1161
Kurucz, R. L. 1993, SYNTHE spectrum synthesis programs and line data
Laskar, J. 1993, Physica D Nonlinear Phenomena, 67, 257
Laskar, J., Froeschlé, C., & Celletti, A. 1992, Physica D Nonlinear Phenomena, 56, 253
Li, J., Tenenbaum, P., Twicken, J. D., et al. 2019, PASP, 131, 024506
Lopez, E. & Fortney, J. J. 2013, in AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 45, AAS/Division for Planetary Sciences Meeting Abstracts #45, 200.08
Lucy, L. B. & Sweeney, M. A. 1971, AJ, 76, 544
Mamajek, E. E. & Hillenbrand, L. A. 2008, ApJ, 687, 1264
Mayor, M., Pepe, F., Queloz, D., et al. 2003a, The Messenger, 114, 20
Mayor, M., Pepe, F., Queloz, D., et al. 2003b, The Messenger, 114, 20
Mazeh, T., Holczer, T., & Faigler, S. 2016, A&A, 589, A75
Moses, J. I., Line, M. R., Visscher, C., et al. 2013, ApJ, 777, 34
Nielsen, L. D., Gandolfi, D., Armstrong, D. J., et al. 2020, MNRAS, 492, 5399
Otegi, J. F., Bouchy, F., & Helled, R. 2020a, A&A, 634, A43
Otegi, J. F., Dorn, C., Helled, R., et al. 2020b, A&A, 640, A135
Owen, J. E. & Jackson, A. P. 2012, MNRAS, 425, 2931
Pepe, F., Mayor, M., Rupprecht, G., et al. 2002, The Messenger, 110, 9
Plotnykov, M. & Valencia, D. 2020, MNRAS, 499, 932
Rein, H. & Liu, S. F. 2012, A&A, 537, A128
Rein, H. & Spiegel, D. S. 2015, MNRAS, 446, 1424
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Santos, N. C., Israelian, G., & Mayor, M. 2004, A&A, 415, 1153
Santos, N. C., Sousa, S. G., Mortier, A., et al. 2013, A&A, 556, A150
Scott, N. J. 2019, in AAS/Division for Extreme Solar Systems Abstracts, Vol. 51, AAS/Division for Extreme Solar Systems Abstracts, 330.15
Skrutskie, M. F. et al. 2006, AJ, 131, 1163
Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, PASP, 124, 1000
Sneden, C. A. 1973, PhD thesis, THE UNIVERSITY OF TEXAS AT AUSTIN.
Sousa, S. G. 2014, [arXiv:1407.5817] [arXiv:1407.5817]
Sousa, S. G., Santos, N. C., Adibekyan, V., Delgado-Mena, E., & Israelian, G. 2015, A&A, 577, A67
Speagle, J. S. 2020, MNRAS, 493, 3132
Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, AJ, 153, 136
Stassun, K. G., Corsaro, E., Pepper, J. A., & Gaudi, B. S. 2018, AJ, 155, 22
Stassun, K. G. & Torres, G. 2016, AJ, 152, 180
Stassun, K. G. & Torres, G. 2018, ApJ, 862, 61
Stevenson, K. B., Harrington, J., Nymeyer, S., et al. 2010, Nature, 464, 1161
Stixrude, L. & Lithgow-Bertelloni, C. 2011, Geophysical Journal International, 184, 1180
Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, ApJ, 809, 77
Szabó, G. M. & Kiss, L. L. 2011, ApJ, 727, L44
Tamayo, D., Rein, H., Shi, P., & Hernandez, D. M. 2019, Monthly Notices of the Royal Astronomical Society, 491, 2885–2901
Torres, G., Andersen, J., & Giménez, A. 2010, A&A Rev., 18, 67
Twicken, J. D., Catanzarite, J. H., Clarke, B. D., et al. 2018, PASP, 130, 064502
Twicken, J. D., Chandrasekaran, H., Jenkins, J. M., et al. 2010, in Proc. SPIE, Vol. 7740, Software and Cyberinfrastructure for Astronomy, 77401U
Venturini, J. & Helled, R. 2017, ApJ, 848, 95
Wang, H. S., Liu, F., Ireland, T. R., et al. 2019, MNRAS, 482, 2222
Wright, E. L. et al. 2010, AJ, 140, 1868

Appendix A: HARPS spectroscopy

Table A.1: HARPS spectroscopy obtained between 31 August 2019 and 11 November 2019.

Time [BJD − 2400000]	RV [m/s]	σ_{RV}	S_{MW}	σ_S	FWHM [m/s]	σ_{FWHM}
			–			
58726.77	−4.63	1.19	0.2308	0.0035	6521.73	9.7
58726.83	−3.58	1.39	0.2205	0.0051	6516.21	15.2
58727.78	−4.91	1.19	0.2218	0.0032	6514.78	9.6
58727.86	−2.92	1.19	0.2223	0.0035	6520.8	10.4
58728.71	−4.13	1.11	0.2305	0.0026	6511.39	7.3
58728.76	−3.05	1.03	0.2284	0.0021	6513.01	5.9
58729.76	−5.6	1.16	0.2246	0.0031	6511.48	9.0
58729.86	−5.54	1.1	0.226	0.0029	6511.24	8.2
58730.74	−0.72	1.49	0.2145	0.0051	6510.23	16.0
58730.82	1.18	1.49	0.2139	0.0055	6515.79	17.1
58744.77	−1.12	1.68	0.2143	0.0068	6522.73	21.2
58745.74	−16.3	3.27	0.2023	0.0179	6534.68	61.2
58746.69	0.09	1.54	0.2148	0.0055	6517.99	17.2
58747.68	6.78	2.19	0.2099	0.01	6505.19	30.9
58747.7	4.5	1.73	0.2089	0.0068	6513.01	22.0
58748.75	7.97	1.56	0.2267	0.0056	6507.91	16.2
58750.76	−2.32	1.4	0.2289	0.0046	6522.1	13.1
58760.76	−1.13	1.79	0.1976	0.008	6507.77	28.3
58760.85	1.43	1.34	0.2064	0.005	6511.14	16.5
58761.72	−1.0	1.22	0.2259	0.0037	6515.05	10.7
58762.71	−2.72	1.33	0.2152	0.0042	6512.07	13.1
58763.78	3.58	1.13	0.2196	0.003	6506.89	9.1
58764.67	0.32	1.21	0.219	0.0031	6509.88	9.5
58765.71	−43.31	4.43	0.23	0.0271	6569.43	76.2
58766.69	−13.89	1.11	0.2193	0.0026	6502.88	7.7
58773.73	−2.55	1.66	0.202	0.0069	6513.39	23.6
58774.69	−10.79	1.31	0.2135	0.0039	6507.43	12.3
58774.8	−13.66	1.57	0.1942	0.0062	6509.0	22.5
58775.67	−10.02	1.31	0.2203	0.0042	6507.57	12.5
58775.81	−11.1	1.89	0.2004	0.0084	6519.12	29.1
58776.74	−0.04	2.33	0.2159	0.0118	6514.25	36.6
58776.8	−8.42	2.18	0.2083	0.011	6509.26	36.0
58777.66	1.5	1.65	0.2212	0.0065	6519.42	19.3
58777.78	−4.05	2.79	0.2323	0.015	6512.9	41.6
58791.6	−8.57	1.61	0.2249	0.0061	6511.53	17.7
58791.78	−6.63	1.85	0.1908	0.0084	6506.83	31.5
58792.62	−1.06	1.31	0.2249	0.004	6506.84	11.6
58792.76	−2.15	1.28	0.2163	0.0044	6506.41	13.5
58793.62	5.46	1.26	0.2296	0.0033	6508.11	9.3
58793.78	6.2	1.22	0.2152	0.004	6510.24	12.5
58794.6	2.03	1.36	0.2122	0.0041	6510.28	13.1
58794.71	2.13	1.3	0.2181	0.004	6507.05	12.2
58795.59	−2.05	1.53	0.2051	0.0046	6506.22	15.3
58795.72	−1.34	1.28	0.213	0.0037	6506.98	11.6
58796.6	−1.28	1.17	0.2217	0.0029	6505.65	8.7
58796.73	−0.69	1.2	0.2181	0.0034	6502.47	10.3
58797.6	−2.84	1.31	0.2107	0.0036	6503.46	11.6
58797.73	−0.47	1.41	0.2095	0.0046	6507.76	15.0
58798.67	−5.87	1.21	0.2189	0.003	6503.02	9.2
58799.75	−16.63	1.53	0.203	0.0058	6509.75	19.8
58799.86	−23.92	1.56	0.1928	0.0064	6525.7	23.5

Table A.2: HARPS spectroscopy obtained between 20 November 2019 and 17 January 2020.

Time [BJD − 2400000]	RV [m/s]	σ_{RV}	S_{MW}	σ_S	FWHM [m/s]	σ_{FWHM}
			–			
58808.67	−0.17	1.29	0.2389	0.0041	6538.41	10.9
58810.65	4.55	1.29	0.2362	0.0038	6530.36	10.4
58811.71	−0.94	1.27	0.228	0.004	6527.1	11.3
58816.73	−7.0	1.1	0.2152	0.003	6504.0	9.4
58819.7	−6.58	1.4	0.2183	0.0044	6501.38	13.3
58821.75	−3.34	1.41	0.2111	0.0047	6496.4	14.9
58823.73	−12.12	1.5	0.2006	0.0053	6497.09	18.5
58830.59	−3.48	0.98	0.2136	0.002	6495.06	6.2
58830.69	−4.1	1.05	0.2117	0.0027	6496.99	8.7
58831.54	−7.06	1.02	0.2134	0.0019	6490.0	6.1
58831.67	−9.5	1.15	0.2176	0.0033	6498.19	9.9
58832.55	−10.45	1.63	0.2077	0.0057	6500.63	18.7
58832.66	−9.7	1.75	0.2083	0.0069	6509.54	20.8
58833.61	−1.92	2.39	0.1895	0.0109	6499.71	41.1
58834.58	9.53	1.31	0.2155	0.0039	6500.35	12.2
58834.68	9.52	1.28	0.2154	0.0042	6505.94	12.9
58835.58	11.45	1.06	0.2275	0.0022	6514.18	6.4
58835.69	8.87	1.2	0.2284	0.0038	6507.52	10.7
58836.55	7.75	1.05	0.2285	0.0022	6518.42	6.2
58836.69	7.74	1.13	0.2274	0.0031	6516.43	8.8
58837.57	10.35	1.23	0.2329	0.0034	6519.22	9.4
58837.69	8.19	1.42	0.2273	0.0051	6528.12	14.6
58838.66	12.94	1.33	0.2383	0.0044	6522.37	11.8
58839.55	9.99	1.12	0.2455	0.0025	6531.33	6.4
58839.61	9.62	1.14	0.2524	0.0029	6533.14	7.1
58840.63	4.6	1.09	0.2506	0.003	6536.72	7.4
58841.68	9.26	1.15	0.243	0.0035	6536.59	9.1
58842.63	14.97	1.21	0.2412	0.0038	6537.23	9.9
58843.56	10.7	1.03	0.2515	0.0021	6535.97	5.2
58844.59	0.13	1.16	0.2475	0.0031	6532.33	7.8
58845.67	−4.45	1.34	0.2389	0.0046	6523.74	12.3
58847.65	−8.6	1.22	0.2252	0.0038	6512.18	11.1
58848.6	−12.47	1.39	0.2146	0.0046	6510.43	14.4
58848.61	−13.08	1.44	0.2291	0.0047	6510.57	13.4
58849.64	−9.39	1.23	0.219	0.0038	6507.37	11.6
58850.58	0.44	1.16	0.2219	0.0029	6501.77	8.4
58852.57	−3.63	1.03	0.22	0.0024	6503.71	7.1
58853.61	−6.25	1.18	0.2178	0.0034	6498.99	9.7
58855.64	2.87	2.48	0.1971	0.0119	6509.76	42.5
58865.55	−5.45	1.2	0.2276	0.0035	6511.98	9.9

3.2.2.2 Mass determinations of the three mini-Neptunes transiting TOI-125: *Nielsen L. D., et al. 2020a, MNRAS, 492, 5399*

In this paper Nielsen et al. (2020a) confirmed the detection of three mini-Neptunes around TOI-125 using HARPS RVs measurements under the NCORES program. There are 82 co-authors in the paper, and my contribution was awarded with the 16th position. TOI-125b, TOI-125c and TOI-125d have similar radii; 2.726±0.075$R_\oplus$, 2.759±0.10$R_\oplus$, and 2.93±0.17$R_\oplus$, respectively. However, the three planets have significantly different masses (9.50±0.88$M_\oplus$, 6.63±0.99$M_\oplus$ and 13.6±1.2$M_\oplus$). This gives a high-low-higher pattern in terms of density when moving outward in the system. From the TESS light curves there are two additional planet candidates which are not confirmed by RV. We derived upper mass limits of 1.6$M_\oplus$ and 2.7$M_\oplus$ for TOI-125.04 and TOI-125.05, respectively.

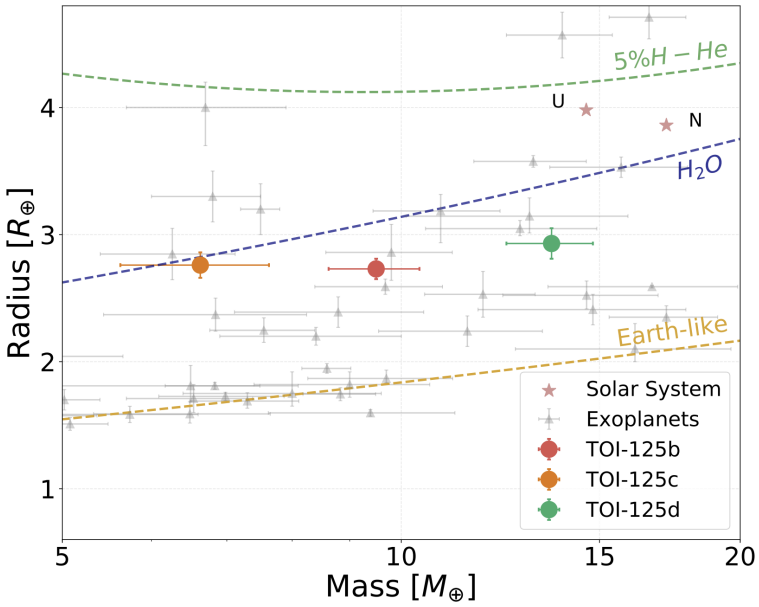

Figure 3.8: Mass-radius diagram of exoplanets from the PlanetS catalog. Also shown are the composition lines of an Earth-like planet, pure-water and 95% H2O + 5% H He. The Figure is taken from Nielsen et al. (2020a)

The system has an interesting architecture, with the two inner planet slightly interior to the 2:1 Mean Motion Resonance (MMR) and the two outer planets slightly external to the 2:1 MMR. TOI125b and TOI-125d both show significant orbital eccentricities of 0.194± 0.04 and 0.168± 0.08, respectively. A dynamical analysis of the system using N-body simulations shows that planetary orbits are stable despite the high eccentricities.

Figure 3.8 shows the M-R diagram of the planets in the TOI-125 system with the observed exoplanet population and various composition lines. We characterize the internal compositions of these three planets and conclude that they are likely likely retain H-He atmospheres and a significant water layer. We estimate the H-He mass fractions of TOI-125b, TOI-125c, TOI-125d to be 0.02±0.01%, 0.02±0.01% and 0.04±0.01%, respectively.

3.2.2.3 TESS Reveals a Short-period Sub-Neptune Sibling (HD 86226c) to a Known Long-period Giant Planet: *Teske J. et al. 2020, AJ, 96, 160*

In this work by Teske et al. (2020) we present the detection and characterization of a short-period (~4 days) planet transiting TOI-652, which was previously known to host a long period (~1600 days) giant non-transiting planet. My contribution was rewarded with the 6th position in a co-author list of 47. We get a radius of 2.16±0.08$R_\oplus$ and a mass of 7.25±1.15$M_\oplus$ for the transiting planet, and update the parameters of the longer-period gas giant. There is a growing evidence that ~30-40% of stars hosting small planets also host a larger long-period planet, which is interesting because it suggests that the presence of outer gas giant planets does not avoid the formation of inner smaller planets, and it may even facilitate the growth of small planets. Whether there are differences in the properties of small planets that have or do not have giant planet companions is thus an interesting open question (see Section 2.3.3).

We explore the range of possible compositions of the transiting planet. We assumed that the Fe/Si and Mg/Si ratios inside the planet are the same as the ratios observed in the stellar photosphere. We estimated a H-He envelope accounting for 6.10^{-3}% of the planetary mass, and with a thickness of 0.39$R_\oplus$. We also find that the other constituents of the planet have relative mass fractions between 32% and 35% with large uncertainties. As found in Otegi et al. (2020b), this regime of the M-R diagram is strongly degenerate, and therefore even with more precise mass and/or radius measurements, it would not be possible to significantly improve the estimate of the mass ratio between the core, mantle, and water layer. Atmospheric characterization will be crucial to better constrain the volatile envelope properties and composition.

3.2.2.4 A remnant planetary core in the hot-Neptune desert: *Armstrong D. J. et al. 2020, Nature, 39-42, 583*

Another system that we characterized with our interior models is TOI-849 (Armstrong et al. 2020), which was already briefly discussed in Section 2.3.1. My position in the co-author list is the 16th out of 94. With a mass of $39.1^{+2.8}_{-2.6}M_\oplus$, radius of 3.45±0.14$R_\oplus$ and equilibrium

temperature of 1800K, it is one of the most remarkable planet discoveries during in the last years. Figure 3.9 shows TOI-849b in the context of the hot-Neptune desert. We can see that it is relatively isolated in the parameter space, suggesting that it is somewhat unique and could have been subjected to unusually aggressive removal of the initial H-He envelope.

We used our models to infer the planetary interior of TOI-849b. The set the core-to-mantle fraction is set by the stellar abundance [Fe/Si] of the host star, and estimated the maximum possible mass of an H–He envelope by assuming a planet without water. The minimum H-He fraction was estimated by a assuming a large fraction of water of 70% by mass, which corresponds to a water rich planet. We obtained that the H-He mass fraction is at minimum $2.9^{+0.8}_{-1.0}$% and at maximum $3.9^{+0.8}_{-0.9}$%, suggesting that the heavy element mass is higher than $38 M_{\oplus}$.

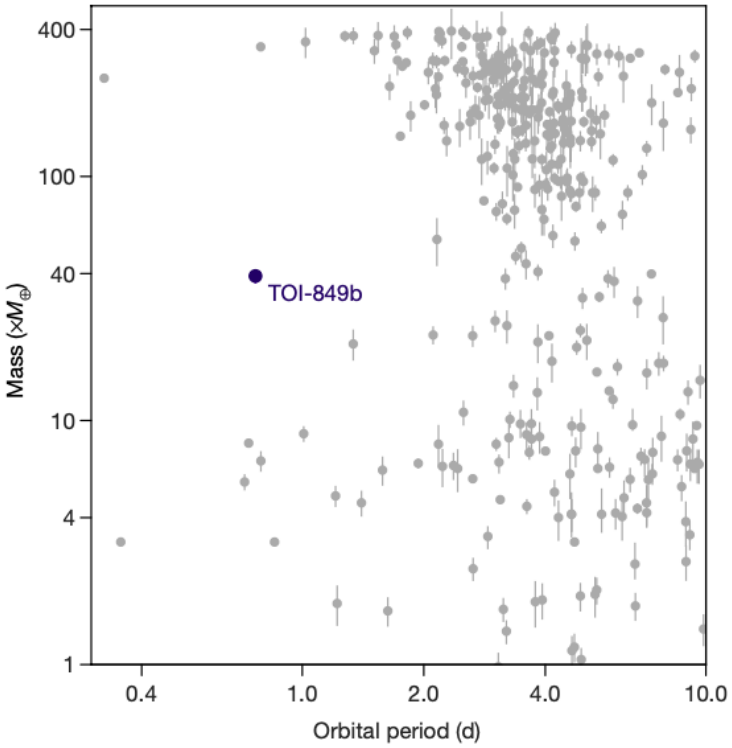

Figure 3.9: Position of TOI-849b in the mass-orbital period diagram. The grey exoplanets correspond to know exoplanets from the NASA Exoplaner Archive. The Figure is taken from Armstrong et al. (2020)

3.2.2.5 TOI-824 b: A New Planet on the Lower Edge of the Hot Neptune Desert: *Burt, J. A. et al. 2020, AJ, 87, 162*

In this study Burt et al. (2020b) reports the detection of a hot Neptune transiting TOI-824. My contribution to the work was awarded with the 11th position out of 54. Using TESS photometry and HARPS RVs under the NCORES program we determine a radius of 2.93±0.20$R_\oplus$ and a mass of 18.47±1.84$M_\oplus$. It has a short orbital period of 1.393 days, and an equilibrium temperature of 1253K. The planet's bulk density is $4.03^{+0.98}_{-0.78}$ g cm^{-3}, making it more than twice as dense as Neptune. We model its interior and estimate a H-He mass fraction of 2.8±0.08%, which is a lower-bound since enriched H-He atmospheres are more compressed, and can therefore increase the planetary H-He mass fraction. Assuming a rock and iron core, thermal evolutions models suggest a H-He mass fraction of $2.4^{+1.1}_{-1.7}$%, which is consistent with the results of our interior models. The core, mantle, and water layer have relative mass fractions of 27%, 38%, and 31% with large uncertainties, respectively.

TOI-824b is interesting in the context of the hot Neptune desert since it appears to lie at the lower edge of the desert. However, its mass and radius suggest that it possesses a significant primary atmosphere. The obtained H-He mass is within the typical range of the warmest Neptune planets discovered by Kepler, even if these planets are much less irradiated than TOI-824b. In the paper we explore several planet evolution and escape models that may be able to explain how can this planet have retained such a significant gaseous envelope despite receiving extreme irradiation. TOI-824 b is the most irradiated small planet at the edge of the hot Neptune desert that has retained its atmosphere, and further studies of its atmosphere could lead to a more detailed characterization.

3.2.2.6 A hot mini-Neptune in the radius valley orbiting solar analogue HD 110113: *Osborn H. et al. 2020, MNRAS, 4842, 502*

In Osborn et al. (2021b) we reported the descovery of a mini Neptune transiting TOI-755. There were 56 people involved in the discovery and I was awarded with the 11th position. The HARPS RVs revealed strong activity on TOI-755, with a rotation period of 20.8± days. After removing this rotation period using a co-fitted gaussian processes (GPs) with the S-index and the FWHM, two Keplerian signals are found at 2.541 and 6.744 days. The inner planet has a mass 4.55±0.62$M_\oplus$ and a radius of 2.05±0.62$R_\oplus$. The outer planet is not transiting, and has a minimum mass of 10.5±1.2$M_\oplus$.

The estimated density of TOI-755b is $2.90^{+0.75}_{-0.59}$ g cm^{-3}, far lower than would be expected

from a rocky core. By modelling its interior we were able to discard a composition without volatiles, suggesting that it has between 0.07 and 1.5% H-He by mass. This is surprising given its position in the radius valley between gaseous mini Neptunes and rocky super-Earths (see Figure 3.10). It is likely that TOI-755b started with a thicker atmosphere of H-He, which, due to both evaporative and core-powered mass-loss, was lost over time. However, typically planets which lose the majority of their gas lose it all Owen & Wu (2017). We suggest two possibilities for the unexpected planetary low density: either TOI-755b has a water-rich core and secondary atmosphere, or it began with a thick H-He envelope and managed to retain a small fraction of it despite significant evaporation and/or heating.

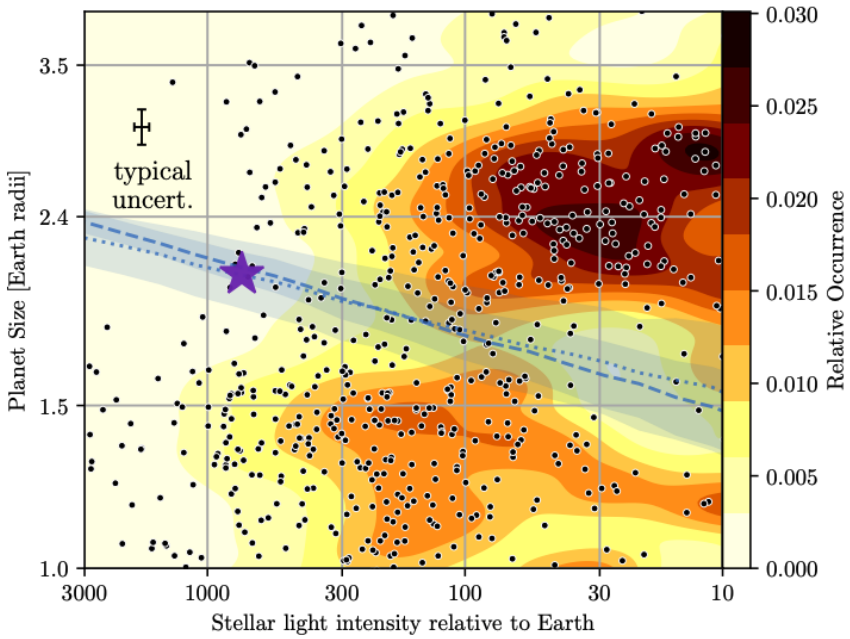

Figure 3.10: Distribution of Kepler planets by both insolation and planetary radius plot, with underlying occurrence distributions adapted from Martinez et al. (2019). TOI-755b is represented as a purple star. The best fit positions of radius valleys are plotted in blue from Martinez et al. (2019). The Figure is taken from Osborn et al. (2021a).

3.2.2.7 TOI-431/HIP 26013: a super-Earth and a sub-Neptune transiting a bright, early K dwarf, with a third RV planet: *Osborn, A., et al. 2021, MNRAS, 2782, 507*

In this work Osborn et al. (2021) presented the discovery of three new planets around the TOI-431 system. TOI-431b and TOI-431d are a super-Earth and a sub-Neptune characterized with photometry from TESS and NGTS and RVs from HARPS, FEROS, Minerva-Australis, and iSHELL. There were 124 people involved in this work, and my position in the list was

the 18th. The inner planet has a ultra-short period of 0.49 days, mass of 3.07±0.35$M_\oplus$, and a radius of 1.28±0.04$R_\oplus$. TOI-431d has a period of 12.46 days, mass of 9.9±1.5$M_\oplus$, and a radius 3.29±0.09$R_\oplus$. TOI-431c is found in the HARPS RVs with a period of 4.84 days and a minium mass of $2.83^{+0.41}_{-0.34}M_\oplus$, but it is not seen to transit. Figure 3.11 shows the location of TOI-431b,d in the M-R diagram against the observed exoplanet population and some composition lines. We run our interior models using the stellar Fe/SI and Mg/Si as a proxy for the planet, and find that TOI-431b has a negligible H-He envelope. The larger companion TOI-431d is expected to have a significant volatile layer of H-He and/or water or about 3.6% or 33% of its total mass, respectively. However, the nature of the volatile is degenerate.

The architecture is this system is intriguing since it is unusual that the middle planet is not transiting, while the inner and outer planets are both seen to transit. It is also a good target for studying planetary evolution, since both TOI-431b and TOI-431d reside in either side of the radius-period valley described in Fulton et al. (2017). We find that the negligible envelope of TOI-431b is likely due to substantial atmospheric loss via photoevaporation. TOI-431d is found to have suffered a much less intense mass loss rate, allowing it to retain a substantial envelope.

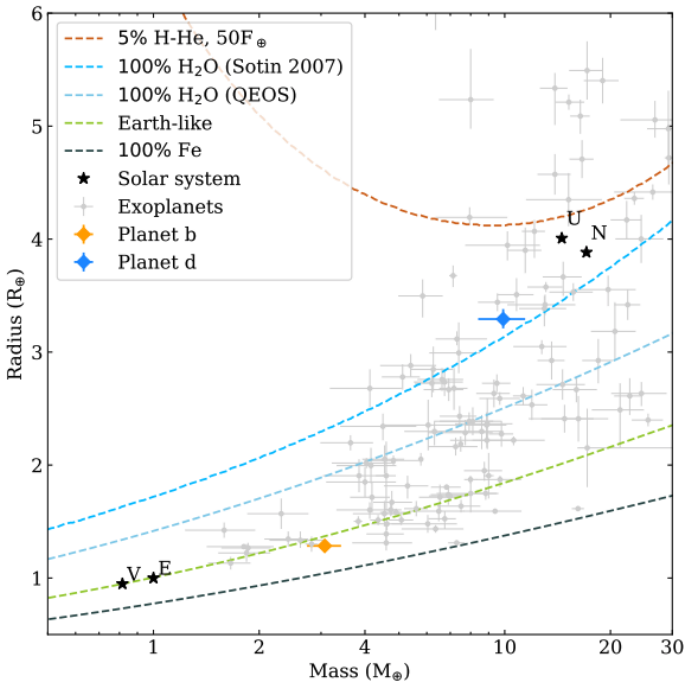

Figure 3.11: M-R diagram of known planets with mass determination better than 4σ from the NASA Exoplanet Archive. TOI-431b (orange) and d (blue) are represented by diamonds, and the black stars represent Solar System planets. The Figure is taken from Osborn et al. (2021)

3.2.2.8 TOI-269 b, an eccentric sub-Neptune transiting a M2 dwarf revisited with ExTrA: *Cointepas M. et al. 2021, A&A, A145, 650*

In this study Cointepas et al. (2021) presented the discovery and characterization of a sub-Neptune transiting the M dwarf TOI-269. My contribution was highly appreciated and awarded with the 7th position out of 62 co-authors. The mass and radius were determined using photometric data from TESS and ground-based follow-up from ExTra, and RVs from HARPS. We obtained an orbital period of 3.698 days, mass of $8.8\pm1.4 M_{\oplus}$ and radius of $2.77\pm0.12 R_{\oplus}$. In addition, we obtained a remarkable eccentricity of 0.43 ± 0.08. As shown in Figure 3.12 the measured eccentricity is almost the highest value for planets with periods shorter than 10 days. Since TOI-269 is likely few billion years old, TOI-269b probably did not acquire its eccentricity recently. The most likely explanation is that it reached its present orbit with planet-planet migration and reached this high eccentricity in the process.

The estimate bulk density of $2.28^{+0.48}_{-0.42}$ g cm^{-3} is significantly lower than the typical density of rock planets, and indicates the presence of a volatile envelope. Our internal models suggest that, although strongly degenerated, TOI-269b has a H-He mass fraction between 0.4% and 4%.

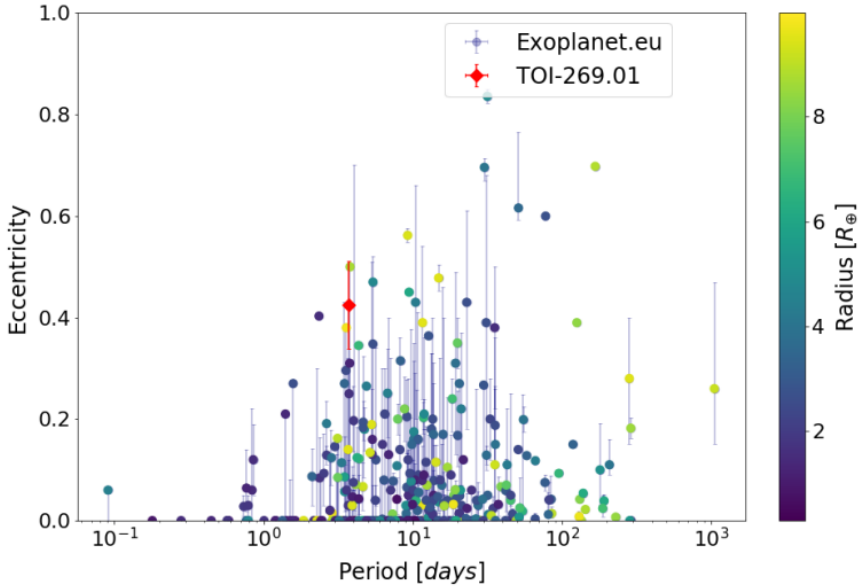

Figure 3.12: Orbital eccentricity against the orbital period from the Exoplanet.eu catalog. The color indicates the planetary radius. The Figure is taken from Cointepas et al. (2021)

3.2.2.9 The HD 137496 system: A dense, hot super-Mercury and a cold Jupiter: *Silva T. A. et al. 2022, Submitted in A&A*

In Silva et al. (2021) we report the discovery giant planet together with an inner super-Mercury orbiting around HD 137496b. My contribution to the interior modeling was highly appreciated and properly awarded with the 5th position out of 34 co-authors. With photometry from K2 and RVs from HARPS and CORALIE we find that the inner planet has a mass of $4.04\pm0.55 M_{\oplus}$ and a radius of $1.31\pm0.05 R_{\oplus}$. This results in an unusually high density of $10.49^{+2.08}_{-1.82}$ g cm^{-3}. Our internal models suggest that HD 137496b needs to have a large iron core, and estimate iron mass fraction (IMF) of 0.73 ± 0.12. If we assume that the ratios of Fe/Si and Mg/Si measured in the host star reflect the planetary ratios, we are not able to match the observed mass and radius of the star. This is shown in Figure 3.13, where we compare the planet's inferrend iron mass fraction (IMF_{planet}) to its stellar counterpart (IMF_{star}). Clearly, the IMF_{planet} is similar to Mercury and the 2σ regions of both IMF_{planet} and IMF_{star} do not overlap. This means that the probability of the planet reflecting the stellar refractory abundances is less than 0.04%.

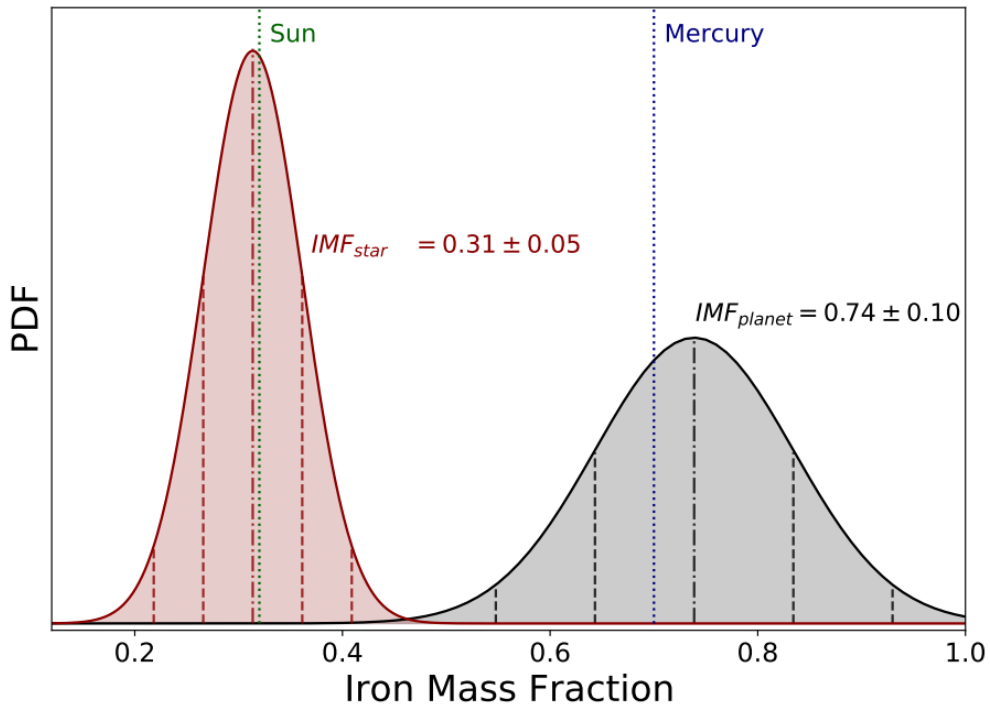

Figure 3.13: Posterior distribution for the estimated iron mass fractions as inferred from the stellar refractory abundances (red) and the planet mass and radii (gray). The Figure is taken from Silva et al. (2021).

HD 137496b is an interesting target for testing formation theories and density enhancing mechanisms. Phenomena that could totally or partly explain the observed density include a giant impact, photophoresis (Wurm et al. 2013), rocklines (Aguichine et al. 2020), magnetic erosion (Hubbard 2014), or magnetic boosts (Kruss & Wurm 2018, 2020).

3.2.2.10 TOI-177, a bright M-dwarf transited by a mini-Neptune amenable to atmospheric characterization: *Almenara J.M. et al. 2022, Submitted in A&A*

We confirm TOI-177b, which is a mini-Neptune in a 2.9 day orbit transiting a M dwarf. My position in the co-author list currently is the 4th out of 41. We model the transit light-curves from TESS and HARPS RVs and obtain a planetary mass of 3.34±0.72$M_\oplus$ and radius of 2.13±0.11$R_\oplus$. TOI-177b lies in the transition between the populations of super-Earths and sub-Neptunes. Figure 3.14 shows that TOI-177 b is located in the upper part of the exoplanet envelope that starts at around 3 . Most of the exoplanets around M-dwarfs are located in the upper part of the envelope, which suggests that it could be due to the low incoming irradiation that allows them to keep most of the H-He atmospheres. TOI-177 b sits above the Earth-like curve, implying that the planet must hold some volatile-envelope that is likely to account for a significant percentage of the planetary radius.

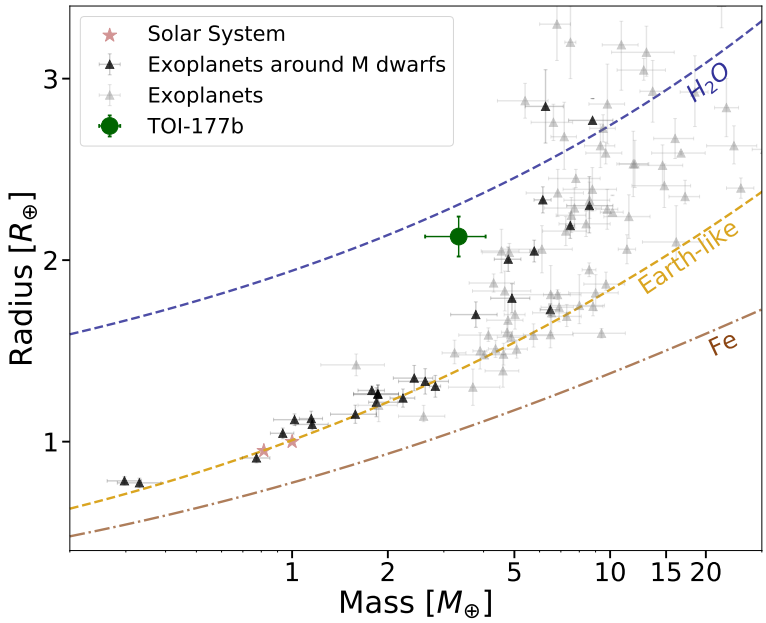

Figure 3.14: Mass-radius diagram of exoplanets from the PlanetS catalog. Exoplanets around M dwarfs are highlighted in black. Also shown are the composition lines of iron, Earth-like and pure water (at T_{eq}=693K).

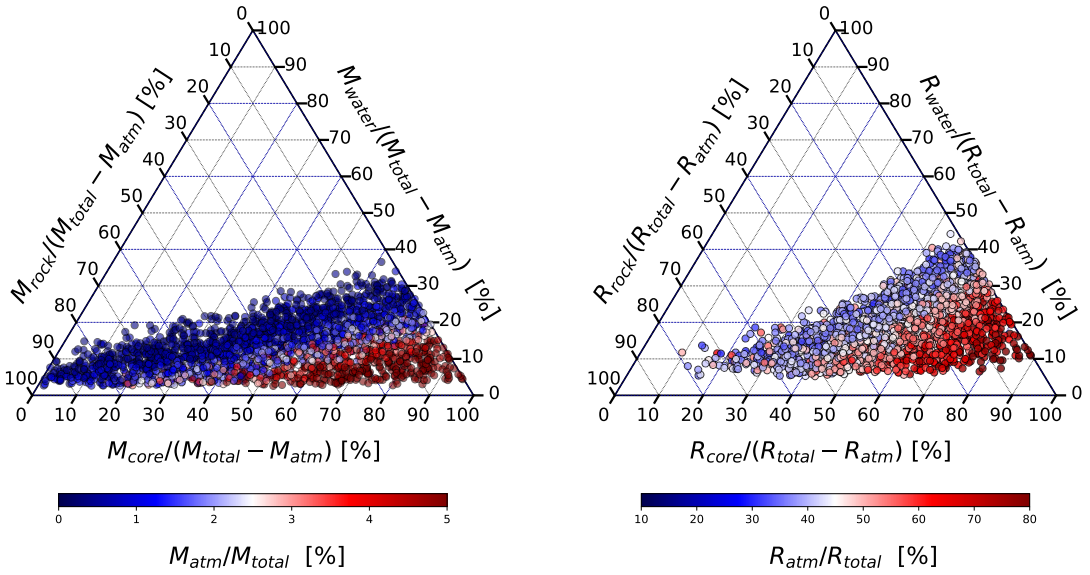

Figure 3.15: Ternary diagram of the inferred internal composition of TOI-177b. The parameter space covered by the posterior distributions in mass are shown on the left, and in radius on the right.

We use our interior models to study the composition of TOI-177b, and Figure 3.15 shows the inferred composition. The ternary diagram shows the degeneracy associated with the determination of the composition of exoplanets with measured mass and radius. We find a median H-He mass fraction of 1.9%, which to a lower-bound since enriched H-He atmospheres are more compressed and, therefore, increase the planetary H-He mass fraction. Nevertheless, we find a large degeneracy between the core, silicate mantle, and water layer, and it does not allow accurate estimates of the masses of these constituents.

Interior models cannot distinguish between water and H-He as the source of low density material, so we also run a 3-layer model without the H_2O and H-He envelopes. In the water-free case we find that the planet is 4.3% H-He, 45% iron and 51% rock by mass, which set maximum limits for the atmospheric and core mass. However, a model without H-He would require 57% water. Another possible scenario would be an iron-poor formation. An iron-free model gives a median atmospheric mass fraction of 0.9%, mantle of 84% and water-layer of 15%. Since any iron added to this model would increase the H-He mass fraction and decrease the mantle mass fraction, these values set an minimum and maximum limit respectively.

3.2.2.11 Final thoughts

Figure 3.16 shows the M-R diagram with the characterized exoplanets. Our interior models have allowed to set useful constraints on the interior structure of the planets of these ten systems. Given planetary mass and radius, we are able to set relatively good estimates of maximum amount of volatiles. However, in most of the cases we get very high uncertainties for the core, mantle, and water masses, due to their strong degeneracy. Overall, we find that a significant degeneracy in the internal structure of most of the sub-Neptune population remains. The planets that we can best characterize are the ones with either high densities (like HD 137496b and TOI-431b) or low densities (like TOI-431d or TOI-125c). In some cases we have used the stellar abundance ratios as proxy for the bulk abundances of the planet (Dorn et al. 2015; Adibekyan et al. 2021). However, there is evidence that the composition of the rocky planets spans a wider range than stars (Plotnykov & Valencia 2021), and whether the stellar abundances actually reflect the planetary ones is still under debate. Therefore, it is crucial to search for additional constraints to reduce the degeneracy of internal structures leading to identical mass and radius. Nowadays the atmospheric composition is assumed or modelled, but future space telescopes as JWST or ARIEL will put further constraints that will help reducing the degeneracy.

The models used in the characterization of the exoplanets have been rather simple. The core has been assumed to be composed only of iron, neglecting lighter elements, and for the mantle Earth-like composition has been assumed. Shah et al. (2021) have recently used reliable prescriptions for the hydration of the core and mantle showing that it may have a significant effect on the observed mass and radius. In addition, Mousis et al. (2020) presented a water model including a supercritical state that may play an important role for highly irradiated planets, which is not included in our models. The addition of these mentioned complexities to our internal models would make it more realistic, but they would add more free parameters to solve in the inference method. We therefore need to keep a balance between having a realistic planetary model while keeping a number of free parameters that allow a good determination.

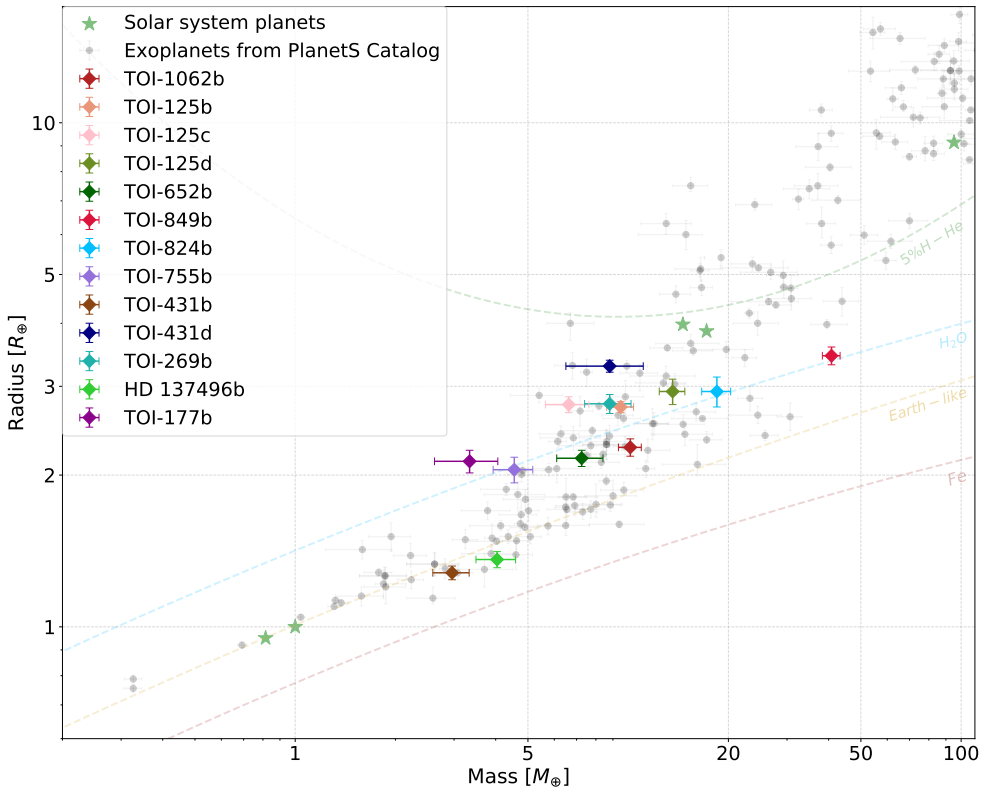

Figure 3.16: Mass-Radius diagram with the exoplanets for which we have studied the internal structure. The gray points show the exoplanets in the PlanetS Catalog, and the green star shaped markers correspond to the Solar System planets. We also display composition lines of pure-iron (brown), Earth-like (light-brown), pure-water (blue), and 5% H-He.

The choices on the theoretical assumptions of our internal models have a crucial importance. In Section 3.1.1 we have seen that the choice of the EOS of the core has an impact on the planetary radius of the order of few percent, and a similar impact was determined for the water EOS in Haldemann et al. (2020). Interior models usually make strong assumptions to reduce the number of free parameters, introducing an uncertainty that often is not taken into consideration. The observational uncertainties are well quantified, and inference models usually propagate the observational uncertainties to the uncertainties in the internal parameters, ignoring the contribution related to the model assumptions. Since observational uncertainties are expected to decrease significantly in the upcoming years, it is important to explore whether the theoretical uncertainties related to the model assumptions could dominate compared to the observational uncertainties. Along with the great efforts to improve the data, I consider that similar efforts should be made on the theoretical front.

3.2.3 On the formation of super-Mercuries by Giant Impacts

Another paper in preparation I am involved is leaded by C. Reinhardt, and studies the formation of iron-rich planets. One of the most popular explanations to Mercury's unusually large iron core is that a giant impact early in the planet's history stripped a large fraction of its mantle (Benz et al. 2007; Asphaug & Reufer 2014; Chau et al. 2018). Several metal-rich exoplanets have been discovered with extreme compositions similar to Mercury (Rowe et al. 2014; Sinukoff et al. 2016; Adams et al. 2017; Kosiarek et al. 2019; Santerne et al. 2018; Astudillo-Defru et al. 2020; Silva et al. 2021). These planets have robust mass and radius measurements, and meet the selection criteria for the PlanetS catalog except the limit in radius uncertainty (in this work we allow radius uncertainty up to 10% instead of the 8% limit of the PlanetS catalog). All these planets are very close to the star, and therefore high velocity impacts during the formation and early evolution may be common.

In the context of super-Earth, Marcus et al. (2009) investigated mantle stripping giant impact using Smoothed Particle Hydrodynamics and derived scaling laws for the mass of the largest fragment as well as the iron mass fraction. Using these scaling laws Marcus et al. (2010) then derive a maximum iron content for observed exoplanets assuming a maximum possible impact velocity of 80km/s. Based on the assumption that the initial pre-impact mass of a super-Earth does not exceed $10M_{\oplus}$ they predict that Mercury-like exo-planets have a maximum mass of $5M_{\oplus}$. However, several more massive, close-in metal-rich planets have been observed (Rowe et al. 2014; Sinukoff et al. 2016; Kosiarek et al. 2019; Astudillo-Defru et al. 2020; Silva et al. 2021). Due to their proximity to the central star their orbital velocities exceed 80km/s, allowing much larger collision velocities than assumed. In addition, we consider that the maximum mass of a pre-impact super-Earth is underestimated in Marcus et al. (2010).

In this work we use 3D hydro simulations to investigate the mantle stripping in giant impacts between super-Earths of up to $19M_{\oplus}$. We derive new scaling laws for the mass of the largest fragment, its iron mass fraction and the critical specific impact energy for the catastrophic disruption. In addition, we use our interior models to estimate the iron mass fraction of the densest observed planets to date, and compare them with the output of the 3D hydro simulations. An example is shown in Figure 3.17, where the post-impact iron mass fraction is plotted against the fragment mass. The coloured circles represent the simulation data, and each colour corresponds to the planet's pre-impact mass. The grey shaded areas mark a restricted impact parameter space. Dark gray corresponds to pre-impact mass below $10M_{\oplus}$ and impact velocities between 20km/s and 100km/s, while the light gray envelope corresponds

to pre-impact mass below $20M_{\oplus}$ and impact velocities between 20km/s and 160km/s. The goal of the Figure is to analyze under with conditions the models can explain the observed data of the dense exoplanets, which is represented with colored errorbars. We see that using target masses of $10.5M_{\oplus}$ we are not able to explain the iron mass fraction of all the observed exoplanets through giant impacts, but it is possible with larger target masses. Similarly, we need to consider impact velocities up to ~160km/s in order to get iron mass fractions compatible with the observations.

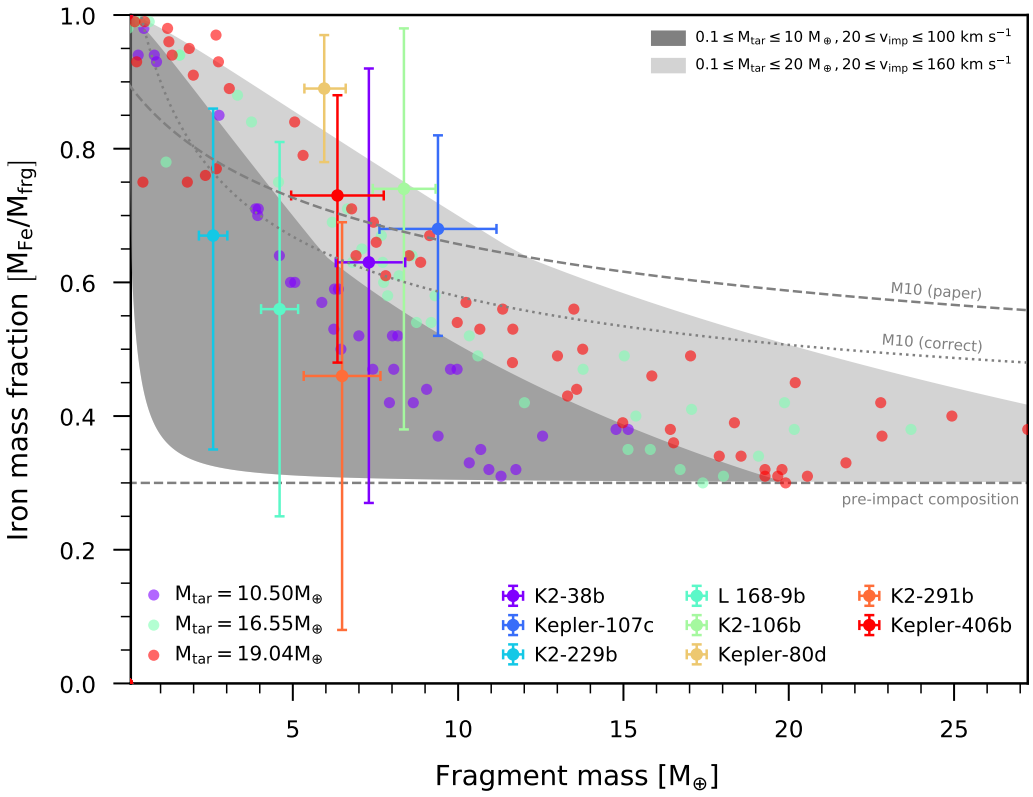

Figure 3.17: The post-impact iron mass fraction against the fragment mass. The circles represent the simulation data and each colour corresponds to the planet's pre-impact mass (violet: 10, cyan: 16.55 and red: 19.04). The dashed and dotted black curves correspond to the original and corrected curve for maximum stripping derived in Marcus et al. (2010). The grey shaded area mark a restricted impact parameter space. The circles with errorbars correspond to our sample of observed iron-rich planets.

Chapter 4

Conclusions

4.1 Main results of the thesis

During the four years of my Ph.D. project, the large number and diversity of discovered exoplanets has allowed to identify and study multiple planet populations. My research has focused on how to characterize the structure of these exoplanets, especially super-Earths, sub-Neptunes, and planets in the transition from terrestrial to gas giants.

A significant part of my work has been dedicated to the investigation of the exoplanet demography, which is key to understand the physical processes affecting planet formation and evolution theories. Demographic analyses heavily depend on the chosen planetary sample, and the main exoplanet catalogs available are built using different criteria and information sources. The resulting M-R diagram, for instance, significantly depends on the chosen exoplanet catalog (e.g. Bashi et al. 2018). In the first study presented in Chapter 2.1 we present an updated exoplanet catalog based on robust mass and radius measurements. Using the NASA Exoplanet Archive as a starting point, we performed a careful analysis of all the planets less massive than $120 M_\oplus$ in order to build as reliable as possible exoplanet catalog. The resulting M-R diagram from the revisited catalog clearly showed two distinct populations, corresponding to rocky exoplanets and volatile-rich exoplanets. We found that the rocky exoplanet populations shows a relatively small density variability and ends at a mass of ~25 $M_\oplus$, possibly indicating the maximum core mass that can be formed. In this work we also presented new empirical M-R relations, and used the composition line of pure water to divide the rocky and volatile-rich

regimes. We show the limitations of other M-R relations in the literature to properly describe the rocky and volatile-rich populations. After the publication of this study we have continued adding the new discovered planet in the revisited catalog and extended it to the gas giant regime.

The large diversity observed in exoplanets also extends to the architecture of multi-planetary systems. Understanding the properties of planets within a system provides hints on the underlying physical processes. The so-called 'peas in the pod' trend has been extensively explored in terms of radius and period ratios, but it was still poorly understood in terms of mass. In Chapter 2.2 we used our revisited reliable exoplanet catalog to perform an in-depth analysis of the similarity in mass, radius and density of multi-planetary systems. We find that planetary systems tend to be more similar in radius than in mass, which could be linked to the stronger impact of the radios on the density and, hence, the planetary composition. We also find a sharp transition in the similarity of systems at $100M_{\oplus}$ and $10R_{\oplus}$, with planets below these limits being significantly more uniform. Finally, although dependence of the similarity with other quantities must be further investigated, we also find some hints that systems around more massive stars tend to be less uniform in mass and radius. However, further studies with more observational data are needed to confirm these trends.

In addition to demographic studies, another substantial part of my work has been focused in the characterization of the internal structure of individual planets, focusing in super-Earths and sub-Neptunes. Building on the work of Dorn et al. (2017), I updated several parts of the core, water and volatile envelope models, as described in Chapter 3.1. This model was used to characterize the internal structure of several discovered exoplanets, including super-Mercuries (Silva et al. 2021), super-Earths (Teske et al. 2020; Osborn et al. 2021), sub-Neptunes (Nielsen et al. 2020a; Burt et al. 2020b; Osborn et al. 2021b; Cointepas et al. 2021) and a very massive core (Armstrong et al. 2020) (see Figure 3.16). In addition, we used our internal models to study several aspects that affect the characterization of super-Earths and sub-Neptunes, such as observational uncertainties, location on the M-R diagram, additional constraints as bulk abundances, or model assumptions. We find that a decrease in observational uncertainties for planets below the Earth-like composition line and above the pure-water composition line leads to a better determination of the core mass and atmospheric mass, respectively. Between these two composition lines the degeneracy is stronger and mass and radius alone do not significantly improve the determination of any internal structure parameter. We also find that relatively small uncertainties on the temperature profiles of the planet lead to significant uncertainties in the observed radius, and we emphasize importance of the theoretical uncertainties related to model assumptions.

Finally, I have also been involved in the target selection, observation, and data analysis of several RV follow-up programs. I have built a program for CORALIE which aims to perform long-term follow-up of planets in the Neptunian desert. This way we aim to find distant companions and/or improve the RV fit and refine the planetary parameters of the observed planets. I have also contributed to the NCORES program for HARPS, which aims to study planets subjected to strong irradiations in which the outer layer has been stripped off by photoevaporation. I leaded the discovery paper of two sub-Neptunes using data from this program and TESS photometry. The RV data were affected by strong stellar activity, so we modeled it using GPs. We found that the inner planet is a mini-Neptune with a small volatile envelope of 0.35% of the mass at most. We also found a non-transiting planet on the RV data in close 2:1 motion resonance with its inner companion. Finally, I have also contributed to the target selection and development of the scientific case of two ESPRESSO proposals aiming to characterize warm mini-Neptunes and mini-Neptunes around M-dwarfs. These two ESPRESSO proposals have been accepted for period 109 (April-September 2022) and will continue exploring two planetary populations that play a very important role in constraining planetary formation and evolution models and, yet, are still poorly understood.

4.2 Future Prospects

The key objective of this thesis was to provide a better understanding of the internal structure of super-Earths and sub-Neptunes. In recent years, our understanding of these planets has increased significantly thanks to the increasing quantity and quality of exoplanet data, but there are many open questions that have yet to be resolved.

The discovery of new exoplanets and the development of theoretical models describing the planetary interior will certainly allow to improve characterization of exoplanet in the near future. The exoplanet demography is often studied in the mass-radius or orbital period-radius planes. However, in order to reach a more complete understanding of the exoplanet demography it is necessary to study the multi-layer nature of the M-R diagram. The amount of exoplanets with well constrained mass and radius has hugely increased thanks to Kepler and TESS, allowing to study the correlations between planetary mass, radius and other quantities. A preliminary research on this has been presented in Section 2.1.3, but more data is needed in order to extract robust conclusions on the relation between planetary and stellar properties. For instance, thanks to recent TESS data we have observed some hints indicating that M-dwarfs may host more rocky and denser planets than FGK stars. This trend is against sev-

eral theoretical predictions, and will soon be confirmed or discarded with new observational data. Ongoing and future space missions like TESS, CHEOPS, or PLATO will rapidly populate the M-R diagram allowing a more complete understanding of exoplanetary demographics.

Precise constraints on the planetary mass and radius are crucial to accurately infer the interior composition. Despite the intrinsic degeneracy of compositions leading to the same mass and radii, in Otegi et al. (2020b) we showed that the interiors of volatile-rich and very dense planets is better constrained with decreasing observational uncertainties. We have a significant amount of time on ESPRESSO for the period 109 to observe long-period, low irradiated planets and M dwarfs to characterize warm mini-Neptune, ocean, icy planets, and/or planets in the habitable zone. These new observations will help understading the nature and origin of these highly interesting exoplanet populations. In addition, new generation instruments like Near-InfraRed Planet Searcher (NIRPS, Bouchy et al. 2017), which is currently under commissioning, will continue playing a relevant role in the characterization of small planets. It will target M-dwarfs, which are challenging to observe due to their intrinsic faintness and high activity. Nevertheless, NIRPS is expected reach a precision of 1 m s^{-1}, comparable to that of HARPS, and will be assisted by adaptive optics. M-dwarfs are the most common type of stars in the solar neighborhood (Henry et al. 2018), and tend to host more rocky planets than FGK stars (Cloutier & Menou 2020). Precise characterization of dense rocky planets will allow a better understanding of its interior structures and composition. In addition, the PLATO mission (Rauer et al. 2014) is set to launch in 2026 and will monitor 250000 bright stars with unprecedented precision. Its goal will be to determine planetary radii with precision of 3%, and masses up to 10%. The large amount of exoplanets with precise mass and radius measurements will allow in-depth analysis of their internal internal structure and their formation and evolution paths.

To further improve the quality of the characterization of exoplanets it is necessary to search for additional constraints on their composition. These constraints will by provided by space telescopes like JWST or ARIEL, which will study their atmospheric composition. In addition, PLATO will derive the age of the systems, adding an extra parameter can be crucial to further constrain the composition and formation paths of the planets, and, therefore, opening a new dimension for planet characterization. For instance, including exoplanet evolution models to the characterization and knowing the age of the system would allow to check whether the envelope mass could be retained over the planet's lifetime, and exclude unphysical configurations. It is also necessary to further investigate whether the stellar abundances are a good proxy for the planet bulk abundances. Even if this assumption can play an important role reducing the degeneracy, the validity of the assumption is still under debate (Adibekyan et al.

2021; Plotnykov & Valencia 2021) and should be further investigated. Additional constraints may eventually also come from galactic kinematics, since it has been found that the amount of small rocky planets depends on the velocity of the stellar total velocity (Bashi & Zucker 2019).

The interplay between the data coming from the new generation instruments and the development of theoretical models is crucial to advance in the understanding of planetary interiors. Observational data will set constraints on the theoretical models, while theoretical models will drive future observations. My research represents another step in advancing our understanding of planets beyond the Solar System.

LIST OF PUBLICATIONS

During my PhD I have contributed to the following papers:

- "*A Jovian planet in an eccentric 11.5 day orbit around HD 1397 discovered by TESS*".
  A&A, 623, A100 (2019)
  Nielsen, Bouchy, Turner, Giles, Mascareño, Lovis, Marmier, Pepe, Ségransan, Udry, **Otegi**, Ottoni, Stalport, Ricker, Vanderspek, Latham, Seager, Winn, Jenkins, Kane, Wittenmyer, Bowler, Crossfield, Horner, Kielkopf, Morton, Plavchan, Tinney et al.

- "*Revisited mass-radius relations for exoplanets below 120 $M_\oplus$*".
  A&A, 634, A43 (2020)
  Otegi, Bouchy, Helled.

- "*Mass determinations of the three mini-Neptunes transiting TOI-125* ".
  MNRAS, 492, 5399 (2020)
  Nielsen, Gandolfi, Armstrong, Jenkins, Fridlund, Santos, Dai, Adibekyan, Luque, Steffen, Esposito, Meru, Sabotta, Bolmont, Kossakowski, **Otegi**, Murgas, Stalport, Rodler, Díaz, Kurtovic, Ricker, Vanderspek, Latham, Seager, Winn, Jenkins et al.

- "*Three short-period Jupiters from TESS. HIP 65Ab, TOI-157b, and TOI-169b* ".
  A&A, 639, A76 (2020)
  Nielsen, Brahm, Bouchy, Espinoza, Turner, Rappaport, Pearce, Ricker, Vanderspek, Latham, Seager, Winn, Jenkins, Acton, Bakos, Barclay, Barkaoui, Bhatti, Briceño, Bry-

ant, Burleigh, Ciardi, Collins, Collins, Cooke, Csubry, dos Santos, Eigmuller, Fausnaugh, Gan, Gillon, Goad, Guerrero, Hagelberg, Hart, Henning, Huang, Jehin, Jenkins, Jordan, Kielkopf, Lavia, Law, Lendl, Lovis, Mann, Marmier, McCormac, Mori, Moyano, Narita, Osip, **Otegi** et al.

- "*A remnant planetary core in the hot-Neptune desert*".
  Nature, 583, 39 (2020)
  Armstrong, Lopez, Adibekyan, Booth, Bryant, Collins, Deleuil, Emsenhuber, Huang, King, Lillo-Box, Lissauer, Matthews, Mousis, Nielsen, Osborn, **Otegi**, Santos, SOusa, Stassun, Veras, Ziegler, Acton, Almenara, Anderson, Barrado, Barros et al.

- "*Impact of the measured parameters of exoplanets on the inferred internal structure*".
  A&A, 640, A135 (2020)
  Otegi, Dorn, Helled, Bouchy, Haldemann, Alibert.

- "*TESS Reveals a Short-period Sub-Neptune Sibling (HD 86226c) to a Known Long-period Giant Planet*".
  AJ, 160, 96 (2020)
  Teske, Diaz, Luque, Mocnik, Seidel, **Otegi**, Feng, Jenkins, Palle, Segransan, Udry, Collins, Eastman, Ricker, Vanderspek, Latham, Seager, Winn, Jenkins, Anderson, Barclay, Bouchy, Burt, Butler Caldwell, Collins, Crane, Dorn Flowers et al.

- "*TOI-824 b: A New Planet on the Lower Edge of the Hot Neptune Desert*".
  AJ, 160, 153 (2020)
  Burt, Nielsen, Quinn, Mamajec, Matthews, Zhou, Seidel, Huang, Lopez, Soto, **Otegi**, Stassun, Kreidberg, Collins, Eastman, Rodriguez, Vanderburg, Halverson, Teske, Wang, Butler, Bouchy, Dumusque, Segransen, Shectman, Crane, Feng, Montet et al.

- "*A sub-Neptune and a non-transiting Neptune-mass companion unveiled by ESPRESSO around the bright late-F dwarf HD 5278 (TOI-130)*".
  A&A, 648, A75 (2021)
  Sozzetti, Damasso, Bonomo, Alibert, Sousa, Adibekyan, Zapatero, Gonzalez, Barros, Lillo-Box, Stassun, Winn, Cristiani, Pepe, Rebolo, Santos, Allart, Barclay, Bouchy, Cabral, Ciardi, Di Marcantonio, D'Odorico, Ehrenreich, Fasnaugh, Figueira, Haldemann, Jenkins, Latham, Lavie, Lo Curto, Lovis, Martins, Megevand, Mnehner, Micela, Molar, Nunes, Oshagh, **Otegi** et al.

- "*A hot mini-Neptune in the radius valley orbiting solar analogue HD 110113*".
  MNRAS, 502, 4842 (2021)
  Osborn, Armstrong, Adibekyan, Collins, Delgado-Mena, Howell, Hellier, King, Lillo-Box, Nielsen, **Otegi**, Santos, Ziegler, Anderson, Briceno, Burke, Bayliss, Barrado, Bryant, Brown, Barros, Bouchy, Caldwell, Conti, Diaz, Dragomir, Deleuil et al.

- "*TOI-269 b: an eccentric sub-Neptune transiting a M2 dwarf revisited with ExTrA*".
  A&A, 650, A145 (2021)
  Cointepas, Almenara, Bonfils, Bouchy, Astudillo-Defru, Murgas, **Otegi**, Wyttenbach, Anderson, Artigau, Canto Martins, Charbonneau, Collins, Correia, Curaba, Delboulbe, Delfosse, Diaz, Dorn, Doyon, Feautrier, Figueira, Forveille, Gaisne et al.

- "*TOI-220 b: a warm sub-Neptune discovered by TESS*".
  MNRAS, 505, 3361 (2021)
  Hoyer, Gandolfi, Armstrong, Deleuil, Acuna, de Medeiros, Goffo, Lillo-Box, Delgado Mena, Lopez, Santerne, Sousa, Fridlund, Adibekyan, Collins, Serrano, Cortes, Howell, Deeg, Aguichine, Barragan, Bryant, Canto, Cooke, Diaz, Esposito, Furlan, Hojjatpanah, Jackman, Jenkins, Jensen, Latham, Leao Matson, Nielsen, Osborn **Otegi** et al.

- "*TESS and HARPS reveal two sub-Neptunes around TOI 1062*".
  A&A, 653, A105 (2021)
  Otegi, Bouchy, Helled, Armstrong, Stalport, Psaridi, Delisle, Stassun, Delgado-Mena, Santos, Hara, Collins, Gandhi, Dorn, Brogi, Fridlund, Osborn, Hoyer, Udry, Hojjatpanah, Nielsen, Dumusque, Adibekyan, Conti, Schwarz, Wang, Figueira et al.

- "*TOI-431/HIP 26013: a super-Earth and a sub-Neptune transiting a bright, early K dwarf, with a third RV planet*".
  MNRAS, 507, 2782 (2021)
  Osborn, Armstrong, Cale, Brahm, Wittenmyer, Dai, Crossfield, Bryant, Adibekyan, Cloutier, Collins, Delgado Mena, Fridlund, Hellier, Howell, King, Lillo-Box, **Otegi**, Sousa, Stassun, Matthews, Ziegler, Ricker, Vanderspek, Latham, Seage, Winn et al.

- "*The HD 137496 system: A dense, hot super-Mercury and a cold Jupiter* ".
  A&A, 657, A68 (2022)
  Silva, Demangeon, Barros, Armstrong, **Otegi**, Bossini, Delgado Mena, Sousa, Adibekyan, Nielsen, Dorn, Lillo-Box, Santos, Hoyer, Stassun, Almenara, Bayliss, Barrado,

Boisse, Brown, Diaz, Dumusque, Figueira, Hadjigeorghiou, Hojjatpanah, Mousis, Osborn, Santerne, Strom, Udry, Wheatley.

- "*The similarity of multi-planet systems*".
  Accepted in A&A (2022)
  Otegi, Helled, Bouchy

Bibliography

Adams, E. R., Jackson, B., Endl, M., et al. 2017, , 153, 82

Adibekyan, V., Dorn, C., Sousa, S. G., et al. 2021, arXiv e-prints, arXiv:2102.12444

Agol, E., Dorn, C., Grimm, S. L., et al. 2021, , 2, 1

Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, , 359, 567

Aguichine, A., Mousis, O., Devouard, B., & Ronnet, T. 2020, , 901, 97

Aigrain, S., Pont, F., & Zucker, S. 2012, , 419, 3147

Alibert, Y. & Benz, W. 2017, Astronomy Astrophysics, 598, L5

Alibert, Y., thiabaud, a., marboeuf, u., et al. 2015, in AAS/Division for Extreme Solar Systems Abstracts, Vol. 47, AAS/Division for Extreme Solar Systems Abstracts, 115.02

Anderson, K. R., Storch, N. I., & Lai, D. 2016, , 456, 3671

Andrews, S. M. 2020, , 58, 483

Andrews, S. M., Huang, J., Pérez, L. M., et al. 2018, The Messenger, 174, 19

Armstrong, D. J., Lopez, T. A., Adibekyan, V., et al. 2020, Nature, 583, 39–42

Asphaug, E. & Reufer, A. 2014, Nature Geoscience, 7, 564

Astudillo-Defru, N., Cloutier, R., Wang, S. X., et al. 2020, , 636, A58

Atreya, S. K., Mahaffy, P. R., Niemann, H. B., Wong, M. H., & Owen, T. C. 2003, , 51, 105

Auclair-Desrotour, P., Leconte, J., Bolmont, E., & Mathis, S. 2019, in EPSC-DPS Joint Meeting 2019, Vol. 2019, EPSC–DPS2019–1352

Auvergne, M., Bodin, P., Boisnard, L., et al. 2009, , 506, 411

Baglin, A. 2003, Advances in Space Research, 31, 345

Bakos, G., Noyes, R. W., Kovács, G., et al. 2004, , 116, 266

Ballard, S. & Johnson, J. A. 2016, , 816, 66

Baraffe, I., Alibert, Y., Chabrier, G., & Benz, W. 2006, Astronomy Astrophysics, 450, 1221–1229

Bashi, D., Helled, R., & Zucker, S. 2018, Geosciences, 8, 325

Bashi, D. & Zucker, S. 2019, , 158, 61

Basilevsky, A. T. & Head, J. W. 2002, Geology, 30, 1015

Beaugé, C. & Nesvorný, D. 2013, , 763, 12

Becker, A., Lorenzen, W., Fortney, J. J., et al. 2014, , 215, 21

Benz, W., Anic, A., Horner, J., & Whitby, J. A. 2007, Space Science Reviews, 132, 189

Bollengier, O., Brown, J. M., & Shaw, G. H. 2019, , 151, 054501

Bolmont, E., Breton, S. N., Tobie, G., et al. 2020, , 644, A165

Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

Bouchet, J., Mazevet, S., Morard, G., Guyot, F., & Musella, R. 2013, , 87, 094102

Bouchy, F., Doyon, R., Artigau, É., et al. 2017, The Messenger, 169, 21

Bower, D. J., Kitzmann, D., Wolf, A. S., et al. 2019, , 631, A103

Brugger, B., Mousis, O., Deleuil, M., & Deschamps, F. 2017, , 850, 93

Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, , 564, A125

Burn, R., Schlecker, M., Mordasini, C., et al. 2021, arXiv e-prints, arXiv:2105.04596

Burt, J. A., Dragomir, D., Mollière, P., et al. 2021, , 162, 87

Burt, J. A., Nielsen, L. D., Quinn, S. N., et al. 2020a, TOI-824 b: A New Planet on the Lower Edge of the Hot Neptune Desert

Burt, J. A., Nielsen, L. D., Quinn, S. N., et al. 2020b, TOI-824 b: A New Planet on the Lower Edge of the Hot Neptune Desert

Caillabet, L., Canaud, B., Salin, G., Mazevet, S., & Loubeyre, P. 2011, , 107, 115004

Carpenter, B., Gelman, A., Hoffman, M. D., et al. 2017, Journal of Statistical Software, 76, 1–32

Carter, P. J., Leinhardt, Z. M., Elliott, T., Walter, M. J., & Stewart, S. T. 2015, , 813, 72

Carter, T. & Humi, M. 2012, Celestial Mechanics and Dynamical Astronomy, 112, 385

Chabrier, G., Mazevet, S., & Soubiran, F. 2019, The Astrophysical Journal, 872, 51

Chabrier, G. & Potekhin, A. Y. 1998, Physical Review E, 58, 4941–4949

Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, , 529, L45

Chau, A., Reinhardt, C., Helled, R., & Stadel, J. 2018, The Astrophysical Journal, 865, 35

Chen, H. & Rogers, L. A. 2016, , 831, 180

Chen, J. & Kipping, D. 2017, , 834, 17

Cloutier, R., Charbonneau, D., Deming, D., Bonfils, X., & Astudillo-Defru, N. 2021, , 162, 174

Cloutier, R. & Menou, K. 2020, The Astronomical Journal, 159, 211

Cointepas, M., Almenara, J. M., Bonfils, X., et al. 2021, , 650, A145

Connolly, J. A. D. 2009, Geochemistry, Geophysics, Geosystems, 10, Q10014

Courcol, B., Bouchy, F., & Deleuil, M. 2016, , 461, 1841

Deeg, H. J. & Alonso, R. 2018, Transit Photometry as an Exoplanet Discovery Method, ed. H. J. Deeg & J. A. Belmonte, 117

Dorn, C., Khan, A., Heng, K., et al. 2015, , 577, A83

Dorn, C., Venturini, J., Khan, A., et al. 2017, , 597, A37

Dziewonski, A. M. & Anderson, D. L. 1981, Physics of the Earth and Planetary Interiors, 25, 297

Díaz, M. R., Jenkins, J. S., Gandolfi, D., et al. 2020, Monthly Notices of the Royal Astronomical Society, 493, 973–985

Elkins-Tanton, L. T. & Seager, S. 2008, , 688, 628

Espinoza, N., Brahm, R., Henning, T., et al. 2020, , 491, 2982

Esposito, M., Armstrong, D. J., Gandolfi, D., et al. 2019, , 623, A165

Fabrycky, D. C., Lissauer, J. J., Ragozzine, D., et al. 2014, , 790, 146

Fegley, Bruce, J. & Lodders, K. 1994, , 110, 117

Feistel, R. & Wagner, W. 2006, Journal of Physical and Chemical Reference Data, 35, 1021

Fletcher, L. N., Drossart, P., Burgdorf, M., Orton, G. S., & Encrenaz, T. 2010, , 514, A17

Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008a, , 678, 1419

Fortney, J. J., Marley, M. S., Saumon, D., & Lodders, K. 2008b, , 683, 1104

Freedman, R. S., Lustig-Yaeger, J., Fortney, J. J., et al. 2014, The Astrophysical Journal Supplement Series, 214, 25

French, M. & Redmer, R. 2015, , 91, 014308

Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, The Astronomical Journal, 154, 109

Gordon, S., McBride, B., & Zeleznik, F. J. 1984, Computer program for calculation of complex chemical equilibrium compositions and applications. Supplement 1: Transport properties

Grasset, O., Castillo-Rogez, J., Guillot, T., Fletcher, L. N., & Tosi, F. 2017, , 212, 835

Grasset, O., Schneider, J., & Sotin, C. 2009, , 693, 722

Guillot, T. 2010, , 520, A27

Gupta, A. & Schlichting, H. E. 2019, , 487, 24

Hadden, S. & Lithwick, Y. 2017, , 154, 5

Hakim, K., Rivoldini, A., Van Hoolst, T., et al. 2018, , 313, 61

Haldemann, J., Alibert, Y., Mordasini, C., & Benz, W. 2020, Astronomy Astrophysics, 643, A105

Hastings, W. K. 1970, Biometrika, 57, 97

Hatzes, A. P. & Rauer, H. 2015, The Astrophysical Journal, 810, L25

Haywood, R. D., Collier Cameron, A., Queloz, D., et al. 2014, , 443, 2517

Helled, R., Bodenheimer, P., Podolak, M., et al. 2014, in Protostars and Planets VI, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 643

Heng, K. & Lyons, J. R. 2016, , 817, 149

Henry, T. J., Jao, W.-C., Winters, J. G., et al. 2018, , 155, 265

Howell, S. B., Sobeck, C., Haas, M., et al. 2014a, , 126, 398

Howell, S. B., Sobeck, C., Haas, M., et al. 2014b, , 126, 398

Hoyer, S., Gandolfi, D., Armstrong, D. J., et al. 2021, , 505, 3361

Hubbard, A. 2014, , 241, 329

Ichikawa, H. & Tsuchiya, T. 2020, Minerals, 10, 59

Ida, S. & Lin, D. N. C. 2008, , 673, 487

Ikoma, M., Nakazawa, K., & Emori, H. 2000, , 537, 1013

Jackson, B., Miller, N., Barnes, R., et al. 2010, Monthly Notices of the Royal Astronomical Society, 407, 910–922

Jiang, C.-F., Xie, J.-W., & Zhou, J.-L. 2020, , 160, 180

Jin, S. & Mordasini, C. 2018, , 853, 163

Jin, S. & Mordasini, C. 2018, The Astrophysical Journal, 853, 163

Jin, S., Mordasini, C., Parmentier, V., et al. 2014, , 795, 65

Johnson, J. D. & Lyon, S. P. 1985, EOS for polycrystalline quartz

Journaux, B., Brown, J. M., Pakhomova, A., et al. 2020, Journal of Geophysical Research (Planets), 125, e06176

Keller, H. 1992

Kosiarek, M. R., Blunt, S., López-Morales, M., et al. 2019, , 157, 116

Kruss, M. & Wurm, G. 2018, , 869, 45

Kruss, M. & Wurm, G. 2020, , 1, 23

Kubyshkina, D. & Vidotto, A. A. 2021, Monthly Notices of the Royal Astronomical Society, 504, 2034–2050

Léger, A., Rouan, D., Schneider, J., et al. 2009, , 506, 287

Lissauer, J. J., Dawson, R. I., & Tremaine, S. 2014, , 513, 336

Lodders, K. 2003, , 591, 1220

Lodders, K. & Fegley, B. 2002a, , 155, 393

Lodders, K. & Fegley, B. 2002b, , 155, 393

Lodders, K., Palme, H., & Gail, H. P. 2009, Landolt Börnstein, 4B, 712

Lopez, E. D. & Fortney, J. J. 2013, , 776, 2

Lopez, E. D. & Fortney, J. J. 2013, The Astrophysical Journal, 776, 2

Lopez, E. D. & Fortney, J. J. 2014, , 792, 1

Lopez, E. D., Fortney, J. J., & Miller, N. 2012, , 761, 59

Lopez, E. D. & Rice, K. 2018, , 479, 5303

Lovis, C. & Fischer, D. 2010, Radial Velocity Techniques for Exoplanets, ed. S. Seager, 27–53

Lozovsky, M., Helled, R., Dorn, C., & Venturini, J. 2018, , 866, 49

Lucy, L. B. & Sweeney, M. A. 1971, , 76, 544

Madhusudhan, N. 2012, , 758, 36

Marcus, R. A., Sasselov, D., Hernquist, L., & Stewart, S. T. 2010, , 712, L73

Marcus, R. A., Stewart, S. T., Sasselov, D., & Hernquist, L. 2009, The Astrophysical Journal Letters, 700, L118

Marcy, G. W., Isaacson, H., Howard, A. W., et al. 2014, , 210, 20

Martinez, C. F., Cunha, K., Ghezzi, L., & Smith, V. V. 2019, , 875, 29

Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20

Mayor, M. & Queloz, D. 1995, , 378, 355

Mazeh, T., Holczer, T., & Faigler, S. 2016, Astronomy Astrophysics, 589, A75

Mazevet, S., Licari, A., Chabrier, G., & Potekhin, A. Y. 2019, , 621, A128

Mazzola, G., Helled, R., & Sorella, S. 2018, , 120, 025701

McDonald, G. D., Kreidberg, L., & Lopez, E. 2019, , 876, 22

McDonough, W. F. & Sun, S. s. 1995, Chemical Geology, 120, 223

Melosh, H. J. 2007, , 42, 2079

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, , 21, 1087

Miguel, Y., Guillot, T., & Fayon, L. 2016, VizieR Online Data Catalog, J/A+A/596/A114

Militzer, B. 2008, in APS April Meeting Abstracts, APS Meeting Abstracts, 11HE.00

Militzer, B. & Hubbard, W. B. 2013, in AGU Fall Meeting Abstracts, Vol. 2013, P11B–05

Militzer, B., Vorberger, J., & Hubbard, W. 2006, in AGU Fall Meeting Abstracts, Vol. 2006, P52A–03

Millholland, S., Wang, S., & Laughlin, G. 2017, , 849, L33

Min, M., Dullemond, C. P., Kama, M., & Dominik, C. 2011, , 212, 416

Mordasini, C. 2020, , 638, A52

Mordasini, C., Alibert, Y., & Benz, W. 2009, , 501, 1139

Mordasini, C., Alibert, Y., Klahr, H., & Henning, T. 2012, Astronomy Astrophysics, 547, A111

More, R. M., Warren, K. H., Young, D. A., & Zimmerman, G. B. 1988, Physics of Fluids, 31, 3059

Mousis, O., Deleuil, M., Aguichine, A., et al. 2020, , 896, L22

Moutou, C., Deleuil, M., Guillot, T., et al. 2013, , 226, 1625

Muñoz, D. J., Lai, D., & Liu, B. 2016, , 460, 1086

Müller, S., Helled, R., & Cumming, A. 2020, , 638, A121

Neil, A. R. & Rogers, L. A. 2020, , 891, 12

Nettelmann, N., Becker, A., Holst, B., & Redmer, R. 2012, , 750, 52

Nettelmann, N., Holst, B., Kietzmann, A., et al. 2008, , 683, 1217

Nielsen, L. D., Gandolfi, D., Armstrong, D. J., et al. 2020a, , 492, 5399

Nielsen, L. D., Gandolfi, D., Armstrong, D. J., et al. 2020b, , 492, 5399

Osborn, A., Armstrong, D. J., Cale, B., et al. 2021, Monthly Notices of the Royal Astronomical Society, 507, 2782–2803

Osborn, H. P., Armstrong, D. J., Adibekyan, V., et al. 2021a, , 502, 4842

Osborn, H. P., Armstrong, D. J., Adibekyan, V., et al. 2021b, , 502, 4842

Otegi, J. F., Bouchy, F., & Helled, R. 2020a, , 634, A43

Otegi, J. F., Bouchy, F., Helled, R., et al. 2021a, , 653, A105

Otegi, J. F., Dorn, C., Helled, R., et al. 2020b, , 640, A135

Otegi, J. F., Helled, R., & Bouchy, F. 2021b, arXiv e-prints, arXiv:2112.07413

Owen, J. E. & Lai, D. 2018, , 479, 5012

Owen, J. E. & Wu, Y. 2013, , 775, 105

Owen, J. E. & Wu, Y. 2017, The Astrophysical Journal, 847, 29

Parmentier, V., Showman, A. P., & Lian, Y. 2013, , 558, A91

Paxton, B., Bildsten, L., Dotter, A., et al. 2010, The Astrophysical Journal Supplement Series, 192, 3

Paxton, B., Cantiello, M., Arras, P., et al. 2013, , 208, 4

Paxton, B., Schwab, J., Bauer, E. B., et al. 2018, , 234, 34

Pepe, F., Mayor, M., Rupprecht, G., et al. 2002, The Messenger, 110, 9

Petrovich, C. 2015, , 799, 27

Plotnykov, M. & Valencia, D. 2021, in Bulletin of the American Astronomical Society, Vol. 53, 1236

Pollacco, D. L., Skillen, I., Collier Cameron, A., et al. 2006, , 118, 1407

Queloz, D., Bouchy, F., Moutou, C., et al. 2009, , 506, 303

Queloz, D., Mayor, M., Weber, L., et al. 2000, , 354, 99

Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249

Ricard, Y., Šrámek, O., & Dubuffet, F. 2009, Earth and Planetary Science Letters, 284, 144

Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Rivera, E. J., Lissauer, J. J., Butler, R. P., et al. 2005, , 634, 625

Rogers, L. A. & Seager, S. 2010, , 716, 1208

Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, , 784, 45

Saar, S. H. & Donahue, R. A. 1997, , 485, 319

Santerne, A., Brugger, B., Armstrong, D. J., et al. 2018, Nature Astronomy, 2, 393

Santerne, A., Díaz, R. F., Moutou, C., et al. 2012, Astronomy Astrophysics, 545, A76

Saumon, D., Chabrier, G., & van Horn, H. M. 1995, , 99, 713

Schmitt, J. R., Agol, E., Deck, K. M., et al. 2014, , 795, 167

Schoonenberg, D., Liu, B., Ormel, C. W., & Dorn, C. 2019, Astronomy Astrophysics, 627,

A149

Schöttler, M. & Redmer, R. 2018, Journal of Plasma Physics, 84, 755840401

Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, The Astrophysical Journal, 669, 1279–1297

Seager, S. & Mallen-Ornelas, G. 2003, The Astrophysical Journal, 585, 1038–1055

Shah, O., Alibert, Y., Helled, R., & Mezger, K. 2021, , 646, A162

Silva, T. A., Demangeon, O. D. S., Barros, S. C. C., et al. 2021, The HD 137496 system: A dense, hot super-Mercury and a cold Jupiter

Sinukoff, E., Howard, A. W., Petigura, E. A., et al. 2016, , 827, 78

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, American Institute of Physics Conference Series, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405

Smith, B. A. & Terrile, R. J. 1984, Science, 226, 1421

Sotin, C., Grasset, O., & Mocquet, A. 2007, , 191, 337

Spiegel, D. S., Fortney, J. J., & Sotin, C. 2014, Proceedings of the National Academy of Science, 111, 12622

Spiegel, D. S. & Turner, E. L. 2011, Proceedings of the National Academy of Sciences, 109, 395–400

Stassun, K. G., Collins, K. A., & Gaudi, B. S. 2017, , 153, 136

Steffen, J. H. 2016, , 457, 4384

Stixrude, L. & Lithgow-Bertelloni, C. 2011, Geophysical Journal International, 184, 1180

Strom, K. M., Strom, S. E., Edwards, S., Cabrit, S., & Skrutskie, M. F. 1989, , 97, 1451

Szabó, G. M. & Kiss, L. L. 2011, , 727, L44

Székely, G. J., Rizzo, M. L., & Bakirov, N. K. 2007, The Annals of Statistics, 35

Tanaka, H., Takeuchi, T., & Ward, W. R. 2002, , 565, 1257

Teske, J., Díaz, M. R., Luque, R., et al. 2020, , 160, 96

Thompson, S. L. & Lauson, H. S. 1972, Improvements in the Chart D Radiation-Hydrodynamic Code. III: Revised Analytic Equations of State, Albuquerque

Umemoto, K. & Hsu, H. 2017, in AGU Fall Meeting Abstracts, Vol. 2017, MR41C–0416

Unterborn, C. T., Desch, S. J., Hinkel, N. R., & au2, A. L. J. 2018, Inward Migration of the TRAPPIST-1 Planets as Inferred From Their Water-Rich Compositions

Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007, , 656, 545

Valletta, C. & Helled, R. 2020, , 900, 133

van Dishoeck, E. F., Bergin, E. A., Lis, D. C., & Lunine, J. I. 2014, in Protostars and Planets VI, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 835

Vazan, A., Kovetz, A., Podolak, M., & Helled, R. 2013, Monthly Notices of the Royal Astronomical Society, 434, 3283–3292

Vazan, A., Ormel, C. W., Noack, L., & Dominik, C. 2018, , 869, 163

Venturini, J., Alibert, Y., Benz, W., & Ikoma, M. 2015, , 576, A114

Venturini, J., Ronco, M. P., & Guilera, O. M. 2020, , 216, 86

Visscher, C. & Fegley, Bruce, J. 2005, , 623, 1221

Visscher, C. & Moses, J. I. 2011, , 738, 72

Wagner, F. W., Sohl, F., Hussmann, H., Grott, M., & Rauer, H. 2011, , 214, 366

Wagner, W. & Pruß, A. 2002, Journal of Physical and Chemical Reference Data, 31, 387

Walker, J. C. G., Hays, P. B., & Kasting, J. F. 1981, , 86, 9776

Weiss, L. M. & Marcy, G. W. 2014, , 783, L6

Weiss, L. M., Marcy, G. W., Petigura, E. A., et al. 2018, , 155, 48

Weiss, L. M., Marcy, G. W., Rowe, J. F., et al. 2013, , 768, 14

Weiss, L. M. & Petigura, E. A. 2020, , 893, L1

Weiss, L. M., Rogers, L. A., Isaacson, H. T., et al. 2016, , 819, 83

Wetherill, G. W. 1985, Science, 228, 877

Wetherill, G. W. 1990, , 88, 336

Wiggins, P. 2008, PLoS ONE, 3, e1406

Winn, J. N. 2010, Exoplanet Transits and Occultations, ed. S. Seager, 55–77

Wolfgang, A., Rogers, L. A., Ford, E. B., & Laughlin, G. 2015, in AAS/Division for Extreme Solar Systems Abstracts, Vol. 3, AAS/Division for Extreme Solar Systems Abstracts, 501.04

Wolszczan, A. & Frail, D. A. 1992, , 355, 145

Workman, R. K., Hauri, E., Hart, S. R., Wang, J., & Blusztajn, J. 2005, in AGU Fall Meeting Abstracts, Vol. 2005, V22A–05

Wurm, G., Trieloff, M., & Rauer, H. 2013, , 769, 78

Zeng, L., Jacobsen, S. B., & Sasselov, D. D. 2017, Research Notes of the American Astronomical Society, 1, 32

Zhu, W. 2020, , 159, 188