

Dissertation
submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany,
for the degree of
Doctor of Natural Sciences

Put forward by
MARTIN SCHLECKER
born in Augsburg, Germany

Oral examination: July 21, 2021

THE ARCHITECTURES OF PLANETARY SYSTEMS:
POPULATION SYNTHESIS MEETS OBSERVATIONS

REFEREES:

PROF. DR. THOMAS HENNING

DR. SABINE REFFERT

Abstract

Advancing our understanding of planet formation is a prime motivation for the search for exoplanets. While it is now widely recognized that multiple planets per system are common, their mutual relationships are still largely unexplored. This thesis investigates such relationships by confronting simulated planet populations with observed planetary systems.

To draw conclusions about the formation environments of planets that we observe today, the causal connections between their properties and those of their natal protoplanetary disks must first be established. In a data-driven approach, I identify the most predictive initial conditions of a planet formation model and show that N -body interactions affect primarily low-mass planets. These insights are then used to study the relations between super-Earths on short orbits and outer giant planets. I find a connection between the composition of simulated planets and the architectures of their systems. This gives rise to the testable hypothesis that high-density inner super-Earths point to a giant companion in the same system. The analysis also suggests that dynamically active giant planets frequently destroy systems of inner super-Earths. This is compatible with the discovery of one of the most eccentric warm Jupiters known that I present in this thesis. I demonstrate in a tidal evolution analysis that this planet is not the progenitor of a hot Jupiter during its high-eccentricity migration.

To explore variations of these trends as a function of stellar host mass, I confront the CARMENES M dwarf survey with a synthetic population of planets around low-mass stars. A striking discrepancy is the observed existence of giant planets around very low-mass stars, which can not be reproduced by our model. Future planet formation theories must explain also this peculiar finding.

Zusammenfassung

Ein wesentliches Ziel der Suche nach Exoplaneten ist es, ein besseres Verständnis von Planetenentstehung zu gewinnen. Während es inzwischen anerkannt ist, dass mehrere Planeten pro System die Regel sind, sind deren gegenseitige Beziehungen noch weitgehend unerforscht. In dieser Dissertation werden solche Beziehungen untersucht, indem simulierte Planetenpopulationen mit beobachteten Planetensystemen konfrontiert werden.

Um Rückschlüsse auf die Entstehung von heute beobachteten Planeten zu ziehen, müssen zunächst die kausalen Zusammenhänge zwischen ihren Eigenschaften und denen der protoplanetaren Scheiben, aus denen sie entstanden sind, ermittelt werden. Durch maschinelles Lernen ermittle ich die Anfangsbedingungen eines Planetenentstehungsmodells mit der höchsten Voraussagekraft und zeige, dass N -Körper-Wechselwirkungen vorwiegend massearme Planeten beeinflussen. Diese Erkenntnisse werden dann genutzt, um die Beziehungen zwischen Supererden auf kurzen Umlaufbahnen und äußeren Riesenplaneten zu untersuchen. Ich finde einen Zusammenhang zwischen der Zusammensetzung simulierter Planeten und den Architekturen ihrer Systeme aus der sich die überprüfbare Hypothese ableitet, dass das Auftreten einer Supererde hoher Dichte auf einen Riesenplaneten im selben System hindeutet. Die Untersuchung zeigt auch, dass dynamisch aktive Riesenplaneten häufig Systeme von inneren Supererden zerstören. Dies ist kompatibel mit der hier vorgestellten Entdeckung eines warmen Jupiters auf einer besonders exzentrischen Umlaufbahn. Eine Analyse der Gezeitenentwicklung des Planeten zeigt, dass er kein Vorläufer eines heißen Jupiters während seiner hochexzentrischen Migration ist.

Um Variationen dieser Beziehungen in Abhängigkeit von der Sternenmasse zu erkunden, konfrontiere ich die CARMENES M-Stern-Durchmusterung mit einer synthetischen Population von Planeten um massearme Sterne. Eine auffällige Diskrepanz ist das beobachtete Auftreten von Riesenplaneten um sehr massearme Sterne, das von unserem Modell nicht reproduziert werden kann. Zukünftige Planetenentstehungstheorien werden auch diesen unerwarteten Befund erklären müssen.

Contents

| | | |
|----------|---|----------|
| 1 | Motivation | 1 |
| 1.1 | State of the field | 1 |
| 1.1.1 | Exoplanet demographics | 1 |
| 1.1.2 | System architectures | 2 |
| 1.1.3 | Planets around low-mass stars | 3 |
| 1.1.4 | Planet population synthesis | 3 |
| 1.1.5 | Open questions | 4 |
| 1.2 | Outline of this thesis | 5 |
| 2 | Foundations | 7 |
| 2.1 | Protoplanetary disks | 7 |
| 2.1.1 | Matter inventory | 8 |
| 2.1.2 | Radial structure | 9 |
| 2.1.3 | Vertical structure | 11 |
| 2.1.4 | Disk masses | 12 |
| 2.1.5 | Orbital velocities | 13 |
| 2.1.6 | Disk evolution | 14 |
| 2.1.7 | Scalings with stellar mass | 16 |
| 2.2 | Planet formation in the core accretion paradigm | 18 |
| 2.2.1 | Stages of planetesimal growth | 19 |
| 2.2.2 | Planetesimal sizes | 21 |
| 2.2.3 | The planetesimal isolation mass | 21 |
| 2.2.4 | The growth timescale | 23 |
| 2.2.5 | Terrestrial planet formation | 24 |
| 2.2.6 | Giant planet formation | 25 |

Contents

| | | |
|----------|---|-----------|
| 2.2.7 | Pebble accretion | 26 |
| 2.2.8 | Planet migration | 27 |
| 2.3 | Planet observations | 28 |
| 2.3.1 | Characterization of planets in the Solar System and beyond | 28 |
| 2.3.2 | Orbital parameters | 28 |
| 2.3.3 | The radial velocity technique | 29 |
| 2.3.4 | Transit photometry | 31 |
| 2.4 | Exoplanet demographics | 35 |
| 2.4.1 | Biases and caveats | 35 |
| 2.4.2 | A census of exoplanets | 37 |
| 2.4.3 | Features in the exoplanet population | 38 |
| 2.4.4 | Architectural trends | 39 |
| 2.4.5 | Planetary systems around M dwarfs | 41 |
| 2.5 | Population synthesis with global models of planet formation | 41 |
| 2.5.1 | The Generation III Bern model | 45 |
| 2.5.2 | M dwarf population synthesis | 52 |
| 3 | Predetermination of Planet Types in Global Core Accretion Models | 57 |
| 3.1 | Motivation | 58 |
| 3.2 | Planet population synthesis | 60 |
| 3.3 | Cluster analysis | 62 |
| 3.3.1 | Data preparation | 62 |
| 3.3.2 | Model selection and hyperparameters | 62 |
| 3.3.3 | Detected planet clusters | 64 |
| 3.3.4 | Model validation | 66 |
| 3.3.5 | Planet clustering as a function of simulation time | 67 |
| 3.4 | Prediction of planet clusters | 69 |
| 3.4.1 | Data preparation, hyperparameters, and training | 69 |
| 3.4.2 | Error and performance analysis | 70 |
| 3.4.3 | Results of planet predictions | 70 |
| 3.4.4 | Differences between single and multi-planet systems | 74 |
| 3.5 | Discussion | 77 |
| 3.5.1 | What determines the type of a planet? | 77 |
| 3.5.2 | Disk mass and embryo distance as predictors for planet type | 79 |
| 3.5.3 | Oligarchic growth of giant planets | 80 |
| 3.5.4 | The influence of N -body interactions | 83 |

| | | |
|----------|--|------------|
| 3.6 | Conclusions | 84 |
| 4 | Warm Super-Earths and Cold Jupiters | 87 |
| 4.1 | Motivation | 88 |
| 4.1.1 | Super-Earths, cold Jupiters, and theoretical predictions | 88 |
| 4.1.2 | Observational findings | 89 |
| 4.1.3 | An imperative to confront theory and observation | 91 |
| 4.2 | Methods | 92 |
| 4.2.1 | Synthetic planet sample | 92 |
| 4.2.2 | Occurrence rates and fraction of planet hosts | 93 |
| 4.2.3 | Observed planet sample | 94 |
| 4.2.4 | Detection limit | 95 |
| 4.3 | Results: Synthetic population | 95 |
| 4.3.1 | System classes | 95 |
| 4.3.2 | Occurrence rates in period-radius | 100 |
| 4.3.3 | Occurrence rates in period-mass | 101 |
| 4.3.4 | Planet radii | 102 |
| 4.3.5 | Relation between metallicity and planet radius | 103 |
| 4.3.6 | Relation to disk properties | 104 |
| 4.4 | Results: comparison with observations | 108 |
| 4.4.1 | Fractions of planet hosts | 109 |
| 4.4.2 | Removal of super-Earths | 112 |
| 4.4.3 | Ice mass fractions | 114 |
| 4.4.4 | Host star metallicity: effects on planet occurrences | 120 |
| 4.4.5 | Possibility of reduced multiplicity in cold Jupiters | 122 |
| 4.5 | Discussion | 123 |
| 4.5.1 | Observed and theoretical host star fractions | 123 |
| 4.5.2 | The missing super-Earths | 126 |
| 4.5.3 | Volatile-poor super-Earths could be proxies for giant planets | 128 |
| 4.5.4 | A negative metallicity correlation for super-Earths in cold-Jupiter hosting systems | 130 |
| 4.6 | Conclusions | 132 |
| 5 | Discovery of A Highly Eccentric Warm Jupiter | 135 |
| 5.1 | Motivation | 136 |

Contents

| | | |
|----------|--|------------|
| 5.2 | Observations | 137 |
| 5.2.1 | TESS photometry | 137 |
| 5.2.2 | Photometric follow-up with CHAT | 137 |
| 5.2.3 | Additional photometry from LCOGT | 138 |
| 5.2.4 | High precision spectroscopy with FEROS | 138 |
| 5.2.5 | Contamination | 138 |
| 5.3 | Analysis | 139 |
| 5.3.1 | Stellar Parameters | 139 |
| 5.3.2 | Joint modeling | 140 |
| 5.3.3 | Approximation of the planetary equilibrium temperature | 151 |
| 5.3.4 | Secondary eclipses, phase curve modulations, and Rossiter-MacLaughlin effect | 153 |
| 5.4 | Discussion | 156 |
| 5.4.1 | TIC 237913194b’s place in the exoplanet population | 156 |
| 5.4.2 | The atmosphere of TIC 237913194b | 157 |
| 5.4.3 | TIC 237913194b’s large eccentricity | 158 |
| 5.5 | Conclusions | 159 |
| 6 | Planetary Systems Around Low-Mass Stars | 161 |
| 6.1 | Motivation | 162 |
| 6.2 | The CARM ₁₂₅ observational sample | 163 |
| 6.3 | Observing the synthetic population | 163 |
| 6.4 | Observed and theoretical occurrence rates | 165 |
| 6.5 | Planetary mass function | 168 |
| 6.6 | Discussion | 170 |
| 6.6.1 | Giant planets around M dwarfs | 170 |
| 6.6.2 | Planetary mass function | 172 |
| 7 | Conclusions and Future Work | 173 |
| 7.1 | Summary | 173 |
| 7.2 | Conclusions | 175 |
| 7.3 | Outlook | 176 |
| 7.3.1 | Extensions of the formation model | 177 |
| 7.3.2 | Fitting for system architectures | 178 |
| 7.3.3 | The architectures of M dwarf planetary systems | 179 |

| | |
|--|------------|
| A Appendix to Chapter 3 | 183 |
| A.1 The choice of a clustering algorithm | 183 |
| A.1.1 Clustering algorithms | 183 |
| A.1.2 Validation metrics and choice of method | 184 |
| A.1.3 Model selection | 185 |
| A.2 Boundary conditions for giant planet formation | 185 |
| B Appendix to Chapter 5 | 189 |
| B.1 Joint fit posteriors | 189 |
| List of Abbreviations | 191 |
| List of Figures | 193 |
| List of Tables | 197 |
| List of Own Publications | 199 |
| Bibliography | 207 |
| Acknowledgements | 239 |

1

Motivation

Let's say a new student is just starting their PhD to carry on my research. If they happen to live on another planet with an awful communication link and I was able to pass on only a single statement to them, what should this statement be? Maybe one could transmit: *Most systems host several planets; their respective properties are mutually dependent and can be related to their formation phase.* If there are still a few bytes left, I would add: *And mind the stellar mass!* For the Earth-bound reader, a slightly more verbose presentation of my results may be found in this thesis. But what were the original questions to this answer? Here, I provide an overview of the foundations my research has built on and of the open questions in the field as my doctorate started.

1.1 State of the field

When the study of exoplanets evolved into a serious line of research only some 20 years ago, its focus was on discovering new exoplanets and improving the techniques to do so (Scott Gaudi et al., 2020). The parameter space these methods were sensitive to was narrow, and thus the utility of this new observational evidence to test planet formation theories was limited. At the time, theoretical models were strongly oriented toward explaining the Solar System planets (e.g., Boss, 1995; Tsiganis et al., 2005; Morbidelli et al., 2015). The steadily increasing number of exoplanet discoveries soon indicated not only that planets are abundant, but also that they occur in a larger variety than expected from studies of our home system (Benz et al., 2014).

1.1.1 Exoplanet demographics

Only recently, the sample of known planets has grown sufficiently large to enable meaningful statistical analyses and to search for structures in the planetary parameter space: the field of exoplanet *demographics* was born. From an observer's perspective, its main goal is

1 Motivation

to determine the frequency of exoplanets and the distribution of their parameters relevant for our understanding of planet formation. This endeavor remains challenging, as each detection method has its own set of intrinsic biases, detection sensitivities, and planet properties it can measure. The currently known planetary population is thus strongly skewed toward larger and more massive planets on short orbits (e.g., [Hsu et al., 2019](#)), and future demographic studies will undoubtedly seek to integrate results from various methods ([Bennett et al., 2019](#)). To complicate things further, it is expensive to determine the properties of host stars such as their mass, which is regarded as a decisive parameter for the planetary systems they harbor ([Scott Gaudi et al., 2020](#)). Despite the various observational challenges and setbacks, several trends that offer clues for planet formation have been found in the observed exoplanet population, and I will review the most relevant ones in Sect. 2.4.

1.1.2 System architectures

While constraints on properties such as radius, mass, and occurrence rates have helped to shed light on the rich variety of exoplanets (e.g., [Fulton et al., 2017](#); [Hsu et al., 2019](#)), the effect of planet multiplicity was largely neglected. Indeed, NASA's *Kepler* mission ([Borucki et al., 2010](#)) has shown that multi-planet systems are the rule rather than the exception ([Lissauer et al., 2011](#); [Latham et al., 2011](#)). Yet it is only recently that progress is being made in evolving from planet-level properties to exploring planetary systems as a whole (e.g., [Ballard & Johnson, 2016](#); [Weiss et al., 2018](#); [He et al., 2020b](#); [Gilbert & Fabrycky, 2020](#)). Much attention has been paid to patterns of intra-system uniformity in *Kepler*'s multi-planet systems. In particular, it has been suggested that planets are like "peas in a pod" - similar in size, on evenly spaced orbits, and ordered by size so that larger planets are further out ([Millholland et al., 2017](#); [Wang, 2017](#); [Weiss et al., 2018](#); [He et al., 2020b](#)). After much debate, there is a growing consensus that at least the size similarities and the uniform orbit spacings are likely real ([Gilbert & Fabrycky, 2020](#)). These architectural insights show that the often quoted diversity of exoplanets is probably rather a diversity of planetary *systems*.

Another still largely unresolved challenge is the determination of conditional occurrence probabilities that concern questions like: given a planet with properties X , what is the probability to find another planet with properties Y in the same system? Such questions could be helpful in testing different planet formation theories, as these are able to provide predictions of this kind ([Winn, 2018](#)). A particularly relevant tentative observation, an increased joint occurrence of inner super-Earths and outer giant planets ([Zhu & Wu, 2018](#)), will be investigated in this thesis.

1.1.3 Planets around low-mass stars

Low-mass stars receive growing interest from the exoplanet community. *M dwarfs*, which comprise the coolest and smallest stars on the main sequence, are also the stellar spectral type most abundant in the Universe (e.g., Salpeter, 1955; Chabrier, 2003). Since the most commonly employed exoplanet detection techniques are more sensitive to less massive (Doppler method, see Sect. 2.3.3) and smaller (transit method, see Sect. 2.3.4) stars, M dwarfs provide a unique opportunity to detect even small, rocky planets whose signals would currently remain undetected in the light from sunlike stars. The sensitivities of both methods also decline with orbital distance. It is therefore a welcomed feature of M dwarfs that due to their smaller luminosity, their habitable zones (commonly defined as the orbital region where a planet with an Earth-like atmosphere would be able to maintain liquid water on its surface, Hart (1979)) are much closer to the star. While, for my aforementioned successor outside the Solar System, Earth has an a priori (geometrical) probability to transit the Sun of 0.47%, it is an encouraging 2.7% for a planet in the habitable zone around an M8V dwarf¹ (Charbonneau & Deming, 2007; Shields et al., 2016, Eqn 2.28). For these reasons, significant progress has been made in observing exoplanets around M dwarfs (e.g., Nutzman & Charbonneau, 2008; Gillon et al., 2016; Quirrenbach et al., 2010), and efforts are underway to discover and characterize more of these planets (e.g., Gillon et al., 2018; Sebastian et al., 2021; Gibbs et al., 2020). Results collected so far are reassuring: the NASA exoplanet archive² lists more than 1000 confirmed planets around stars less massive than $0.6 M_{\odot}$, some of which are rocky planets located in the circumstellar habitable zone (e.g., Anglada-Escudé et al., 2016; Gillon et al., 2017; Zechmeister et al., 2019; Dreizler et al., 2020; Vanderburg et al., 2020). Population-level studies based on these discoveries have shown that low-mass stars are the most abundant planet hosts in the solar neighborhood (Hsu et al., 2020) and host more small, potentially habitable planets than solar-type stars (Mulders et al., 2015b; Dressing & Charbonneau, 2015).

1.1.4 Planet population synthesis

The quest to explain the above mentioned distributions and trends was the incentive for planet population synthesis studies, which aim at simulating planetary systems that can be directly compared to observables of exoplanets and their overall population. The method assumes that the diversity of exoplanets is caused by a diversity of their original environ-

¹“Habitable zone” is here approximated by calculating the orbital distance at which a planet with Earth’s albedo has the same equilibrium temperature as Earth. Many other factors have to be accounted for to assess the actual prospects of a planet to host life, for example tidal and radiative interaction with its host star (e.g., Shields et al., 2013; kumar Kopparapu et al., 2016; France et al., 2020; Turbet & Selsis, 2021).

²<https://exoplanetarchive.ipac.caltech.edu>, queried on 2021-05-06.

ments, that is, of the protoplanetary disks in which the planets have formed (Mordasini, 2018). With increasing demands from exoplanet demographics, population synthesis evolved from the pioneering studies by Ida & Lin (2004a) to ever more sophisticated models. Successively, additional physical mechanisms such as planet migration (Ida & Lin, 2008; Dittkrist et al., 2014), the structural evolution of planet interiors (Mordasini et al., 2009a), atmospheric escape (Jin et al., 2014), long-term planetary evolution (Mordasini et al., 2012c), dust evolution and disk inhomogeneities (Hasegawa & Pudritz, 2011; Alessi et al., 2017), disk chemistry (Cridland et al., 2016; Thiabaud et al., 2014), pebble accretion (Bitsch et al., 2015; Ndugu et al., 2018), and stellar mass dependencies (Ida & Lin, 2005; Alibert et al., 2011; Liu et al., 2019, Burn et al., in press) were covered. A major advancement was the step from simulating only one planet per disk to considering multiple planets that gravitationally interact (Thommes et al., 2008; Ida & Lin, 2010; Alibert et al., 2013). This component has proven particularly important for producing realistic parameter distributions of small planets (Mordasini, 2018), which in nature seem to often huddle in compact orbital configurations (e.g., Mayor et al., 2011). However, previous studies based on such multi-planet populations have mostly focused on attributes on the planetary level, neglecting quantitative statistical analyses of the intra-system relationships between planets (Thommes et al., 2008; Alibert et al., 2013; Mordasini, 2018; Emsenhuber et al., 2020b; Burn et al., 2021). First attempts to investigate such trends in multi-planet systems have only recently been made or are in preparation (Mulders et al., 2020, Mishra et al., in prep.).

1.1.5 Open questions

Undoubtedly, the observational and theoretical efforts outlined here and in more detail in Chapter 2 have provided important foundations for our understanding of planetary system architectures and their connection to planet formation. At the same time, these advances have produced some intriguing new puzzles. When I started my dissertation, a number of key questions were waiting to be answered:

1. **Are the increasingly complex multi-planet models still deterministic?**
2. **Is the formation of one planet type or another already set by primordial properties of the disk and host star?**
3. **What are the relations between inner systems of rocky planets and outer giant planets, and can the core accretion scenario explain these relations?**

4. Are giant planets on wide orbits favorable or detrimental for systems of inner rocky planets?
5. Is there observational evidence for interactions among these planet types?
6. Are the bulk properties of planets connected to the architectures of their planetary systems?
7. Can population synthesis models reproduce the planetary systems around low-mass stars?

1.2 Outline of this thesis

The goal of this thesis is to link the architectures of planetary systems to physical processes involved in their formation. I approach this problem by confronting theoretical simulations of planet formation and evolution with empirical exoplanet demographics.

Chapter 2 lays the foundations necessary to understand and motivate the subsequent parts. I review the current state of the art in planet formation and discuss different paradigms. This involves mainly the physical appearance and processes in protoplanetary disks and the *core accretion* scenario of planet formation, which is subject of this thesis. I also present relevant observational techniques and some important findings in the demography of exoplanets. The chapter ends with an introduction of planet population synthesis and a description of the theoretical model I use.

Chapter 3 provides a close look inside this model in its nominal version assuming solar-type host stars. In a data-driven approach, I scrutinize the deterministic links between the model's initial conditions, which are essentially disk properties, and the properties of the simulated planets it produces. I use machine learning algorithms to first identify clusters of similar planets, which I then try to predict from initial conditions. A special focus lies in the differences between a single-planet approximation and the full consideration of N -body interactions. This chapter also includes a brief comparison of the synthetic planet population with observed *Kepler* occurrence rates.

Chapter 4 investigates, based on the same model, the relations between small planets on short orbits ("super-Earths") and giant planets on wide orbits ("cold Jupiters") that inhabit the same system. To enable a statistical confrontation of observed and simulated

1 Motivation

systems, I make a crude correction for the detection bias inherent in the observed sample. A previously observed trend that inner super-Earths frequently coexist with cold Jupiters is then being compared with theoretical predictions. I further investigate the role of giant planets on eccentric intermediate orbits (“warm Jupiters”) in controlling the magnitude of this trend. I discuss the connection between the appearance of such a planet and the initial conditions of the protoplanetary disk. This results in the testable prediction of a link between the composition of inner super-Earths and the architecture of their systems.

Chapter 5 presents the discovery of a warm giant planet on a highly eccentric orbit around a solar-type star. Originally detected in space-based photometry of the *Transiting Exoplanet Survey Satellite* (TESS, [Ricker et al., 2014](#)), the planetary nature of TIC 237913194b is confirmed via ground-based photometric and spectroscopic observations from three different telescopes. By jointly modeling all datasets, I precisely constrain its mass, radius, and orbital properties. Based on a tidal evolution analysis, I discuss potential origins of the planet’s extreme dynamical state. The chapter also addresses prospects for observing its atmosphere, which might be subject to chemical disequilibrium processes due to the high eccentricity of the orbit.

Chapter 6 specifically addresses planets around low-mass stars. I statistically compare a synthetic planet population generated with a formation model adapted to low-mass stars to exoplanets from the CARMENES ([Quirrenbach et al., 2010](#)) survey, which searches for planets around nearby M dwarf stars with the Radial Velocity (RV) method. For this purpose, I first generate a synthetic M dwarf population with a stellar mass distribution corresponding to that of the CARMENES sample. Then, I derive planetary occurrence rates and parameter distributions and confront them with those of the detected CARMENES planets. I conclude with a discussion of the differences in the theoretical and observed planetary mass functions.

Chapter 7 summarizes and interprets the findings of this dissertation. I discuss their implications in the context of the above science questions and provide some suggestions for future studies.

Chapters 3 to 5 are based on articles that I have published in peer-reviewed scientific journals or that are accepted for publication at the time of writing. Chapter 6 presents work in progress whose expansion will appear in a future publication.

2

Foundations

Before the first discovery of an exoplanet around a solar-type star in 1995 (Mayor & Queloz, 1995), attempts to formulate a theory of planet formation have mostly focused on the Solar System (Wuchterl, 2007). Most of the basic features of these theories are however valid until today, namely the formation of planets as a by-product of star formation from the collapse of a gravitationally unstable cloud fragment, the emergence of a circumstellar disk of gas and solid material (i.e., dust), and the genesis of all planets in a system in and from this disk. In the following, I briefly discuss the current picture of the formation and physical structure of these *protoplanetary disks* and the most promising avenues for the appearance of protoplanets. This shall motivate my approach of simulating the formation of planetary systems, which I present in Sect. 2.5.

2.1 Protoplanetary disks

From an observational point of view, young stellar objects (YSO) have been classified into four evolutionary stages according to the shape of their spectral energy distribution (SED) (Lada & Wilking, 1984; Lada, 1987).

- For *Class 0* objects, the SED resembles those of a single-temperature blackbody and peaks in the far-infrared (FIR) or mm domain. This spectral distribution is explained by Class 0 objects being protostars fully enveloped by an optically thick cloud of gas and dust. The envelope re-emits optical photons from the protostar at longer wavelengths (Andre et al., 1993).
- *Class I* objects show an SED that is either flat or with a positive slope in the near-infrared (NIR) and mid-infrared (MIR). Such emission appears when the infalling material has already formed a disk due to angular momentum conservation. This disk is still embedded in an envelope.

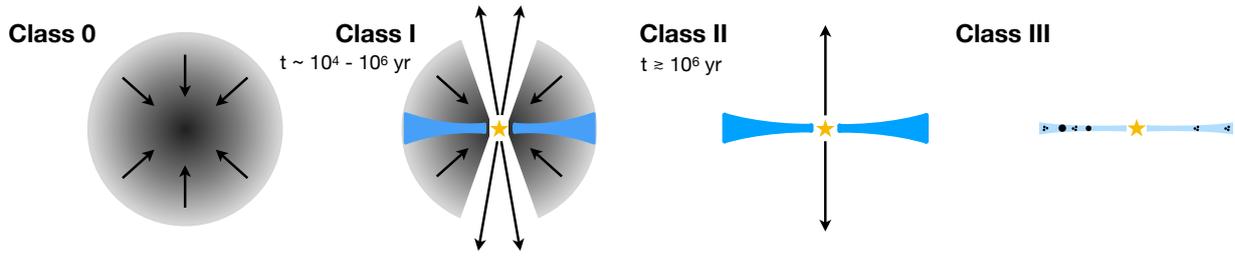


Figure 2.1: Evolutionary stages of YSOs. According to the “Lada sequence”, YSOs can be classified into gravitationally collapsing cloud fragments (Class 0), accretion disks that are embedded in an envelope (Class I), exposed accretion disks potentially with embedded protoplanets (Class II), and debris disks with no or very little gas (Class III). Figure reproduced from Pohl (2018) with kind permission by the author.

- Once this envelope has dispersed through accretion and molecular outflows, our line of sight is clear to the disk, and we refer to *Class II* objects. Their spectra are broader than blackbody distributions and flat or have negative slopes at wavelengths $> 2 \mu\text{m}$.
- *Class III* spectra can be fit with reddened blackbody functions representative of reddened stellar photospheres near the zero age main sequence (ZAMS). At this stage, even the gas in the disk has dispersed and a pre-main-sequence (PMS) star surrounded by a disk of debris and, possibly, a planetary system remains.

Figure 2.1 illustrates these four known evolutionary stages of YSOs. The research results presented in this thesis concern only the Classes I–II and do not take into account protostellar envelopes. Hence, in the following sections I will completely ignore such envelopes and focus on the cradles of planetary systems: circumstellar disks of gas and solid material. The following brief outline of the physical structure of planet-forming disks loosely follows the descriptions in Williams & Cieza (2011) and Armitage (2019).

2.1.1 Matter inventory

At the time of its initial formation, a protoplanetary disk’s composition is inherited from the interstellar medium (ISM). Thus, the bulk of its mass consists of molecular hydrogen, and the solid component consists of sub-micrometer sized dust particles. These particles are mainly silicates and carbonaceous materials (Draine, 2003). One of the key parameters for subsequent planet formation is the ratio between the gaseous and the solid component, i.e., the dust-to-gas ratio $\zeta_{\text{d,g}}$. Its magnitude, too, is passed on from the ISM and values around $\zeta_{\text{d,g}} = 0.01$ are typically assumed (Bohlin et al., 1978). During the evolution of the protostar and disk, a variety of physical processes such as irradiation, sublimation and recondensation, grain growth, turbulent mixing, or thermal annealing alter the initial properties of dust grains (Henning, 2010). The silicates are believed to subdivide depending

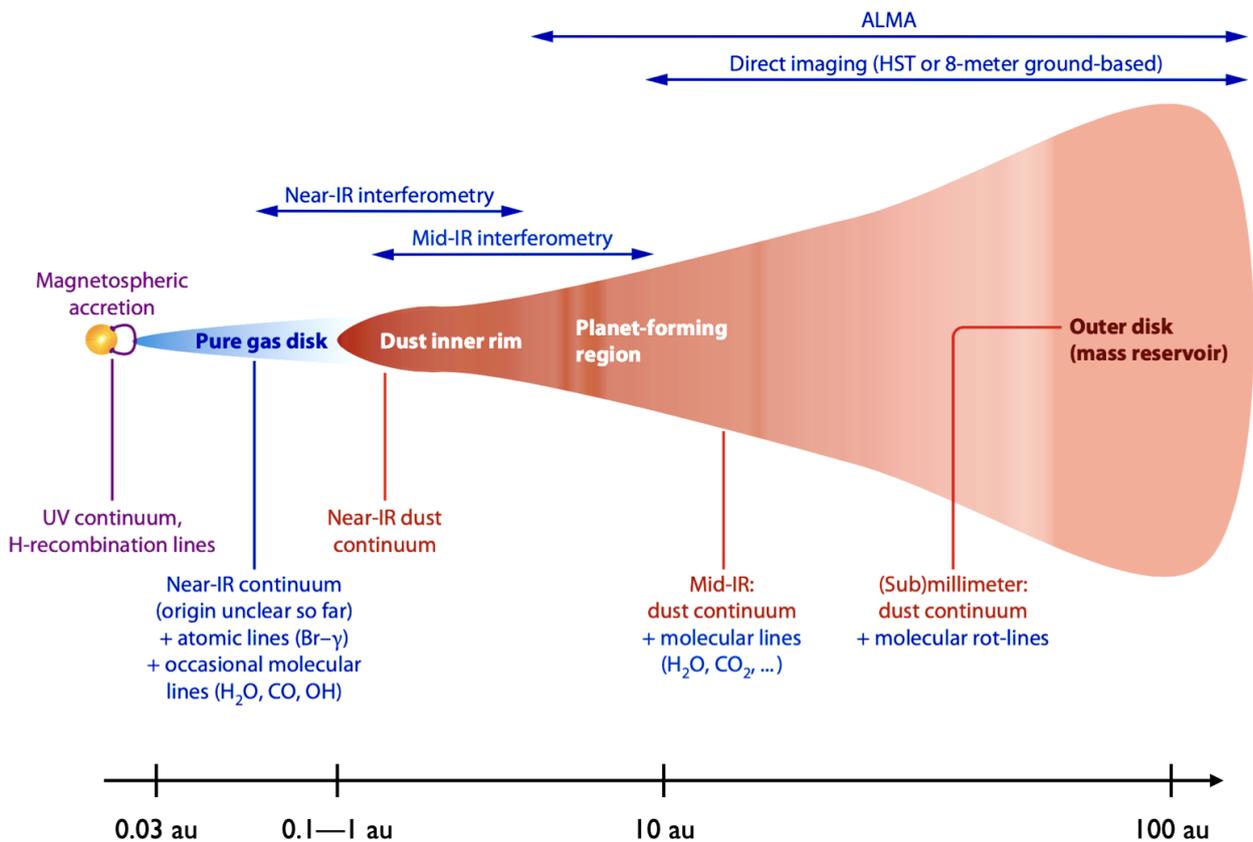


Figure 2.2: Schematic diagram of a protoplanetary disk. From inside out, the disk can roughly be subdivided into a magnetospheric cavity, a pure gas disk where dust sublimation temperatures are exceeded, and a combined gas and dust disk that can extend out to hundreds of au. Above the pictogram, the observational techniques that are sensitive to certain regions are shown. Below it, the kind of emission of each region is shown. Image adapted from [Dullemond & Monnier \(2010\)](#) with kind permission by the authors.

on their orbital distance: while the outer disk regions still contain amorphous interstellar silicates, the grains become annealed in closer parts of the disk and finally form crystalline dust with equilibrium compositions in the innermost regions ([Gail, 2004](#)).

2.1.2 Radial structure

Figure 2.2 shows a schematic diagram that subdivides a prototypical protoplanetary disk into several regions that I touch on in the following. For details about observations of these regions, the reader is referred to related reviews such as [Dullemond & Monnier \(2010\)](#); [Zhao et al. \(2020\)](#); or [Andrews \(2020\)](#). One of the better constrained properties of protoplanetary disks are the radial position and morphology of their outer rims. The most precise measurements in this regard are nowadays performed in the (sub-)millimeter range, primarily by using the Atacama Large Millimeter/Submillimeter Array (ALMA). There is

2 Foundations

no sharp edge where the millimeter continuum flux drops instantly to zero – one typically defines the outer radius as the radial position that encloses a specified percentage of the total flux (Trapman et al., 2019). Typical outer radii of tens to hundreds of au seem to be the norm (Isella et al., 2009; Tripathi et al., 2017; Ansdell et al., 2018). Notably, it was found that the gas component, as inferred from measurements of CO rotational lines, exceeds the size of the dust disk measured in the continuum emission (Isella et al., 2007). In combined observations of the same disks in CO and continuum, Ansdell et al. (2018) found that gas disks are larger by roughly a factor of two. This size difference has been attributed to radial drift of the \sim millimeter-sized particles these observations are sensitive to, followed by dust depletion of the outer regions (Trapman et al., 2019, see Sect. 2.1.5). However, similar signatures can be produced by optical depth differences between continuum and line emission (e.g., Facchini et al., 2017) or radially varying grain growth efficiencies (e.g., Guilloteau et al., 2011; Birnstiel et al., 2010, 2012).

To describe the radial density profiles of gas disks, typically a self-similar solution to the viscous evolution equations is applied (Lynden-Bell & Pringle, 1974; Hartmann et al., 1998; Andrews et al., 2009). The resulting profile is a power law tapered by an exponential decay at the outer boundary and takes the form

$$\Sigma(r) = \Sigma_0 \left(\frac{r}{R_c} \right)^{-\beta_g} \exp \left[- \left(\frac{r}{R_c} \right)^{2-\beta_g} \right], \quad (2.1)$$

where Σ_0 is a normalization constant corresponding to the surface density at a reference radius, β_g is the slope of the profile, and R_c defines the location of the exponential decrease.

This description does not take into account an inner disk edge. Such a gap between the stellar surface and the disk is motivated by material leaving the disk plane and being accreted onto the star along poloidal magnetic field lines (Bouvier et al., 2007). This is assumed to happen where the field lines are co-rotating with and hence able to magnetically couple to ionized gas in the disk (Günther, 2013). The co-rotation radius resides at only a few stellar radii, corresponding to angular separations of milliarcseconds even for closeby protoplanetary disks, which is a major impediment to spatially resolved observations of this region. The temperatures at such close distances to the star exceed the sublimation temperatures of silicates (~ 1500 K, Hillenbrand et al., 1992). Therefore, the inner rim of the dust component lies further away from the star at several 0.1 au (Eisner et al., 2005, 2009, 2010; Dullemond & Monnier, 2010; Flock et al., 2016, 2017, 2019a).

2.1.3 Vertical structure

Ignoring self-gravitating effects, the vertical structure of the gas density ρ_g in a planet-forming disk is determined by hydrostatic equilibrium, that is, a balance between the gradient of the gas pressure P and the vertical component of the stellar gravity g_z ,

$$\frac{dP}{dz} = -\rho_g g_z. \quad (2.2)$$

In the simplified picture of a thin, vertically isothermal disk, $\partial T/\partial z = 0$, and assuming an ideal gas equation of state, a Gaussian density profile develops. The resulting gas density profile in the vertical direction z is then given by

$$\rho_g = \frac{\Sigma_g}{\sqrt{2\pi}H} \exp\left[-\frac{z^2}{2H^2}\right], \quad (2.3)$$

where H is the vertical scale height of the disk. The value of H is a function of the temperature of the gas: the higher it is, the more “puffed up” is the disk and the lower is the density in the disk midplane. Possible sources of heat include irradiation by the host star or stars in the neighborhood, viscous heating, and energy released by infalling gas (Chiang & Goldreich, 1997; Nakamoto & Nakagawa, 1994). H can be expressed as the ratio between the isothermal sound speed c_s and the Keplerian angular frequency Ω_K ,

$$H = \frac{c_s}{\Omega_K} = \left(\frac{c_s^2}{GM_\star}\right)^{1/2} r^{3/2}, \quad (2.4)$$

where $c_s = \sqrt{\frac{k_B T}{\mu m_H}}$ with the Boltzmann constant k_B , the mean molecular weight μ , and the mass of a hydrogen atom m_H . Nakamoto & Nakagawa (1994) estimate the midplane temperature as

$$T_m^4 = \frac{1}{2\sigma} \left(\frac{3\kappa_R}{8} \Sigma(r, t) + \frac{1}{2\kappa_P \Sigma(r, t)} \right) \dot{E}_\nu + T_{\text{irr}}^4. \quad (2.5)$$

Here, σ is the Stefan-Boltzmann constant, κ_R is the Rosseland mean opacity, κ_P is the Planck opacity, and $\dot{E}_\nu = \frac{9}{4} \Sigma(r, t) \nu \Omega_K^2$ is the viscous energy dissipation rate at Keplerian frequency Ω_K and with an effective viscosity ν (see Sect. 2.1.6 below). The effective temperature T_{irr} contains contributions from stellar and external irradiation (Hueso & Guillot, 2005; Fouchet et al., 2012).

From Eqn 2.4 it is apparent that if the temperature decreases radially less steeply than r^{-1} , the *relative* height of the disk H/r will increase with radius (Kenyon & Hartmann, 1987). Under typical assumptions, this is the case and protoplanetary disks show the *flared*

profile indicated in Fig. 2.2 (Chiang & Goldreich, 1997; D’Alessio et al., 1998; Dullemond et al., 2002).

2.1.4 Disk masses

A key parameter for subsequent planet formation in protoplanetary disks is their initial mass, i.e., the amount of material available for planet formation. What “initial” means, however, is not straight-forward to answer as it is challenging to put exact time stamps onto the evolutionary stages outlined above, and leftover matter infall onto the protostellar system may still occur during the Class I phase (McKee & Ostriker, 2007). Most YSOs are observed in the Class II stage, and the majority of them already show significant signs of evolutionary processes (e.g., Andrews, 2020).

While most of a disk’s mass is in the gaseous phase, a considerable number of studies approached the mass determination problem by measuring the dust continuum emission (e.g., Andrews & Williams, 2005; Andrews et al., 2010, 2013; Ansdell et al., 2016; Pascucci et al., 2016; Barenfeld et al., 2016). A prerequisite for this technique, where total dust masses are calculated from integrated continuum fluxes, is that the emission is optically thin. By assuming a dust-to-gas ratio, the result is then converted into a total disk mass. Challenges of this method are the temperature-dependent conversion from fluxes to dust masses, unknowns in the optical properties and evolution of dust grains, and the practicality of adopting dust-to-gas ratios of the interstellar medium for protoplanetary disks (e.g., Molyarova et al., 2017; Pohl, 2018; Lenz, 2020).

A more direct approach is to observe the gas phase. The majority of the gas consists of molecular hydrogen and helium (Solomon & Werner, 1971; Herbst & Klemperer, 1973), but observations of these species are very challenging and limited to specific regions of the disk (e.g., Carmona et al., 2011). Other molecules are more accessible and can be translated into total gas disk masses by using conversion factors. Promising tracer molecules are hydrogen deuteride (HD) and carbon monoxide (CO). HD abundances are high and it is well coupled to molecular hydrogen, but it poses practical challenges for disk observers (Bergin et al., 2013). CO, on the other hand, facilitates observations through its low-energy rotational transitions and its high abundance in circumstellar environments (Williams & Cieza, 2011; Öberg & Bergin, 2021). Often, less abundant isotopologues of CO have to be observed to fulfill the optical thickness requirement (e.g., Zhang et al., 2017). Another issue is that freeze-out of CO in cold regions of disk mid-planes can render it unobservable (Henning & Semenov, 2013). Nevertheless, CO and its isotopologues remain the preferred molecular tracers for total gas masses (Molyarova et al., 2017).

Submillimeter surveys of star-forming regions with an age of a few Myr revealed typical gas disk masses of $10^{-3} M_{\odot} - 10^{-1} M_{\odot}$ (often lower than $1 M_{\text{J}}$, Andrews et al., 2010; Ansdell et al., 2016; Miotello et al., 2016; Long et al., 2017). Measured dust disk masses range from a few M_{\oplus} to $\sim 1 M_{\text{J}}$ and vary with stellar mass (e.g., Pascucci et al., 2016; Barenfeld et al., 2016; Manara et al., 2018).

2.1.5 Orbital velocities

The different components within a disk, i.e., the gas component and solid components of varying sizes, are subject to different forces regarding their orbital motion. A more in-depth derivation of the following relations can be found in Burn (2020). For the gas component, the radial forces are governed by the Euler momentum equation

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} = -\frac{1}{\rho} \nabla P - F, \quad (2.6)$$

where ρ is the gas density, P is its pressure, and F are the forces acting on a parcel of gas. In a simplified picture, F represents only the gravitational force by the central star. In this case, the radial component of Eqn 2.6 can be expressed as

$$\frac{v_{\phi}^2}{r} = g_r + \frac{1}{\rho} \frac{\partial P}{\partial r}, \quad (2.7)$$

where $g_r = \frac{GM_{\star}r}{(r^2+z^2)^{3/2}}$ is the radial component of the gravitational acceleration. Any radial dependence on the gas pressure will thus act as a force term that lets the gas velocities deviate from a Keplerian velocity v_{K} . The left-hand side of this equation can be identified as the centrifugal term. Expressed in terms of the angular velocity $\Omega_g(r, z) = \frac{v_{\phi}}{r}$, Eqn 2.7 yields

$$r\Omega_g^2 = \frac{GM_{\star}r}{(r^2+z^2)^{3/2}} + \frac{1}{\rho} \frac{\partial P}{\partial r}. \quad (2.8)$$

The solid components, on the other hand, do not directly feel a gas pressure. For large particles that do not couple to the gas, the pressure gradient term can be dropped, yielding

$$r\Omega_K^2 = \frac{GM_{\star}r}{(r^2+z^2)^{3/2}} \quad (2.9)$$

with the angular velocity $\Omega_K(r, z)$ of a circular Keplerian orbit. As a consequence, the gaseous and solid components generally move on different azimuthal velocities. Under typical assumptions for unperturbed disks with a smooth surface density profile, both the gas density and the temperature decrease with r , leading to a negative $\frac{\partial P}{\partial r}$ term in Eqn 2.8.

This causes the gas to orbit at slightly sub-Keplerian velocities. The difference is not large: For a temperature of 30 K and at $r \approx 500$ au, the gas moves at roughly a factor 0.99 compared to an ideal Keplerian orbit (Armitage, 2007; Rosenfeld et al., 2013). Nevertheless, the effect has considerable consequences for particle dynamics and the radial distribution of solid material in protoplanetary disks. While very small grains ($\sim \mu\text{m}$ and with Stokes numbers $St \ll 1$) move with the gas on its sub-Keplerian orbit, larger particles with Stokes numbers around unity (corresponding to roughly mm to cm sizes) will experience a gas drag. This “headwind” causes the grains to lose angular momentum and drift radially inward (Whipple, 1972; Nakagawa et al., 1986; Weidenschilling, 1977; Takeuchi & Lin, 2002; Birnstiel et al., 2010, 2012). Very large bodies such as planetesimals ($St \gg 1$) are also subject to gas drag, but because of their large momentum they do not lose significant orbital energy on timescales relevant for planet formation. For $St \approx 1$ particles, the drift timescale can be as short as 100 yr, which poses a challenge for the formation of larger bodies (see Pohl, 2018). On the other hand, the mobility of solid material through the disk is key for an emerging paradigm of planet formation that I will present in Sect. 2.2.7.

2.1.6 Disk evolution

Protoplanetary disks are no static objects. They evolve in time through various physical mechanisms, for instance the settling of dust toward the midplane (D’Alessio et al., 1999) and its coagulation (e.g., Birnstiel et al., 2010), or dynamical interactions with massive (proto-)planets or stellar companions. I will largely ignore these effects and focus on two other processes that drive disk evolution, since they are of fundamental relevance for the models I employ in the following chapters. These processes are *viscous accretion* and *photoevaporation*.

The suggestion of a viscosity in protoplanetary disks is a direct consequence of the rather high observed accretion rates of classical T Tauri stars ($10^{-9} M_{\odot} \text{yr}^{-1} - 10^{-7} M_{\odot} \text{yr}^{-1}$, e.g., Gullbring et al., 1998; Alcalá et al., 2014; Manara et al., 2020): how can the required flow rates of disk material be explained? The challenge is to provide a mechanism for rapid loss of significant amounts of angular momentum. If the material possesses some kind of finite viscosity, angular momentum and matter can be redistributed through frictional interaction that occurs due to differential orbital rotation. Common molecular viscosity has been shown to be too inefficient to account for the observed accretion rates (e.g., Pringle, 1981), and the range of alternative explanations has been broad (Perryman, 2018). Weizsäcker (1948) derived equations of motion for flat, rotating bodies of gas and assumed turbulent viscosity as a source of dissipation. On this basis, Lynden-Bell & Pringle (1974); Pringle (1981) derived a solution for the time evolution of a viscous accretion disk. The

resulting partial differential equation for the evolution of the gas surface density Σ yields

$$\frac{d\Sigma}{dt} = \frac{3}{r} \frac{d}{dr} \left(r^{1/2} \frac{d}{dr} (r^{1/2} \nu \Sigma) \right). \quad (2.10)$$

The modern paradigm is still that turbulent flows, driven by various disk instabilities, account for the bulk of a disk's viscosity (e.g., Balbus & Hawley, 1991; Klahr & Bodenheimer, 2003, 2006; Johansen et al., 2007; Barranco et al., 2018; Gerbig et al., 2020; Schäfer et al., 2020). For many purposes, the exact origin of this viscosity is not relevant and one employs a simple parameterization: In their seminal paper, Shakura & Sunyaev (1973) introduce an effective viscosity

$$\nu = \alpha c_s H, \quad (2.11)$$

where H is the pressure scale height defined above, c_s is the local speed of sound, and the dimensionless parameter α measures the efficiency of turbulence in transporting angular momentum. Order of magnitude estimates of $\alpha = 10^{-4} - 10^{-2}$ have been suggested to meet both protostellar accretion rates and the observed disk dissipation times I discuss below (e.g., Bell & Lin, 1994; Hartmann et al., 1998; Armitage, 2018).

Besides viscous accretion, photoevaporation is another main mechanism through which protoplanetary disks lose material and dissipate. It is driven by highly energetic photons in the range of far-Ultraviolet (FUV), extreme-Ultraviolet (EUV), and X-ray. Such photons can be emitted from a disk's host star, but also from nearby massive stars. The common picture of photoevaporation assumes that EUV photons from the stellar chromosphere dissociate or ionize molecular gas in the disk and heat it to $\sim 10^4$ K (Hollenbach et al., 1994). Where the thermal energy is sufficient that the ions exceed the local escape velocity, a *photoevaporative wind* carries material away. For a solar-mass star, the critical radius for this process lies at roughly 10 au (Clarke et al., 2001). An additional source of ionizing radiation impinging the disk are closeby OB stars, which inevitably occur in cluster environments with hundreds or thousands of stars (Lada & Lada, 2003). Their strong UV radiation erodes mainly the less strongly bound outer parts of disks (Matsuyama et al., 2003). Particularly visually stunning evidence for this mechanism has been obtained from Hubble Space Telescope (HST) observations of the Orion Nebula (Bally et al., 2000).

The contributions of viscous accretion and photoevaporation to the evolution and depletion of a disk vary with time. At early stages, typical accretion rates for solar-mass stars are as high as $10^{-8} M_\odot \text{ yr}^{-1}$ (Armitage, 2019) and the effect of photoevaporation can be neglected. Once the accretion rate becomes comparable to the photoevaporation rate, presumably $10^{-10} M_\odot \text{ yr}^{-1} - 10^{-9} M_\odot \text{ yr}^{-1}$, the inner disk is being cut off from the outer

reservoir close to the critical radius where photoevaporation is most efficient (Clarke et al., 2001). The inner disk then depletes within a viscous timescale ($\lesssim 10^5$ yr), and a hole of a few au diameter forms. With the outer disk now directly exposed to the high-energy radiation from the star, it rapidly evaporates from the inside out (Gorti & Hollenbach, 2009).

The lifetimes of circumstellar disks, and thus the time available for planet formation within them, are determined through the combined effect of the processes I have just discussed. Constraining this important parameter can not be done on a star-by-star basis; instead, one resorts to comparing the fractions of disk-hosting stars in young stellar clusters of different ages. While different definitions exist of what observational signature traces a protoplanetary disk (Mamajek et al., 2009), the classical approach is to identify infrared excess in the SEDs of the objects in question. When the fractions are plotted against cluster age, the data can be well fitted by an exponential decay with a timescale of 2.5 Myr (Mamajek et al., 2009). Challenges of this method are the notoriously difficult-to-determine ages of young stellar clusters (Soderblom et al., 2014) and disentangling the stellar mass dependence (see e.g., Ribas et al., 2014). More recently, Richert et al. (2018) found that these disk longevity estimates are sensitive to the choice of pre-main sequence evolutionary models and suggest a larger uncertainty than previously assumed. It is instructive to put the obtained values into context with the dynamical times in the disk: disk lifetimes of a few Myr imply that these disks live for several millions of dynamical times in the inner regions and thousands of dynamical times at hundreds of au.

2.1.7 Scalings with stellar mass

In many ways, the physical processes in and properties of protoplanetary disks are connected to the (proto-)stars in their center. It is thus not surprising that correlations between properties of disks and their host stars have been established, both theoretically and observationally. A prominent example is the stellar mass-disk mass relation: By converting mm-wave luminosities to solid disk masses M_{dust} , Andrews et al. (2013) found that M_{dust} tends to increase with increasing stellar mass M_* . In a more recent survey of the Chamaeleon I star-forming region, Pascucci et al. (2016) found a steeper than linear relationship and suggest that $M_{\text{dust}} \propto M_*^{1.3-1.9}$. Barenfeld et al. (2016) reported a similar relation for the YSOs in the Upper Scorpius OB Association. However, different power law slopes have been found in star-forming regions of different ages, which could point to an evolutionary process (Pascucci et al., 2016; Ansdell et al., 2017).

The sizes of protoplanetary disks vary with M_* , too. Andrews et al. (2010) reported a weak positive trend of the effective disk radius with disk mass, which could later be

recovered in a combined sample of 105 disks observed with Submillimeter Array (SMA) and ALMA ([Andrews et al., 2018b](#)). Additional, even less constrained scaling relations with stellar mass are suspected in the disk lifetimes ([Ribas et al., 2015](#)) and the positions of the inner disk edges ([Irwin et al., 2008](#)). The stellar mass-dependencies mentioned here will find consideration in Sect. 2.4.5, where I introduce the M dwarf version of our planet formation model.

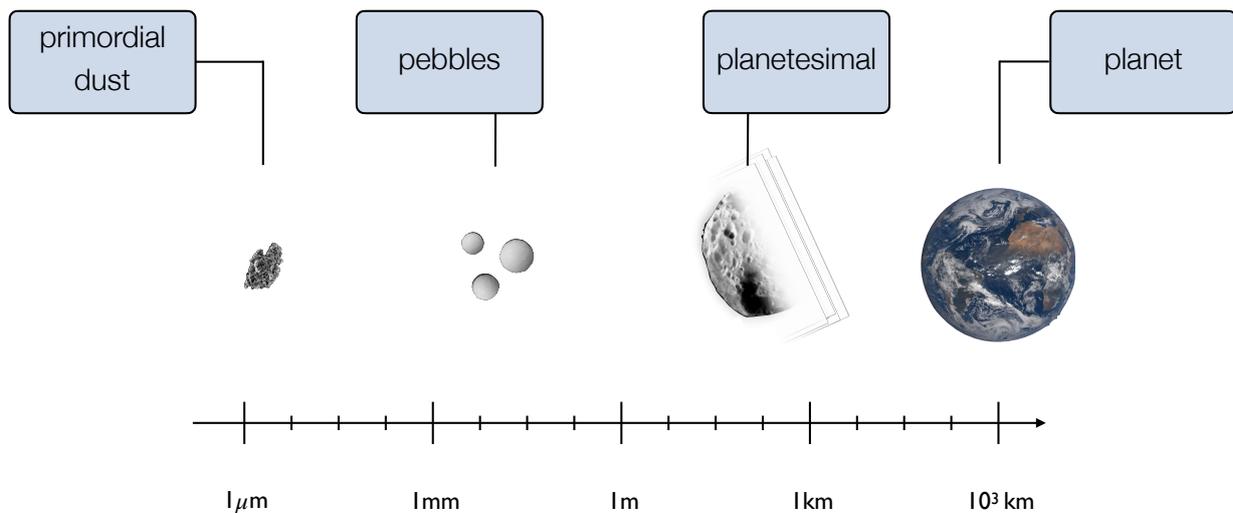


Figure 2.3: Schematic summary of the steps from dust to planets. Primordial dust grains stick together to form larger aggregates. mm to cm-sized pebbles have special aerodynamic properties under typical disk conditions and are observable. Larger compounds evade observations. In the classical core accretion picture, planetesimals in the km size range are the main building blocks to form planetary embryos, which ultimately grow into terrestrial planets or the cores of gas giants. In total, this evolution covers around 13 orders of magnitude in size and 40 orders of magnitude in mass. Image credits: NASA (adapted).

2.2 Planet formation in the core accretion paradigm

The transition from the dilute material in protoplanetary disks to compact, dense planets requires a compression of ten orders of magnitude (Wuchterl, 2007). Two hypotheses to bring about this compression have been put forward:

1. *Gravitational instability*, a “top down” approach that assumes the direct gravitational collapse of disk fragments similar to the Jeans instability (Jeans, 1902) in star formation (Cameron, 1978), and
2. *Core accretion*, where rocky planets and giant planet cores form “bottom up”, and solid planetary embryos act as a seed for additional gravitational growth (Pollack et al., 1996).

The latter is the currently favored scenario and the only one able to explain the full range of the planetary mass function (e.g., Maoz, 2016; Mulders, 2018; Parker, 2020; Adams et al., 2021). In this thesis, I consider only the core accretion paradigm, which I introduce in the following sections.

In a nutshell, planets form in a two-stage process: first, through subsequent coagulation of circumstellar dust to ever larger compounds, kilometer-sized *planetesimals* and, ultimately, a solid planetary core¹ forms (see Fig. 2.3). While for small planets akin to the terrestrial planets in the Solar System this is the end of the story, cores growing to sufficient masses will shift to the next gear: they accrete and retain a significant gaseous envelope. If this envelope reaches a critical mass, additional disk gas is rapidly accreted in a runaway process that is only limited by the local gas reservoir (Perri & Cameron, 1974; Mizuno et al., 1978; Mizuno, 1980; Bodenheimer & Pollack, 1986; Pollack et al., 1996). Core growth must be efficient enough to reach this critical mass within the disk’s lifetime, and this is strongly dependent on the amount of solid material that is available locally (Kokubo & Ida, 1998). This in turn depends on the total mass of solids in a given disk, which can be expressed as a global dust-to-gas ratio (see Sect. 2.1.1). Since host star and disk originate from the same molecular cloud core, a tracer of the solid inventory is the metallicity of the star: stars of higher metallicity form with disks of higher solid content, and thus are more likely to host giant planets. This “metallicity effect” is well established from an observational perspective as an increased giant planet occurrence rate with increasing stellar metallicity (e.g., Gonzalez, 1997; Santos et al., 2001; Fischer & Valenti, 2005). Numerical simulations following the core accretion scheme have reproduced the finding on numerous occasions (e.g., Ida & Lin, 2004b; Mordasini et al., 2009a; Emsenhuber et al., 2020b).

A large body of literature has investigated the various growth stages from primordial dust to the first planetesimals and the challenges in overcoming several growth barriers (e.g., Johansen et al., 2007; Morbidelli et al., 2009; Birnstiel et al., 2012; Drazkowska et al., 2016; Lenz et al., 2019; Hyodo et al., 2019; Klahr & Schreiber, 2020; Ida et al., 2020). In this thesis, I will unabashedly evade these complications and consider only the subsequent phases of planet formation, starting with the growth of already existing planetesimals.

2.2.1 Stages of planetesimal growth

Following the classical paradigm discussed in Raymond et al. (2014), planetesimal growth happens in several subsequent stages. At first, growth rates are low and all planetesimals in a system form concurrently with a mass ratio of order unity. The growth rate of a (large) planetesimal of mass M and radius R that accretes smaller field planetesimals of

¹Note that “core” refers to the complete solid component of a planet; in the case of a giant planet this is everything not consisting of hydrogen and helium. This is not to be confused with the geophysical meaning of “core”, which includes the central iron core but not other components such as a silicate mantle (Mordasini et al., 2012b).

2 Foundations

mass m ; $m < M$, can be written as

$$\frac{dM}{dt} \simeq n_m \pi R^2 \left(1 + \frac{v_{\text{esc}}^2}{v_{\text{rel}}^2} \right) v_{\text{rel}} m. \quad (2.12)$$

Here, n_m is the number density of field planetesimals, v_{rel} is the relative velocity between the large planetesimal and the field planetesimals, and v_{esc} is the escape velocity from the large planetesimal's surface (Kokubo & Ida, 1996). In this equation, the $\frac{v_{\text{esc}}^2}{v_{\text{rel}}^2}$ quantifies the gravitational focusing effect: for sufficiently massive planetesimals, the collisional cross-section is enhanced compared to the geometric cross-section.

A second growth mode occurs when the velocity dispersion of the field planetesimals is smaller than the escape velocity and thus gravitational focusing is efficient. Eqn 2.12 reduces to

$$\frac{dM}{dt} \propto \Sigma_{\text{solid}} M^{4/3} v^{-2}, \quad (2.13)$$

with the surface density and velocity dispersion of planetesimals Σ_{solid} and v (where $v \simeq v_{\text{rel}}$) (Greenberg et al., 1991). As long as the mass in small planetesimals still dominates the system and back-reactions of the growing planetesimal on Σ_{solid} and v can be ignored,

$$\frac{1}{M} \frac{dM}{dt} \propto M^{1/3}, \quad (2.14)$$

leading to *runaway growth* of planetary embryos.

In later stages, known as *oligarchic growth*, a few tens or hundreds of planetary embryos interact gravitationally with each other and dominate the dynamics (Kokubo & Ida, 1998, 2000; Raymond et al., 2014). When the mass of such a massive core exceeds roughly 100 times the average planetesimal mass, the random velocities of the planetesimals in its vicinity increase to $v \propto M^{1/3}$, and the relative growth rate from Eqn 2.14 becomes

$$\frac{1}{M} \frac{dM}{dt} \propto \Sigma_{\text{solid}} M^{-1/3}. \quad (2.15)$$

As more and more planetesimals are accreted and M increases, Σ_{solid} decreases (Lissauer, 1987). The relative growth rate is now a decreasing function in M ; in other words, the growth is self-limiting. This slowing down of the runaway growth leads to the formation of similarly sized planetary embryos (Ida & Makino, 1993). Through the dynamical effect of orbital repulsion (Kokubo & Ida, 1995), two adjacent embryos with masses M_1, M_2 maintain an orbital separation of roughly 10 mutual Hill radii $R_{\text{H,m}}$ to each other (Kokubo & Ida, 1998), where $R_{\text{H,m}} = 1/2 (a_1 + a_2) [(M_1 + M_2) / (3M_\star)]^{1/3}$.

2.2.2 Planetesimal sizes

A debated parameter in the oligarchic growth regime is the dominant size of the planetesimals being accreted (Morbidelli et al., 2009; Weidenschilling, 2011; Krivov & Wyatt, 2021; Klahr & Schreiber, 2021; Venturini et al., 2020b; Voelkel et al., 2020; Burn, 2020). As the growth timescale is sensitive to this parameter, it has a rather large impact on planetary growth. Smaller planetesimals are more affected by gas drag, which damps their orbital inclinations and eccentricities, and the relative velocities between an embryo and the field planetesimals are small. The gravitational focusing effect is enhanced. On the other hand, large planetesimals do not couple strongly to the gas and the effect of damping is negligible. The high relative velocities resulting from gravitational interactions lead to lower accretion efficiencies. Planetesimals formed via the streaming instability have been shown to form large with preferred sizes of roughly 100 km (Klahr & Schreiber, 2020), consistent with impact craters on Pluto and Charon (Singer et al., 2019) and the size frequency distribution (SFD) of asteroids (Morbidelli et al., 2009). However, these results have been challenged by Morbidelli et al. (2021), who combined new data from the *New Horizons* flyby of Pluto, Charon (Singer et al., 2019), and the cold classical Kuiper belt object (KBO) Arrokoth (Spencer et al., 2020) to derive a less top-heavy SFD. These new findings favor the hypothesis that the present-day KBO distribution is the result of significant collisional processing and not indicative of the initial size distribution (Schlichting et al., 2013).

Theoretical simulations have been successful in forming giant planets with large (~ 100 km) planetesimals, as long as they form in situ (Fortier et al., 2009). When planet migration is taken into account, however, planets do not reach sufficient masses before they are accreted onto the star due to the short migration timescales (Fortier et al., 2013, see Sect. 2.2.8). These effects can be consolidated with a planetesimal radius of 300 m (Fortier et al., 2013), which enables the formation of gas giants in disks of sufficient metallicity. We assumed this planetesimal radius in our model (Emsenhuber et al., 2020a). I note that a single planetesimal size is a stark simplification, and thus this poorly constrained quantity should be seen as a *tuning parameter* that we chose to achieve realistic growth timescales.

2.2.3 The planetesimal isolation mass

A very useful concept to study the properties of planetary embryos growing via oligarchic growth is the planetesimal isolation mass M_{iso} (Lissauer, 1993). The idea is that a growing protoplanet, if given enough time, will eventually accrete all planetesimals within an annulus surrounding its orbit, the so-called *feeding zone*. The planet will thereby adopt a mass M_{iso} that depends on the mass of available solid material within this zone, hence on the local planetesimal surface density Σ_{solid} . Given a feeding zone of width $b \simeq 10r_H$,

2 Foundations

where $r_H = a \left(\frac{M_P}{3M_\star} \right)^{1/3}$,

$$M_{\text{iso}} = 2\pi ab \Sigma_{\text{solid}} \simeq 2\pi a 10a \left(\frac{M_P}{3M_\star} \right)^{1/3} \Sigma_{\text{solid}}, \quad (2.16)$$

where a is the semi-major axis of the planetary orbit and M_P is the current mass of the planet.

Assuming for the planetesimal surface density Σ_{solid} a specific functional form, perhaps a truncated power law as in Eqn 2.1, one can estimate the total disk mass needed to reach a predefined core mass at a given orbital distance. In the following, I derive this quantity, which I use in Sect. 3.5.3 to characterize the parameter space in which giant planets can form. Setting the planetary mass to the planetesimal isolation mass, $M_P \equiv M_{\text{iso}}$, yields

$$M_{\text{iso}} = \left(\frac{20\pi}{3^{1/3}} \right)^{3/2} a^3 \Sigma_{\text{solid}}^{3/2} M_\star^{-1/2}. \quad (2.17)$$

To get an estimate on which initial solid mass content is required to reach a certain isolation mass, I may express this as

$$\Sigma_{\text{solid}} = \left(\frac{3^{1/3}}{20\pi} \right) \frac{M_\star^{1/3} M_{\text{iso}}^{2/3}}{a^2}. \quad (2.18)$$

For the power law disk profile used in our model (Andrews et al., 2009, compare Eqn 2.1),

$$\Sigma(r) = \Sigma_0 \left(\frac{r}{r_0} \right)^{-\beta} \exp \left[- \left(\frac{r}{r_{\text{cut,g}}} \right)^{(2-\beta)} \right], \quad (2.19)$$

I consider the outer disk radii $r_{\text{cut,g}}$ and $r_{\text{cut,s}}$ for the gas and solid disk, respectively. The radial slope of Σ_{solid} is characterized by the power law index β , and Σ_0 is the surface density at a reference orbital distance $r_0 = 5.2$ au. Then, the total mass of the planetesimal disk is

$$M_{\text{solid}} = \frac{2\pi \Sigma_0}{r_0^{-\beta}} \frac{r_{\text{cut,s}}^{2-\beta}}{2-\beta}, \quad (2.20)$$

where $r_{\text{cut,s}} = 0.5 r_{\text{cut,g}}$ (following findings from dust disk observations, Ansdell et al., 2018) and $\beta = 1.5$ (motivated by planetesimal formation models, Lenz et al., 2019). Sub-

stituting Eqn 2.18 into Eqn 2.20, the total solid mass required to reach M_{iso} is given by

$$M_{\text{solid}}(M_{\text{P}} = M_{\text{iso}}) = \frac{3^{1/3}}{10} \frac{r_{\text{cut,s}}^{2-\beta}}{2-\beta} \frac{M_{\star}^{1/3} M_{\text{iso}}^{2/3}}{a^{2-\beta}} \exp \left[- \left(\frac{a}{r_{\text{cut,s}}} \right)^{2-\beta} \right]^{-1}. \quad (2.21)$$

2.2.4 The growth timescale

While the isolation mass is a plausible concept, it can only predict the final planet core masses in domains where the time to accrete the solid material is sufficient. The available time to do so is mainly given by the lifetime of the protoplanetary disk (typically a few Myr, see Sect. 2.1.6), which competes with the time needed to accrete all material in a planet's feeding zone. In this section, I briefly introduce this *growth timescale* τ_{grow} and show under what conditions it can be the limiting factor for the final planetary mass.

The time needed by an isolated, static (non-migrating) planet on a circular orbit to clear the annulus of material available to it depends on several factors: the width of that annulus or feeding zone b , which is proportional to the orbital semi-major axis a ; the solid surface density Σ_{solid} at the planet's location; the planetesimal accretion rate, which is in particular a function of the planetesimal size (compare Sect. 2.2.2 Fortier et al., 2013); and the host star mass M_{\star} . Based on N -body simulations by Kokubo & Ida (2002), Mordasini (2018) approximate τ_{grow} for the oligarchic growth regime by

$$\begin{aligned} \tau_{\text{grow}} \approx & 1.2 \times 10^5 \text{yr} \left(\frac{\Sigma_{\text{solid}}}{10 \text{ g cm}^{-2}} \right)^{-1} \left(\frac{a}{1 \text{ au}} \right)^{1/2} \left(\frac{M_{\text{c}}}{M_{\oplus}} \right)^{1/3} \left(\frac{M_{\star}}{M_{\odot}} \right)^{-1/6} \\ & \times \left[\left(\frac{\Sigma_{\text{g}}}{2400 \text{ g cm}^{-2}} \right)^{-1/5} \left(\frac{a}{1 \text{ au}} \right)^{1/20} \left(\frac{M_{\text{pla}}}{10^{18} \text{ g}} \right)^{1/15} \right]^2. \end{aligned} \quad (2.22)$$

Similarly to the derivation of $M_{\text{solid}}(M_{\text{P}} = M_{\text{iso}})$ above, one can derive the solid disk mass needed to reach a specific core mass in the outer disk regions, where growth is mainly limited by the growth timescale τ_{grow} . Solving Eqn 2.22 for Σ_{solid} and substituting into Eqn 2.20 gives

$$\begin{aligned} M_{\text{solid}}(a, \tau_{\text{grow}}) = & 7.54 \text{ g cm}^{-2} \frac{r_0^{\beta} r_{\text{cut,s}}^{2-\beta}}{2-\beta} \left(\frac{M_{\text{c}}}{M_{\oplus}} \right)^{1/3} \left(\frac{M_{\star}}{M_{\odot}} \right)^{-1/6} \\ & \times \left[\left(\frac{\Sigma_{\text{g}}(a)}{2400 \text{ g cm}^{-2}} \right)^{-1/5} \left(\frac{M_{\text{pla}}}{10^{18} \text{ g}} \right)^{1/15} \right]^2 \left(\frac{\tau_{\text{grow}}}{1 \text{ Myr}} \right)^{-1} \left(\frac{a}{1 \text{ au}} \right)^{3/5}, \end{aligned} \quad (2.23)$$

where Σ_{g} was computed using the median of the reference surface density $\Sigma_{0,\text{gas}}$ in the

fiducial $1 M_{\odot}$ synthetic planet population I use in this thesis (see Sect. 2.5.1). For the cutoff radii of the gas and solid disk, I proceeded in the same way and assumed the population median, respectively. For the planetesimal mass M_{pla} , I assumed a density of 1 g cm^{-3} , which results in $M_{\text{pla}} = 1.13 \times 10^{11} \text{ kg}$ for the planetesimals in our model (Emsenhuber et al., 2020a). I further adopted a core mass M_c of $10 M_{\oplus}$.

I will use the analytic boundary conditions involving M_{iso} and τ_{grow} that I derived here in Sect. 3.5.3, where I compare them to the regions in disk parameter space where certain types of planets occur in our formation model.

2.2.5 Terrestrial planet formation

For massive oligarchs the size of Mercury or the Moon, the story is not over with the oligarchic growth phase. En route to terrestrial planets such as Mercury, Venus, Earth, and Mars, a final stage of formation involves long-term dynamical instabilities (Armitage, 2007). As long as the oligarchs have little mass compared to the surrounding smaller bodies, a balance exists between viscous stirring, where eccentricities and inclinations in the planetesimal disk are increased (Ohtsuki et al., 2002), and dynamical friction, which lowers the relative velocities (and thus the eccentricities and inclinations) of the largest bodies (Stewart & Wetherill, 1988). As the surface density of oligarchs becomes comparable to that of the field planetesimals, dynamical friction can no longer compensate viscous stirring (Goldreich et al., 2004). The random velocities, orbital eccentricities, and inclinations of the planets increase and orbit crossings may occur. In the inner parts of the system, this leads to frequent encounters; planets collide and merge. At large orbital distances, ejections of oligarchs from the system can occur. This *giant impact phase* is the main reason for including computationally expensive N -body calculations in planet formation models, and for running them as long as is practical beyond the disk phase. While the minimum numerical time for realistic results remains debatable, the stage effectively ends when the number of large bodies in the system is reduced enough such that their gravitational interactions do not contribute anymore to further instability. Finally, dynamical friction with remaining planetesimals damps the eccentricities and inclinations of the planets, and circular, coplanar orbits as in the Solar System can occur. Compared to the previous stages, the giant impact phase is chaotic - similar initial conditions can lead to very different outcomes. I will study this issue in detail in Chapter 3, which deals with the links between the initial properties and final architectures of simulated planetary systems.

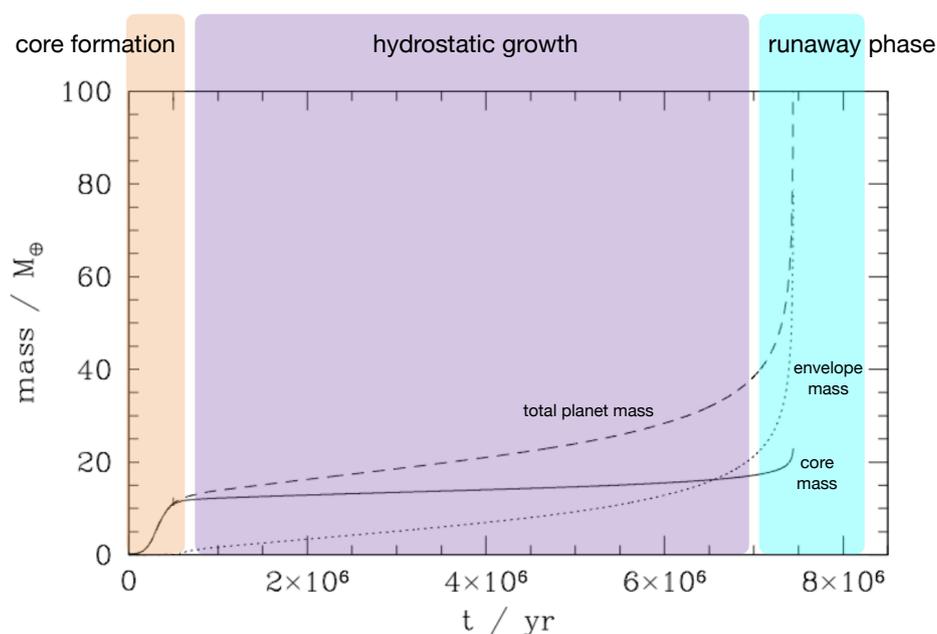


Figure 2.4: Accretion phases of giant planets in the core accretion paradigm. After the solid core has formed on a timescale $\lesssim 1$ Myr, a long phase of slow gas and planetesimal accretion in hydrostatic equilibrium follows. Roughly when the envelope mass becomes comparable to the core mass, a positive accretion feedback of the contracting envelope triggers *runaway gas accretion*. Figure adapted from Pollack et al. (1996) and lecture notes by Philip J. Armitage².

2.2.6 Giant planet formation

Where planetesimal accretion can overcome the limiting factors outlined above, a core can grow massive enough to accrete large amounts of gas from the disk. Significant amounts of gas start to be accreted roughly at embryo masses of $10 M_{\oplus}$ (Piso & Youdin, 2014), and once the gas mass becomes comparable to the core mass, hydrostatic equilibrium is lost and *runaway gas accretion* sets in (Perri & Cameron, 1974; Mizuno et al., 1978; Mizuno, 1980; Bodenheimer & Pollack, 1986; Pollack et al., 1996). In this growth phase, illustrated in Fig. 2.4, the gaseous envelope contracts under its own gravity, releasing energy. Additional gas can be accreted within the Bondi radius, further increasing the mass and gravitational pull of the planet (see Raymond & Morbidelli, 2020). A positive feedback arises and the gas envelope mass increases exponentially with time, allowing the accretion of hundreds of M_{\oplus} and forming a gas giant planet. The accretion ceases only when the disk disperses (see Sect. 2.1.6) or when the planet forms a deep gap in the disk that shields it from the gas supply (Lin & Papaloizou, 1979, 1980; Crida et al., 2006).

²<https://solarsystemorigins.wordpress.com/2013/11/10/lecture-12-giant-planet-formation> received on Apr. 13, 2021.

2.2.7 Pebble accretion

Reaching the critical core mass for runaway gas accretion via the accretion of planetesimals can be challenging, especially when planet migration is taken into account (see Sect. 2.2.8). Depending on the choice of planetesimal size distribution, planets are at risk of migrating into the star before accreting sufficient solids (Mordasini et al., 2009a; Fortier et al., 2013; Ronco et al., 2017). An emerging paradigm of core accretion might provide a solution to this problem: as I motivated in Sect. 2.1.5, particles with Stokes numbers around unity (\gtrsim mm in size) tend to drift radially inwards for typical radial disk profiles. An approximately stationary planet thus sees a flux of these *pebbles* passing by. Since particles of this size are prone to aerodynamic drag, they efficiently settle within a large fraction of a planet’s Hill sphere and are accreted (Ormel & Klahr, 2010; Ormel, 2017). This is in contrast to larger planetesimals, which experience a much smaller collisional cross section. Depending on the (aerodynamic) situation of the pebbles in their immediate environment (see Ormel, 2017), the accretion of pebbles can be considerably faster than that of planetesimals. Pebble accretion can thus significantly decrease the time needed to grow a planetary core (Lambrechts & Johansen, 2012, 2014; Johansen et al., 2019).

Why, then, do we discover so many super-Earths and Neptune-like planets of a few M_{\oplus} (see Sect. 2.4) that appear to have stopped their growth at intermediate core masses? Several mechanisms can halt the further accretion of solids onto a planet of such a mass. If another planet forms concurrently further out, it can become a sink for solid material that would otherwise drift to the inner system. This happens when the planet becomes so massive that the fraction of the pebble flux accreted by the planet ϵ_{PA} approaches unity (Ormel et al., 2017). This fraction can be approximated by

$$\epsilon_{\text{PA}} \approx 0.1 \times \left(\frac{q}{10^{-5}} \right)^{\frac{2}{3}}, \quad \text{with } q = \frac{M_{\text{Planet}}}{M_{\star}}, \quad (2.24)$$

and thus the criterion is well in range of, for instance, super-Earths around low-mass stars. However, since the timescale for accreting some M_{\oplus} of pebbles is only on the order of $\sim 10^5$ yr (Bitsch et al., 2019a), this process is rather sensitive on the timing of the flux cut-off. In a scenario with less temporal constraints, the planet in question reaches its *pebble isolation mass*, where it locally alters the gas pressure gradient such that the inward drift of pebbles is halted, which in turn starves the planet from further solid accretion (Morbidelli & Nesvorniy, 2012; Lambrechts et al., 2014; Bitsch et al., 2018b; Ataiee et al., 2018).

Generally spoken, a key difference between classical planetesimal accretion models and pebble accretion is the mobility of solid material within the disk. While the planetesimal

surface density a planet can tap on is largely set by the initial disk profile, the radial transport (or obstruction) of solids in the pebble accretion picture has a profound impact on the development of planetary systems. Taking into account or ignoring any of the two processes thus leads to different outcomes of the final system architectures (e.g., Brügger et al., 2020). In Chapter 4, I investigate the relations between close-in and distant planets, and I will discuss the potential influence of pebble accretion on the results of this analysis.

2.2.8 Planet migration

A variety of different planet-disk interactions have been studied, ranging from the opening of gaps in gas disks (Lin & Papaloizou, 1979; Papaloizou & Lin, 1984) or dust disks (e.g., Paardekooper & Mellema, 2006), to planet-induced spiral density waves (e.g., Goldreich & Tremaine, 1979; Goldreich & Tremaine, 1980). In this thesis, I shall confine myself to a case with particularly far-reaching consequences for the architectures of planetary systems: the *migration* of planetary orbits through angular momentum exchange with the surrounding disk. Planets may change their orbital semi-major axes due to gravitational torques between the gas disk and the planet, which occur at the Lindblad and corotation resonances (Goldreich & Tremaine, 1979). Low-mass planets ($\sim 0.1 M_{\oplus} - 10 M_{\oplus}$) perturb the disk only weakly and experience migration in the *Type I* regime, where the torques scale as $\propto M^2$ (see Armitage, 2018) and the resulting drift timescale is inversely proportional to their mass (Ward, 1997). Lindblad torques are proportional to the local gas surface density and contribute exclusively to *inward* migration. On the other hand, corotation torques have a complex dependency on the structure and thermodynamics of a disk and, if they outperform the Lindblad torques, can lead to outward migration (Paardekooper et al., 2011).

More massive planets ($\gtrsim 3(H/r)^3 M_{\star}$, Armitage, 2018) exert a torque on the disk inside and outside of their location, which in sum pushes gas away from their orbits (e.g., Kley & Nelson, 2012). The resulting gap is maintained if this torque is larger than the viscous torque that tends to refill the void (Lin & Papaloizou, 1979, 1986). Trapped in the gap, the planet migrates in the *Type II* regime with a rate similar to the viscous timescale (Lin & Papaloizou, 1986; Ward, 1997; Dürmann & Kley, 2017, however, also see Kanagawa et al. (2018)).

As planet migration can lead to large-scale displacements of planets during their formation phase, it has considerable impact on their final semi-major axes, masses, and compositions (e.g., Kley & Nelson, 2012; Venturini et al., 2020b).

2.3 Planet observations

2.3.1 Characterization of planets in the Solar System and beyond

Now that we have a basic understanding of how planets come into being, what do we see of them? Only for a single planetary system one could argue that it was characterized “in detail”, even though an infinity of open questions prevail. For all planets in the Solar System, we have resolved images in reflected light at the minimum, and some of them were investigated in situ in various ways and levels of detail (e.g., [Fjeldbo et al., 1966](#); [Neugebauer & Snyder, 1966](#); [Avduevskij et al., 1971](#); [Toulmin et al., 1977](#); [Bolton et al., 2017](#)). It will still be a while for the first geologist to set foot on another planet than Earth, but the Apollo program has yielded a rich treasure of samples at least from our own Moon ([Papanastassiou et al., 1970](#); [Jolliff & Robinson, 2019](#)). Beyond that, in situ measurements at and samples from minor bodies such as asteroids and comets provide insights about the building blocks of the Solar System (e.g., [Sierks et al., 2015](#); [Galiano et al., 2020](#)). Unsurprisingly, our conception of planet formation had been based on the observations in the Solar System, before the first exoplanets could be detected. Unfortunately, the data we can obtain about these planets is much less detailed, which is of course due to their much greater distance from us observers. As pointed out already quite a long time ago by [Giordano Bruno \(1584\)](#), these then hypothetical planets around the fixed stars are too small and faint compared to their hosts to be observed. While today this spell is partially lifted and exoplanets can be directly imaged under specific conditions (e.g., [Marois et al., 2008](#); [Lagrange et al., 2010](#); [Macintosh et al., 2015](#); [Chauvin et al., 2017](#)), the majority of them are detected and characterized via indirect methods not involving resolved images. On the following few pages, I will introduce those methods relevant for the research presented in this thesis before showcasing some insights from *statistical* investigations of the exoplanet population as a whole in Sect. 2.4.

2.3.2 Orbital parameters

The properties of a planet’s orbit are among the most sought-after attributes of exoplanets, and to a large degree they determine whether it is detectable or not. Different ways of defining orbital elements for Keplerian orbits exist; I follow here the scheme in the review by [Perryman \(2018\)](#). A Keplerian orbit in three-dimensional space is characterized by seven parameters³:

³In some definitions, the relation between a and P is implied and the number of required orbital elements reduces to six.

Orbital period P

The time it takes the planet to complete one orbit. P is a direct observable in methods relying on the detection of cyclic signals such as planetary transits or RV signatures.

Semi-major axis a

The semi-major axis of the ellipse the orbit follows. a is related to the orbital period P through Kepler's third law.

Eccentricity e

A unit-less parameter describing the shape of the orbit, in particular its non-circularity. For bound orbits, $0 \leq e < 1$.

Epoch t_p

Position of an object on its orbit at a reference time, e.g., its periastron passage. The following angular quantities project a physical orbit into an apparent one; all three depend only on the orientation of an observer relative to the orbit.

orbit inclination i

The angle between the reference plane (plane of the sky) and the orbital plane. $i = 0^\circ$ for a face-on orbit.

Longitude of the ascending node Ω

Angle to the ascending node, which is the point on the orbit where the object crosses the reference plane while moving away from the observer.

Argument of pericentre ω

angular position of the pericenter with respect to the ascending node in the orbital plane. It is measured in direction of motion. ω becomes arbitrary for $e = 0$.

2.3.3 The radial velocity technique

When a planet orbits a star, it causes a reflex motion around the common center of mass or *barycenter*. This periodic movement of the star can be measured in three ways, all corresponding to their own exoplanet detection techniques:

1. Measuring the change of the arrival time of a periodic reference signal that is caused by the change in light travel time, e.g. *pulsar timing*. In fact, the first extrasolar planets, albeit orbiting a rather exotic host, were found with this method (Wolszczan & Frail, 1992). They orbit the millisecond pulsar PSR-B1257+12 b in orbits with periods of 98.2 and 66.6 d.

2 Foundations

2. Obtaining time series of the angular position on the sky plane (*astrometry*). This technique, while reserved for nearby stars, has the unrivaled advantage of enabling the solution of the complete orbital elements, including information on the coplanarity of multi-planet systems. While only few planets have been discovered with this method to date, the astrometric detection of presumably thousands of planets by the Gaia mission is imminent through its microarcsec-precision all-sky survey (Perryman et al., 2014).
3. A complementary method to astrometry is measuring the star's displacement *along* the line of sight, the Doppler or *RV* technique. A large fraction of the exoplanets detected so far have been discovered with this method, and it is often used to confirm transiting planet candidates.

Here, I will focus on the *RV* method, as it will be relevant for Chapters 4, 5, and 6. As a star with a planetary companion moves around their common center of mass, its velocity along our line of sight varies, and its spectral lines will be subject to periodic Doppler shifts. The detection method therefore involves repeated spectroscopic observations from which one extracts *radial velocity* measurements. For the purpose of planet detection, the total values of these measurements are irrelevant - what one is after are periodic signals in the *RV* data, which may reveal an orbiting companion (but several false positive scenarios exist). While in the case of a planet on a circular orbit such signals are sinusoidal, for eccentric orbits they can assume a range of shapes depending on the specific values of e and ω (Perryman, 2018). In other words, the signal shape also depends on the orientation of the viewer. If the star hosts a *planetary system*, its *RV* time series are a superposition of the radial motions induced by all orbiting planets, where each signal has its own amplitude and period.

The radial velocity semi-amplitude

How strong is this signal for realistic orbital and planetary parameters? In its most common expression, the radial velocity semi-amplitude is given by

$$K = \left(\frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_\star + M_p)^{2/3}} \frac{1}{(1 - e^2)^{1/2}}, \quad (2.25)$$

where M_\star and M_p are the stellar and planetary mass, respectively, G is the gravitational constant, P is the period of the planet's orbit, e is its eccentricity, and i is its inclination relative to our line of sight (Cumming et al., 1999). If the mass of the star M_\star can be estimated from its spectral type, $M_p \sin i$ can be determined. However, since the inclination

can not be constrained by measuring radial velocity alone, an unknown factor $\sin i$ remains and only a lower limit to the planet mass can be obtained. In the special case of a circular orbit and $M_p \ll M_*$, the amplitude of the then sinusoidal velocity variations can be expressed as (Torres et al., 2008)

$$K = 28.4 \text{ ms}^{-1} \left(\frac{P}{1 \text{ yr}} \right)^{-1/3} \left(\frac{M_p \sin i}{M_J} \right) \left(\frac{M_*}{M_\odot} \right)^{-2/3}. \quad (2.26)$$

As is evident from this relation, the amplitude of a planetary Doppler signal is sensitive on a planet's mass, inclination, and orbital distance, as well as on the stellar mass. A hypothetical Jupiter-mass planet on a 1 d orbit around a solar-type star induces a signal with an amplitude of 0.2 km s^{-1} for $i = 90^\circ$ (Struve, 1952).

Periodogram analysis

A central step in the search of planetary signals in RV time series is the search for periodic signals in the data, which are typically unevenly spaced and noisy. One well-known method is the Lomb-Scargle (LS) periodogram (Lomb, 1976; Scargle, 1982), which can be described as an equivalent to least-square fitting of sine waves. However, since RV measurements have finite errors and do not generally follow a sinusoidal curve with zero mean, the method has been shown to be non-robust for planet searches (Cumming et al., 1999). The Generalized Lomb-Scargle periodogram (GLS, Zechmeister & Kürster, 2009) avoids some of its shortcomings and includes weights for measurement errors as well as a treatment for data offsets (Pinamonti et al., 2017). It provides a means for period searches of the best-fitting Keplerian orbit, even if it is eccentric. Currently, the GLS is the most widely used method for period searches of planetary signals in RV time series, and I use it in the periodogram analysis of TIC 237913194 in Chapter 5.

2.3.4 Transit photometry

As of the writing of this thesis, the most successful exoplanet detection technique by numbers has been the *transit method*. In contrast to the RV technique, it does not rely on measuring the *color* of the light emitted by a star, but rather on precisely estimating its *amount*. A transit occurs when an opaque body passes between a star and an observer, who will notice a decrease in brightness of the star. When this body is an orbiting planet, it causes periodic repetitions of these dimmings, and they show a characteristic profile defined by the shapes, sizes, and orbital configurations of the two bodies. How well this light curve profile can be reproduced depends primarily on the photometric noise level and on the ca-

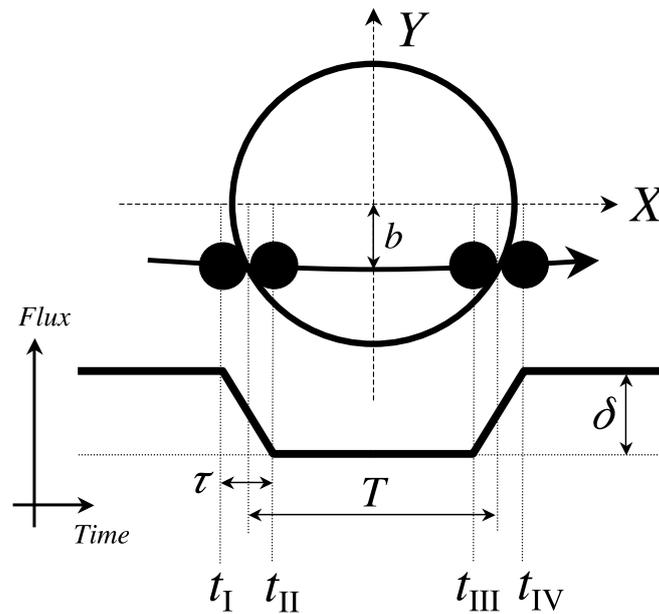


Figure 2.5: Light curve profile of a transiting planet. The transit shape is characterized by four *contact points* I to IV, at which the time derivative of the flux has a discontinuity. The transit duration T and the durations of the ingress and egress phases τ are determined by the impact parameter b , which depends on the orbital inclination. For a planet in full transit ($b < 1/2 R_*$), the transit depth δ is governed by the ratio of the planetary and stellar radius. For Earth transiting the Sun, $\delta \approx 84$ ppm (Koch et al., 2010). Figure reproduced from Winn (2010) with kind permission by the author.

dence (the average time interval between successive exposures) of the observations. The key requirement for observing a transit at all is that we observe the planetary orbit edge-on, that is, with an inclination $i \approx 90^\circ$. The probability for such a configuration is small and a strong function of orbital distance ($p \propto a^{-1}$, Borucki & Summers, 1984).

Observables

While transit photometry has been a fruitful planet discovery method, it can also constrain planetary and orbital parameters. Figure 2.5 is a schematic representation of a light curve showing a single transit. This is the default case: the planet is significantly smaller than the star and makes a full transit, that is, there exists a phase when the planetary disk lies completely within the stellar disk⁴. The shape of the transit light curve can then be characterized by the contact points I to IV. At the points I and IV, the sky-projected distance

⁴Since we consider exoplanet transits, it is also reasonable to assume that the distance between star and planet is much smaller than the distance to us observers. The transit depth is thus independent of the orbital distance of the planet.

between planet and star equals the sum of their radii, $R_* + R_P$, and at the points II and III the distance equals their difference, $d = R_* - R_P$ (Winn, 2010). Obviously, the amount of stellar flux removed by the planet is a function of its size. However, from the transit depth δ one can determine only the *ratio* of the planetary and stellar radii R_P/R_* . This highlights the importance of precise stellar radius determinations. Expressed in terms of Earth radii R_\oplus and Solar radii R_\odot , the transit depth

$$\delta = 84 \text{ ppm} \left(\frac{R_P}{R_\oplus} \right)^2 \left(\frac{R_*}{R_\odot} \right)^{-2}. \quad (2.27)$$

In real observations, the light curve shape is altered by stellar limb darkening, which decreases the flux toward the edge of the stellar disk. Now, the transit depth in full transit is not constant as in Fig. 2.5, but instead the bottom of the trough is rather U-shaped. Limb darkening thus has to be taken into account when fitting light curve models to transit photometry. The stellar intensity profile can be approximated by a variety of functional forms; linear (Schwarzschild & Villiger, 1906), quadratic (Kopal, 1950), square-root (Diaz-Cordoves & Gimenez, 1992), logarithmic (Klinglesmith & Sobieski, 1970), and four-parameter nonlinear laws (Claret, 2000) are common and the best choice largely depends on the photometric noise level (Espinoza & Jordán, 2016).

An additional transit observable is the time of mid-transit T_0 , which fixes the phase of the planet on its orbit. If several transit events of the same planet are observed, the time interval between them directly gives the orbital period P . For an in-depth discussion of transit light curves and higher-order photometric effects, see Perryman (2018).

Eccentric orbits

When the orbit of a transiting planet has non-zero eccentricity e , the shape of the transit profile becomes a function of orbital phase. In particular, the transit duration changes by a factor $(1 + e \cos \omega)/(1 - e^2)^{1/2}$ compared to an equivalent circular orbit (Ford et al., 2008). Depending on the argument of periapsis ω at which we observe the transit, this factor can be less than unity (close to pericenter) or larger than unity (close to apocenter). Unless the transit occurs at peri- or apocenter, the transit light curve will further be asymmetric due to different ingress and egress durations. However, even under favorable conditions such as for HD 147506b ($e = 0.507$, transit at true anomaly $f_0 = 4.64$ rad (Bakos et al., 2007)) the asymmetry is of order 10^{-6} and barely detectable. Furthermore, planets on eccentric orbits are more likely to transit by a factor of $(1 - e^2)^{-1}$ (Barnes, 2007).

Light curve fitting

To extract physical parameters from photometric observations, a theoretical light curve model is fitted to the photometric time series. In the common Bayesian approach, posterior probability distributions of these parameters are obtained by approximating a likelihood function numerically with a random sampling algorithm such as Markov Chain Monte Carlo (MCMC) or nested sampling. One can also combine datasets from multiple observational techniques and instruments and simultaneously fit parameters of a combined model (e.g., Espinoza et al., 2016; Günther & Daylan, 2021). This global approach can even be extended by jointly fitting for stellar properties (e.g. Beatty et al., 2012; Eastman et al., 2013; Hartman et al., 2019) and the underlying photometric and RV noise model (e.g. Sarkis et al., 2018; Cañas et al., 2019; Luque et al., 2019; Kossakowski et al., 2019; Espinoza et al., 2020; Schlecker et al., 2020a; Kemmer et al., 2020, and Chapter 5).

The TESS telescope

The photometric data leading to the exoplanet discovery I present in Chapter 5 was obtained with the *Transiting Exoplanet Survey Satellite* (*TESS*, Ricker et al., 2014). In contrast to *Kepler*, which observed a single 115 square degree field of view throughout its primary mission (Borucki et al., 2010), *TESS* has observed almost the entire sky at 30 min cadence and with observing baselines of 27 d – 351 d in its all-sky survey. The survey was split into 26 partly overlapping sectors of 24×96 degree, each of which was observed for about one month by the spacecraft’s four telescopes. Figure 2.6 illustrates this scanning strategy and the resulting monitoring times of different celestial regions. The primary goal of this survey was to provide a sample of small, transiting planets around bright stars for spectroscopic follow-up observations that provide constraints on their masses and atmospheres. Perhaps the largest contribution of *TESS* to our understanding of planet formation will be a significant increase in the number of planets smaller than $4 R_{\oplus}$ with measurements of both their mass and their radius, which is accompanied by constraints on their bulk density (Barclay et al., 2018). Besides this primary objective, the *TESS* all-sky survey and its extended mission is contributing to the field of exoplanet demographics by covering a wide range of stellar properties (Scott Gaudi et al., 2020).

To detect and characterize transit signals, which have typical durations on the order of hours, a sufficient sampling rate is required. The stars in the *TESS* Input Catalog (TIC, Stassun et al., 2018) have each been measured every ~ 2 min during observations of their sector. Photometry of the remaining majority of stars is available from full-frame images, which were observed at a 30 min cadence (Guerrero et al., 2021). The *TESS* data presented

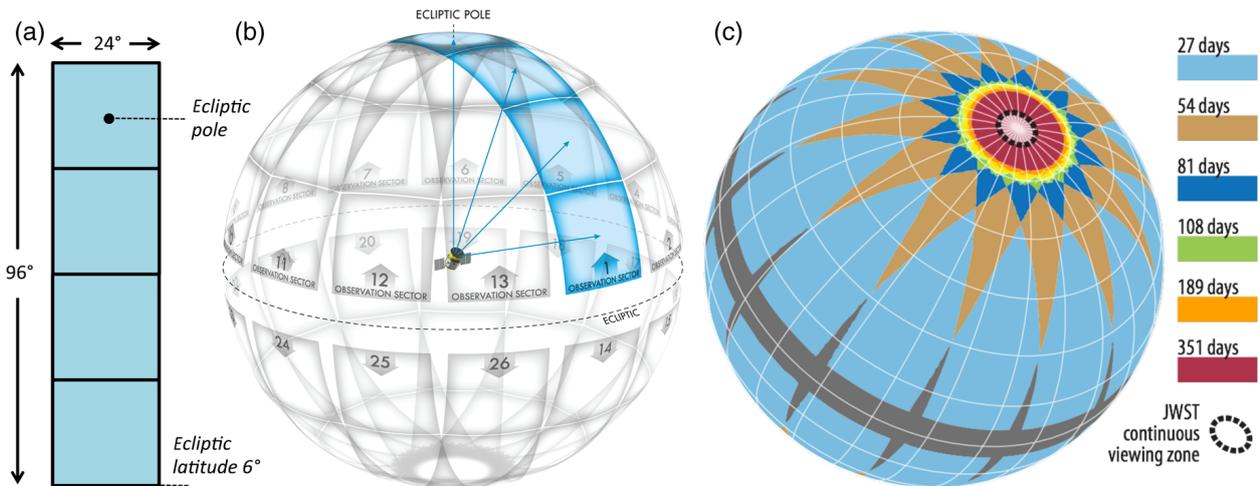


Figure 2.6: Field of view and scanning strategy of the *TESS* primary survey. a) and b): four cameras cover each of 26 observation sectors, which partially overlap and are each observed for 27 d. c): duration of the observations as a function of sky coordinates. The coverage depends on the number of overlapping sectors and is highest in the continuing viewing zone (dashed line) of the *James Webb Space Telescope* (*JWST*, [Beichman et al., 2014](#)), where it will be able to observe at any time. Figure originally published in [Ricker et al. \(2014\)](#) under a Creative Commons Attribution 3.0 Unported License.

in Chapter 5 is based on this observing mode.

2.4 Exoplanet demographics

The exoplanet discovery and characterization methods outlined above are able to provide constraints on the properties of planets, their orbits, and their host stars. While interesting on their own, the real power of these parameters for exploring planet formation physics comes into play when their statistical distributions can be determined. This is the craft of exoplanet demographics, some of which most interesting findings I would like to present here.

2.4.1 Biases and caveats

Several known caveats limit the validity of demographic studies, and most of them can be categorized as either *selection effects* that arise from the selection of targets with certain properties, or *detection biases* where the likelihood of a detection varies with (astrophysical or instrumental) properties. Selection effects that may impede or bias demographic studies include:

- targeting of inactive stars with low “RV noise”, particularly closeby and/or popular stars (e.g., [Anglada-Escudé et al., 2016](#); [Ribas et al., 2018](#))

2 Foundations

- selection of stars with higher a-priori occurrence or detection probability, e.g., low-mass stars where RV signals are enhanced (Kürster et al., 2003, compare Eqn 2.25) and where transit signals are stronger due to the smaller stellar radius
- avoidance of giant stars due to pulsations that can mimic planetary signals (but see Frink et al., 2002; Zechmeister et al., 2008). The fact that highly eccentric planets can be better disentangled from such stellar signals may add another bias in the eccentricity distribution.
- avoidance of binary systems
- changes of the observing strategy (e.g., increasing the number of measurements when a candidate signal is found or stopping to observe a star that turned out to be active)
- non-consideration of the Galactic environment, i.e., stellar clustering (Winter et al., 2020)
- non-consideration or biased sampling of stellar age, which can affect demographic features (e.g., David et al., 2020)
- biases introduced by the subsequent data analysis, for example at “detrending” photometric or RV time series.

Most of these issues can be mitigated by a well-defined (e.g., volume complete) target sample, or at least by a clear definition of the selection function and rigorous tracking (or better avoidance) of strategic decisions during an ongoing survey.

Detection biases, on the other hand, are inherent to the discovery method. I list here only some of the most impactful effects on exoplanet demography. For instance, the transit method is strongly affected by the geometrical transit probability, which for a planet on a randomly oriented, circular orbit around a star with radius R_* amounts to (Borucki & Summers, 1984)

$$p = \frac{R_*}{a} \simeq 0.005 \left(\frac{R_*}{R_\odot} \right) \left(\frac{a}{1\text{au}} \right)^{-1}. \quad (2.28)$$

This probability drops steeply with orbital distance, which gives rise to a strong bias toward short-period planets. Adding to this is a bias introduced by the time baseline of the observations, as usually three transits are required for a detection. This affects in particular ground-based surveys due to non-continuous monitoring (e.g., Perryman, 2018; Gibbs

et al., 2020). In the dimension of planetary radius, the sensitivity is limited through the minimum difference in relative flux that can still be detected in photometric data (Gaidos & Mann, 2013, compare Eqn 2.27).

RV searches, too, are biased toward massive planets on short orbits. The amplitude of the signal we are after scales with the period as $K \propto P^{-1/3}$ (compare Eqn 2.25), and the available time baseline of the observations is further limiting the maximum period at which a planet can be securely detected. The RV semi-amplitude is also linearly dependent on the planetary mass. Assuming a measurement precision of 1 m s^{-1} (Fischer et al., 2016), the just detectable threshold mass of a planet on a circular orbit around a Sun-like star that we observe edge-on ($i = 90^\circ$) is $10 M_\oplus$ (Reiners et al., 2018b). RV discoveries are therefore, similar to transits, biased toward more massive planets on short orbits.

An additional complication inherent to the RV method is the degeneracy between a planet’s mass and its orbital inclination – without complementary information, the M_P and $\sin i$ terms in Eqn 2.25 cannot be separated (Perryman, 2018). If planetary orbits are distributed isotropically, the true mass is unlikely to be much larger than the observable minimum mass (e.g., Mordasini, 2008). However, while an isotropic *prior* distribution of orbit orientations may be justified, this does not mean that the *posterior* distribution of $\sin i$ for a particular observed system is an isotropic one as it depends on the intrinsic distribution of M_P (Ho & Turner, 2011). Care must thus be taken when measured minimum masses are translated to true masses.

2.4.2 A census of exoplanets

One of the first relations found in the then still sparsely populated exoplanet sample was a correlation between the metallicity of a star and its likelihood to host a giant planet (e.g., Santos et al., 2004; Fischer & Valenti, 2005; Johnson et al., 2010): the occurrence rate of relatively short-period giant planets increases with increasing metallicity. Core accretion theory readily explains this “**metallicity effect**” with an enhanced core growth in metal-rich disks, which in turn facilitates reaching the critical core mass for runaway gas accretion (e.g., Mordasini et al., 2009b).

With the current census of more than 4000 confirmed exoplanets it is now feasible to statistically evaluate their observables. Figure 2.7 shows parameter distributions of exoplanets discovered by the most common exoplanet detection techniques. The RV method is sensitive to a planet’s mass and populates most of the mass-period diagram. On the other hand, the transit method provides planetary radii and contributes the majority of planets in the radius-period diagram. Most of the transiting planets have been discovered by the *Kepler* mission.

2 Foundations

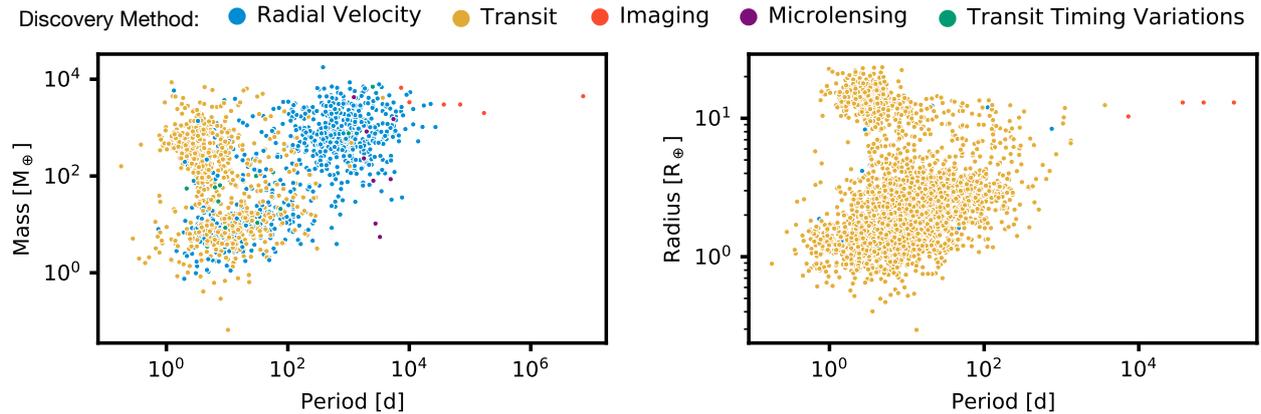


Figure 2.7: Distributions of currently confirmed exoplanets⁵ in mass-period space (left, includes all planets with a mass measurement, and for RV-inferred minimum masses, $\sin i = 1$ was assumed) and radius-period space (right, includes all planets with a radius measurement). With a current total planet count of 4375, the large diversity of planets is apparent. These diagrams are subject to observational biases that depend on the discovery method. For instance, the emptiness of the lower-right corners is entirely due to the low sensitivities in this parameter region. Similarly, the apparent abundance of “hot Jupiters” (upper left corners) is merely caused by the strong decline of the transit probability with orbital distance and these planets are strongly over-represented in the detections (Scott Gaudi et al., 2020).

While the distribution of detected exoplanets is strongly affected by observational biases and is not representative of the true underlying planet population, several *real* features can immediately be identified from the plots:

1. The detected exoplanet sample shows a large variety.
2. There is structure in the distributions.
3. Different detection methods cover different parameter regions.

2.4.3 Features in the exoplanet population

In orbital period ranges where the transit and RV methods are sensitive, both diagrams show a scarcity of intermediate masses and sizes around $\sim 10 M_{\oplus} - 100 M_{\oplus}$ and $\sim 4 R_{\oplus} - 10 R_{\oplus}$, respectively (Mayor et al., 2011). This region of low occurrence was predicted from core accretion models, where this mass range corresponds to planets with ongoing runaway gas accretion (Mordasini et al., 2009b). As this process happens on a short timescale and planets would only attain such final masses if their parental gas disk disperses within a narrow time window, their occurrence is reduced. It should be

⁵Based on data from the NASA exoplanet archive, <https://exoplanetarchive.ipac.caltech.edu>, queried on 2021-04-29.

noted that microlensing surveys, which are most sensitive to larger orbital distances of a few au (Scott Gaudi et al., 2020), are in contrast with these findings. Indeed, based on the Microlensing Observations in Astrophysics II (MOA-II) survey, Suzuki et al. (2016) determined a large occurrence of planets with masses similar to Neptune ($\approx 17 M_{\oplus}$). In Suzuki et al. (2018), a comparison of the same dataset with planet population syntheses challenges the predicted depth and orbital distance-dependency of the desert and points to possible gaps in our current theoretical understanding of runaway gas accretion. While reducing the disk viscosity seems to help reconciling theory and observations (Szulágyi et al., 2014; Suzuki et al., 2018), the nature and cause of the “sub-Saturn desert” remains under debate (e.g., Bennett et al., 2021).

Another distinct feature in the radius-period plot is the almost empty triangular region on short orbital distances and intermediate radii, often termed (hot) “Neptune desert” (e.g., Szabó & Kiss, 2011; Mazeh et al., 2016). Its origin is not fully understood; scenarios include the combined effect of gas accretion and planet migration (e.g., Cimerman et al., 2017; Dawson & Johnson, 2018), migration to inner disk regions where gas accretion comes to a halt (e.g., Flock et al., 2019b), high-eccentricity migration (e.g., Matzakos & Königl, 2016), and photoevaporation of planetary atmospheres due to the high irradiation on close orbital distances (e.g., Owen & Wu, 2013; Lopez & Fortney, 2014; Jin et al., 2014).

Atmospheric photoevaporation has also been linked to a demographic feature not as readily seen in the default radius-period plots: based on a selected sub-sample of planets with precise radius measurements, a bimodal distribution in the radii of small planets has been discovered (Fulton et al., 2017; Fulton & Petigura, 2018; Hsu et al., 2018; Van Eylen et al., 2018; Mordasini, 2020). This “radius valley” splits the population of close-in planets roughly at $R_P = 2 R_{\oplus}$ (see Fig. 2.8).

2.4.4 Architectural trends

Besides statistical features on the level of single planets, correlations between planets orbiting in the same system have been suggested. One of the most controversial such trend in recent years is the so-called “peas in a pod” pattern. This finding from multi-planet systems in *Kepler* data suggests that the planets within a given system tend to be of similar size, to have evenly-spaced orbits, and to be ordered in size (Millholland et al., 2017; Wang, 2017; Weiss et al., 2018; He et al., 2020b). The topic has sparked an extended debate, as the patterns were suspected to result merely from an observational bias (Zhu, 2020; Murchikova & Tremaine, 2020). In a recent study, Gilbert & Fabrycky (2020) investigated the intra-system sizing and spacing trends from an information theoretic perspective and

2 Foundations

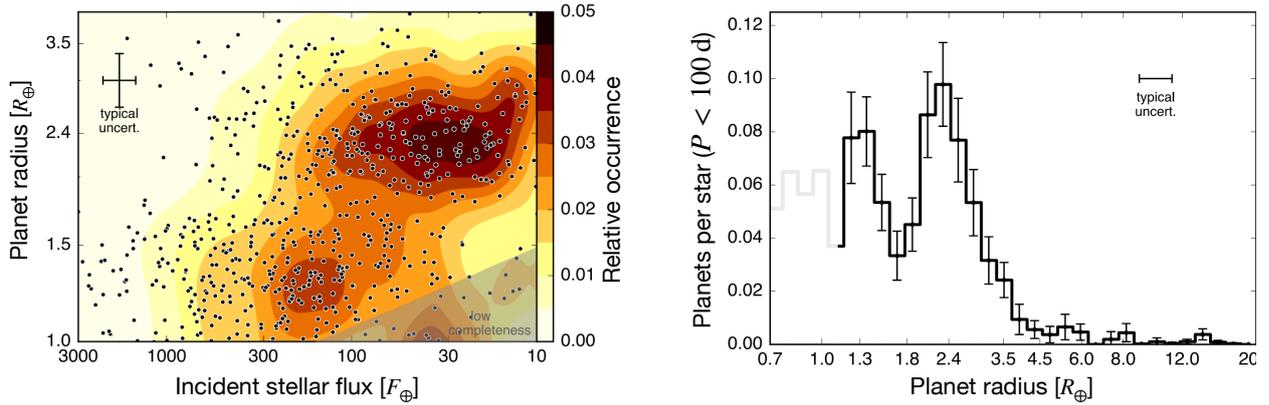


Figure 2.8: Valley in the radius distribution of small planets. *Left:* Radii of planets with precise radius measurements as a function of their incident stellar flux relative to Earth’s. Two populations of planets are separated by a gap believed to be sculpted by atmospheric mass loss. *Right:* Completeness-corrected radius distribution of planets on orbits with periods shorter than 100 d. A bimodal distribution separates bare planetary cores at radii less than $\sim 2 R_{\oplus}$ from larger planets with significant atmospheric envelopes. Both plots originally presented in [Fulton et al. \(2017\)](#) and adapted.

compared *Kepler* data to synthetic populations from the forward models EPOS ([Mulders et al., 2018](#)) and SysSim ([Hsu et al., 2018](#)). They confirm that planets within the same system tend to be similar in size and find strong evidence for uniform orbital spacing. A size ordering effect could not be established conclusively.

Particularly interesting for constraints of planet formation theories are conditional occurrence rates or occurrence probabilities. An occurrence of planets dependent on the existence of other planets in the same system would indicate that planet formation is a coordinated process rather than a purely random one. While this is a rather unexplored area, some claims regarding intra-system relationships have been made. [Huang et al. \(2016\)](#) reported that warm Jupiters (giant planets with orbital periods 10 d – 200 d) tend to have smaller companions, whereas hot Jupiters with periods shorter than 10 d have none. This was taken as an indication that warm Jupiters might form in situ. A similar question was addressed by [Schlaufman & Winn \(2016\)](#), who found that hot Jupiters have comparable probabilities to have a distant giant companion as warm Jupiters and regard this as evidence against the high-eccentricity scenario of hot Jupiter formation (e.g., [Rasio & Ford, 1996](#)). More recently, [Zhu & Wu \(2018\)](#) suggested that the conditional probability of finding a cold Jupiter ($P > 400$ d) in a system that also hosts a super-Earth is enhanced compared to the general probability of finding a cold Jupiter. If solidified, this suspicion would offer an opportunity to discriminate between competing planet formation scenarios. For example, pure pebble accretion models would, in contrast, predict an anti-correlation be-

2.5 Population synthesis with global models of planet formation

tween these planet types (e.g., [Ormel et al., 2017](#); [Owen & Murray-Clay, 2018](#)). Chapter 4 presents an in-depth investigation of [Zhu & Wu \(2018\)](#)'s claim.

2.4.5 Planetary systems around M dwarfs

Of special interest are planetary systems around low-mass stars of spectral class M. Not only are they the most abundant stars in the Milky Way (e.g., [Salpeter, 1955](#); [Chabrier, 2003](#)); they also facilitate the detection of small and low-mass planets due to their small size and mass ([Dressing & Charbonneau, 2013](#)). Many planet search programs are currently specifically targeting M dwarfs, for instance CARMENES ([Quirrenbach et al., 2010](#); [Reiners et al., 2018b](#)), EDEN ([Gibbs et al., 2020](#)), MEarth ([Nutzman & Charbonneau, 2008](#)), SPECULOOS ([Gillon et al., 2018](#)), or TRAPPIST ([Gillon et al., 2016](#)), to name just a few. The number of available data for demographic studies is therefore rapidly increasing and has already reached a level where several trends can be made out. One of the earliest such relation is a correlation of giant planet occurrence with stellar mass as well as with the host star metallicity ([Johnson et al., 2010](#); [Bonfils et al., 2013](#)). These trends support the core accretion paradigm, as more massive and metal-rich stars host more massive disks enriched in dust (e.g., [Asplund et al., 2005](#); [Mordasini et al., 2009a](#); [Pascucci et al., 2016](#); [Barenfeld et al., 2016](#); [Ansdell et al., 2016](#)), which provide more solid material for efficient planetary core formation. The same trend is not apparent for the population of super-Earths and sub-Neptunes, which have been abundantly detected in *Kepler* data. This suggest that the formation of small and intermediate-sized planets is efficient around low-mass stars, regardless of their metallicity. A somewhat puzzling finding was that the occurrence rates of these planets are even enhanced around stars less massive than the Sun ([Mulders et al., 2015a](#)). An explanation for this trend has yet to be provided, but a relation to planet migration has been hypothesized ([Mulders, 2018](#)). Tentative evidence now suggests that this trend also holds within the spectral class M and that planetary occurrence rates are higher around later M dwarfs ([Hardegree-Ullman et al., 2019](#)).

2.5 Population synthesis with global models of planet formation

As I have motivated in previous sections, planet formation happens in multiple stages that cover a vast range of spatial and temporal scales. Multiple subfields of physics are employed to understand the processes happening during these stages, ranging from (magneto-)hydrodynamics to explain the evolution of gaseous accretion disks to gravitational interactions between the forming planets. None of the processes I introduced above

2 Foundations

is fully understood, and for many of them, direct observational constraints are not attainable. The main “fix points” of planet formation models are the observable domains of protoplanetary disks and the detection of fully assembled planetary systems, both with their intrinsic caveats and biases (see Sect. 2.4). If these observational constraints can be linked somehow, perhaps with a comprehensive theoretical model, one can hope to gain understanding about the physical processes evading our telescopes. This is the promise of *planet population synthesis*: based on carefully selected initial conditions and on the current theoretical understanding, it produces synthetic planet populations that can be statistically compared to observed planet samples (Mordasini et al., 2015).

What are the requirements for a model to achieve this goal? A major challenge is that not all relevant processes happen independently of each other, but instead can be coupled and must be considered simultaneously. One example is the relation between a planet’s gas accretion and the opacity of its envelope, which is sensitive on the amount and properties of solid material it accretes (e.g., Podolak, 2003). Any model that aims to provide an end-to-end description of planet formation, as required for a self-consistent planet population synthesis, must take into account all relevant physical processes. However, this is currently not feasible in the degree of detail that one can employ in simulations focusing on a single such process. To keep the computational costs within a viable range, the common approach is to approximate the results of those detailed studies while ensuring that the effects on the planet formation process are still adequately captured. For instance, the evolution of the protoplanetary disk can be simplified by solving the one-dimensional equations for axisymmetric disks (see 2.29), and its viscosity can be parametrized (Eqn 2.11, Shakura & Sunyaev, 1973).

Varying degrees of model complexity and a range of different assumptions were employed in past population synthesis studies (see, e.g., Ida & Lin, 2004b,a, 2005, 2008; Thommes et al., 2008; Mordasini et al., 2009a,b; Hasegawa & Pudritz, 2011; Alibert et al., 2011; Mordasini et al., 2012b; Bitsch et al., 2015; Ndugu et al., 2018; Johansen et al., 2019; Bitsch et al., 2019a; Izidoro et al., 2019; Lambrechts et al., 2019; Mulders et al., 2020; Alessi et al., 2017, 2020; Miguel et al., 2020).

Figure 2.9 summarizes various families of planet formation models based on the core accretion paradigm. Without claim for completeness, these models have all been utilized in population synthesis studies, where initial conditions are randomly drawn from distributions motivated by observations. All models use more or less simplified approximations for the gas disk model, which typically assumes a radial power law distribution motivated by the Minimum Mass Solar Nebula (MMSN) (Weidenschilling, 1977; Hayashi,

2.5 Population synthesis with global models of planet formation

| | Ida & Lin | Hasegawa & Pudritz | pebble accretion I | pebble accretion II | Gen. I Bern model | Gen. II Bern model | Gen. III Bern model |
|------------------------|--|---|-----------------------------------|---|--|--|--|
| Reference | <i>Ida&Lin2004a</i> | <i>Hasegawa & Pudritz 2011, 2012</i> | <i>Bitsch+2015, Ndugu+2018</i> | <i>Izidoro+2019, Lambrechts+2019, Bitsch+2019</i> | <i>Alibert+2005</i> | <i>Alibert+2013</i> | <i>Emsenhuber+2020a</i> |
| Gas disk model | power law | power law, planet traps | fit to 2D hydro simulations | fit to 3D hydro simulations, inner edge | 1+1D truncated power law, inner edge | 1+1D truncated power law, inner edge | 1D truncated power law, inner edge |
| Star-disk interaction | — | irradiation | irradiation, evolution | irradiation, evolution | — | irradiation | irradiation, evolution |
| Embryos per disk | 1 | 1 | 1 | 60 | 1 | ~10 | 50-100 |
| Planetesimal accretion | runaway, oligarchic | runaway, oligarchic | — | — | runaway, 100km | oligarchic, 300m | oligarchic, 300m |
| Pebble accretion | — | — | analytical | analytical** | — | — | — |
| Planet migration | Type II | Type I, Type II | Type I, Type II | Type I, Type II | Type I [Ⓡ] , Type II | Type I, Type II | Type I, Type II |
| Gas accretion | Kelvin-Helmholtz contraction | Kelvin-Helmholtz contraction | analytical (Piso&Youdin 2014) | analytical (Piso&Youdin 2014) | 1D structure, equilibrium flux limited | 1D structure, non-equilibrium flux limited | 1D structure, local reservoir limited |
| Envelope evolution | — | — | — | — | — | — | D-burning, photoevap. |
| Notes | + Type I migration [Ⓡ] (Ida&Lin2008) + <i>N</i> -Body, 16 emb. (Ida&Lin2010) similar approach in Miguel+2016 | also see Alessi+2017, Alessi&Pudritz 2018 | + ext. heating by stellar cluster | | photoevap. sink term included. + evolution, core structure, D-burning, photoevap. (Mordasini+2012a) | | variable stellar mass (Burn+, subm.) in dev.: embryo formation (Voelkel+, in prep.) |

[Ⓡ] reduced Type I migration

** free parameter in Lambrechts+2019

used in this thesis



Figure 2.9: Comparison of different population synthesis models and the physical prescriptions they include.

1981, see Sect. 2.1.2). The power law is sometimes truncated by an outer exponential term and/or a term producing an inner disk edge. All gas disk models assume a viscously evolving disk. Except for the pioneering models by [Ida & Lin \(2004b\)](#) and the Generation I Bern Model ([Alibert et al., 2005](#)), all models adopt some prescription of disk irradiation by the central star. The latter included a sink term for photoevaporation in the disk evolution but the disk temperature structure is not affected. Where stellar irradiation is included, the evolution of the star is only taken into account in some of the models. The single-embryo-per-disk approximation has been popular, as N -body treatments that account for gravitational interactions between planets are notoriously expensive. However, this has been acknowledged as a major shortcoming for models aiming at reproducing systems of small planets, which frequently occur in multiple systems and often in compact configurations (e.g., [Lissauer et al., 2011](#); [Latham et al., 2011](#); [Mayor et al., 2011](#)). Their mutual interaction may lead to orbital excitations, but also to giant collisions and ejections from the system. It also affects orbital migration through captures into mean motion resonances, which under certain conditions can even reverse the migration direction ([Masset & Snellgrove, 2001](#)). This characteristic was leveraged in the “Grand Tack” model to reproduce the final layout of planets in the Solar System ([Walsh et al., 2011](#)). As computing power became more accessible, most authors have turned to employing multi-planet models. Solid accretion is in most of the current models planetesimal-based, and differences exist in which accretion regime is being modeled (compare Sect. 2.2.1). Some modelers have implemented the emerging paradigm of pebble accretion (see Sect. 2.2.7) using various prescriptions. Remarkably, there is still a lack of comprehensive population syntheses that take into account both solid accretion modes (but see [Coleman et al. \(2019\)](#); [Brügger et al. \(2020\)](#) for groundbreaking work). All currently employed models take into account both major types of planet migration: the typically rapid Type I migration due to excitation of density waves in the disk, and the slower Type II migration experienced by massive planets that have opened a disk gap (see Sect. 2.2.8). Some older models used an artificial reduction factor for Type I migration to prevent too many planets from migrating into the star. Throughout the list, gas accretion onto planets is not modeled in state-of-the-art detail (e.g., [Marleau et al., 2017](#); [Moldenhauer et al., 2021](#)), which would be intractable for a Monte Carlo approach, but instead it is approximated. A common way to do this is to calculate an accretion rate depending on the envelope contraction given by the Kelvin-Helmholtz timescale ([Ikoma et al., 2000](#)). The more complex approach is to numerically integrate one-dimensional structure equations ([Bodenheimer & Pollack, 1986](#)), which is followed by the family of Bern models ([Alibert et al., 2005, 2013](#); [Mordasini et al., 2012b](#);

Emsenhuber et al., 2020a). Additional effects that are not covered by every model are the evolution of the internal structure of planetary atmospheres, and in particular their loss in the evolutionary phase after the disk has dissipated (Mordasini, 2018). The Bern models include photoevaporative mass losses from its second generation on (Jin et al., 2014). In its latest version, also Deuterium burning in massive planetary envelopes as well as variational stellar masses (Burn et al., in press) are covered. A new development already in preparation will be a self-consistent formation treatment of planetary embryos (Voelkel et al., in prep.).

In this thesis, I utilize the *Generation III Bern global planet formation and evolution model* (Emsenhuber et al., 2020a) to simulate planetary systems. The resulting synthetic planet populations are investigated in Chapters 3, 4, and 6. The following summary of the inner workings of the model is based on Schlecker et al. (2020b).

2.5.1 The Generation III Bern model

This semi-analytical model couples the evolution of a viscously-spreading protoplanetary disk with planet formation following the core accretion paradigm (Perri & Cameron, 1974; Mizuno et al., 1978; Mizuno, 1980) and a planet migration scheme (both Type I (Paardekooper et al., 2011) and Type II (Dittkrist et al., 2014)). Solids are accreted via planetesimals in the oligarchic regime (Ida & Makino, 1993; Ohtsuki et al., 2002; Thommes et al., 2003). Multiple planets can form in the same disk and their mutual gravitational interaction is modeled via an N -body integrator. The model is based on earlier work, where Alibert et al. (2005); Mordasini et al. (2009a) simulated the single embryo case, Mordasini et al. (2012c,b) combined the formation phase with long-term evolution, and Alibert et al. (2013) included a formalism for concurrent formation of multiple protoplanets inside a single disk. For a thorough description of the Generation III Bern Model and an outline of recent advancements of the framework, we refer to Emsenhuber et al. (2020a). Additional reviews are provided in Benz et al. (2014) and Mordasini (2018).

Disk model

The gas disk model describes the evolution of a viscous accretion disk (Lüst, 1952; Lynden-Bell & Pringle, 1974) using an α parametrization as in Shakura & Sunyaev (1973, see Eqn 2.11) for the viscosity ν . We chose a viscosity parameter of $\alpha = 2 \times 10^{-3}$ that is constant throughout the disk and time-independent. This specific value provides realistic stellar accretion rates (Mulders et al., 2017; Manara et al., 2019). The combination of this viscous accretion with photoevaporative mass loss results in disk lifetimes that are in agreement with observations (Haisch, Jr. et al., 2001; Fedele et al., 2010; Richert et al.,

2018).

The vertical structure was computed following the approach of [Nakamoto & Nakagawa \(1994, compare Sect. 2.1.3\)](#) including stellar irradiation under an evolving luminosity of the star and viscous heating. The stellar luminosity was obtained from tabulated values ([Baraffe et al., 2015](#)).

The initial profile of the gas surface density is based on Eqn 2.1, but with an additional term to produce an inner disk edge. The resulting profile is given by

$$\Sigma_{\text{g}}(r, t = 0) = \Sigma_0 \left(\frac{r}{R_0} \right)^{-\beta_{\text{g}}} \exp \left[- \left(\frac{r}{R_{\text{cut,g}}} \right)^{2-\beta_{\text{g}}} \right] \left(1 - \sqrt{\frac{R_{\text{in}}}{r}} \right), \quad (2.29)$$

forming a power law with an exponential decrease at the outer edge. Here, Σ_0 is the initial gas surface density at Jupiter's semi-major axis, $R_0 = 5.2 \text{ au}$, β_{g} dictates the slope of the profile, $R_{\text{cut,g}}$ defines the location of the exponential decrease, and R_{in} marks the inner edge of the disk. We considered only the radial dimension and assumed azimuthal symmetry. For R_{in} , we adopted inner disk boundaries distributed like the assumed co-rotation radii of young stars ([Venuti et al., 2017](#)). The surface density is described by the radial power law in Eqn 2.29 with a fixed exponent $\beta_{\text{g}} = 0.9$ ([Andrews et al., 2010](#)).

The gas surface density is affected by gas accretion of the protoplanets, which is calculated by solving one-dimensional internal structure equations following [Bodenheimer & Pollack \(1986\)](#), see [Emsenhuber et al. \(2020a\)](#). In addition, photoevaporation acts on the gas disk and dominates its evolution in the final phase of its lifetime. We model photoevaporation due to high-energy photons by the host star based on the mechanism described in Sect. 2.1.6. [Clarke et al. \(2001\)](#) provide photoevaporative mass loss rates as a function of orbital distance, which we consider in our disk evolution model ([Mordasini et al., 2012c](#)). In addition, external photoevaporation due to irradiation of hot, nearby stars in the environment of the protoplanetary system is modeled using the simplified prescription in [Matsuyama et al. \(2003\)](#). Combining viscous accretion, planetary accretion, and photoevaporation, the overall evolution of the gas disk is governed by the viscous diffusion equation ([Lynden-Bell & Pringle, 1974](#)),

$$\frac{\partial \dot{\Sigma}_{\text{g}}}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left[3r^{1/2} \frac{\partial}{\partial r} \left(r^{1/2} v \Sigma_{\text{g}} \right) \right] - \dot{\Sigma}_{\text{g,photo}} - \dot{\Sigma}_{\text{g,planet}}, \quad (2.30)$$

where $\dot{\Sigma}_{\text{g,photo}}$ and $\dot{\Sigma}_{\text{g,planet}}$ are the sink terms for photoevaporation and accretion onto planets, respectively. The latter removes gas from the planetary feeding zone. At a planetary gas accretion rate \dot{M}_{env} , the gas surface density within the feeding zone depletes

as

$$\dot{\Sigma}_{g, \text{planet}} = \frac{\dot{M}_{\text{env}}}{\pi(r_2^2 - r_1^2)} \quad (2.31)$$

with the feeding zone boundaries r_1 and r_2 that obey $r_1 < r_2$ (Emsenhuber et al., 2020a).

The accretion of solid material is treated similarly. A planetesimal disk provides material for planetary cores and evolves depending on the accretion behavior of forming protoplanets in the disk. It is modeled in a fluid-like description where the dynamical state of the planetesimals evolves depending on the stirring due to other planetesimals and protoplanets (Thommes et al., 2003; Chambers, 2006; Fortier et al., 2013). Gravitational back-reactions of planetesimals on planets are not taken into account. Following Chambers (2006), the planetesimal accretion rate can be approximated as

$$\dot{M}_{\text{core}} = \Omega \bar{\Sigma}_s R_H^2 p_{\text{coll}}, \quad (2.32)$$

where $\bar{\Sigma}_s$ is the mean planetesimal surface density in the planetesimal feeding zone, R_H is the Hill radius, and p_{coll} is the collision probability (Inaba et al., 2001).

To account for the radial drift of particles with low to intermediate Stokes numbers (e.g., Weidenschilling, 1977; Birnstiel & Andrews, 2014, see Sects. 2.1.2, 2.2.7), our solid disk is more compact than the gas disk. This is imposed by a steeper slope in the planetesimal surface density as well as a decreased solid disk size. We chose a slope index $\beta_s = 1.5$, which is similar to the minimum mass solar nebula (MMSN) (Weidenschilling, 1977; Elbakyan et al., 2020). The exponential cutoff radius, which defines the outer edge of the solid disk, was set to half the radius of the gas disk (Ansdell et al., 2018).

Planet formation model

Growth of the protoplanetary cores occurs via two channels: planetesimal accretion and giant impacts (collisions between protoplanets). The planetesimal accretion rate is obtained following the approach of Fortier et al. (2013). The collision probability takes into account the eccentricity and inclination distributions of the planetesimals following Ohtsuki et al. (2002). Although any unique planetesimal size represents a strong simplification, we assumed a uniform value of 300 m. Planetesimals of this size experience sufficient damping by gas drag to ensure viable relative velocities while exhibiting typical drift timescales that are longer than the disk lifetime (Fortier et al., 2013; Burn et al., 2019). They have further been shown to provide realistic accretion efficiencies to reproduce the observed population of exoplanets across various planetary mass regimes (Fortier et al., 2013).

The gravity of a protoplanet's core causes the attraction of gas. Initially, the gas accretion rate of the planetary envelope is governed by the ability to radiate away the binding energy released by the accretion of both solids and gas (Pollack et al., 1996). To determine the envelope mass and its structure, the model solves the one-dimensional internal structure equations following Bodenheimer & Pollack (1986) (also see Alibert & Venturini, 2019).

During this stage, known as the "attached" phase, the boundary between planet envelope and surrounding disk is continuous. The efficiency of cooling improves with increasing planet mass, so that the gas accretion rate increases with time. This can eventually result in a runaway accretion of gas, where the accretion rate exceeds the amount of gas that can be supplied by the disk. When this happens, the envelope is no longer in equilibrium with the surrounding disk and contracts (Bodenheimer et al., 2000). In this "detached" phase, the internal structure equations determine the planet radius (Mordasini et al., 2012c).

Orbital migration

While embedded in a disk, planets undergo orbital migration per exchange of angular momentum with the surrounding gas (Goldreich & Tremaine, 1980). Low-mass planets embedded in the disc migrate in the Type I regime, while massive planets open a gap and migrate in the Type II regime. We consider non-isothermal Type I migration following Paardekooper et al. (2011) with a reduction of the co-rotation torques due to the planet's eccentricity and inclination following Coleman & Nelson (2014). For the Type II migration rate, we follow Dittkrist et al. (2014) and use the fully suppressed, non-equilibrium radial velocity of the gas. To determine the point for transition from Type I to Type II, we use the gap opening criterion by Crida et al. (2006). No artificial reduction factors are applied.

Long-term evolution

After the dispersal of the protoplanetary disk due to the combination of viscous accretion and photoevaporation, we modeled the thermodynamical evolution of each survived planet until a simulation time $t = 10$ Gyr. The evolution module starts with a planet's internal structure at the end of the formation phase to calculate how it cools and contracts in the long term, including the effects of atmospheric escape, bloating, and stellar tides (Mordasini et al., 2012c). Besides contraction and accretion, an additional luminosity term arises from radioactive decay of long-lived nuclides (Mordasini et al., 2012b). We further took into account compositional changes of the planetary core and envelope following the method in Thiabaud et al. (2015).

N-body interactions

Being a multi-planet simulation, the Bern model includes gravitational interaction among the growing planets during the formation phase. We employed the Mercury N -body integrator (Chambers, 1999, 2012) to compute the orbital evolution of a system and detect collisions of planets. Orbital migration as well as a damping of eccentricities and inclinations were included as additional forces. For practical reasons, we stopped the computationally expensive N -body calculations after a simulation time of 20 Myr.

Monte Carlo sampling

Reflecting the natural variation in disk properties, we performed our simulations as a Monte Carlo experiment: we repeatedly ran our model with a different set of initial conditions, which we sampled randomly from continuous distributions. This approach enabled us to make quantifiable statistical assessments despite the complex nature of the planet formation process.

We sampled from among four disk initial parameters: the initial gas disk mass, M_{gas} , the dust-to-gas ratio, $\zeta_{\text{d,g}}$, the mass loss rate due to photoevaporative winds, \dot{M}_{wind} , and the inner disk edge, R_{in} . In addition, we randomly drew the starting locations of planetary embryos, which are instantly initialized at the beginning of the simulation.

Constraints that can be imposed on the distribution of these parameters through observations of protoplanetary disks are limited. Where they were not available, we were left with theoretical arguments. Section 2.5.1 describes the chosen ranges for each random variable and we show their distributions in Fig. 2.10.

Initial conditions

In this section, I present the motivations behind the chosen distributions for our Monte Carlo parameters. The specific choices of values refer to the synthetic planet population NG76, which can be considered our nominal population and which I investigate in Chapters 3 and 4. In the model used for studying M dwarf planetary systems, some of these Monte Carlo parameters are scaled with stellar mass (see Sect. 2.5.2).

Initial gas disk mass M_{gas}

The gas content largely governs the mass of a protoplanetary disk (Ansdell et al., 2016, however, see Miotello et al. (2016)). While the intrinsic distribution of gas disk masses is poorly constrained, observations point to masses ranging between 0.1% and 10% of

2 Foundations

the stellar host mass (Andrews et al., 2010; Manara et al., 2016). In contrast to previous population synthesis generations, I consider **gas disk masses** of earlier times to constrain this Monte Carlo variable. The masses are distributed following a survey of Class I objects, i.e., of disks that are still engulfed in an outer envelope (Tychoniec et al., 2018). To compute the gas mass of our computational disk M_{gas} , I drew $\log \frac{M_{\text{gas}}}{M_{\odot}}$ from a normal distribution $\mathcal{N}(\mu = -1.49, \sigma^2 = 0.123)$ (Tychoniec et al., 2018).

In combination with the disk radius $R_{\text{cut,g}}$ (see below), the resulting initial gas surface density at a reference radius of 5.2 au, Σ_0 , varies log-normally with a median of 132 g cm^{-2} . Note that the range of M_{gas} was cut to avoid both extremely massive and very low-mass disks.

Dust-to-gas ratio $\zeta_{\text{d,g}}$

I assumed that the bulk metallicity of the disk $[M/H]_{\text{disk}}$ equals the heavy-element abundances of the protostar and modeled it as a normal distribution with a mean $\mu = -0.03$ and a standard deviation $\sigma = 0.20$ that follows the dispersion of stellar metallicities in the Solar neighborhood (Santos et al., 2005, compare Fig. 2.11).

$[M/H]_{\text{disk}}$ is then readily converted to the dust-to-gas ratio of the disk via

$$\zeta_{\text{d,g}} = Z_0 \cdot 10^{[M/H]_{\text{disk}}}, \quad (2.33)$$

where $Z_0 = 0.0149$ (Lodders, 2003). Its distribution is shown in Fig. 2.10. In combination with the disk's gas content, the dust-to-gas ratio determines the amounts of solids available for forming planets.

Photoevaporative mass loss rate \dot{M}_{wind}

A constraining factor for a system's ability to build planets is the lifetime of its protoplanetary disk. Large IR surveys of young stellar clusters enabled to identify the fraction of stars with disks as a function of their age, indicating a disk lifetime of a few Myr (Haisch, Jr. et al., 2001; Mamajek et al., 2009; Fedele et al., 2010). It is believed that photoevaporation from high-energy photons, in combination with viscous accretion, is the main mechanism for the rapid dispersal of disks at the end of their lifetimes (Clarke et al., 2001; Owen et al., 2012, see Sect. 2.1.6). I parameterized this effect with a mass loss rate \dot{M}_{wind} , which I varied such that the distribution of synthetic disk lifetimes is similar to the observed distribution in Fedele et al. (2010). The parameter was normalized to a hypothetical disk extending to 1000 au and thus does not equal the absolute mass loss rates of our disks.

Inner disk edge R_{in}

The physical motivation for an inner edge of the gas disk is the development of a magnetospheric cavity (Bouvier et al., 2007), which is thought to extend to the corotation radius, i.e. the location where the angular velocity of the stellar magnetic field and of the orbiting gas are equal (e.g., Günther, 2013). For the numerical disk, I adopted the orbit radius corresponding to the rotation period of its host star for R_{in} . I drew these periods from a distribution based on recent measurements in the young stellar cluster NGC 2264, which has an estimated age of 3 Myr (Venuti et al., 2017). The resulting log-normal distribution has a mean period of 4.74 d and a standard deviation of 0.31 dex.

Disk radius $R_{\text{cut,g}}$

The sizes of protoplanetary disks are both constrained by observations (e.g., Andrews et al., 2010, 2018b; Ansdell et al., 2018) and by analytically solving the viscous accretion disk problem (Lynden-Bell & Pringle, 1974). In the model, the radial extent of the gas disk $R_{\text{cut,g}}$ is not an independent Monte Carlo variable but I computed it from M_{gas} using a scaling relation derived from millimeter continuum emission sizes (Andrews et al., 2018b). Our disk radii range from roughly 20 au to 150 au.

The solid material is represented by a continuous disk of solids. To take into account spatial concentration due to inward drift of solid material (Weidenschilling, 1977), its size is initially half of that of the gas disk. This choice of scaling is motivated by millimeter observations of the Lupus star-forming region, which suggest a factor of two difference between gas and dust disks (Ansdell et al., 2018).

Starting location of planet embryos a_{start}

For each system, I inserted 100 planetary seeds with a starting mass of $0.01 M_{\oplus}$ into the disk. The initial location of these embryos were randomly drawn from a log-uniform distribution in semi-major axis. This follows N -body simulations of planetary embryos that found oligarchs spaced by a few Hill spheres, that is, their separations are proportional to their orbital distance (Kokubo & Ida, 2000). The roughly Moon-mass seeds were distributed from the inner edge of the disc R_{in} up to 40 au with the additional constraint that no embryo was placed closer than 10 Hill radii to another.

For a more detailed description of the observational and theoretical grounds of these Monte Carlo parameters, I refer to Mordasini et al. (2009a,b), Mordasini et al. (2012b) and Emsenhuber et al. (2020b).

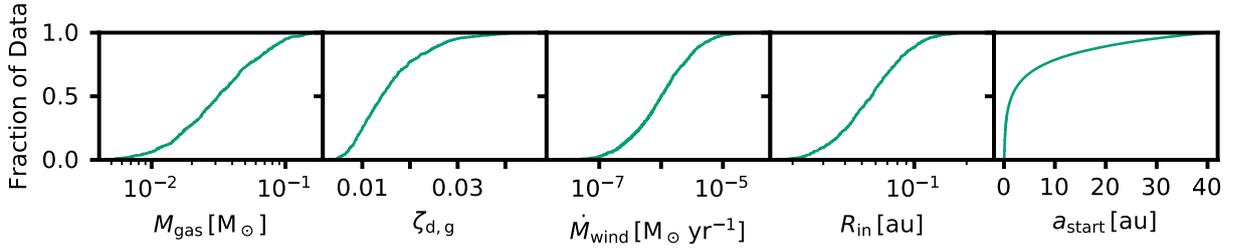


Figure 2.10: Cumulative distribution of Monte Carlo parameters. The disk initial conditions, which were drawn randomly from these distributions for each simulation, comprise the initial gas disk mass M_{gas} , the dust-to-gas ratio $\zeta_{\text{d,g}}$, the photoevaporative mass loss rate \dot{M}_{wind} , the inner disk edge R_{in} , and the starting locations of the planetary embryos a_{start} .

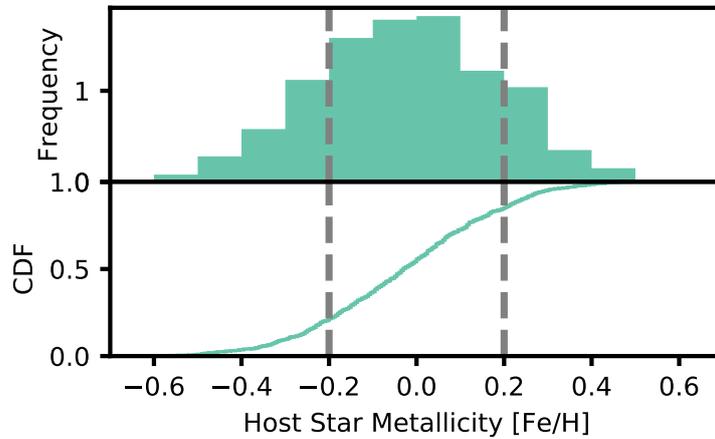


Figure 2.11: Distribution of host star metallicities. *Upper panel:* Frequency normalized to one. *Lower panel:* Empirical distribution function showing the fraction of host stars that are less than or equal to the specified metallicity. Dashed lines denote the borders of the metallicity bins used in this study.

Table 2.1 shows the distributions of Monte Carlo parameters chosen for our simulations, as well as additional parameters that were kept fixed.

2.5.2 M dwarf population synthesis

Planet-targeted observations of low-mass stars are gaining momentum and promise valuable insights due to the possibly different formation conditions of the planets they host. It is thus very timely to match these observational advances with dedicated theoretical studies of M dwarf planetary systems. In Chapter 6, I present a statistical analysis focusing on planets around low-mass stars of spectral type M. For this purpose, our collaboration adapted the Generation III Bern Model to the M dwarf regime of host stars (Burn et al., in press).

Table 2.1: Model parameters for the NG76 multi-planet population

| Parameter | Symbol | Distribution | Range or Median ^{+84%} _{-16%} |
|---------------------------------------|-------------------------|--------------------|--|
| Fixed Parameters | | | |
| Stellar Mass | | – | 1 M _⊙ |
| Disk Viscosity | α | – | 2×10^{-3} |
| Power Law Index (Gas) | β_g | – | 0.9 |
| Power Law Index (Solids) | β_s | – | 1.5 |
| Radius of Planetesimals | | – | 300 m |
| Number of Planet Seeds | | – | 100 |
| Mass of Planet Seeds | | – | 0.01 M _⊕ |
| Monte Carlo Parameters | | | |
| Host Star Metallicity | [Fe/H] | normal | -0.03 ± 0.20 |
| Initial Gas Surface Density at 5.2 au | Σ_0 | log-normal | $132_{-27}^{+37} \text{ g cm}^{-2}$ |
| Inner Disk Radius | R_{in} | log-normal | $4.74_{-2.42}^{+4.94} \text{ d}$ |
| Gas Disk Cutoff Radius | $R_{\text{cut,g}}$ | log-normal | $56_{-21}^{+36} \text{ au}$ |
| Solid Disk Cutoff Radius | $R_{\text{cut,s}}$ | log-normal | $R_{\text{cut,g}}/2$ |
| Photoevaporation Efficiency | \dot{M}_{wind} | log-normal | $(1.0_{-0.7}^{+2.2}) \times 10^{-6} \text{ M}_{\odot}/\text{yr}$ |
| Starting Position of Planet Seeds | a_{start} | uniform in log a | R_{in} to 40 au |
| Derived Parameters | | | |
| Initial Gas Disk Mass | M_{gas} | log-normal | $0.03_{-0.02}^{+0.04} \text{ M}_{\odot}$ |
| Initial Solid Disk Mass | M_{solid} | \sim log-normal | $95_{-55}^{+147} \text{ M}_{\oplus}$ |
| Dust-to-gas Ratio | $\zeta_{\text{d,g}}$ | log-normal | $0.02_{-0.01}^{+0.01}$ |
| Disk Dispersal Time | t_{disk} | – | $(3.2_{-1.0}^{+1.9}) \times 10^6 \text{ yr}$ |

Notes. Upper panel: Initial conditions that are fixed for each simulation. Middle panel: Monte Carlo parameters that are drawn randomly. Lower panel: Quantities that are derived from or controlled by other parameters. Upper and lower limits denote 84th and 16th percentiles, respectively.

Scalings with stellar mass

Besides the mass and luminosity profile of the host star, we scaled the following parameters with the mass of the star:

- As outlined in Sect. 2.1.2, we assume that the **position of the inner disk edge** is proportional to the equatorial rotation rate of the host star, such that its position corresponds to the co-rotation radius. Scalings of T Tauri rotation rates with stellar mass are under debate; we therefore adopt a mass-independent log-normal distribution of rotation periods from [Venuti et al. \(2017\)](#).
- Multiple sources report a stellar mass-scaling of the **gas disk masses** of protoplanetary disks (e.g., [Pascucci et al., 2016](#); [Ansdell et al., 2016](#); [Barenfeld et al., 2016](#)). While determining the specific shape of the scaling law is non-straightforward (compare

2 Foundations

Burn et al., in press), we adopted a linear scaling $M_{\text{gas},0} \propto M_{\star}$.

- Since we do not consider a stellar-mass dependency of the stellar metallicity, we computed the total **planetesimal disk mass** by multiplying the gas disk masses with spectroscopically measured metallicities in the solar neighborhood (Santos et al., 2003).
- The **gas disk radii** are coupled to the disk masses and thus scale with host star mass. Following Andrews et al. (2010, 2018b), we adopt a scaling relation $M_{\text{gas},0} \propto R_{\text{disk}}^{1.6}$.
- As in the nominal solar-type model, our **dust disk radii** are by a factor of 0.5 smaller than the gas disk radii (Ansdell et al., 2018).
- We initially place 50 **planetary embryos** into each disk. Their mass does not scale with stellar mass and is kept fixed at $10^{-2} M_{\oplus}$. As in the model for solar analogs, the starting position of each embryo is drawn from a log-uniform distribution in orbital radius with an additional requirement of a distance > 10 mutual Hill radii to the next embryo. Embryos are placed between the inner disk edge and an outer limit, which is at 40 au for solar-type stars and is scaled as $(M_{\star}/M_{\odot})^{1/3}$.
- The distribution of **disk lifetimes** remains the same as in the solar-mass models, since there are currently no conclusive constraints on this parameter as a function of stellar mass.

Besides these Monte Carlo parameters, there are additional processes we model that are affected by a difference in stellar mass. These are:

- the gas disk evolution, as viscous heating and irradiation vary with M_{\star} (Hueso & Guillot, 2005)
- the evolution of e and i of the planetesimals, which depend on planetesimal Hill radii (Rafikov, 2004)
- the planetesimal accretion rate, which scales linearly with planetary and planetesimal mutual Hill radius (Fortier et al., 2013)
- the fraction of ejected planetesimals, which depends on the escape velocity
- the orbital evolution of the planets due to planet-disk interactions (see Burn et al., in press)

- tidal migration rates, which are lower for lower stellar masses (Benítez-Llambay et al., 2011)

The NGM synthetic planet population

Using the adapted planet formation model outlined above, we computed a dedicated M dwarf planet population termed “NGM”. It is organized as a grid in stellar mass, where we simulated 1000 systems for each of the five host star masses $0.1 M_{\odot}$, $0.3 M_{\odot}$, $0.5 M_{\odot}$, $0.7 M_{\odot}$, and $1.0 M_{\odot}$. Due to some failed simulations, the final NGM population consists of 4996 systems (for a detailed breakdown, see Tab. 6.1). Each of these systems started with 50 planetary embryos initially injected into the disk; a number that reduces during the simulations due to collisions, ejections from the system, and accretion onto the star. At a simulation time of $t = 5$ Gyr, we “observe” the remaining synthetic population, which then consists of roughly 20,000 planets. Table 2.2 gives some insight into the composition of the

Table 2.2: Fraction of systems harboring different planet types in the NGM M dwarf planet population. Table adapted from Burn et al. (in press)

| Planet type | Stellar mass [M_{\odot}] | | | | |
|-------------------------------|------------------------------|------|------|------|------|
| | 0.1 | 0.3 | 0.5 | 0.7 | 1.0 |
| $M_{\text{P}} > 1 M_{\oplus}$ | 0.44 | 0.77 | 0.88 | 0.91 | 0.95 |
| Earth-like | 0.70 | 0.88 | 0.89 | 0.89 | 0.84 |
| Super-Earth | 0.19 | 0.54 | 0.71 | 0.78 | 0.79 |
| Neptunian | 0.01 | 0.08 | 0.17 | 0.22 | 0.27 |
| Sub-giant | 0.00 | 0.00 | 0.02 | 0.03 | 0.05 |
| Giant | 0.00 | 0.00 | 0.02 | 0.09 | 0.19 |

final population by listing, for each stellar mass bin, the fractions of systems that host certain planet types. These fractions are defined as the number of systems containing *at least one* respective planet type divided by the total number of systems in the sub-population. They do not take into account intra-system multiplicities and thereby differ from planetary occurrence rates. We distinguish between: planets more massive than Earth; Earth-like planets with masses $0.5 M_{\oplus} - 2 M_{\oplus}$; super-Earths with masses $2 M_{\oplus} - 10 M_{\oplus}$; Neptunian planets with masses $10 M_{\oplus} - 30 M_{\oplus}$; sub-giants with masses $30 M_{\oplus} - 100 M_{\oplus}$; and giant planets with masses $M_{\text{P}} > 100 M_{\oplus}$ (Burn et al., in press). The most pronounced features and trends with stellar mass are:

- a high fraction of systems with planets more massive than Earth that increases with higher stellar mass,

2 Foundations

- a high fraction of intermediate-mass M dwarfs hosting Earth-like planets,
- an increasing fraction of systems containing super-Earths or Neptunian planets with increasing stellar mass,
- a strong stellar mass-dependency of the formation of sub-giants and giant planets, which do not occur around stars less massive than $0.5 M_{\odot}$,
- a general scarcity of these planet types.

The *NGM* population is discussed in greater detail in a dedicated publication by Burn et al. (in press).

3

Predetermination of Planet Types in Global
Core Accretion Models

The content of this chapter is based on the publication [Schlecker et al. \(2021, in press\)](#), in *Astronomy & Astrophysics*. It is part of the paper series “The New Generation Planetary Population Synthesis (NGPPS)”.

Details of authorship: I am the corresponding author of the study, which resulted from an internship of the second author Dang Pham at the Max Planck Institute for Astronomy (MPIA). I devised the original project, served as direct advisor for Dang Pham and eventually prepared the manuscript and figures. Dang Pham performed the comparison of clustering methods and their hyperparameters and wrote the code for this task. He further contributed Figure 3.1 and helped drafting the sections on clustering algorithms and their selection (Appendix A.1). Remo Burn aided in interpreting the results, deriving the boundary conditions for giant planet formation (Sect. 2.2), and structuring the manuscript. Yann Alibert and Christoph Mordasini provided direction and coordinated the study within the paper series of NGPPS. Alexandre Emsenhuber is the lead author of the updated planet formation model this study utilizes and helped formulating the description of the model. Thomas Henning, Hubert Klahr, and Lokesh Mishra provided advice during the project and gave critical feedback on the manuscript.

3.1 Motivation

One of the most remarkable findings in recent years of exoplanetology has been the enormous diversity of planetary systems (e.g., Ribas & Miralda-Escudé, 2007; Howard et al., 2012; Fressin et al., 2013; Petigura et al., 2013a; Mulders et al., 2015b; Hobson & Gomez, 2017; Brewer et al., 2018; Owen & Murray-Clay, 2018; Hsu et al., 2019; Bryan et al., 2019; He et al., 2020a). The rapidly increasing number of confirmed planets improves our ability to explore this diversity and to understand its origins. To this end, a variety of physical mechanisms that influence the formation and evolution of planetary systems, and therefore shape their demographics, have been investigated. I have reviewed the ones considered most important at present in Chapter 2. Intensively studied mechanisms include the evolution of accretion disks (e.g., Lüst, 1952; Lynden-Bell & Pringle, 1974; Pringle, 1981), their interaction with embedded planets that may result in orbital migration (e.g., Goldreich & Tremaine, 1979; Tanaka et al., 2002; D’Angelo et al., 2003; Paardekooper et al., 2011; Dittkrist et al., 2014), how these protoplanets form and grow by accreting solid components and gas (e.g., Bodenheimer & Pollack, 1986; Ida & Makino, 1993; Pollack et al., 1996; Thommes et al., 2003; Fortier et al., 2013), their gravitational interaction among each other (e.g., Chambers et al., 1996; Raymond et al., 2009), photoevaporation of both protoplanetary disks (Hollenbach et al., 1994; Clarke et al., 2001; Alexander et al., 2014) and planetary atmospheres (Lammer et al., 2003; Owen & Jackson, 2012; Jin et al., 2014), and the long-term evolution of planets and their atmospheres (e.g., Bodenheimer & Pollack, 1986; Guillot, 2005; Fortney & Nettelmann, 2010; Mordasini et al., 2012c). While all these processes leave an imprint on the final planetary systems, observing them while they are in action has proven to be very challenging and was possible only in rare cases (e.g., Kepler et al., 2018). Global models of planet formation can mitigate this shortcoming by combining as many relevant physical processes as possible and simulating the growth and evolution of planets in an end-to-end fashion. Thereby, they provide a link between properties of disks and observables of the resulting planets. When employed within a Monte Carlo experiment with distributions of initial conditions, synthetic planet populations can be produced and statistically evaluated (e.g., Ida & Lin, 2004a; Mordasini et al., 2009a; Ndugu et al., 2018). Such population synthesis frameworks are increasingly able to produce different kinds of planets, from terrestrial-sized rocky planets to gas giants, using the same formation model.

The core accretion scenario (Perri & Cameron, 1974; Mizuno et al., 1978; Mizuno, 1980), in which a solid planetary core forms that may subsequently accrete a gaseous envelope, has been recognized as the most common planetary formation avenue. Concerning the

problem of how this solid core grows, two different approaches have emerged: commonly, the growth of the solid component has been modeled as the accretion of \sim km-sized planetesimals (e.g., [Ida & Makino, 1993](#); [Thommes et al., 2003](#)). Under this assumption, the thresholds in the disk properties responsible for the emergence of different planet types are determined by the availability of planetesimals at the position of a growing planet and by the timescale for accreting them ([Lissauer, 1987, 1993](#); [Kokubo & Ida, 2000](#)). In recent years, a growing body of literature includes the accretion of mm to cm-sized “pebbles”, whose motions are decoupled from the gas disk ([Ormel & Klahr, 2010](#); [Lambrechts & Johansen, 2012](#); [Ndugu et al., 2017](#), see Sect. 2.2.7). Here, the resulting radial motion of the particles causes an interrelation between the inner and outer regions of the disk ([Morbidelli & Nesvorný, 2012](#); [Lambrechts & Johansen, 2014](#); [Ormel et al., 2017](#)).

Both approaches have allowed the unambiguous predetermination of planetary parameters from initial conditions (e.g., [Kokubo & Ida, 2002](#); [Ida & Lin, 2004b](#); [Lin et al., 2018](#)). However, with ever more sophisticated models of increasing complexity, it is uncertain whether these relationships persist. In particular, the inclusion of N -body treatment of protoplanets could destroy these connections due to the chaotic component it introduces. A number of studies have addressed this problem in different ways, either by categorizing the outcomes of simulations with different initial conditions ([Mordasini et al., 2009a, 2012a](#); [Bitsch et al., 2015, 2018a](#); [Miguel et al., 2020](#)), or by relating synthetic populations to the observed sample of exoplanets ([Mordasini et al., 2009a](#); [Chambers, 2018](#); [Fernandes et al., 2019](#); [Mulders et al., 2020](#)) or transitional disks ([Chaparro Molano et al., 2019](#)). A main limitation of these advances has been their restriction to a particular region of the planetary parameter space.

Recent advancements of our formation model ([Emsenhuber et al., 2020a](#)) now allow for an extension of these investigations to the full range of currently known planet types. Therefore, in this study, I statistically assess the relations between a number of relevant disk properties and the emerging planet types in the context of the core accretion paradigm. To this end, I investigate synthetic planet populations computed with the Generation III *Bern* Model of planet formation and evolution ([Emsenhuber et al., 2020a](#)). The publication this chapter is based on ([Schlecker et al., 2021](#), in press) is part of the NGPPS series of papers. Previous studies in this series have presented populations from this model with different numbers of planets per system ([Emsenhuber et al., 2020b](#)) and varying host star masses ([Burn et al., in press](#)). Here, I focus on two populations of systems around solar-type stars: *NG73* for isolated single planets, and *NG76* with 100 planetary embryos growing concurrently ([Emsenhuber et al., 2020b](#)). I thereby take care to follow a purely

data-driven approach and do not presuppose planet types motivated by observations or theoretical arguments.

This chapter is divided into six sections. In Sect. 3.2, I describe the formation model and introduce the synthetic planet populations. I then present a cluster analysis performed on these populations in Sect. 3.3. Section 3.4 investigates to what degree the identified clusters of similar planets can be predicted from properties of protoplanetary disks. In Sect. 3.5, I interpret our results and discuss their implications for planet formation. I conclude by summarizing our findings in Sect. 3.6.

3.2 Planet population synthesis

This work analyzes synthetic planet populations for solar-mass host stars from the Generation III Bern global model of planet formation and evolution (Emsenhuber et al., 2020a, see Sect. 2.5.1). The formation part of the model combines the evolution of a protoplanetary disk with both gas and solids components, the growth and determination of the internal structure of protoplanets, their dynamical interactions and gas-driven planetary migration.

The gas disk is modeled as a viscously accreting disk (Lüst, 1952; Lynden-Bell & Pringle, 1974; Pringle, 1981) with an α -parametrization (Shakura & Sunyaev, 1973) for the turbulent viscosity. The vertical structure is computed following Nakamoto & Nakagawa (1994) and Hueso & Guillot (2005) under an evolving luminosity of the star (Baraffe et al., 2015). The solid disk component is modeled in a fluid-like description where the dynamical state of planetesimals is given by the stirring due to other planetesimals and protoplanets (Thommes et al., 2003; Chambers, 2006; Fortier et al., 2013).

The formation of protoplanets follows the core accretion paradigm (Perri & Cameron, 1974; Mizuno et al., 1978; Mizuno, 1980) with planetesimal accretion in the oligarchic regime (Ida & Makino, 1993). We calculate the structure of the planetary envelopes by directly solving one-dimensional internal structure equations (Bodenheimer & Pollack, 1986). Initially, gas accretion is limited by the ability of the planet to radiate away the gravitational energy release by accretion of solids and gas (Pollack et al., 1996; Lee & Chiang, 2015). At this stage, the internal structure is used to compute the gas accretion rate. Once a planet exhausts the supply from the gas disk (either because cooling becomes efficient or because the disk disperses), the envelope is no longer in equilibrium with the disk and contracts (Bodenheimer et al., 2000). In this *detached* phase, the internal structure equations are used to determine the planet's radius. The formation stage also includes gas-driven planetary migration in the Type I (Paardekooper et al., 2011) and Type II (Dittkrist

et al., 2014) regimes.

The planetary seeds start with a mass of $0.01 M_{\oplus}$ and are inserted with random initial orbital distances a_{start} drawn from a log-uniform distribution between the inner disk edge and 40 au. When multiple embryos are present in the same disk, their gravitational interactions are modeled during the first 20 Myr using the Mercury N -body integrator (Chambers, 1999). After this time, the model switches to the evolutionary stage. Here, the thermodynamical evolution is calculated for each planet individually up to a simulation time of 10 Gyr. This stage includes atmospheric loss via photoevaporation (Jin et al., 2014) and tidal migration. As a result, the model is able to compute the planets' masses, radii, and luminosities as a function of time.

For a thorough description of the Generation III Bern Model and an outline of recent advancements of the framework (Alibert et al., 2005; Mordasini et al., 2009a, 2012c,c; Alibert et al., 2013), we refer to Emsenhuber et al. (2020a).

Synthetic planet populations are produced by running the model in a Monte Carlo scheme, where initial conditions are drawn randomly from distributions motivated by observational (Santos et al., 2003; Lodders, 2003; Andrews et al., 2010; Venuti et al., 2017; Ansdell et al., 2018; Tychoniec et al., 2018) or theoretical constraints (Drażkowska et al., 2016; Lenz et al., 2019). The distributed variables include the initial gas disk mass M_{gas} , the inner edge of the disk r_{in} , its dust-to-gas ratio $\zeta_{\text{d,g}}$, the mass loss rate due to photoevaporative winds \dot{M}_{wind} , and the starting locations of the planetary seeds a_{start} . The values or distributions of all model parameters are listed in Tab. 2.1 and are motivated in detail in Emsenhuber et al. (2020a) and Emsenhuber et al. (2020b).

Our goal is to uncover characteristic links between these properties and the emerging planet types, which requires to robustly define the latter first. This step may be impaired by the stochasticity of an N -body treatment that smears the boundaries between clusters of similar planets. We thus examine both a population with a single planet per system and a population with multiple planets per system. For the single-planet population, called NG73, 30,000 systems were simulated. In 29,455 systems the planet was not accreted onto the star and is still present after 5 Gyr, which we consider as time of observation.

To consider the impact of gravitational interactions among planets, we investigate the multi-planet population NG76 and compare it to the single-planet case. In each of its systems, an initial set of 100 protoplanets competed for material and interacted gravitationally. All other boundary conditions were left the same, and the Monte Carlo parameters were drawn from the same distributions. The N -body module integrated for 20 Myr to cover the entire formation phase with planets still embedded in the disk, as well as an appro-

priate subsequent evolutionary era without disk interaction (Emsenhuber et al., 2020a). Out of the 1000 simulated systems, 32,030 planets survived until $t = 5$ Gyr. For detailed descriptions of both planet populations, see Emsenhuber et al. (2020b) and Schlecker et al. (2020b).

3.3 Cluster analysis

A cluster analysis aims at identifying groups of entities that share similar properties in a specific set of parameters. In our case, we aim to explore which distinct planet species emerge from our planet formation model and how they compare to observed (exo-)planet types. Accordingly, we chose as training features three parameters typically obtained from exoplanet observations: the orbital semi-major axis a , the planet mass M_P , and the planet radius R_P . Our clustering was done in a purely data-driven fashion and without any prior knowledge on existing or expected planet types. The only information our clustering model received was a snapshot of our synthetic planet population at a simulation time of 5 Gyr.

3.3.1 Data preparation

In general, clustering methods are not scale-invariant (Jain & Dubes, 1988). The application of cluster algorithms to unevenly scaled data sets can thus lead to compromised results. Based on the distribution of the parameters of interest in our data set, we rescaled the features a , M_P , and R_P by applying a \log_{10} .

3.3.2 Model selection and hyperparameters

We performed the clustering using Gaussian Mixture Models (GMM, McLachlan, 1988), a class of hierarchical, probabilistic clustering algorithms. A GMM consists of multiple components $i = 1 \dots N$ of multivariate normal distributions, each characterized by its weight ϕ_i , its mean μ_i , and its covariance matrix Σ_i . The model then takes the form

$$\sum_{i=1}^N \phi_i \mathcal{N}(\mu_i, \Sigma_i). \quad (3.1)$$

During training on a data set, the parameters ϕ_i , μ_i , and Σ_i are updated using the expectation-maximization (EM) algorithm (Hartley, 1958). A free hyperparameter is the number of Gaussian components N , that is, the number of Gaussian distributions the data points are assumed to be generated from. The trained GMM gives each data point a set of N probabilities, corresponding to the probability that the data point belongs to a specific component i . When we classified our data, we assigned each planet the component (i.e.,

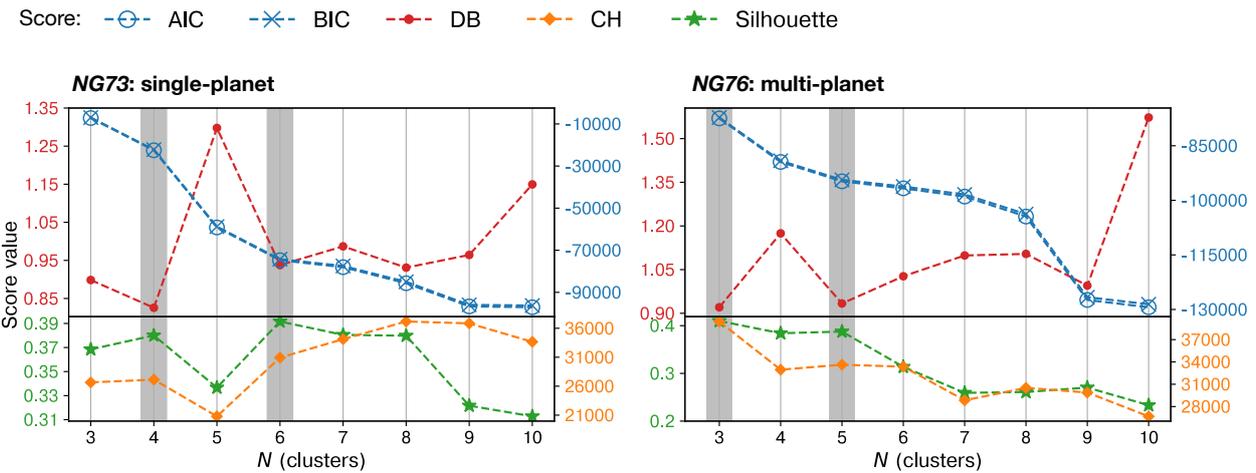


Figure 3.1: Validation scores for Gaussian Mixture Models with different numbers of components N . For AIC, BIC, and DB (top panels), lower values are preferred; and for Silhouette score and CH (bottom panels), higher values are preferred. AIC and BIC generally show indistinguishable values. Based on these scores, sensible choices are $N = 4$ and $N = 6$ for NG73, and $N = 3$ and $N = 5$ for NG76 (highlighted in gray). Note the different y-axis scales.

the planet cluster) with the highest probability.

Since GMMs, and clustering algorithms in general, are unsupervised methods, the selection of a “best” model has to be seen in the context of the goal we want to achieve. We aimed at identifying groups of planets based on overdensities in the planetary parameter space, regardless of their shape. With this goal in mind, we have explored several other algorithms in addition to GMMs and found that they consistently performed worse on our data set (see Appendix A.1). Using the `scikit-learn` (Pedregosa et al., 2011) implementation of GMMs with default arguments, the only free hyperparameter was the number of clusters in the data N . In finding the optimal choice of N , we were aided by several validation metrics. We considered the Akaike Information Criterion (AIC, Akaike, 1973; Cavanaugh & Neath, 2019), the Bayesian Information Criterion (BIC, Schwarz, 1978), the Davies-Bouldin score (DB, Davies & Bouldin, 1979), the Caliński-Harabasz score (CH, Caliński & Harabasz, 1974), and the Silhouette score (Rousseeuw, 1987). These metrics assess clustering performance with different approaches, and due to the complex structures in our data they can contradict each other. We provide a detailed description of the different metrics in Appendix A.1.2. For now, it is important to note that AIC, BIC, and DB should be minimized, and CH and the Silhouette score should be maximized. In Fig. 3.1, we show the different scores for GMMs with $N \in [3, \dots, 10]$ upon applying them to our single-planet (NG73) and multi-planet (NG76) population, respectively. For NG73, two potential choices stick out, $N = 4$ and $N = 6$. To decide between these options, we pro-

3 Predetermination of Planet Types in Global Core Accretion Models

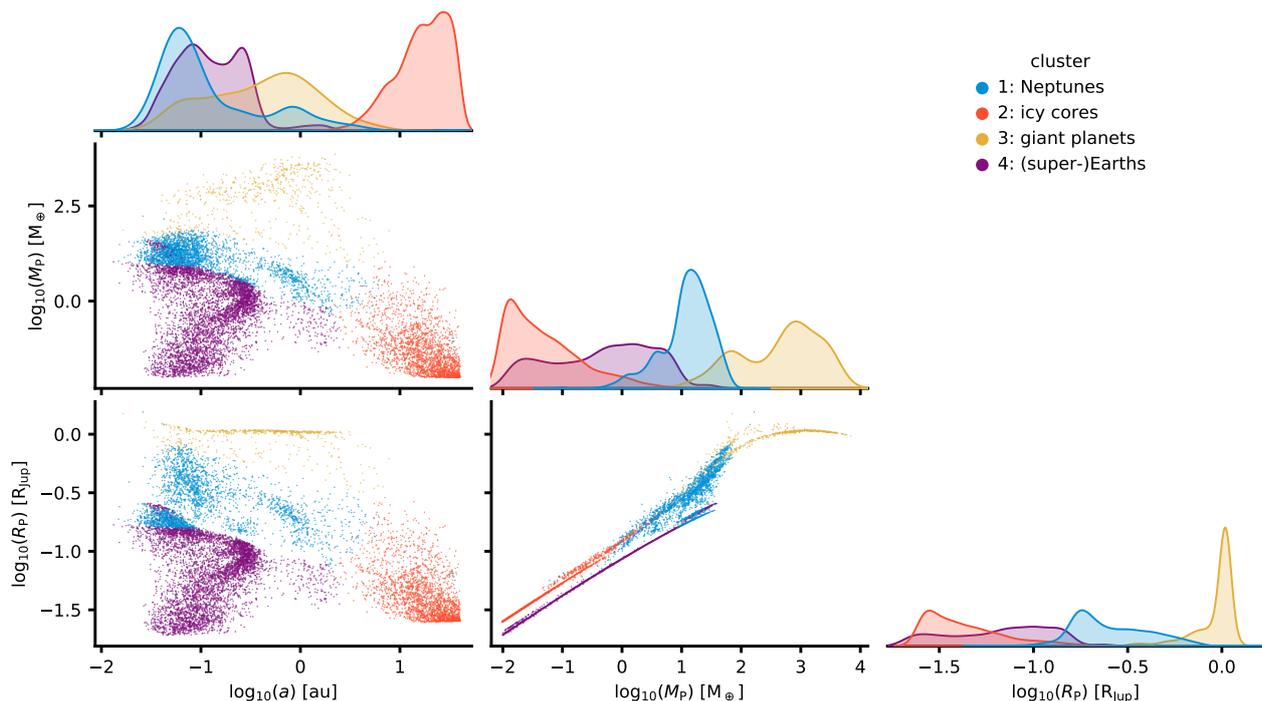


Figure 3.2: Planet clusters in a 5 Gyr old synthetic planet population with a single planet per system. For all combinations of planet observables a , M_P , and R_P , the different colors denote clusters identified by a four-component Gaussian Mixture Model (GMM). On the diagonal, we show Kernel Density Estimates of the distributions. Without any information about the physics in our formation model, the GMM identified four planet species roughly corresponding to (sub-)Neptunes (blue), icy cores (red), giant planets (yellow), and (super-)Earths (purple).

duced diagnostic scatter plots where all possible 2D projections of the planetary parameter space are shown with planets color-coded by cluster affiliation. The plots for the candidate models are shown in Fig. A.2. While human bias might be an issue at this step, we took care to judge the clustering only based on over- and underdensities of planets and not based on where we expected different planet types. We found that the GMM with $N = 4$ performed best. For NG76, both $N = 3$ and $N = 5$ yielded promising scores. By judging the corresponding diagnostic plots, we concluded that $N = 5$ clusters is the preferred mode. With all hyperparameters fixed, we performed the unsupervised training of our nominal GMM on the full data sets and considering full covariance matrices.

3.3.3 Detected planet clusters

In the single-planet case, the clustering algorithm identified four separate planet species in our population. Figure 3.2 shows these clusters in the various projections in $\{a, M_P, R_P\}$ space. In general, we notice clear separations between the clusters in all projections, albeit with visible contaminations. Ordered by ascending planetary mass, the clusters are as

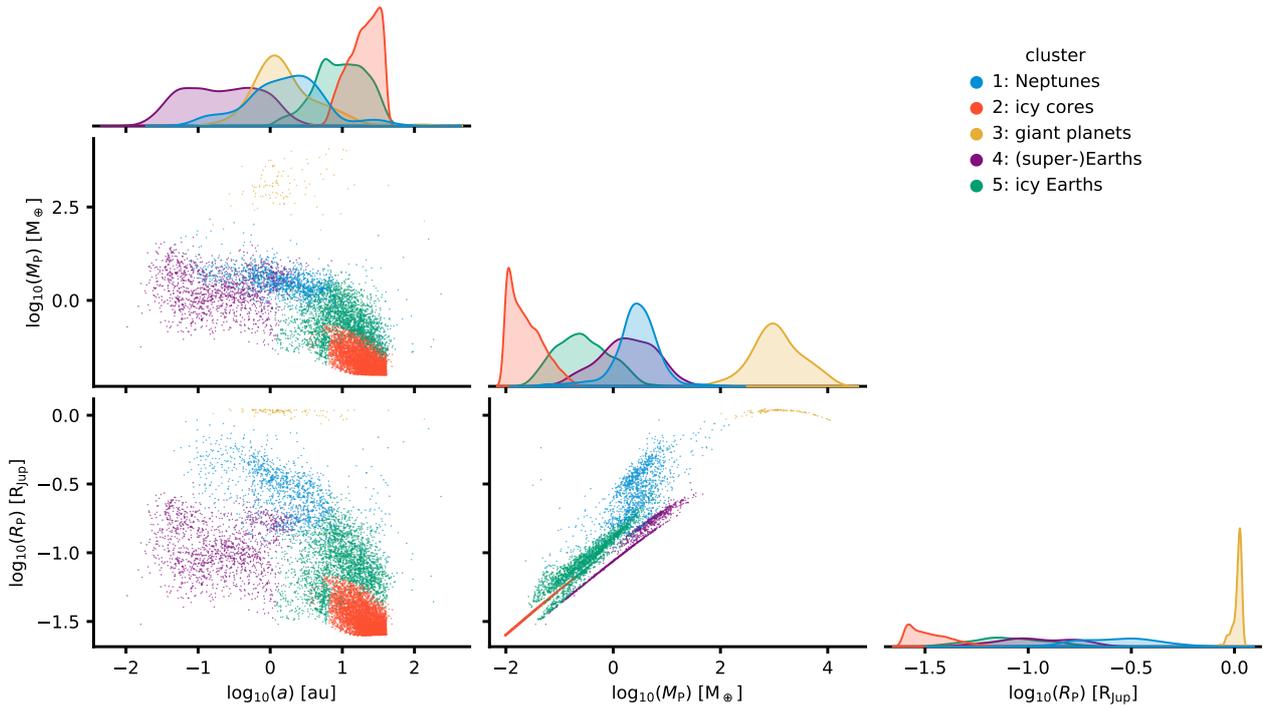


Figure 3.3: Same as Fig. 3.2, but for a multi-planet population. The Gaussian Mixture Model (GMM) prefers solutions including a fifth component of distant, icy planets shown in green. In general, the clusters are less clearly separated than in the single-planet population.

follows: clusters 2 and 4 are populations of bare planet cores without atmospheres, and they are cleanly separated in semi-major axis. Both clusters are separate from cluster 1, which are close-in planets enhanced in gas and with masses of mostly tens of M_{\oplus} . A fourth distinct group of very massive planets ($M_P \gtrsim 100 M_{\oplus}$) is formed by cluster 3 with a clear separation from the other species.

Since the GMM is not aware of the underlying physics these clusters result from, it is of interest to interpret the identified clusters and relate them to known planet types. Cluster 2 corresponds to an unobserved population of distant, low-mass planets. As they formed beyond the water ice line and are rich in volatile species, we refer to this group as “icy cores”. Cluster 4 planets are atmosphere-less and rocky, and thus comparable to the observed population of close-in terrestrial planets and super-Earths (e.g., [Hsu et al., 2019](#)). By simultaneously taking into account all dimensions of the parameter space, the GMM spatially separated icy cores and (super-)Earths in the region of the water ice line (without any information about its existence). This lead to the clean separation of rocky and icy planets in the $M_P - R_P$ diagram (diagonal lines in the plot). Cluster 1 roughly corresponds to the observed population of (sub-)Neptunes. In planet radius space, these planets are mostly located above the radius valley (e.g., [Fulton et al., 2017](#); [Mordasini,](#)

2020, see discussion in Sect. 3.4.4). There is some contamination by cluster 1 planets in the region of the largest and closest super-Earths, which we attribute to the inability of a GMM to fit a deviation from the otherwise extremely straight line of cluster 4 planets in $M_P - R_P$ space. Finally, cluster 3 can be identified as gas giant planets. This becomes especially clear in the $M_P - R_P$ plane, where they occupy the region where in the physical model electron degeneracy occurs. This effect flattens off the mass-radius relation at the high-mass end (e.g., Chabrier et al., 2009).

Figure 3.3 illustrates the clustering in the multi-planet case, during which we ignored the system affiliation of the planets and treated them as independent entities. Based on the scoring scheme described above, the clearest clustering can be achieved with five components. The overall partitioning appears similar to before, and the fifth component not present in the single-planet population covers planets on distant orbits that have intermediate densities and masses of roughly $0.05 M_\oplus - 3 M_\oplus$. We refer to these planets as “icy Earths”. These planets are distributed in a sharp line in mass-radius space, which makes the GMM consider them detached from the more dispersed “icy cores”. Notably, the bulk of the “Neptunes” moved to more distant orbits compared to the single-planet case. This is in line with the observed existence of Neptune-sized planets at orbital distances of several au Suzuki et al. (2016); Kawahara & Masuda (2019). For a comparison of Bern model planets and gravitational microlensing events, we refer to (Suzuki et al., 2018).

3.3.4 Model validation

Unlike supervised machine learning algorithms, unsupervised techniques cannot be tested by applying the trained model to a test set due to the lack of “labeled” data. For validation of the clustering itself, we used the aforementioned performance metrics. To evaluate how robust the detected clustering is, we let the model predict the cluster affiliation of a dataset of similar structure and compared these predictions to the original clustering. To produce these test data, we employed Gaussian Mixtures of 80 components and full covariance matrices as generative probabilistic models. We trained them on the $\{a, M_P, R_P\}$ subspace of the original population synthesis data. The samples drawn from these models show a very similar structure in the whole domain (compare Fig. 3.4). Note that these “planets” are entirely the product of the generative models and have never been in contact with a physical formation model.

For comparison, we also fed our nominal clustering models with samples drawn from log-uniform distributions with boundaries roughly corresponding to the suprema of the population synthesis data, i.e. $a \sim 10^{\mathcal{U}(-1,2)}$, $M_P \sim 10^{\mathcal{U}(-2,4)}$, and $R_P \sim 10^{\mathcal{U}(-2,0)}$. With these pseudo-random data, the models predict clusters that do not resemble the original

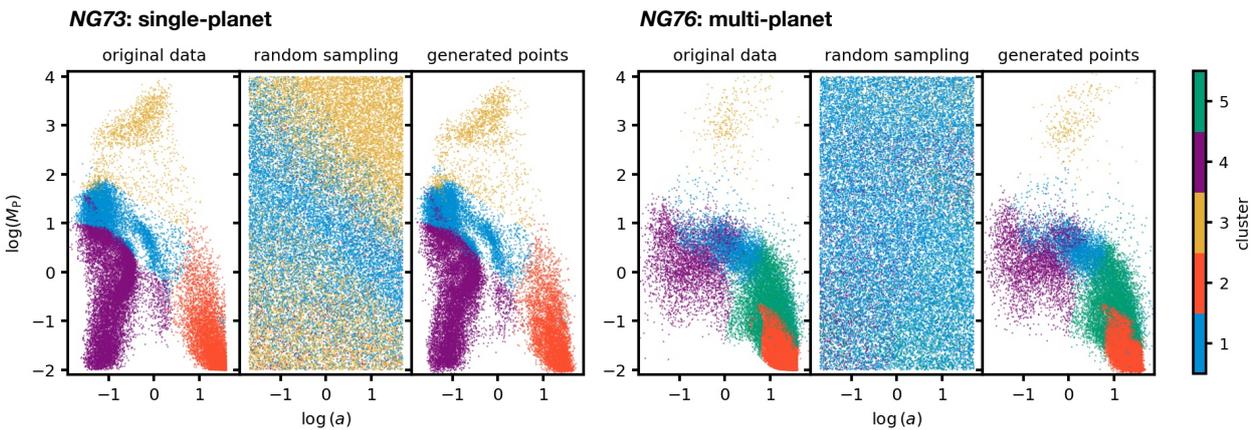


Figure 3.4: Model validation via generative models. For each of the two planet populations, we show the clustering result of our Gaussian Mixture Model on Population Synthesis data (left), random noise (center), and data from a generative model (right). Note that the latter do not stem from a physical formation model but were generated from a high-order GMM that was trained on the original data. The clusters detected in these new data show largely the same structure as the original ones, whereas in the random noise no reliable clusters are found.

structures and they appear in most projections almost random. These two tests show that our trained models neither overfit the dataset, nor do they produce any clear clusters where none are expected. The generative models can also be used to draw a virtually unlimited number of synthetic planets when the computational costs of employing the full formation model are prohibitive (similar to [Mulders et al., 2018](#)).

3.3.5 Planet clustering as a function of simulation time

The cluster analysis took place at a simulation time of $t = 5$ Gyr. We now trace the identified clusters back in time to investigate their past evolution. Figure 3.5 shows their position in semi-major axis-mass space at simulation times 0.1 Myr, 0.3 Myr, 0.6 Myr, 1 Myr, 2 Myr, and 10 Myr. In particular in the single-planet population, the clusters occupy distinct domains already at early times and follow characteristic paths in this parameter space. These paths are set by concurrent accretion and planet migration and their respective timescales.

In the following, we focus on the single-planet case where the evolutionary paths can be traced most clearly. At the beginning, all planets are still of such low mass that migration has little effect. Planet growth is determined by the local planetesimal density, feeding zone size, and orbit timescale, and it is most efficient at intermediate orbital distances ([Emsenhuber et al., 2020b](#)). At a few 10^5 yr, an outward migration zone located at a few au divides the planetary tracks into two branches. On the outer branch, giant planets evolve similarly as the outer wing of Neptunes. They branch off when runaway gas accretion sets in, while Neptunes continue migrating inward with moderate growth.

3 Predetermination of Planet Types in Global Core Accretion Models

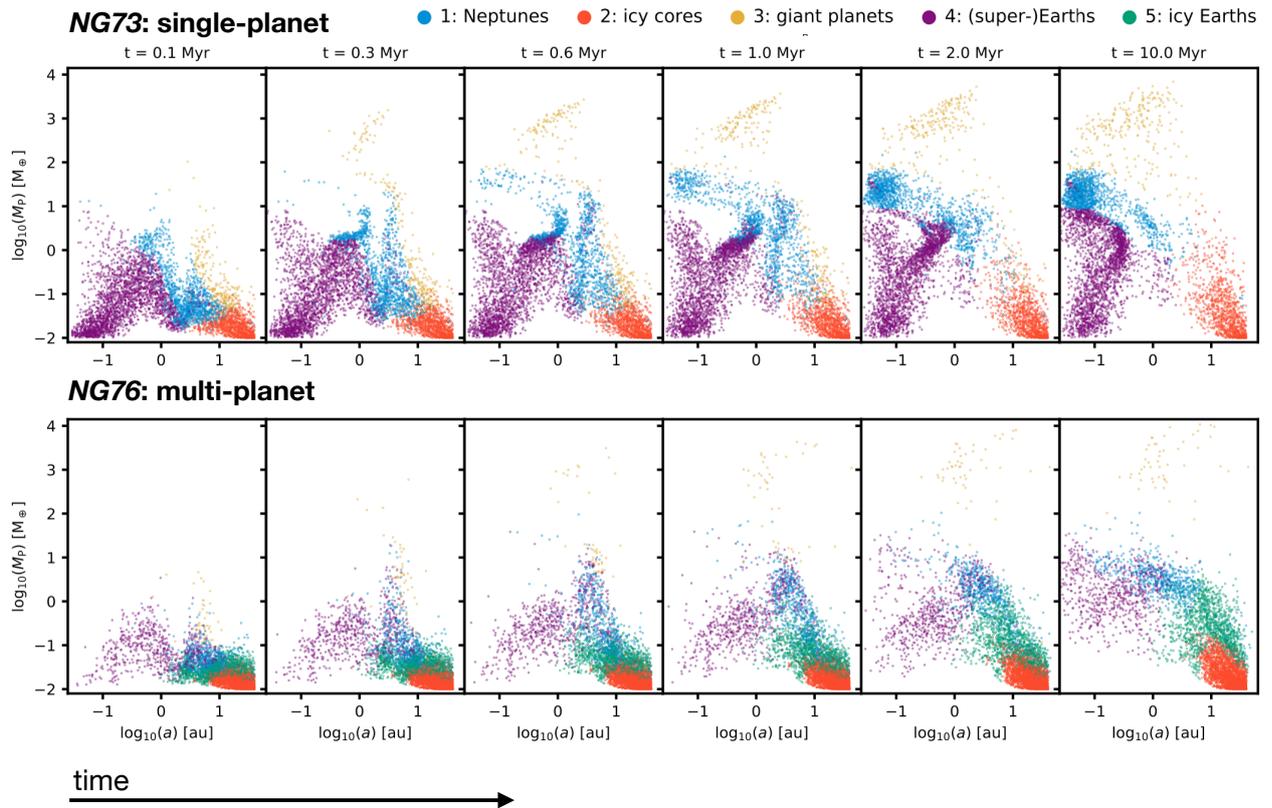


Figure 3.5: Early time evolution of the clusters identified by the Gaussian Mixture Model. Each subplot shows a sample of 5000 planets at their current position in semi-major axis-mass space and color-coded by their future cluster affiliation, which is only determined at $t = 5$ Gyr. Concurrent accretion and migration leads to characteristic evolutionary paths. Distinct groups of planets form already at early simulation times.

At later times, another outward migration zone leads to the underdensity in the cluster of close-in (super-)Earths. Icy cores do not exhibit significant growth and largely remain in their initial domain.

Most of the processes that define the different planet types in this parameter space are finished after a few Myr or, at the latest, when the gas disk disperses. Exceptions are atmospheric photoevaporation, which happens on 100 Myr to Gyr timescales (e.g., [Lopez et al., 2012](#); [King & Wheatley, 2021](#)) and still turns some close-in (sub-)Neptunes into super-Earths, and tidal interaction with the host star affecting some ultra-short period planets. In the case of multiple planets per system, N -body interactions can have an additional long-term impact. A striking result of planet-planet interactions are the significantly lower migration rates compared to the single-planet case, in particular in the Neptunes cluster.

In general, it appears that planet populations form distinct groups very early in the formation process. This begs the question whether the cluster affiliation of a planet can

already be predicted from the initial conditions of the simulation.

3.4 Prediction of planet clusters

Our planet formation model provides a deterministic link between properties of protoplanetary disks and properties of planets. This link could be blurred by N -body interactions between the planets, hence in the following experiment we consider first the single-planet population. Our approach was to employ a Random Forest classifier (Ho, 1998; Breiman, 2001) to predict the cluster of a planet from its corresponding set of disk properties. Random Forests are ensembles of uncorrelated, binary classifiers known as decision trees. Such ensembles achieve strongly improved generalization accuracies compared to single-tree classifiers by constructing trees in pseudorandomly selected feature subspaces (Tin Kam Ho, 1995). The individual trees are further decorrelated by drawing, with replacement, random subsets of the input data during training (“bagging”, Breiman, 1996).

With varying sizes of the individual clusters (for instance, only $\sim 5\%$ of the planets in NG73 are giant planets), our dataset is strongly imbalanced. This is problematic for classification algorithms such as Random Forests, which aim to minimize the overall error rate and thereby tend to neglect minority classes (Chen et al., 2004). To account for this imbalance, we employed a balanced Random Forest classifier as implemented in the `imbalanced-learn`¹ python package. This variant of Random Forest randomly under-samples each bootstrap sample on the individual tree level during training (Lemaître et al., 2017).

3.4.1 Data preparation, hyperparameters, and training

Our classifier learned rules based on four features: the initial gas disk mass $M_{\text{gas},0}$, the initial solid disk mass $M_{\text{solid},0}$, the initial orbital distance of the planetary embryo a_{start} , and the disk lifetime t_{disk} . The solid disk mass is a derived quantity that we computed from the gas disk mass and host star metallicity. We rescaled these features to account for their large differences in scale: t_{disk} and a_{start} were transformed by a \log_{10} function, and $M_{\text{gas},0}$ and $M_{\text{solid},0}$ were modified to roughly Gaussian distributions by the Box-Cox transform (Box & Cox, 1964). The clustering above assigned each synthetic planet a probability to belong to each of the clusters. For the subsequent analysis, we avoided planets that cannot be mapped clearly to a cluster and kept only those with a probability of affiliation > 0.99 . This decreased our sample from 29455 to 23278 planets. Finally, we divided the data into a random subset containing 80% of the initial data for training and a test set with the remaining 20% to determine the performance of the classifier. The resulting training set

¹<https://imbalanced-learn.org>

contains between 1059 (giant planets) and 8486 ((super-)Earths) planets per cluster. We trained an ensemble of 500 fully grown estimators, that is, without reducing the depth of the trees by pruning them, on this set.

3.4.2 Error and performance analysis

To measure the generalization performance of the trained model already during its development, we predicted clusters from the out-of-bag samples, which were never seen by the respective estimator during training. The average of the resulting out-of-bag score produces an estimate for the accuracy of the entire ensemble, and we obtained a score of 98% here. However, classification accuracy is not a sufficient performance measure since we are dealing with a strongly skewed data set. In the following, we investigate the types of errors our model makes and measure its performance.

We computed a confusion matrix using five-fold cross-validation. For this purpose, the dataset was randomly split into five evenly sized folds; the model was trained five times on $5 - 1 = 4$ folds, and then evaluated on the fold it was not trained on. The left panel of Fig. 3.6 shows the confusion matrix produced from the labeled training set and the predictions from cross-validation. Rows correspond to the actual clusters, and columns are the predictions of our model. Each field $x_{i,j}$ in the matrix shows the fraction of times a planet of cluster i was classified as a planet of cluster j . Most planets fall into the diagonal, meaning a correct classification. All clusters are predicted with more than 95% accuracy and the largest errors occur for clusters 2 and 3. The right panel of the figure shows the same matrix with the correct classifications removed and the color map rescaled. It is obvious that the errors are largely symmetric. The highest rate of misclassification occurred between clusters 2 and 3 (3% of icy cores were confused with giant planets and vice-versa). The reason is that the former are frequently progenitors of the latter, and prediction of those planets that just (do not) reach the conditions for runaway gas accretion is difficult (compare Fig. 3.2).

To estimate the generalization error the model makes when applied to data not part of the training set, we measured its performance on the test set of 4656 systems we held out before. Based on five-fold cross-validation, it achieves an overall accuracy of 97% and misclassifications occur between the same clusters as seen in the training set. This shows that the model is not significantly overfitted.

3.4.3 Results of planet predictions

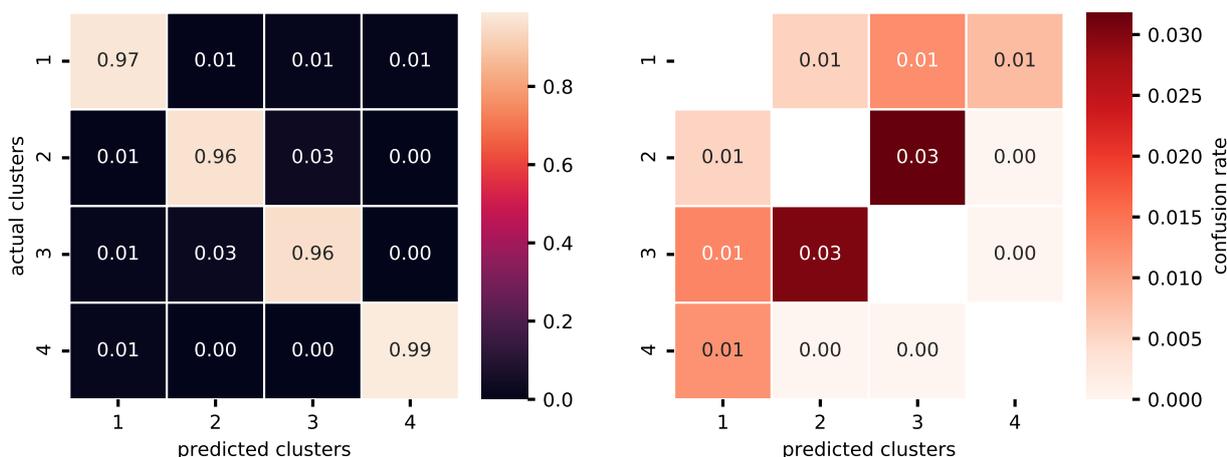


Figure 3.6: Confusion of planet classifications. Left: confusion matrix from five-fold cross-validation. Rows are the actual clusters and columns are the predicted clusters. All clusters are classified with more than 95% accuracy. Right: same, but correct classifications removed to emphasize errors. Most misclassifications occur between clusters 2 and 3, which correspond to icy cores and giant planets.

Correlations with disk properties

For each of the clusters identified in Sect. 3.3.3, we show the distributions and pairwise relationships of their corresponding disk properties in Fig. 3.7. Underdensities in the scatter plots are due to removed planets of ambiguous cluster affiliation. Unsurprisingly, giant planets (yellow) grow in disks with large reservoirs of solid material M_{solid} and high gas mass M_{gas} . It is evident that most of these clusters, which are labeled at “observation time” $t = 5$ Gyr, form groups already in this parameter space, that is, before the simulations started. However, they differentiate distinctly only in the projections involving the start position of planetary embryos a_{start} . The separation is especially clear in $a_{\text{start}} - M_{\text{solid},0}$ space, which shows the least overlap of different clusters. With increasing initial orbital distance, the dominant planet species are (super-)Earths, Neptunes, giant planets, and icy cores.

Feature importance

Our classification model reaches high accuracies for all planet clusters, but it is interesting to see which disk features are most important for a successful classification. This is possible by measuring the feature importance of the dataset given to the model using the Mean Decrease Impurity MDI (Breiman et al., 1984). MDI quantifies to what extent a feature reduces the impurity of the trees in the Random Forest. Put simply, it is a measure

3 Predetermination of Planet Types in Global Core Accretion Models

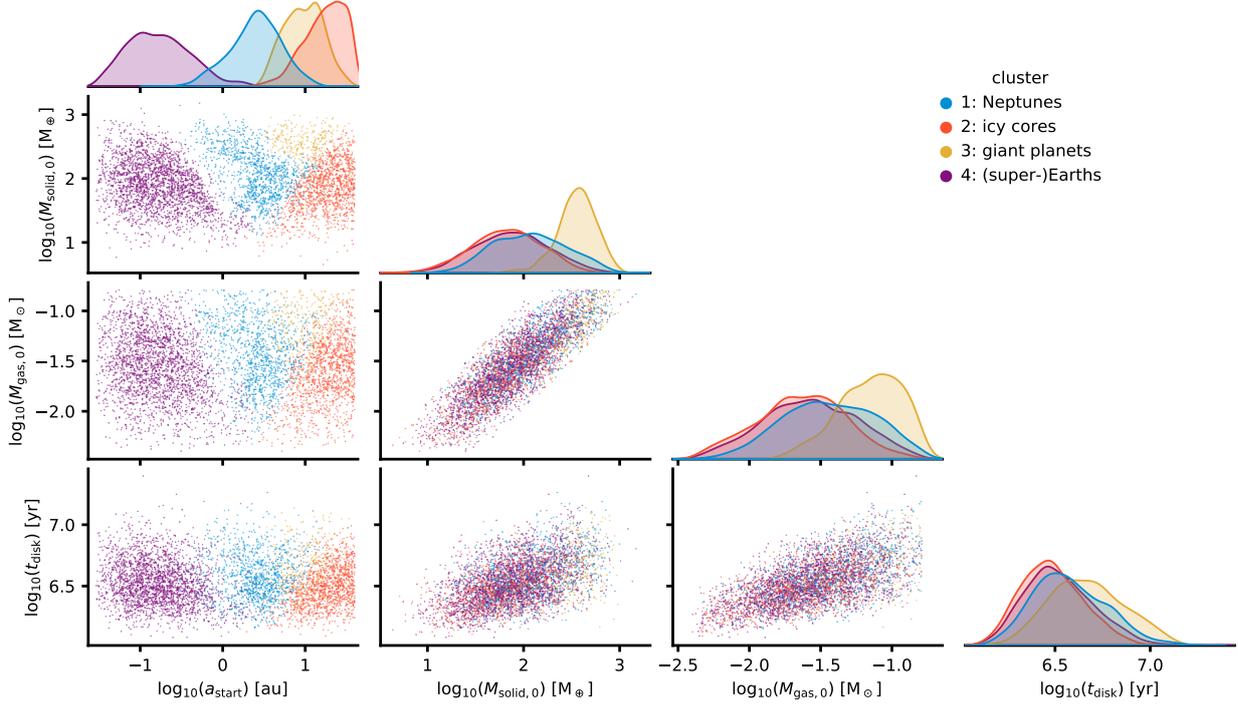


Figure 3.7: Pairwise relationships between all disk parameters, sorted by cluster affiliation. For 5000 randomly sampled planets in the population, each parameter is plotted against every other parameter while the color defines the planet’s cluster. The diagonal panels show the univariate distributions of the respective parameters, again colored by cluster assignment. Planet species most clearly separate in $a_{\text{start}} - M_{\text{solid},0}$ space, and the formation of giant planets (yellow) requires large solid reservoirs and a narrow range of initial orbital distance.

Table 3.1: Feature importances of disk properties

| Input Parameter | $M_{\text{solid},0}$ | $M_{\text{gas},0}$ | a_{start} | t_{disk} |
|-----------------|----------------------|--------------------|--------------------|-------------------|
| <i>MDI</i> | 0.21 | 0.07 | 0.68 | 0.04 |

of how well the nodes can use the feature to split the data set into “pure” child nodes, each containing only data of a single label. A higher score means that the feature is more important for correct classification. We list the *MDI* for each input parameter in Table 3.1. With a score of 0.68, the starting position of the planetary core a_{start} is clearly the parameter most sensitive for predicting a planet’s cluster. The gaseous mass of the disk and its lifetime are the least important features.

However, the degree of dependency on certain disk features varies from cluster to cluster. To get a cluster-specific insight, we multiply for each cluster the mean of each feature

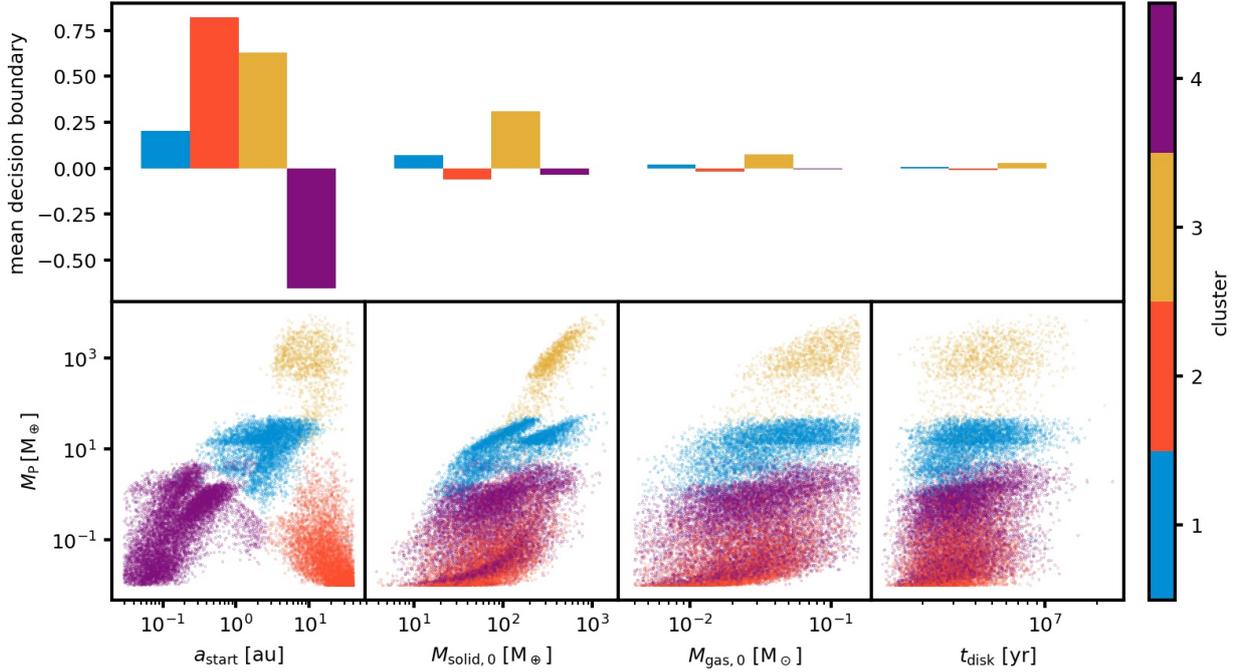


Figure 3.8: Relation between disk features and planet species. Upper panel: Mean decision boundaries of the classifier, indicating the importance of each feature and its preferred magnitude for the different clusters. The starting location of the planet embryo a_{start} shows the largest variance in decision boundary. Giant planets (yellow) are also very sensitive on $M_{\text{solid},0}$ and somewhat sensitive on $M_{\text{gas},0}$. Lower panels: relationship of the input features with planet mass. The starting location of the planet embryo a_{start} shows the strongest correlation with cluster affiliation and planet mass.

with the feature importance. This mean decision boundary

$$D_{c,f} = MDI_f \cdot \langle X_{y=c} \rangle \quad (3.2)$$

denotes for each cluster c the sensitivity of the classifier on feature f . Here, $X_{y=c}$ are the scaled training data with labels y corresponding to cluster c . Fig. 3.8 illustrates all cluster-specific mean decision boundaries. $D_{c,f}$ quantifies the sensitivity on a parameter by its magnitude, as well as the orientation of the decision boundary by its sign. For example, the large negative value of cluster 4 in a_{start} means that these planets prefer small initial orbital distances and their correct classification is very sensitive on this feature.

In the lower panels of Fig. 3.8, we plot all input features against the resulting planet mass at 5 Gyr, which is a proxy for cluster affiliation. Most planet clusters are especially sensitive on the initial orbital distance of the planetary embryo a_{start} . Planets with masses higher than $\sim 10 M_{\oplus}$ are also very sensitive on the solid mass $M_{\text{solid},0}$ and slightly sensitive on $M_{\text{gas},0}$. The disk lifetime t_{disk} shows a weak correlation with planet mass and plays only

a subordinate role.

3.4.4 Differences between single and multi-planet systems

Mutual interactions between planets in the same system introduce a fair amount of stochasticity, and some features that stood out in the single-planet population are smeared out in the multi-planet case. One example is the bimodal distribution of planet radii in the observed exoplanet sample (Fulton et al., 2017; Fulton & Petigura, 2018; Hsu et al., 2018; Van Eylen et al., 2018; Mordasini, 2020, Sect. 2.4), which was theoretically predicted to be caused by photoevaporation of planetary envelopes by high-energy radiation from their host star (Jin et al., 2014; Owen & Wu, 2013; Lopez & Fortney, 2013). Other mechanisms have been proposed to produce this “radius valley” at roughly $2 R_{\oplus}$ as well, including atmospheric loss due to internal heat from cooling planetary cores (Ginzburg et al., 2018; Gupta & Schlichting, 2019), impacts of planetesimals (Wyatt et al., 2019) or other protoplanets (Liu et al., 2015), different internal compositions of planets residing above or below the valley (Zeng et al., 2019; Venturini et al., 2020a), and atmospheric stripping by external radiation sources in stellar cluster environments (Kruijssen et al., 2020). In the Generation III *Bern* Model, photoevaporation by the host star and collisional stripping are taken into account.

The upper panels of Fig. 3.9 show the radius distributions of planets on close orbits ($P < 80$ d) in the single and multi-planet populations, respectively. Overplotted are occurrence rates derived from the *Kepler* mission in Hsu et al. (2019), which we marginalized over the period range 0 d – 80 d. The propagated uncertainties are indicated by vertical bars, and arrows mark upper limits. In our single-planet population, the evaporation valley is much less pronounced in this marginalized radius distribution than in radius-orbital distance space, where it shows a steep negative slope (compare Fig. 3.2). This highlights the importance of characterizing such demographic features in multiple dimensions. Compared to the observed valley at $\sim 2 R_{\oplus}$ (e.g., Fulton et al., 2017; Hsu et al., 2019), the synthetic one is shifted to larger radii. As has been shown in Jin & Mordasini (2018), this is due to atmosphere-less, icy cores that migrated inwards from regions beyond the water ice line. This population is included in the planet cluster representing Neptunes, since the clustering algorithm mainly discriminated between (super-)Earths and Neptunes as rocky and icy planets, respectively.

In the multi-planet population, this is not the case. Here, the different clusters divide close-in planets into bare cores and planets with H/He envelopes, and the emerging radius valley separates the (super-)Earths and Neptunes clusters. Again, the valley is shifted to around $3 R_{\oplus}$. Compared to the single-planet case, the slope of the valley in radius-orbital

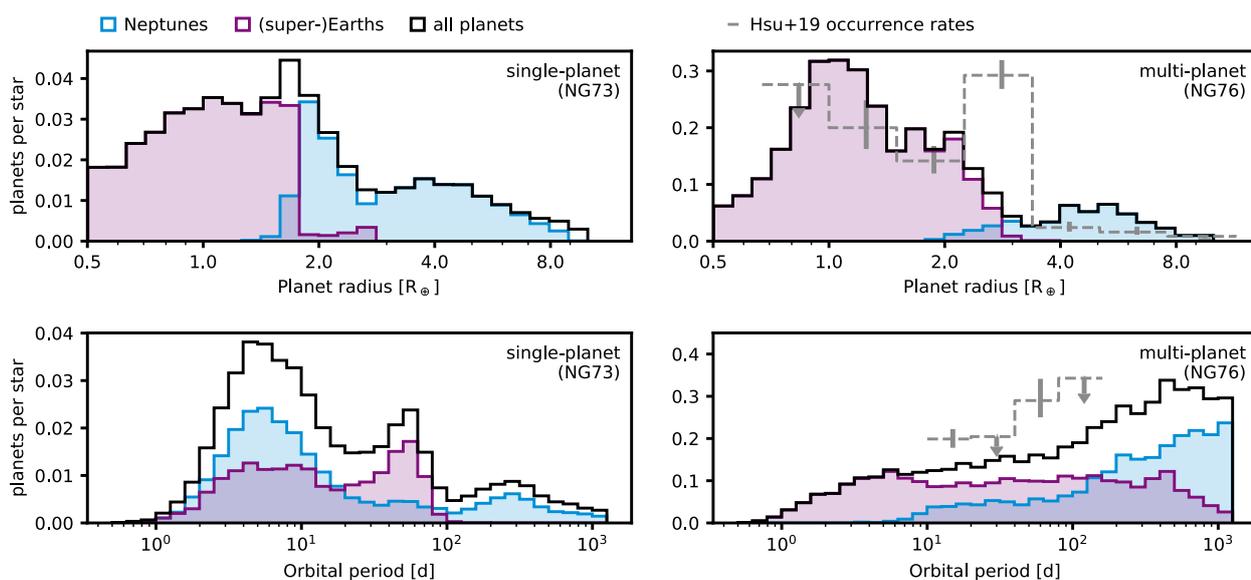


Figure 3.9: Radius and period distributions of Neptunes and (super-)Earths. The contributions by Neptunes and (super-)Earths are shown in blue and purple, respectively. Upper panels: planet radius distribution for planets with periods $P < 80$ d. In the single-planet case (left), a population of migrated, icy cores in the Neptunes cluster shifts the synthetic radius valley to larger radii. In the case of multiple planets per system (right), the minimum in the distribution separates (super-)Earths and Neptunes. Compared to observed occurrence rates from *Kepler* (Hsu et al., 2019, gray), this minimum is shifted towards larger radii.

Lower panels: period distributions of planets $\geq 1 R_{\oplus}$. While the single-planet population (left) shows a multi-modal distribution, the multi-planet population has a continuous slope similar to observed occurrence rates. Note the different normalizations of synthetic and observed planets.

distance is less pronounced, which makes it appear deeper in the one-dimensional radius histogram. Future work within this series will address the synthetic radius valley in a more thorough manner (Mishra et al., in prep.).

Other differences between the single and multi-planet populations can be seen in their period distributions (lower panels of Fig. 3.9). In the single-planet case, the combined contributions from (super-)Earths and Neptunes lead to a multi-modal distribution. On the other hand, the multi-planet population shows a continuous slope. In the range where Hsu et al. (2019) provide reliable occurrence estimates, this slope matches the observed one well. Causes for the difference between the single- and multi-planet case are the displacement of planets in semi-major axis due to gravitational encounters, a lack of close-in “failed cores” due to the high likelihood of such encounters on short orbits, and trapping of planets in resonant chains. In addition, mixed planetary compositions occur as a consequence of merger events. This places the planets into a continuum of bulk densities.

3 Predetermination of Planet Types in Global Core Accretion Models

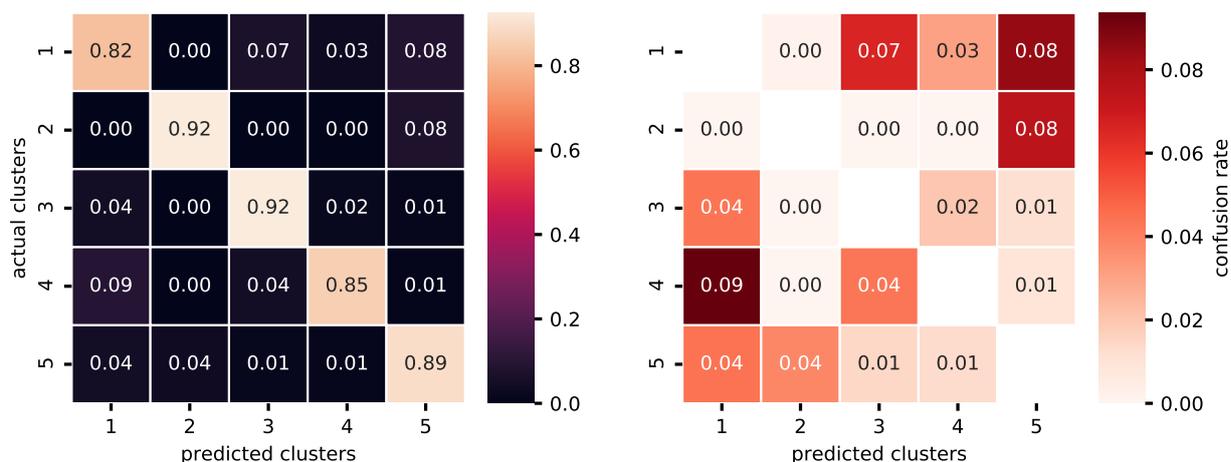


Figure 3.10: Confusion of cluster classifications for a multi-planet population with N -body interactions. Same as Fig. 3.6, but computed for a population with 100 planets per disk that interact gravitationally. Clusters 2 (icy cores) and 3 (giant planets) are predicted most reliably. Due to giant collisions the classifier cannot predict, the super-Earths in cluster 4 are often mistaken for (sub-)Neptunes (cluster 1).

Regardless of this “stochastic processing” of the planets, we attempted to predict their clusters from initial conditions using the same features as in the single-planet case and following the procedure described in Sects. 3.4.1 to 3.4.2. Similar to before, keeping only planets that the GMM assigned to a specific cluster with a probability > 0.99 reduces the set to 21,761 planets. The randomly drawn training set comprising 80 % of the data contains between 252 (giant planets) and 10367 (icy cores) planets per cluster. A balanced Random Forest we trained on this set achieved an accuracy of 89 % based on five-fold cross-validation. The other 4353 systems, which we left out as a test set, are predicted with 86 % accuracy.

Similar to Fig. 3.6, Fig. 3.10 shows the confusion matrix of a Random Forest predicting the planet clusters in the multi-planet population. The ability to predict planet clusters from initial conditions varies across different planet types, with icy cores and giant planets being the most robust species. It can be seen that clusters 1 (Neptunes) and 4 ((super-)Earths), which occupy similar mass ranges, are affected by confusion the most. This is mainly due to the lack of (super-)Earths $\lesssim 0.1 M_{\oplus}$ in the multi-planet case, where they typically fall victim to giant collisions with other planets. Neptunes are frequently mistaken as icy Earths and (super-)Earths are frequently confused to be Neptunes. These three groups of intermediate-mass planets share a similar domain in parameter space.

Figure 3.11 shows the positions of the planets in the multi-planet population in disk property space. Again, the different clusters differentiate the most in solid disk mass and

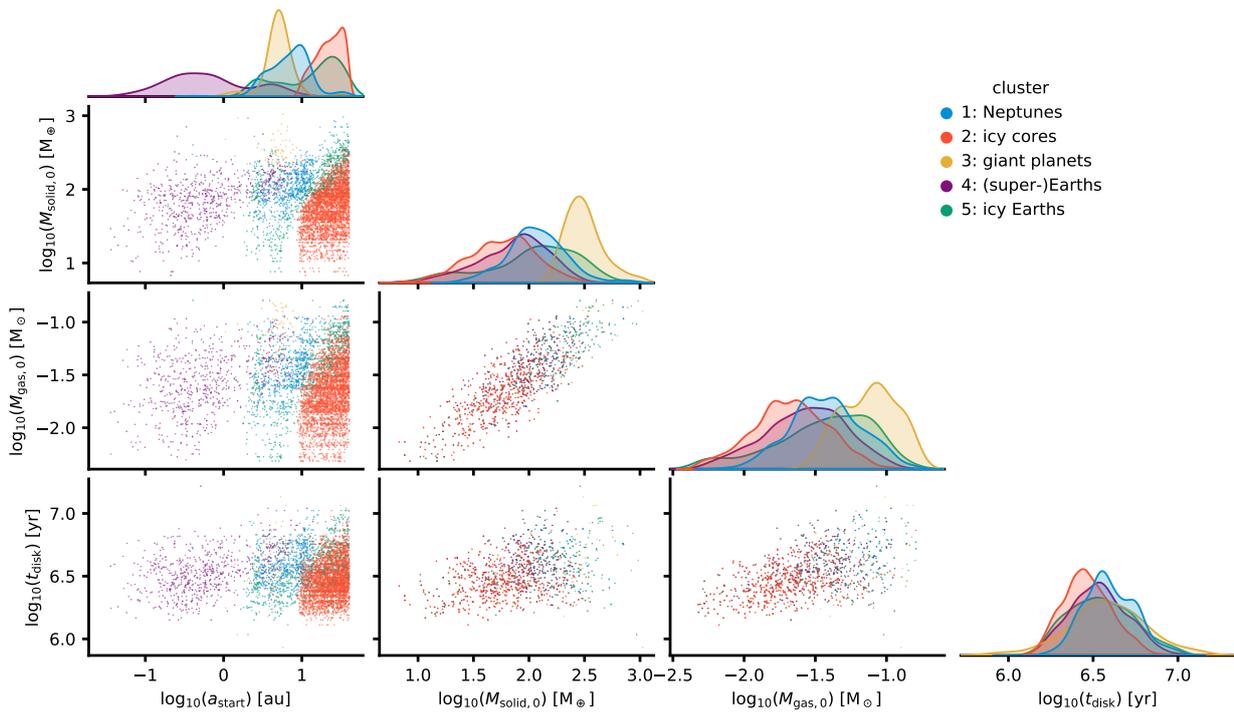


Figure 3.11: Pairwise relationships between all disk parameters, sorted by cluster affiliation. Same as Fig. 3.7, but for a multi-planet population with N -body interactions. The separation of clusters is less pronounced than in the single-planet case.

initial orbital separation. Compared to the single-planet case, the separation of the clusters is less clean. The additional cluster identified in NG76, “icy Earths”, share a lot of parameter space with other planet types.

Using the mean decision boundary defined above (Eqn 3.2), the dependence of different planet clusters on specific initial conditions can be visualized also for the multi-planet population (Fig. 3.12). The relationships largely copy those of the single-planet case: the starting location of the planet embryo shows the largest decision boundary amplitudes and differences among the clusters, and giant planets retain their distinct dependence on high solid and gas reservoirs.

3.5 Discussion

3.5.1 What determines the type of a planet?

By predicting a planet’s cluster from a set of initial conditions of our planet formation model, we were able to establish links between properties of the protoplanetary disk and the corresponding planets (see Sect. 3.4.3). These links can be elucidated by using the

3 Predetermination of Planet Types in Global Core Accretion Models

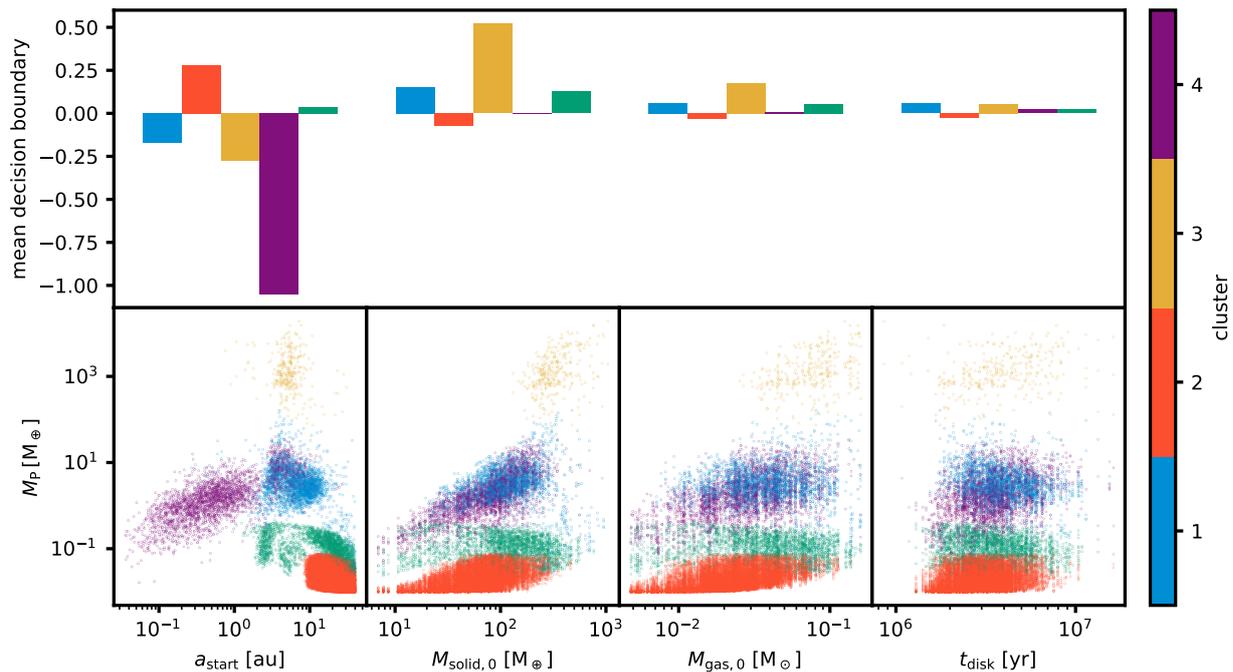


Figure 3.12: Relation between disk features and planet species. Same as Fig. 3.8, but for a multi-planet population with N -body interactions. As in the single-planet case, the starting location of the planet embryo a_{start} shows the largest variance in decision boundary. Giant planets (yellow) form only at high $M_{\text{solid},0}$ and sufficient $M_{\text{gas},0}$.

planet mass M_P as a proxy for the planet cluster and relating it to disk features (see Fig. 3.8). The feature with by far the highest predictive power is the starting location of the emerging protoplanetary embryo a_{start} , which is expected in a core accretion scenario: an embryo at small orbital distance has only a small feeding zone from which it can accrete and thus it will remain small. At very large orbital distance, the dynamical and growth timescales are very large and the disk will have disappeared before a protoplanet can gain significant mass (Lissauer, 1987, 1993; Kokubo & Ida, 2002; Mordasini et al., 2009a). Exactly at what orbital separations efficient planet growth is possible further depends on the amount, size, mass, and aerodynamic properties of planetesimals available there, and thus on the solid disk mass $M_{\text{solid},0}$ (see below for a more detailed discussion on the interplay between orbital distance and local planetesimal density). As can be seen in the lower left panel of Fig. 3.8, intermediate orbits provide the best conditions for rapid growth. These trends are responsible for the clear separation of planet clusters in the $a_{\text{start}}-M_P$ plane. Very small or very large initial orbital separations always lead to “failed cores” (low-mass instances of clusters 2 and 4). Short-period terrestrial planets and super-Earths (cluster 4) start on small orbits less than 1 au. (sub-)Neptunes (cluster 1) require intermediate orbits

of roughly 0.5 au – 10 au. Finally, giant planets (cluster 3) start on distant orbits ($\gtrsim 3$ au).

Other initial parameters show rather diverse importances that depend on the planet type. The mean decision boundaries (Eqn 3.2) of $M_{\text{solid},0}$ and $M_{\text{gas},0}$ are close to zero for all clusters except giant planets, implying a small feature importance of these parameters for most planet types. While these two parameters are correlated in our model, which could in principle spuriously decrease their MDI , their relation to M_{P} (lower panels of Fig. 3.8) reveals indeed only a weak relation to planet type. The picture differs for giant planets, which only form in disks that are rich both in gas ($M_{\text{gas},0} \gtrsim 0.04 M_{\odot}$) and solids ($M_{\text{solid},0} \gtrsim 200 M_{\oplus}$). Given a specific starting location of its core, the efficiency of giant planet formation is strongly governed by $M_{\text{solid},0}$. The reason is this parameter's direct relation to the local planetesimal density in the disk and thus a protoplanet's ability to reach a core mass sufficient for runaway gas accretion. Lastly, the disk lifetime stipulates the time within which planet formation has to conclude. Surprisingly, this parameter shows close to no correlation with the resulting planet type. This shows that most disks provide material long enough (median ≈ 3.4 Myr) to complete planet formation. Within the scope of our model, early disk dispersal is not the preferred pathway to halt planet formation at low and intermediate masses.

We conclude that the occurrence of a certain type of planet is fundamentally related to disk properties, and it depends in particular on the orbital distance where the planetary embryo forms. Currently, we treat this important parameter as a Monte Carlo variable that is distributed based on simple theoretical arguments (Kokubo & Ida, 2000). This is a major shortcoming of our formation model and our findings highlight the importance of a consistent treatment of planetary embryo formation (Voelkel et al., 2021a,b). Another effect we neglected thus far are the gravitational interactions between planets. We address this aspect below by discussing simulations done with the same model but multiple forming planets per disk (see Sect. 3.5.4). Future studies should also take into account the effects of pebble accretion (Ormel & Klahr, 2010; Lambrechts & Johansen, 2012), which influence the efficiency of solid accretion and may lead to a global redistribution of solid material in protoplanetary disks (e.g., Lambrechts & Johansen, 2014; Morbidelli et al., 2015; Ormel et al., 2017; Bitsch et al., 2019a).

3.5.2 Disk mass and embryo distance as predictors for planet type

Now that we have identified the solid disk mass and the initial orbital separation of a planetary embryo as the most important features, we investigate the regions different planet types occupy in the space that these parameters span. Figure 3.7 shows distinct borders between the different clusters that can be explained by the combination of processes our

planet formation model covers. The diagonal border between cluster 1 planets, which correspond to icy and atmosphere-bearing “Neptunes” on close and intermediate orbits, and cluster 4 planets, which are dry (super-)Earths, is shaped by photoevaporation of planetary envelopes: we recall that the clustering algorithm made the separation between these clusters mainly in R_P , which leads to a completely atmosphere-less (super-)Earth cluster and a cluster of Neptunes that predominantly bear H/He envelopes. However, close to all (super-)Earths initially held an envelope that they subsequently lost due to photoevaporation, a fate that the more massive Neptunes were spared. Thus, the more solid material is available at a specific orbital distance, the more likely planets will grow massive enough to retain their atmospheres in the long term. The efficiency of photoevaporation is further a function of orbital distance, leading to the negative slope of the border between clusters 1 and 4 in $a_{\text{start}} - M_{\text{solid},0}$ (Jin & Mordasini, 2018). Cluster 2 (“icy cores”) contains only terrestrial planets and failed cores with high amounts of volatile species and no atmospheres. They formed on distant orbits where the growth timescale is large, preventing them from growing beyond terrestrial size within the lifetime of the protoplanetary disk (Kokubo & Ida, 2000).

3.5.3 Oligarchic growth of giant planets

The giant planets (cluster 3) in our planet population occupy a distinct region at large starting positions and high solid disk masses (see Fig. 3.7). It abruptly cuts off around 4 au, which corresponds to typical water ice line positions at accretion time (Burn et al., 2019). Here, the solid surface density jumps by a factor of four (Mordasini et al., 2012a), and significantly higher total solid disk masses are required to reach runaway gas accretion interior of this orbit. We therefore only considered planets beyond 4 au when we characterized the shape of the giant planet cluster. We did so by determining the hyperplanes in $a_{\text{start}} - M_{\text{solid},0}$ space that best separate these planets from other species. A Support Vector Machine (SVM, Cortes & Vapnik, 1995) maximizes the distance of this plane to planets that belong to the “giant planets” cluster and all those that do not. We used the implementation in scikit-learn (Pedregosa et al., 2011) with a linear kernel and otherwise default hyperparameters, and trained the SVM on the full population. As in logarithmic representation the giant planet cluster has a triangular shape, we can approximate its border by a broken power law. Setting $y = \log_{10}(M_{\text{solid}})$ and $x = \log_{10}(a_{\text{start}})$, we fitted the piecewise linear function

$$y = \begin{cases} k_1 x + y_0 - k_1 x_0 & x \leq x_0 \\ k_2 x + y_0 - k_2 x_0 & x > x_0 \end{cases} \quad (3.3)$$

| x_0 | y_0 | k_1 | k_2 |
|------------------------|------------------------|-------------------------|------------------------|
| $1.04^{+0.01}_{-0.01}$ | $2.22^{+0.01}_{-0.01}$ | $-0.42^{+0.04}_{-0.05}$ | $1.20^{+0.03}_{-0.04}$ |

Table 3.2: Best-fit parameters for the broken power law in Equation 3.3. Uncertainties are 16th and 84th percentiles obtained via bootstrap sampling.

to separation functions found by the SVM. The best-fit values for these parameters are listed in Tab. 3.2. We calculated their uncertainties by the bootstrapping method: we repeatedly drew N random planets with replacement, where N is the total number of planets in our synthetic planet population, and trained the SVM on each of 1000 samples generated this way. In Fig. 3.13, we overlay the so found giant planet boundary onto the planets in $a_{\text{start}} - M_{\text{solid},0}$ space. Generally, giant planets form when $\log_{10} \left(\frac{M_{\text{solid}}}{1 M_{\oplus}} \right) \gtrsim 2.7 - 0.4 \log_{10} \left(\frac{a_{\text{start}}}{1 \text{ au}} \right)$ for cores emerging within ~ 10 au and when $\log_{10} \left(\frac{M_{\text{solid}}}{1 M_{\oplus}} \right) \gtrsim 1.0 + 1.2 \log_{10} \left(\frac{a_{\text{start}}}{1 \text{ au}} \right)$ for cores emerging beyond. We point out that this result is only valid in the context of the assumptions of our model. Plausible limitations that might have influenced this outcome are the assumptions of a single population of planetesimals of the same size and efficient embryo formation throughout the disk, the non-consideration of pebble accretion (Ormel & Klahr, 2010), and the largely featureless numerical disk that does not allow for “planet traps” (Chambers, 2009). Another probable source of error is the omission of gravitational interactions between planets in the same system – the giant planet domain shifts moderately and is more diffuse when multiple concurrently forming planets are assumed (see Sect. 3.5.4). Nevertheless, we focus here on typical outcomes of isolated protoplanets since it allows a more quantitative assessment.

We also compared this boundary to characteristic parameters for planetesimal accretion in the oligarchic growth regime: the planetesimal isolation mass M_{iso} and the growth timescale τ_{grow} (e.g., Kokubo & Ida, 2000; Raymond et al., 2014). On intermediate orbits of a few au, planetary growth is limited by the amount of material that can be accreted. M_{iso} is a useful concept to quantify the maximum attainable core mass given this limit. On the other hand, τ_{grow} gives an estimate for the time needed to reach a certain core mass, and sets the limit for wider orbits. For comparison with the giant planet cluster, we computed the local planetesimal densities corresponding to specific values of M_{iso} and τ_{grow} and translated them into total planetesimal disk masses $M_{\text{solid},0}$. See Sect. 2.2 for derivations of these quantities.

Since our model includes planet migration, planets can accrete solid material beyond their planetesimal isolation mass by moving through the disk. Nevertheless, M_{iso} is a proxy for how much can be accreted at a specific orbital distance and it is instructive to

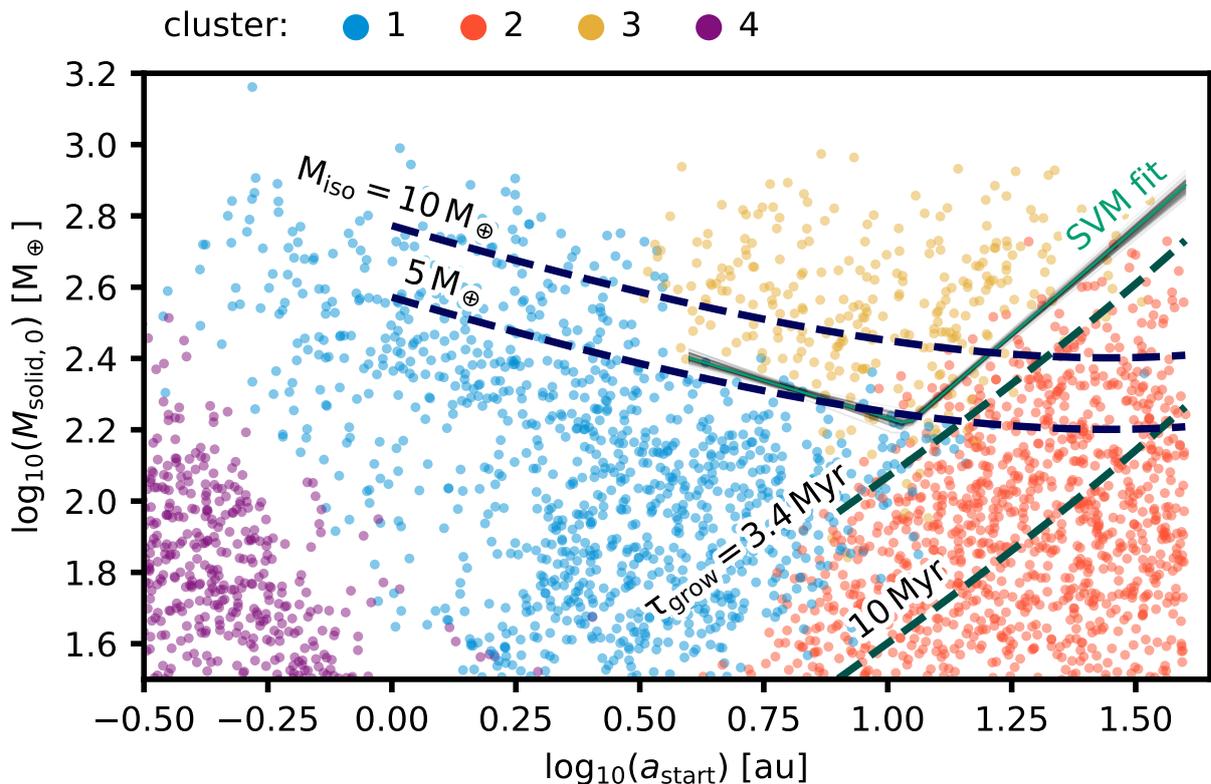


Figure 3.13: The four clusters of planets in $a_{\text{start}} - M_{\text{solid},0}$ space of their nascent protoplanetary disk. The green line is the hyperplane that best separates the giant planet cluster (yellow markers) from the other planets and was obtained by training a Support Vector Machine (SVM). Closeby gray lines show random draws from bootstrap sampling and illustrate the uncertainties. We overplot isolines of planetesimal masses needed to reach specific core masses (blue dashed lines), as well as isolines corresponding to specific growth timescales for reaching a core mass of $10 M_{\oplus}$ (green dashed lines). Their slopes are similar to the SVM fit that encloses the giant planets, indicating that the onset of runaway growth is limited by the locally available planetesimal mass and by the disk lifetime.

compare the shape of the giant planet population in $a_{\text{start}} - M_{\text{solid},0}$ space with the borders between planet clusters. In Fig. 3.13, we overplot isolines of disk solid masses necessary to reach different planetesimal isolation masses as a function of orbital separation (dashed blue lines). The lower border of the giant planet cluster matches well the slope of these lines. This indicates that in intermediate-mass disks with a few hundreds of M_{\oplus} in solids, giant planet formation is limited by the protoplanets reaching M_{iso} , that is, by clearing their feeding zone from solid material. We caution that the proximity of this border to the $M_{\text{iso}} = 5 M_{\oplus}$ isoline does not imply that runaway gas accretion has set in at this mass, as planet migration results in a larger effective feeding zone (Alibert et al., 2005).

Beyond ~ 10 au, the border of the cluster matches the slope of isolines for different growth timescales. At these larger orbital distances, τ_{grow} can reach the order of Myr for low planetesimal surface densities and thus becomes comparable to the lifetime of the protoplanetary disk. In this regime, the growth of a planetary core is limited by the time available to accrete the planetesimals in the domain of a planet's orbit. As can be seen in the plot, the $M_{\text{solid},0}(a)$ isoline where the growth timescale corresponds to the median of the disk lifetime, $\tau_{\text{grow}} \approx 3.4$ Myr, is a good fit to the border between giant planets (yellow) and icy cores (red). Indeed, most of the giant planets close to this threshold formed in long-lived disks (see Fig. A.3). This indicates that for planetesimal densities just sufficient for the formation of massive cores, entering runaway gas accretion depends on the longevity of the host disk.

3.5.4 The influence of N -body interactions

Our cluster analysis and prediction from initial conditions has shown that even in the case of multi-planet systems with gravitational interactions, most of the links between disk and planet properties remain intact (see Sect. 3.4.4). Still, the demographic structures in the multi-planet population are somewhat smeared out compared to the single-planet case, and the strength of this effect is different for individual clusters. We have seen that (super-)Earths and Neptunes are affected the most by this sort of mixing. These planet types cannot be reliably predicted from disk properties if N -body interactions are taken into account. Interestingly, the confusion is asymmetric: planets predicted as Neptunes often become (super-)Earths, while those predicted as (super-)Earths rarely become Neptunes. The reason is something the classifier cannot predict: the misclassified (super-)Earths are typically planets that got stripped of their atmospheres in giant collisions with other planets. From this follows that our model would produce too many Neptunes if such collisions are not taken into account (as is the case in single-planet simulations). This highlights the need for global planet formation models to include a consistent treatment of N -body interactions and giant impacts, as has already been suggested by [Alibert et al. \(2013\)](#) and in [Emsenhuber et al. \(2020a\)](#).

Another difference compared to the single-planet case is that close-in planets with small radii and masses are strongly depleted. This is because they often undergo giant collisions and merge into more massive bodies. The resulting lack of “sub-Earths” provides an interesting prediction for future planet searches that will push beyond the current mass/radius limits. Whether a multitude or a desert of such planets will be found could give valuable clues to the prevalence of planetary collisions.

3.6 Conclusions

We have investigated how different properties of protoplanetary disks relate to the emergence of different planet types in a planetesimal-based core accretion context. By performing a cluster analysis on synthetic planet populations from a global model of planet formation and evolution, we identified clusters of planets in a parameter space of typical exoplanet observables. We examined how well these clusters can be predicted from disk properties and studied the dependencies of different planet types. Our main conclusions are:

1. Planets form distinct groups in $\{a, M_P, R_P\}$ space, especially when dynamical interactions within multi-planet systems are neglected. Without presupposing planet types or their number, we identified four clusters corresponding to (sub-)Neptunes, icy cores, giant planets, and (super-)Earths.
2. These groups differentiate within the first 0.1 Myr of the formation process and show correlations with properties of their host disks. Such associations between disk and planet properties enable the prediction of planet species to high accuracy (98 % in the single-planet case and 89 % in the multi-planet case).
3. The most important predictor for planet clusters is the orbital position of the emerging planetary core, followed by the solid mass available in the disk. The disk lifetime plays a subordinate role, but can be a limiting factor for threshold values of the above mentioned parameters.
4. The position of giant planets in disk parameter space can be associated with known characteristics of oligarchic planetesimal accretion: for limited available amounts of solid material and within ~ 10 au, core growth is limited by planetesimal isolation and giant planets form when $\log_{10} \left(\frac{M_{\text{solid}}}{1 M_{\oplus}} \right) \gtrsim 2.7 - 0.4 \log_{10} \left(\frac{a_{\text{start}}}{1 \text{ au}} \right)$. On more distant orbits, core accretion is limited by the growth timescale and giants emerge when $\log_{10} \left(\frac{M_{\text{solid}}}{1 M_{\oplus}} \right) \gtrsim 1.0 + 1.2 \log_{10} \left(\frac{a_{\text{start}}}{1 \text{ au}} \right)$.
5. When multiple planets form and interact in the same system, for most planet types the associations between disk properties and planet properties remain. However, planets on track to become sub-Neptunes often lose their atmospheres in giant collisions and turn into super-Earths, which impedes predictions for this planet type.

Overall, we have shown that synthetic planet populations from state-of-the-art core accretion models largely mirror the planet types recognized by exoplanet demographics. Our

results highlight the importance of N -body integrations in global planet formation models that aim for reliable predictions in the domain of low-mass planets. Beyond that, constraining the orbital distances at which planetary cores form is of major relevance for the full range of planet types. Population syntheses of the next generation should recognize this by including self-consistent treatments of planetary embryo formation.

4

Warm Super-Earths and Cold Jupiters

A Weak Occurrence Correlation, but With a Strong
Architecture-Composition Link

The content of this chapter is based on the publication [Schlecker et al. \(2020b, in press\)](#) in *Astronomy & Astrophysics*.

Details of authorship: I am the corresponding author of the study and prepared the complete manuscript, including all figures, taking into account input by all co-authors. Thomas Henning and Christoph Mordasini conceived the project and provided conceptual ideas. Yann Alibert, Willy Benz, Christoph Mordasini, Alexandre Emsenhuber, and Remo Burn contributed to implementing the planet formation model that produced the synthetic planet population investigated in the paper. I performed part of the numerical simulations using supercomputing facilities available at the Max Planck Institute for Astronomy and prepared the data for analysis. The subsequent statistical analysis presented in the paper was performed by me. Christoph Mordasini, Thomas Henning, Remo Burn, and Hubert Klahr aided in interpreting the results.

4.1 Motivation

4.1.1 Super-Earths, cold Jupiters, and theoretical predictions

While, in the past, planet formation theories have focused on the Solar System (e.g., Pollock et al., 1996), this focus has since shifted toward the goal of finding explanations for a whole variety of planets and planetary systems. Important sources of constraints for these theories are the occurrence rate (or frequency) of exoplanets as a function of various orbital or physical properties as well as the fraction of stars hosting such planets (e.g., Petigura et al., 2013b; Foreman-Mackey et al., 2014; Hsu et al., 2018; Mulders et al., 2018). In recent years, the growing sample of confirmed exoplanet systems have made such occurrence studies possible, enabling us to statistically compare theory and observations. While the first detected exoplanet around a main-sequence star was a giant planet on a close orbit (Mayor & Queloz, 1995), it has now been established that “cold Jupiters” (CJ) in distant orbits are much more frequent but not as readily detected (Wittenmyer et al., 2020). Aside from spotting these types of giant planets, recent technological and methodological advances have also enabled the discovery of small, terrestrial planets, although our detection sensitivity is still limited to those on close orbits. This development led to the discovery of an unexpected population of planets that are not present in the Solar System: planets with masses higher than that of Earth but substantially below those of our local ice giants, that is, so-called super-Earths (SE, e.g., Mayor et al., 2011). It has been estimated that they orbit 30% – 50% of FGK stars, often in multiplanet systems (Fressin et al., 2013; Petigura et al., 2013b; Zhu et al., 2018; Mulders et al., 2018).

Since cold Jupiters influence their environment due to their large masses, it seems likely that they have an effect on such close-in low-mass planets (e.g., Raymond et al., 2006; Horner & Jones, 2010; Raymond & Izidoro, 2017). The open question concerns exactly how they affect the formation and subsequent evolution of inner planets and if their existence in a system facilitates the formation of super-Earths or excludes it, rather. If hot super-Earths form in situ, there should be a positive correlation between outer giant planets and inner terrestrial systems: whenever favorable conditions enable efficient growth of planetesimals in a protoplanetary disk, both planet types can emerge (Chiang & Laughlin, 2013).

However, in situ formation has been criticized as it is not able to account for the variety of architectures observed in these systems (Raymond & Cossou, 2014), thus, most current core accretion models assume orbital migration as a key ingredient (e.g., Alibert et al., 2005; Emsenhuber et al., 2020a). In these models, planetary cores originate from or-

bits that diverge from their final location through a process that typically involves inward migration. This mechanism predicts an anti-correlation between inner super-Earths and cold giants: due to the strong dependence of accretion timescales on the orbital radius, the innermost core is expected to grow most efficiently (e.g., [Lambrechts & Johansen, 2014](#)), enabling a subsequent runaway accretion of a massive gas envelope ([Pollack et al., 1996](#)). The emerging giant planet now prevents cores that form further out from migrating inward to become hot super-Earths ([Izidoro et al., 2015](#)). On the other hand, planetary cores resulting from giant collisions can reach runaway accretion earlier, which facilitates an early growth of distant giant planets ([Klahr & Bodenheimer, 2006](#)). Models describing the growth of inner planets via pebble accretion ([Ormel & Klahr, 2010](#); [Lambrechts & Johansen, 2012](#)), which relies on a radial flux of mm to cm-sized pebbles to the inner system ([Lambrechts & Johansen, 2014](#)), predict an additional impact from massive outer planets. When they carve a gap into the disk deep enough to generate a local pressure maximum, the inward drift of pebbles is halted just outside of the planetary orbit ([Morbidelli & Nesvorný, 2012](#); [Lambrechts et al., 2014](#)), possibly inhibiting the formation of inner terrestrial planets ([Ormel et al., 2017](#); [Owen & Murray-Clay, 2018](#)). Depending on the timing of this cut-off of pebble flux, a negative effect on the occurrence of inner super-Earths can arise. If this scenario occurs frequently, the existence of both planet types in the same system should be rare and their occurrences anti-correlated.

4.1.2 Observational findings

Despite the theoretical predictions above, a number of recent observational studies that tested the relations between super-Earths (SE) and cold Jupiters (CJ) found, instead, a *positive* correlation. [Zhu & Wu \(2018\)](#) measured the frequency of cold Jupiter-hosting systems in a sample of 31 systems harboring super-Earths that were first discovered by the radial velocity (RV) method. This frequency corresponds to the conditional probability of a system harboring a cold Jupiter, given that there is at least one super-Earth in the system, $P(\text{CJ}|\text{SE})$. They established $P(\text{CJ}|\text{SE}) = 0.29$, which is a strong enhancement compared to the fraction of field stars containing a cold Jupiter $P(\text{CJ}) \sim 0.10$. The opposite case, that is, the conditional probability of any super-Earth given a cold Jupiter in the system $P(\text{SE}|\text{CJ})$ could only be derived indirectly but was found to be even higher with the anti-correlation case definitively excluded. [Herman et al. \(2019\)](#) strengthen this claim by counting five systems of transiting close-in planets in their sample of twelve long-period transiting planets. This trend was confirmed by [Bryan et al. \(2019\)](#) based on a search for long-period giant companions in 65 super-Earth systems, where half of them were originally discovered by the transit method and the other half by the RV method. Applying

4 Warm Super-Earths and Cold Jupiters

different boundaries for mass and period than [Zhu & Wu \(2018\)](#), they find $P(\text{CJ}|\text{SE}) = (34 \pm 7)\%$ and come to the conclusion that close to all cold Jupiter-hosting systems harbor at least one super-Earth.

In light of these independent suggestions of a strong positive correlation, it is surprising that a recent RV survey that searched for super-Earth companions in a sample of 20 Solar-type stars hosting long-period giants detected none ([Barbato et al., 2018](#)). While their sample of 20 systems is small, their null detection is very unlikely if the correlation is indeed as high as reported.

We note that their definition of super-Earths differs from the one in [Zhu & Wu \(2018\)](#) and the survey is not complete in the respective mass-period range. Their super-Earths have $M \sin i$ between 10 and $30 M_{\oplus}$ and reside on orbits with periods less than 150 d. For such planets, they “conservatively” assume a detection sensitivity $P_{\text{detect}} = 0.5$. To mirror the survey in a numerical experiment, we repeatedly drew a pseudo-random number $x \in [0.0, 1.0)$ and counted a “detection” if:

$$x < P_{\text{detect}} \cdot P(\text{SE}|\text{CJ}), \quad (4.1)$$

where P_{detect} is the probability to detect an existing super-Earth system and $P(\text{SE}|\text{CJ})$ is the fraction of systems hosting super-Earths in cold Jupiter-hosting systems. Each round of 20 iterations represents a realization of the survey with corresponding N detections. We repeated this experiment 10 000 times to obtain a probability for each N .

Figure 4.1 shows the detection probabilities of such a survey for $P_{\text{detect}} = 0.5$ and four different occurrence probabilities. For a very low value $P(\text{SE}|\text{CJ}) = 0.1$, the probability to find zero super-Earths is as high as $P(N = 0) = 0.35$. If, on the other hand, the probability is $P(\text{SE}|\text{CJ}) = 0.3$, this value drops to $P(N = 0) = 0.04$. For probabilities of $P(\text{SE}|\text{CJ}) = 0.5$ or higher, $P(N = 0)$ approaches zero. It is extremely unlikely to find zero planets in 20 systems if $P(\text{SE}|\text{CJ}) \gtrsim 0.5$ and $P_{\text{detect}} = 0.5$.

Regardless of the different super-Earth definitions in [Zhu & Wu \(2018\)](#) and [Barbato et al. \(2018\)](#), the latter do not detect any sub-giant planets in their sample systems. If we adopt the numbers reported by [Zhu & Wu \(2018\)](#) for such planets, their conditional super-Earth probability is $P(\text{SE}|\text{CJ}) = 0.9$, and the average sensitivity of the survey must be as low as 0.15 to obtain a probability of 5% for their null result. It is thus difficult to reconcile the results of [Zhu & Wu \(2018\)](#), [Bryan et al. \(2016\)](#), and [Herman et al. \(2019\)](#) with the one presented in [Barbato et al. \(2018\)](#). Recently, [Barbato et al. \(2020\)](#) stated an impaired sensitivity for their survey, which could explain the non-detection.

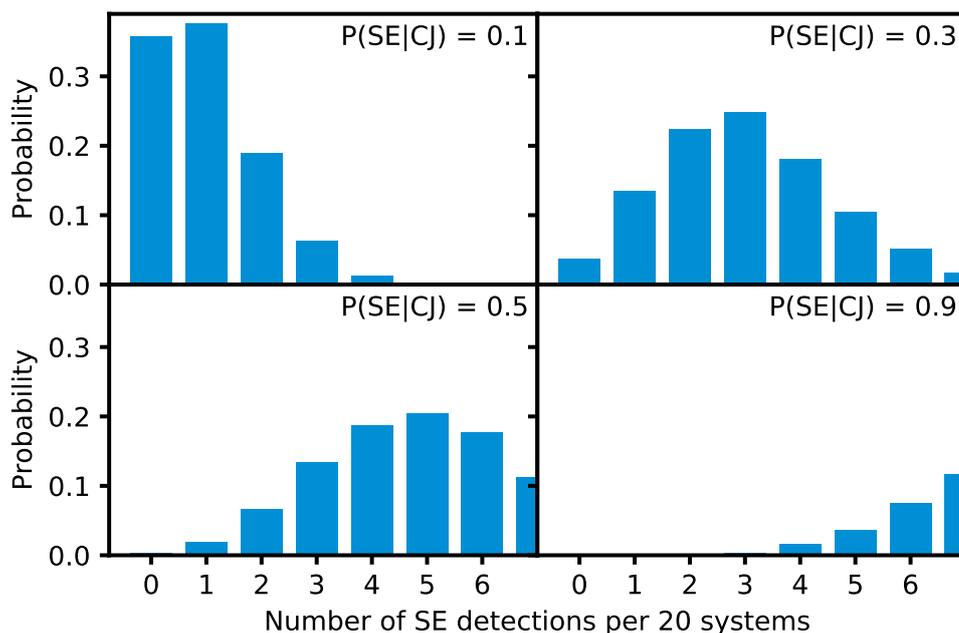


Figure 4.1: Probability of finding different numbers of super-Earths in a survey of 20 target stars. Each panel assumes a different conditional super-Earth probability ranging from $P(\text{SE}|\text{CJ}) = 0.1$ to 0.9 and we adopt a survey sensitivity for this planet type of $P_{\text{detect}} = 0.5$. The probability to find zero planets approaches zero for $P(\text{SE}|\text{CJ})$ greater than ~ 0.5 .

4.1.3 An imperative to confront theory and observation

Given this range of different results and implications for planet formation theory, great potential lies in the search for similar correlations in synthetic populations of planets produced by theoretical models. The purpose of this study is a detailed characterization of the relations between inner super-Earths and cold Jupiters based on the core accretion theory of planet formation. To that end, we use synthetic planetary systems that were obtained with the Generation III *Bern* Model of planet formation and evolution (Emsenhuber et al., 2020a) to investigate the mutual influence of these planet types in and to test the observed trends. In our simulations, we consider planetary systems around solar-type stars. Emsenhuber et al. (2020b), used this model to perform a population synthesis of multi-planet systems from initial conditions representative of protoplanetary disks in star forming regions. Here, we extend that work by applying a generic detection bias and statistically compare the synthetic quantities to measured exoplanet observables.

The chapter is structured as follows: in Sect. 4.2, I introduce our formation model with its initial conditions and describe how I prepared our synthetic data. I present the synthetic population produced with this model, called *NG76*, in Sect. 4.3. Section 4.4 pursues

the comparison of our population with the observed exoplanet sample. In Sect. 4.5, I interpret our findings and discuss their implications. Finally, I conclude this chapter by summarizing our results and predictions in Sect. 4.6.

4.2 Methods

To investigate the relations between inner rocky planets and cold gas giants, we performed a statistical comparison between a synthetic planet population and a sample of observed exoplanets. We focus on planetary systems around solar-type stars and fixed the stellar mass to $1 M_{\odot}$ throughout. The synthetic planetary systems were calculated with the Generation III *Bern* global model of planetary formation and evolution ((Emsenhuber et al., 2020a)), which is described in Sect. 2.5.1. In this section, we explain the definitions we used to classify planets and to compute their occurrence, introduce the observed sample and its biases, and demonstrate the statistical methods we applied.

4.2.1 Synthetic planet sample

The Monte Carlo run of our formation model yielded a population of synthetic planets that live in independent systems. We carried out some preparatory steps before performing the statistical analysis and comparisons with the observed sample: first, we neglected all protoplanets from further analysis that were either accreted onto the star, ejected out of the system, or did not grow beyond a total mass of $0.5 M_{\oplus}$. Next, we computed the orbital period of each remaining planet from its current semi-major axis and assuming a Solar-mass host star using Kepler's Third Law. We then categorized the sample according to the mass and period ranges in Table 4.1 into distinct planet classes, where we considered super-Earths and cold Jupiters using the nominal definitions of Zhu & Wu (2018). Since we are interested in the probability that a given system forms a particular planet species, we counted unique systems instead of planets to compute occurrence probabilities.

There is no general consensus about the limits in radius, mass, or composition that distinguish between different planet classes. To facilitate comparison with observational studies, we defined planet types according to the mass and period limits in Zhu & Wu (2018) and list them in Table 4.1.

We accounted for biases due to orbit inclinations by multiplying synthetic planet masses with an artificial $\sin(i)$ term, where i is the relative inclination between the orbital plane of the innermost planet and the line of sight to an observer. It is reasonable to assume isotropic orientations of orbital planes, we therefore followed Mordasini (2008) and drew

Table 4.1: Planet Classifications

| Classification | Planet Mass [M_{\oplus}] | Orbital Period [d] |
|----------------|--|--------------------|
| super-Earth | $2 M_{\oplus} \leq M_P \sin(i) \leq 20 M_{\oplus}$ | $P < 400$ d |
| cold Jupiter | $M_P \sin(i) \geq 95 M_{\oplus}$ | $P > 400$ d |

$\sin(i)$ from the distribution,

$$f(\sin(i)) = \frac{\sin(i)}{\sqrt{1 - \sin(i)^2}}. \quad (4.2)$$

Our synthetic population consists of 1000 planetary systems. After a simulation time of 5×10^9 yr, a total of 32 030 planets on bound orbits have survived in these systems. Using the selection criteria in Table 4.1, we arrive at a sample of 538 super-Earths in 291 systems and 182 cold Jupiters in 140 systems.

4.2.2 Occurrence rates and fraction of planet hosts

It is crucial to distinguish between the planet occurrence rate, which constitutes a number of planets per star, and the fraction of stars hosting planets. We consider the occurrence rate as a measure for the frequency of planets per domain in the physical parameter space and define it as

$$\eta = \frac{100}{N_{\star}} n_p(\mathbf{x}), \quad (4.3)$$

where N_{\star} is the number of systems in the population and $n_p(\mathbf{x})$ is the number of planets with properties \mathbf{x} that lie in a chosen interval $d\mathbf{x}$ of the parameter space. For the purposes of this chapter, this space is spanned by combinations of orbital period, planet size, planet mass, disk solid mass, and host star metallicity. We normalize η to planets per 100 systems, for convenience.

We further construct the fraction of stars hosting a planet, $P(X)$. Here, X corresponds to a specific planet species that is defined by a parameter interval $d\mathbf{x}$. $P(X)$ is readily obtained by dividing the number of systems containing at least one planet of type X , N_X , by the total number of systems, that is,

$$P(X) = \frac{N_X}{N_{\star}}. \quad (4.4)$$

The probability to form, for instance, a super-Earth system, $P(\text{SE})$, is then the number of unique systems containing at least one super-Earth divided by the number of systems in the population. We note that $P(X)$ is the probability that a planetary system contains at least one instance of planet species, X , regardless of the multiplicity within a given system.

4 Warm Super-Earths and Cold Jupiters

Analogously, we computed probabilities involving non-formations, $P(\bar{X})$, by counting the systems that are lacking a planet of type X.

Conditional probabilities that quantify the fraction of systems with a planet type given that another type is present (or missing) in the system help reveal the effects of simultaneous formation of these planets in the same system. We obtained such conditional probabilities for all possible combinations of planet types. As an example, to compute the probability of having a cold Jupiter in a system hosting at least one super-Earth, $P(\text{CJ}|\text{SE})$, we divided the number of super-Earth systems containing a cold Jupiter by the number of super-Earth systems. We proceeded equally with conditional probabilities of non-formations.

The uncertainties of synthetic probabilities follow a Poisson statistic since the problem is equivalent to counting measurements without errors in a binned statistic. The requirement of independence of the individual measurements is justified since we count systems and not single planets (which could influence each other within the same system). We computed uncertainties of the conditional probabilities using Gaussian error propagation.

4.2.3 Observed planet sample

As this study investigates relations between super-Earths and cold Jupiters, we compared our synthetic population with observational samples that include these planet types. [Zhu & Wu \(2018\)](#) computed a variety of planet host fractions for these species and reported a positive correlation regarding their formation. Where not stated otherwise, we refer to their numbers when using observed quantities. A wide range of values has been reported for the fraction of stars hosting inner super-Earths $P(\text{SE})$, involving different mass/period limits and detection techniques (e.g., [Howard et al., 2010, 2012](#); [Fressin et al., 2013](#); [Petigura et al., 2013b](#); [Zhu et al., 2018](#); [Mulders et al., 2018](#)). For consistency with the super-Earth definitions in [Zhu & Wu \(2018\)](#), we adopted $P(\text{SE}) = 0.30$ from [Zhu et al. \(2018\)](#).

Where quantities were missing in the literature, we obtained them using standard rules of probability theory: the observed fraction of systems that formed no super-Earth and no cold Jupiter, $P(\overline{\text{SE}} \cap \overline{\text{CJ}})$, and the fraction of systems that formed super-Earths but no cold Jupiter, $P(\text{SE} \cap \overline{\text{CJ}})$, were computed by applying the summation rule for probabilities. Using the reported probability for super-Earth systems, $P(\text{SE})$, we obtain:

$$P(\overline{\text{SE}} \cap \overline{\text{CJ}}) = 1 - [P(\text{SE}) + P(\text{CJ}) - P(\text{SE} \cap \text{CJ})] \text{ and} \quad (4.5)$$

$$P(\text{SE} \cap \overline{\text{CJ}}) = P(\text{SE}) - P(\text{SE} \cap \text{CJ}). \quad (4.6)$$

Taking into account that the planetary systems in our observed sample are hosted by main sequence stars and are presumably dynamically stable on Gyr timescales, we ana-

lyzed a snapshot of the synthetic planet population at 5 Gyr. At this age, the protoplanetary disk has long been dispersed and the following evolutionary phase, in which thermodynamic evolution shapes the characteristics of a planet’s envelope, has largely concluded (Mordasini et al., 2012b). It is therefore reasonable to assume that the error we introduced by assuming the same age for all stars in the sample is typically smaller than the observational uncertainties.

4.2.4 Detection limit

Accounting for detection limits in the observed sample, we employed a simple detection limit based on a minimum RV semi-amplitude

$$K = \left(\frac{2\pi G}{P}\right)^{1/3} \frac{M_P \sin i}{(M_P + M_*)^{2/3}} \frac{1}{\sqrt{1 - e^2}}, \quad (4.7)$$

where P denotes the orbital period and e is the eccentricity (Cumming et al., 1999). Zhu & Wu (2018) indicate that they removed all planet candidates with $K < 1 \text{ m s}^{-1}$. However, their sample seems to have a sharp truncation at $K \sim 2 \text{ m s}^{-1}$ (compare their Fig. 1) which is difficult to explain by an intrinsic feature of the population. We suspect that this drop is due to a stronger detection bias than assumed and adopted a more conservative minimum K of 2 m s^{-1} for our synthetic sample to enable a more plausible comparison.

4.3 Results: Synthetic population

4.3.1 System classes

We classified the synthetic planetary systems into four classes: systems with neither super-Earths nor cold Jupiters ($\overline{\text{SE}} \cap \overline{\text{CJ}}$), systems with at least one super-Earth but no cold Jupiters ($\text{SE} \cap \overline{\text{CJ}}$), systems with at least one cold Jupiter but no super-Earths ($\overline{\text{SE}} \cap \text{CJ}$), and systems containing both planet types ($\text{SE} \cap \text{CJ}$). Figures 4.2, 4.3, 4.4, and 4.5 show the time evolution of randomly sampled systems from each of these classes. For each system, we show all planets more massive than $0.5 M_\oplus$, regardless of their detectability. Horizontal bars denote the orbital range of eccentric planets. From left to right, the columns correspond to the systems’ states at simulation times 0.3 Myr, 1 Myr, 3 Myr, the time of disk dispersal t_{disk} , the integration time of the N -body code $t = 20 \text{ Myr}$, and 5 Gyr. At the final time, the dashed line marks an RV detection limit of $K = 2 \text{ m s}^{-1}$ and planets below this threshold are grayed out.

Overall, the systems show diverse architectures even within the same class of systems. The classes where one or both planet types are excluded often contain planets that would

4 Warm Super-Earths and Cold Jupiters

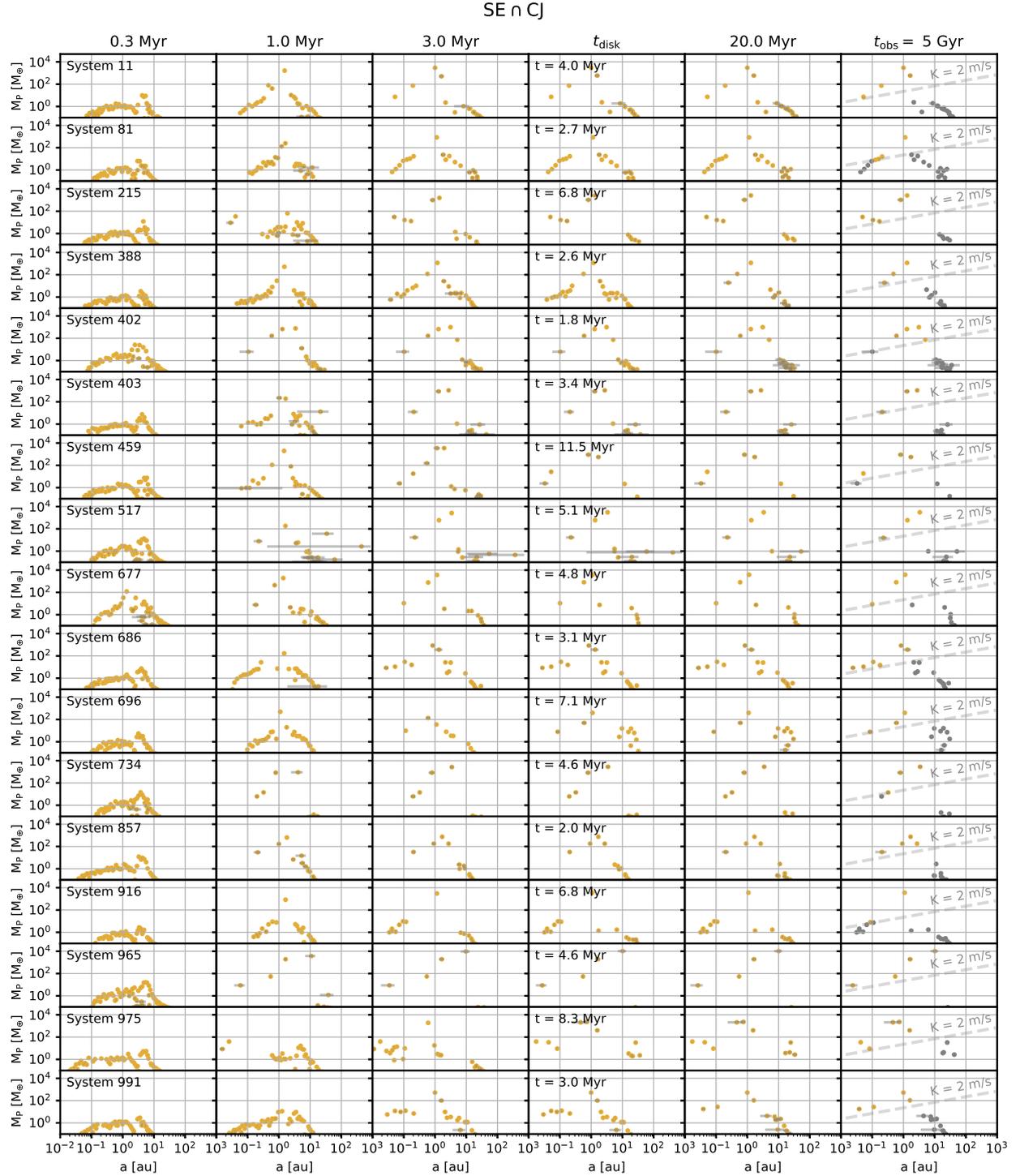


Figure 4.2: Time evolution of systems with cold Jupiters and super-Earths. For a number of randomly sampled systems, we show the mass-semi-major axis relation of the planets at six different times, where t_{disk} is the disk dispersal time and 20 Myr is the integration time of the N -body module. In the last column, unobservable planets are grayed out and the dashed line indicates the detection limit of 2 m/s^{-1} . Horizontal gray lines visualize the orbital range of eccentric planets.

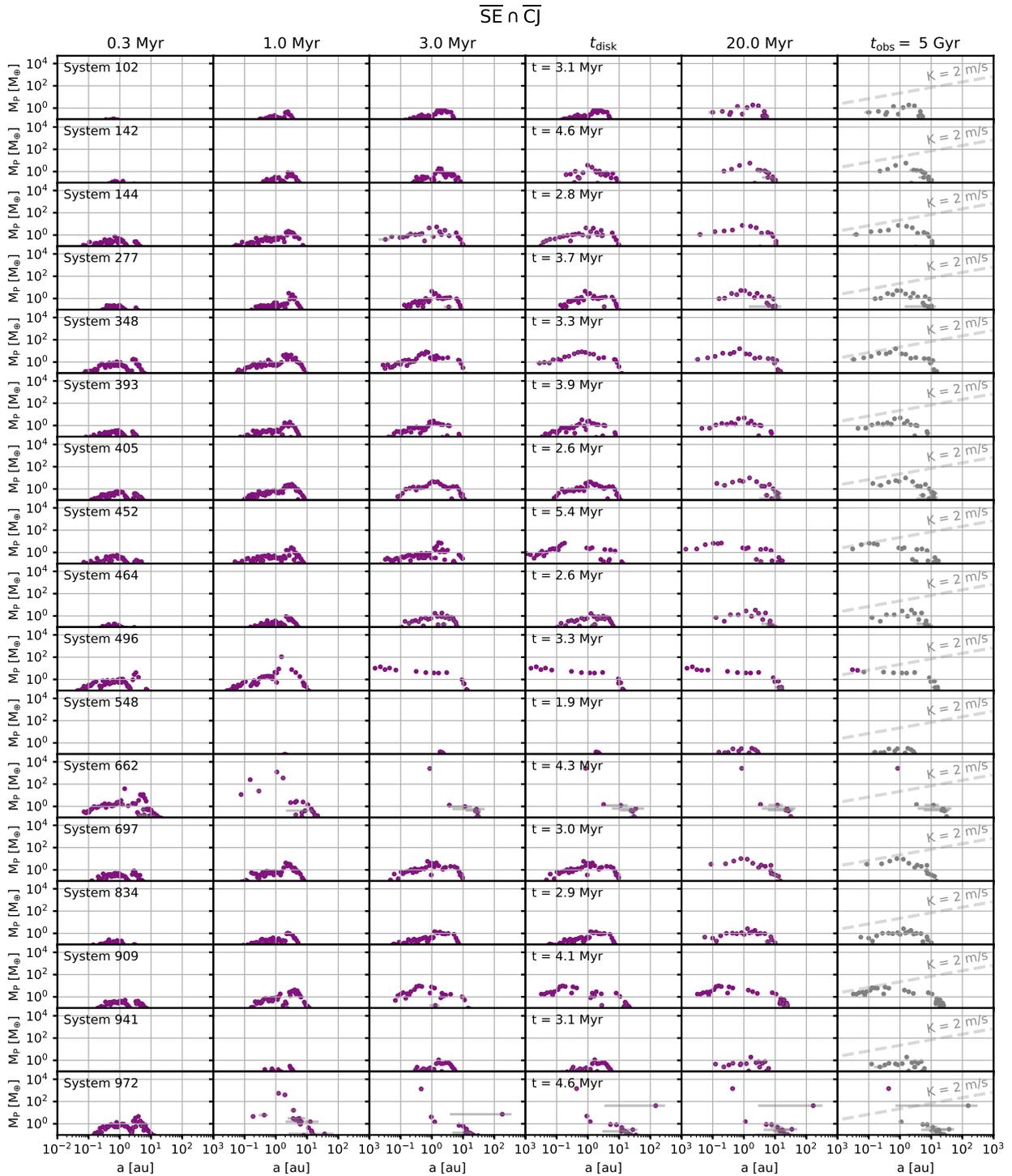


Figure 4.3: Same as Fig. 4.2, but for systems containing neither super-Earths nor cold Jupiters.

4 Warm Super-Earths and Cold Jupiters

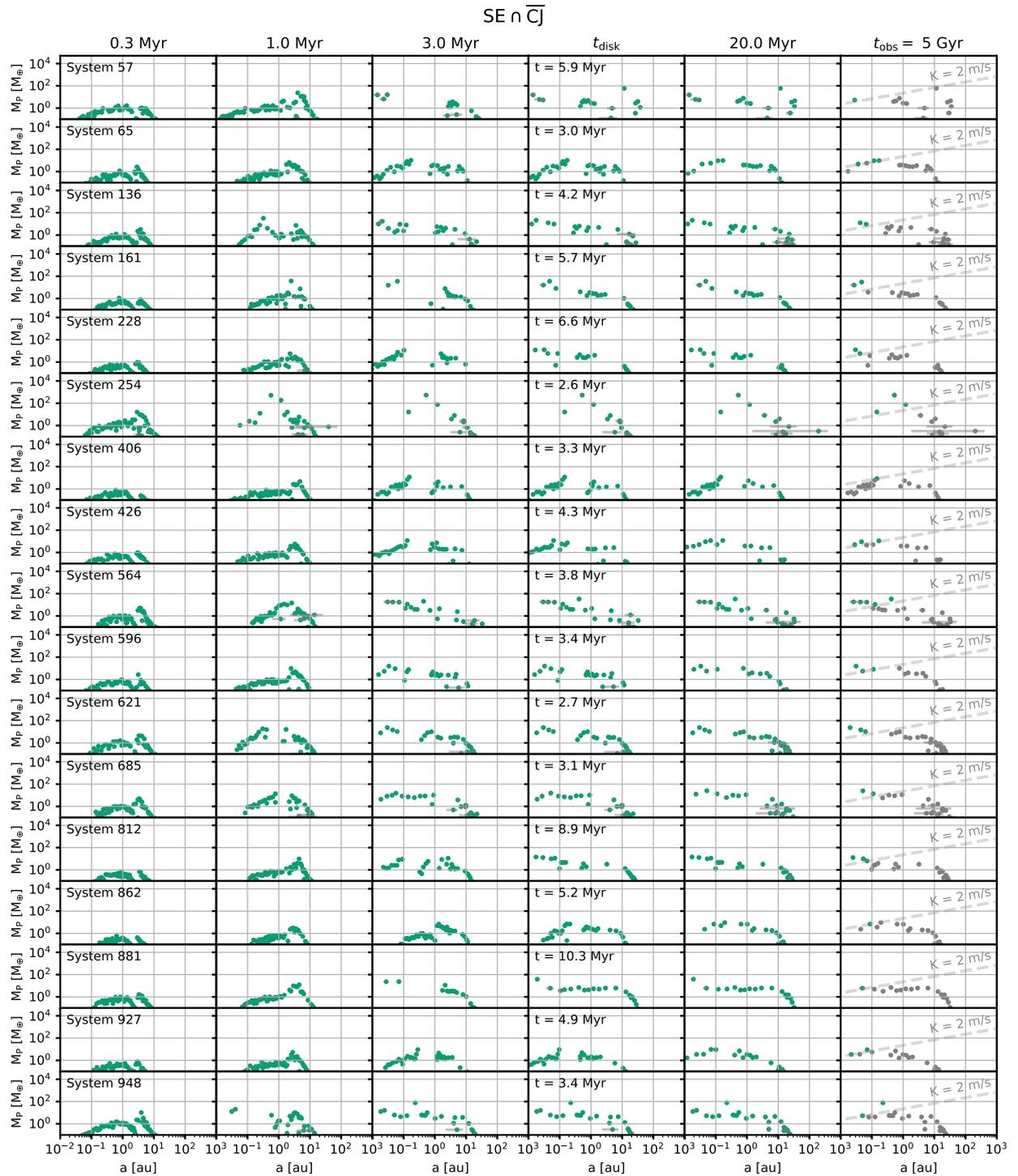


Figure 4.4: Same as Fig. 4.2, but for systems containing super-Earths and no cold Jupiters.

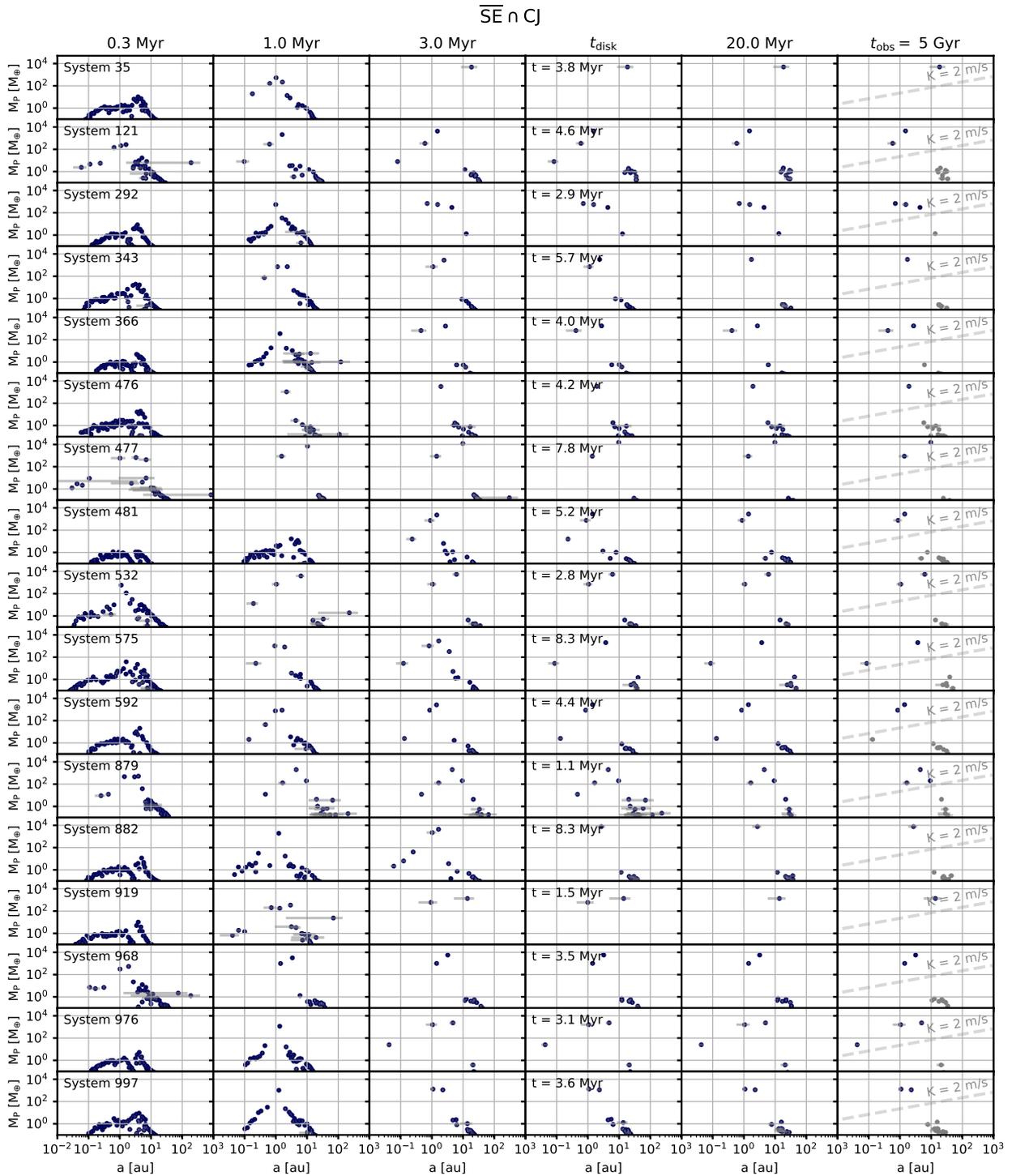


Figure 4.5: Same as Fig. 4.2, but for systems containing cold Jupiters and no super-Earths.

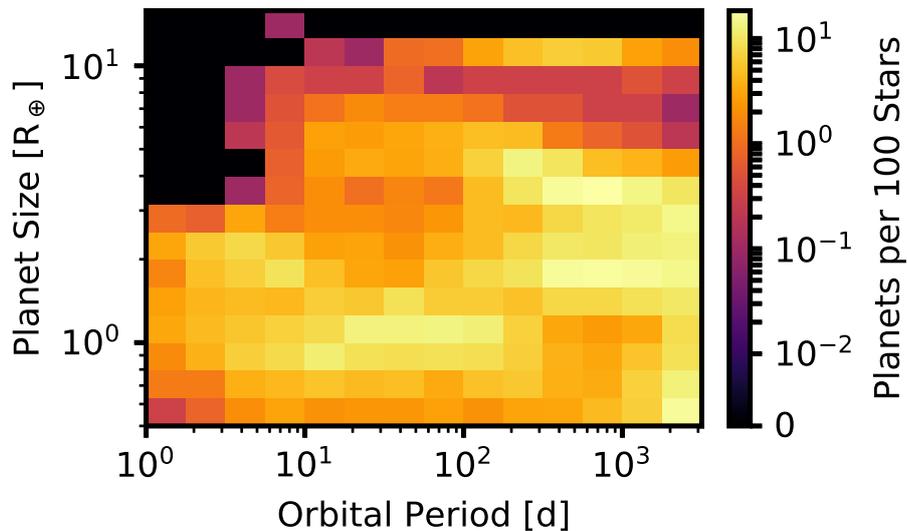


Figure 4.6: Occurrence map of the synthetic population. Planet occurrences are normalized to planets per 100 stars per period-radius bin, where each bin corresponds to 0.25 dex in period and 0.1 dex in planet radius, respectively. Planets with $R_P < 0.5 R_\oplus$ and beyond $P = 3000$ d are not shown. Most planets are of terrestrial size and reside on intermediate orbits. A distinct group of giant planets breaks away from the remaining population at $\sim 10 R_\oplus$.

nominally fulfil the criteria of that planet type. These planets are either not detectable according to our chosen detection limit or were assigned an unfavorable inclination and were thus not classified as a super-Earth or cold Jupiter.

4.3.2 Occurrence rates in period-radius

Figure 4.6 shows the occurrence rate, $\eta(P, R_P)$, in planet radius and orbital period for the full synthetic planet population at an age of 5 Gyr. Each bin covers 0.25 dex in period and 0.1 dex in planet radius, and their counts are normalized to planets per 100 stars. We do not show planets beyond $P = 3000$ d, as our N -body integration time of 20 Myr is too short to account for their large growth timescales. We also exclude objects smaller than $0.5 R_\oplus$, which are not observable with state-of-the-art exoplanet detection techniques (e.g., Dumusque et al., 2011; Cloutier et al., 2018; Reiners et al., 2018b; Bryson et al., 2020; Trifonov et al., 2020). The majority of planets are of terrestrial size and reside on intermediate or wide orbits. A sub-population of Jupiter-sized planets ($R_P \sim 11 R_\oplus$) is clearly differentiated from the contiguous remaining population and preferentially populates the period range of a few hundred to ~ 1000 d.

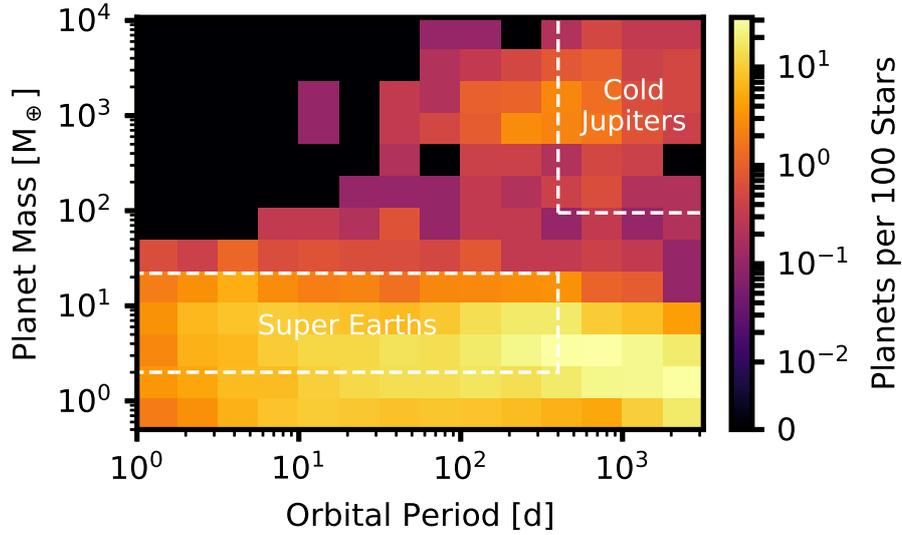


Figure 4.7: Occurrences of the synthetic population in the mass-period plane. Normalization and binning are the same as in Fig. 4.6, except that each planet mass bin corresponds to $1/3$ dex. White lines border the mass-period limits for super-Earths and cold Jupiters, respectively.

4.3.3 Occurrence rates in period-mass

Similarly to Fig. 4.6, Fig. 4.7 shows synthetic planet occurrence rates in the mass-period plane $\eta(P, M_P)$. The occurrence is normalized to number of planets per 100 stars per mass-period interval. Following the same argument as for very small radii, we refrain from considering planets less massive than $0.5 M_\oplus$. As in Fig. 4.6, we exclude planets on periods beyond 3000 d. Unsurprisingly, the distribution is similar to the radius-period diagram with high-mass planets more dispersed, although they still form a distinct population. There are no distinct populations of hot and cold Jupiters, only a small number of giants with $P \sim 10$ d separates from the main group of giant planets at intermediate to large orbital distances. The latter is only partly included in our definition of cold Jupiters owing to the comparability with the observed planet sample. Rocky planets of terrestrial to super-terrestrial mass occupy predominantly periods of hundreds to thousands of days; the bulk of planets populating outer regions falls outside our nominal super-Earth definition. We note that compared to previous population syntheses that lacked intermediate-mass inner planets (e.g., [Mordasini et al., 2009a](#)), our current model produces a significant number of super-Earths: 291 out of 1000 systems harbor a planet that obeys our criteria for a super-Earth (compare Table 4.1). This difference is mainly caused by our improved description of planet migration, which in particular treats the shift from type I to type II migration self-consistently. No artificial inhibition factors for type I migration are necessary to reproduce

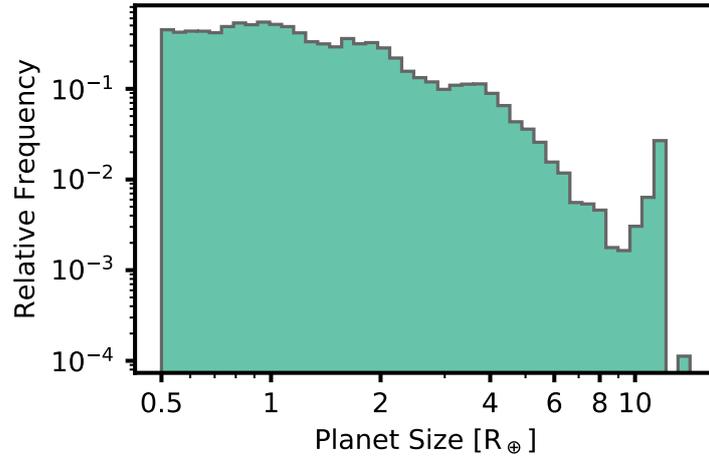


Figure 4.8: Distribution of planet radii in the synthetic population. We exclude planets smaller than $0.5 R_{\oplus}$ and with periods beyond 3000 d. The radius frequencies follow a distinct bimodal distribution with the bulk at its low-size end.

observed period distributions.

4.3.4 Planet radii

Figure 4.8 reveals a bimodal structure in the radius distribution of our synthetic population: most planets are terrestrial or super-Earth-sized, but an additional, shallower local maximum close to $1 R_{\text{Jup}}$ exists in the radius distribution. This bimodality separates giant planets that experienced runaway gas accretion from planets that did not and was seen already in earlier generations of population synthesis models (Mordasini et al., 2012b).

It is noticeable that even though planets with masses far beyond $1 M_{\text{Jup}}$ (see Fig. 4.7) occur in our population, the radius distribution shows a sharp cutoff at $\sim 12 R_{\text{Earth}}$. This feature also appears in the observed exoplanet population (e.g., Chen & Kipping, 2016) and is explained by electron degeneracy in the interior of giant planets (e.g., Chabrier et al., 2009). Close to Jupiter mass, the polytropic index $n \sim 1$ in the equation of state and the radius is independent of the mass. This leads to a wide range of planet masses populating a narrow region in planet radii. In the synthetic population shown here, this effect is enhanced since we assume the same atmospheric opacity for all planets and show all planets at the same age of 5 Gyr (Mordasini et al., 2012b).

A prominent feature in planetary radius-period diagrams is a depleted region separating small super-Earths from larger sub-Neptunes. This “photoevaporation valley” was predicted by formation and evolution models (Jin et al., 2014) and later confirmed observationally (Fulton et al., 2017; Hsu et al., 2018). While originally explained by pho-

to evaporation, alternative processes have also been hypothesized to produce the pattern. Debated mechanisms include core-powered mass loss, where the core's internal luminosity removes the planetary atmosphere (Ginzburg et al., 2016, 2018; Gupta & Schlichting, 2019); different formation pathways of planets above and below the gap (Zeng et al., 2019); and planetesimal impacts (e.g., Liu et al., 2015; Wyatt et al., 2019).

Our synthetic population reproduces the radius valley at most in an attenuated form. The reason for this lies presumably in our simplified treatment of collisional envelope stripping, where we add the impact luminosity of a collision event to the intrinsic planetary luminosity (Emsenhuber et al., 2020a). In contrast to photoevaporation from high-energy photons from the star, this mechanism not only affects the innermost region of the system but also the envelopes of planets further out. Also, more massive planets suffer from atmospheric loss than it is the case with photoevaporation alone. Both effects fill up the radius valley. When the luminosity from impacts is neglected in our model and atmospheres are stripped only by photoevaporation, a significant radius valley emerges (compare Jin & Mordasini (2018)). At least two possible shortcomings of our current prescription would be plausible to explain the observed mismatch with the empirical radius distribution: There might be less collision events than assumed, or they do not remove atmospheres as efficiently as modeled. Further studies will aim at distinguishing these possibilities as well as the contributions of different atmosphere-depletion mechanisms.

4.3.5 Relation between metallicity and planet radius

The frequency histogram in Fig. 4.9 illustrates the dependencies between host star metallicity and planet radius for different planet types. It includes all planets with radii between $0.5 R_{\oplus}$ and $12 R_{\oplus}$. The planet frequency in the metallicity-radius plane reveals a clear positive correlation of gas giant occurrence and stellar metallicity, in agreement with the well-established correlation in the observed exoplanet population (Santos et al., 2004; Fischer & Valenti, 2005; Johnson et al., 2010; Buchhave et al., 2018). Our occurrence density confirms the observed paucity of large planets with sub-Solar metallicity (Petigura et al., 2018). Small terrestrial planets populate a wide range of metallicities and their host stars are not enhanced in $[\text{Fe}/\text{H}]$.

Overplotted is a statistics binned in radius for different planet types; vertical bars denote the standard deviations in each bin. When all planets are considered, their average host star metallicities are consistent with the observed trends in Buchhave et al. (2014) and Narang et al. (2018).

While all types of giant planets are enhanced in metallicity, cold Jupiters between $8 R_{\oplus} - 10 R_{\oplus}$ are more metal-rich than their siblings in the same size range. This differ-

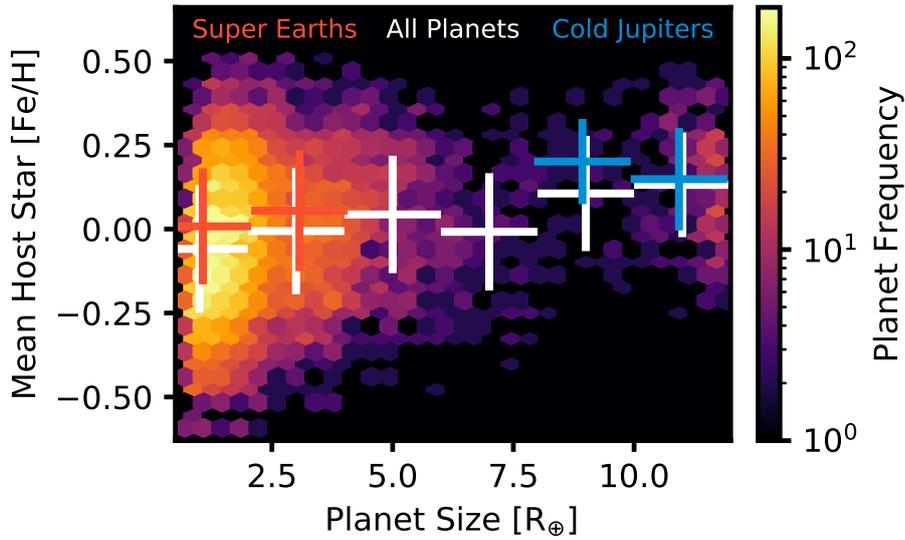


Figure 4.9: Dependence of host star metallicity on planet radius. For each radius bin, the cross denotes the mean $[\text{Fe}/\text{H}]$ and the vertical bar is its standard deviation. The markers are slightly offset horizontally for clarity. Overall, there is a positive metallicity trend with planet size. Hosts of super-Earths show marginally higher metallicity with respect to the overall sample. Cold Jupiters are enhanced in metallicity.

ence is not related to their orbital distance but due to our classification based on their mass (compare Table 4.1), which excludes large planets with $M_P < 95 M_\oplus$ from being classified as cold Jupiters. On average, our cold Jupiters have a higher bulk density and thus contain more solids compared to the entirety of planets in this radius range.

For super-Earths, we find only a weak positive trend with planet radius. Also, the metallicities of stars harboring these planets are not significantly enhanced compared to the full population. This is consistent with findings of [Sousa et al. \(2018\)](#) who report indistinguishable metallicity distributions of Solar neighborhood stars and stars hosting low-mass planets, respectively.

4.3.6 Relation to disk properties

In contrast to studies that focus on observed planet populations, for the synthesized population, we have the full history of each simulated system at hands. This includes the initial properties and evolution of the protoplanetary disk in which the synthetic planets formed (or not). Figure 4.10 reveals the distributions of these features for each of the system classes in Table 4.3 as well as for all planets with masses of $> 0.5 M_\oplus$ (“all”, gray lines). In each case, we show the parameter distributions of all planets that survived the entire formation and evolution phase, for example, the “SE” population contains all surviving

Table 4.2: Initial conditions of different populations

| population | [Fe/H] | $M_{\text{solid}} [M_{\oplus}]$ | $M_{\text{gas}} [M_{\odot}]$ | $R_{\text{cut,g}} [\text{au}]$ | $t_{\text{disk}} [\text{Myr}]$ | $a_{\text{start}} [\text{au}]$ |
|--|---|------------------------------------|--|----------------------------------|--|----------------------------------|
| CJ | 0.13 ^{+0.15} _{-0.17} | 290 ⁺¹²⁷ ₋₅₆ | 0.07 ^{+0.03} _{-0.03} | 95 ⁺²¹ ₋₂₄ | 3.35 ^{+2.26} _{-1.24} | 22 ⁺¹² ₋₁₇ |
| SE | 0.05 ^{+0.19} _{-0.19} | 156 ⁺⁷⁰ ₋₄₇ | 0.04 ^{+0.03} _{-0.02} | 69 ⁺²⁵ ₋₁₈ | 3.90 ^{+1.90} _{-1.28} | 16 ⁺¹⁴ ₋₁₁ |
| SE \cap CJ | 0.07 ^{+0.19} _{-0.12} | 283 ⁺⁸⁴ ₋₅₀ | 0.07 ^{+0.03} _{-0.03} | 95 ⁺²¹ ₋₂₄ | 3.35 ^{+1.96} _{-1.38} | 20 ⁺¹³ ₋₁₅ |
| $\overline{\text{SE}} \cap \overline{\text{CJ}}$ | -0.11 ^{+0.20} _{-0.19} | 54 ⁺⁴¹ ₋₂₉ | 0.02 ^{+0.02} _{-0.01} | 42 ⁺²² ₋₁₄ | 2.84 ^{+1.14} _{-0.81} | 11 ⁺¹⁶ ₋₉ |
| SE \cap $\overline{\text{CJ}}$ | 0.05 ^{+0.19} _{-0.20} | 145 ⁺⁵¹ ₋₃₉ | 0.04 ^{+0.03} _{-0.01} | 66 ⁺²³ ₋₁₆ | 4.07 ^{+1.74} _{-1.43} | 15 ⁺¹⁴ ₋₁₀ |
| $\overline{\text{SE}} \cap \text{CJ}$ | 0.17 ^{+0.12} _{-0.18} | 302 ⁺¹²⁹ ₋₆₆ | 0.07 ^{+0.04} _{-0.03} | 95 ⁺²⁶ ₋₂₅ | 3.41 ^{+2.29} _{-1.29} | 24 ⁺¹¹ ₋₁₈ |
| all | -0.03 ^{+0.20} _{-0.20} | 99 ⁺⁹⁰ ₋₅₀ | 0.03 ^{+0.03} _{-0.02} | 56 ⁺³⁰ ₋₁₉ | 3.40 ^{+1.85} _{-0.98} | 5 ⁺⁵ ₋₄ |
| initial | -0.03 ^{+0.22} _{-0.21} | 95 ⁺¹⁴⁷ ₋₅₅ | 0.03 ^{+0.04} _{-0.02} | 56 ⁺³⁶ ₋₂₁ | 3.23 ^{+1.90} _{-0.98} | 2 ⁺¹³ ₋₁ |

Notes. Initial Conditions for different populations. For each parameter, we quote its median for all combinations of SE and CJ, plus for the entire population of survived planets. Upper and lower limits denote 84th and 16th percentiles, respectively. Compare Fig. 4.10 for a visual representation of the data.

planets in super-Earths-hosting systems and not only their super-Earths. For comparison, the dotted lines (“initial”) denote the distributions for the complete set of simulations. We note that the gas disk radius $R_{\text{cut,g}}$ is not an independent parameter but a unique function of the gas disk mass M_{gas} . Table 4.2 contains the 16th, 50th, and 84th percentiles of these quantities for each system class.

In all physical disk parameters related to available planet material (metallicity, solid and gas mass, and disk size), the same three distinct populations are differentiated: systems without super-Earths or cold Jupiters, systems that formed intermediate-mass planets, such as super-Earths, and systems that formed cold Jupiters. This clustering is particularly illustrative in M_{solid} when retraced from low to high values:

- $\overline{\text{SE}} \cap \overline{\text{CJ}}$: at low solid masses of tens of M_{\oplus} , only low-mass planets occur that do not reach super-Earth mass or beyond.
- “All” class represents all survived planets and thus closely resembles the initial conditions.
- SE and SE \cap $\overline{\text{CJ}}$: in disks of intermediate supplies of solids, cores of several M_{\oplus} can form which result in super-Earths.
- SE \cap CJ, CJ, and $\overline{\text{SE}} \cap \text{CJ}$: from $M_{\text{solid}} \gtrsim 200 M_{\oplus}$, cold Jupiters can form. As shown below, these giant planets can pose a threat to inner super-Earth systems, which are frequently destroyed in $\overline{\text{SE}} \cap \text{CJ}$ systems.

4 Warm Super-Earths and Cold Jupiters

The starting position of the planetary embryo is a particularly decisive feature (compare [Schlecker et al. \(2021, Chapter 3\)](#)) and shows a separate pattern: while the overall population, which is dominated by terrestrial-mass planets, is shifted to small orbits, < 10 au, all other populations follow a more balanced distribution. Again, the cold Jupiter-hosting populations are differentiated. They show a bimodal distribution of initial orbits that divides them into planets we labeled as super-Earths or cold Jupiters and companions in the same systems that are undetectable.

The disk lifetime is rather insensitive to the outcome, but shows a similar clustering of system classes as the other parameters. While $\overline{\text{SE}} \cap \overline{\text{CJ}}$ systems, which consist largely of low-mass planets, have a median disk lifetime of 2.8 Myr, the disks of systems hosting super-Earths but no cold Jupiters last for 4.1 Myr on average. For a more detailed analysis of the links between disk properties and resulting planet types, see Chapter 3.

We now look at correlations between disk initial conditions and planet occurrence on the system level. To do so, we computed Spearman's rank correlation coefficient ρ ([Spearman, 1904](#)) for combinations of disk parameters (compare Table 2.1) and occurrences of a planet type. The coefficient ranges from -1 to $+1$, where a positive (negative) coefficient denotes a positive (negative) rank correlation between two variables and $\rho = 0$ corresponds to no correlation. All synthetic systems were included. The correlation map in Fig. 4.11 includes the initial gas disk mass, M_{gas} , the initial solid disk mass, M_{solid} , the host star metallicity, $[\text{Fe}/\text{H}]$, the exponential cutoff radius of the gas disk, $R_{\text{cut,g}}$, and the disk dispersal time, t_{disk} . It further incorporates occurrence rates for three different planet types: n_{SE} and n_{CJ} are the number of super-Earths and cold Jupiters per system, respectively. n_{tot} is the per-system frequency of all planets more massive than $0.5 M_{\oplus}$. All occurrence rates show positive correlations with M_{gas} , M_{solid} , and $R_{\text{cut,g}}$ (which is itself a function of M_{gas}). n_{CJ} and n_{tot} are also moderately correlated with metallicity, and n_{SE} and n_{tot} show some dependence on t_{disk} . For all occurrence rates, the strongest correlation is obtained with M_{solid} .

In order to identify trends of occurrence rates of the different planet types with this parameter, we compute a rolling mean along the solid mass axis and plot the corresponding mean planet occurrences of cold Jupiters and super-Earths, respectively (Fig. 4.12). The rolling window moves with step size one and consists of 80 systems; this window size is a trade-off between resolution and robustness against random variations. Shaded areas in the plot cover one standard deviation around the mean. For cold Jupiters, we obtain a monotonically increasing mean occurrence rate that starts around $150 M_{\oplus}$ and flattens out arriving at $\sim 1.0 \pm 0.6$ planets for $M_{\text{solid}} \gtrsim 300 M_{\oplus}$.

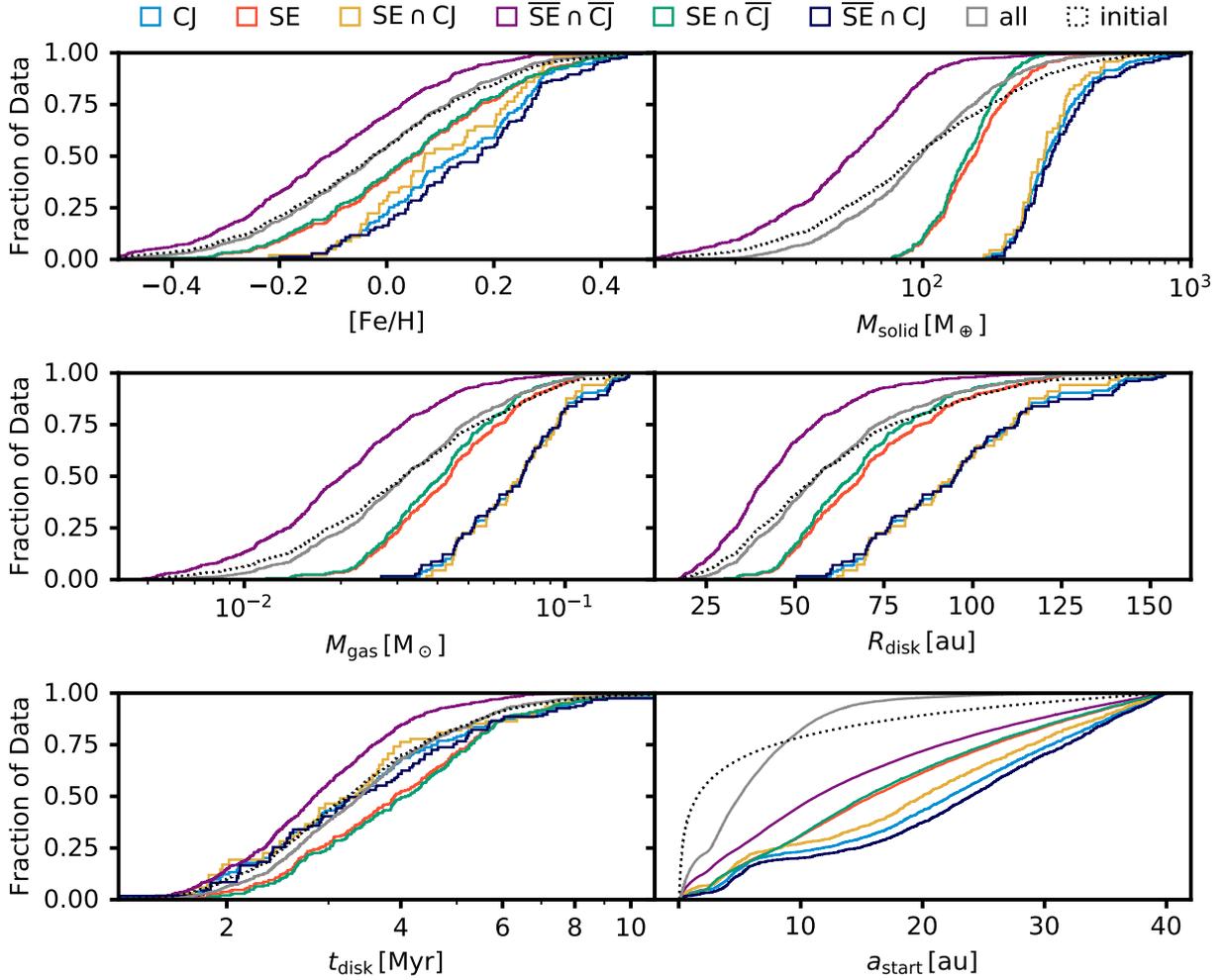


Figure 4.10: Initial conditions for different populations. For each parameter, we show empirical distribution functions for all combinations of SE and CJ, plus for the entire population of survived planets with $M_P > 0.5 M_\oplus$. The dotted lines show the initial distributions for the simulations. The CJ-hosting populations form compact clusters in most parameters, whereas the SE populations spread more depending on the existence of CJ in the systems.

The picture is very different for super-Earths. Their formation starts at lower disk masses with a steep increase in frequency up to a peak of 1.9 ± 1.2 planets per system at $160 M_\oplus$ in solids. At higher disk masses, the occurrence drops below unity and slightly descends beyond $300 M_\oplus$. Interestingly, these strong variations coincide with features in the cold Jupiter occurrence. This points to the destruction of inner rocky planetary systems by emerging outer giants in systems with very high initial solid disk mass M_{solid} . Generally speaking, systems harboring both super-Earths and cold Jupiters require disks with intermediate reservoirs of solids. If they are too small, no giant planets form. If they

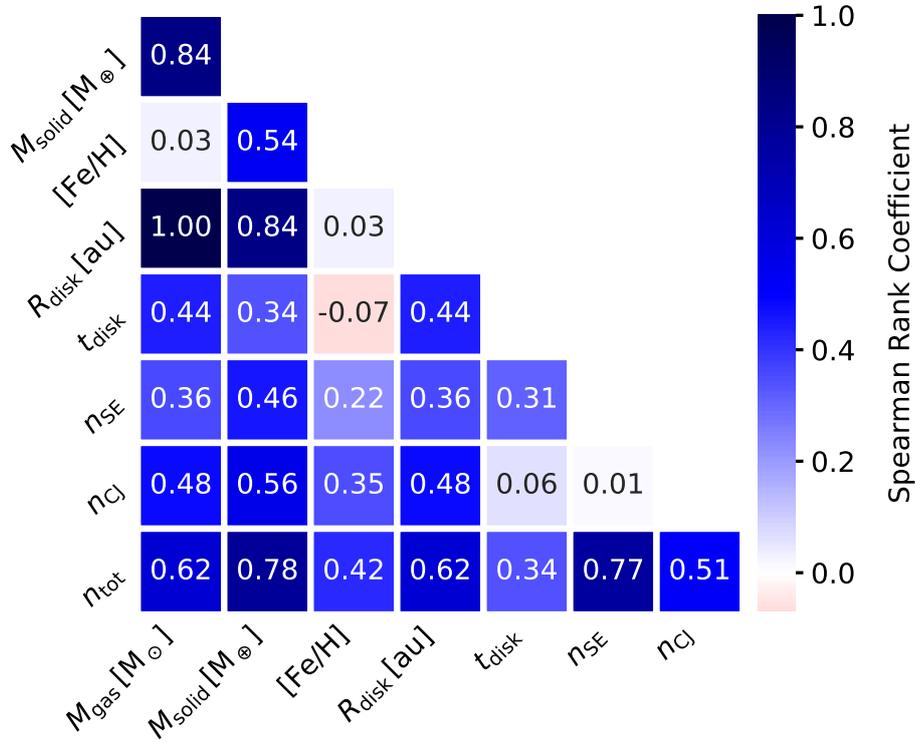


Figure 4.11: Correlation Map of disk properties and planet occurrence rates. For every combination of two quantities, we compute the Spearman rank coefficient to assess mutual correlations. The (identical) upper triangle and the self-correlating diagonal are removed for clarity. Almost all parameters show positive correlations and of all disk features, the solid material supply of the disk shows the highest correlation with all planet occurrences.

are too large, the super-Earths are destroyed and only giants remain.

4.4 Results: comparison with observations

In this section, we aim to compare the planet population *NG76* produced by our formation model with observed exoplanets. We focus on the populations of super-Earths and cold Jupiters and compare them to recent results based on data from ground-based radial velocity measurements and from the *Kepler* mission (Borucki et al., 2010). For a confrontation of observed and calculated planetary bulk densities in Sect. 4.4.3, we compiled our own sample of confirmed exoplanets .

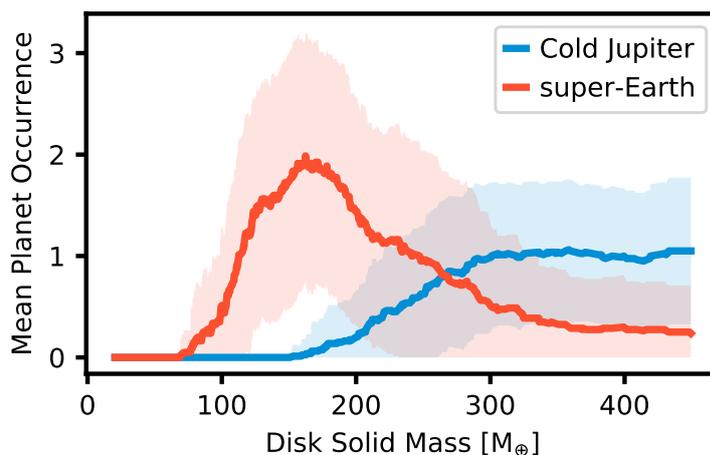


Figure 4.12: Mean planet occurrence per system as a function of initial solid disk mass. A rolling mean occurrence along the solid disk mass axis is shown for each planet type separately. Shaded areas cover ± 1 standard deviation of the rolling mean. Cold Jupiters form only in disks exceeding $M_{\text{solid}} \approx 150 M_{\oplus}$, where their occurrence shows a shallow positive correlation with solid mass. For super-Earths, there is a sharp increase to a maximum occurrence at $\sim 160 M_{\oplus}$, then it drops before flattening out at $\sim 250 M_{\oplus}$.

4.4.1 Fractions of planet hosts

To understand the relations between close-in super-Earths and outer gas giants and to constrain their mutual influence in the formation of planetary systems, we are interested in the fractions of planetary systems that form (and maintain) these planets.

In Fig. 4.13, we show unconditional and conditional probabilities for the existence of super-Earths and cold Jupiters. Table 4.3 lists their numerical values and compares them with their counterparts inferred from observations. These are based on confirmed planets that were initially detected with the RV method (Zhu & Wu, 2018). Where values are missing in their paper, we compute them using standard rules of probability theory (compare Sect. 4.2.2). To maintain sufficient orbital separation between the considered planet classes and for easier comparison, we do not include “Warm Jupiters” but adhere to the criteria in Table 4.1.

Zhu et al. (2018) report a super-Earth frequency of 0.30 based on detections of the *Kepler* survey. On the other hand, cold Jupiters are found in 11 % of systems around solar-type hosts (Cumming et al., 2008). Wittenmyer et al. (2016) derive a frequency of $6.2^{+2.8}_{-1.6}$ % but consider only giants with orbital periods $\sim 5 \text{ yr} - 19 \text{ yr}$. Herman et al. (2019) find an occurrence rate of 15^{+6}_{-5} % for large planets within a period range of $2 \text{ yr} - 10 \text{ yr}$ based on newly detected transit candidates from *Kepler*.

With $P(\text{SE})_{\text{syn}} = 0.29 \pm 0.02$ and $P(\text{CJ})_{\text{syn}} = 0.14 \pm 0.01$, our synthetic population is

4 Warm Super-Earths and Cold Jupiters

Table 4.3: Fractions of Stars hosting super-Earths and cold Jupiters.

| Probability | Observed ^d | Full Population ^e | [Fe/H] < -0.2 | -0.2 < [Fe/H] < 0.2 | 0.2 < [Fe/H] | $\bar{\mu}$ | \bar{e} |
|---|-----------------------|------------------------------|---------------|---------------------|--------------|-------------|-----------|
| P(SE) | 0.30 ^b | 0.29±0.02 | 0.13±0.02 | 0.31±0.02 | 0.43±0.05 | 1.8±0.9 | 0.08±0.10 |
| P(CJ) | 0.11 ^d | 0.14±0.01 | <0.01 | 0.12±0.01 | 0.42±0.05 | 1.3±0.5 | 0.18±0.19 |
| P(SE ∩ CJ) | 0.09 | 0.05±0.01 | <0.01 | 0.05±0.01 | 0.10±0.03 | 3.7±0.9 | 0.11±0.13 |
| P($\overline{\text{SE}} \cap \overline{\text{CJ}}$) | 0.69 ^b | 0.62±0.02 | 0.87±0.07 | 0.62±0.03 | 0.25±0.04 | 1.8±0.8 | 0.09±0.10 |
| P(SE ∩ $\overline{\text{CJ}}$) | 0.21 ^b | 0.24±0.02 | 0.12±0.02 | 0.26±0.02 | 0.32±0.05 | 2.2±1.0 | 0.06±0.10 |
| P($\overline{\text{SE}} \cap \text{CJ}$) | 0.01 ^c | 0.09±0.01 | <0.01 | 0.07±0.01 | 0.32±0.05 | 2.1±0.9 | 0.19±0.18 |
| P(SE CJ) | 0.90±0.20 | 0.34±0.06 | 1.0±1.41 | 0.41±0.09 | 0.25±0.07 | - | - |
| P(CJ SE) | 0.32±0.08 | 0.16±0.03 | 0.04±0.04 | 0.15±0.03 | 0.24±0.07 | - | - |
| P($\overline{\text{SE}}$ CJ) | 0.10±0.20 | 0.66±0.09 | <0.01 | 0.59±0.11 | 0.75±0.14 | - | - |
| P($\overline{\text{CJ}}$ SE) | 0.68±0.08 | 0.84±0.07 | 0.96±0.27 | 0.85±0.09 | 0.76±0.14 | - | - |

Notes. The top part shows the fraction of stars harboring (lacking) super-Earths (SE), cold Jupiters (CJ), and combinations thereof at an age of 5 Gyr. The bottom panel shows conditional probabilities $P(A|B)$ where A denotes the existence of at least one instance of planet type A in a given system and \overline{A} denotes its non-existence. Uncertainties of probabilities are based on Poisson errors. The last two columns list the mean planet multiplicity $\bar{\mu}$ and mean eccentricity \bar{e} with their standard deviations. While for P(SE) (P(CJ)), this takes into account only super-Earths (cold Jupiters); for the other classes we consider all planets with $K > 2 \text{ m s}^{-1}$.

(^a) if not stated otherwise, probabilities are from [Zhu & Wu \(2018\)](#) using their “nominal” super-Earth definition where $M_{\text{P}} \sin i < 20 M_{\oplus}$.

(^b) quoting for P(SE) the fraction of *Kepler* systems hosting super-Earths in [Zhu et al. \(2018\)](#)

(^c) order of magnitude estimate by [Zhu & Wu \(2018\)](#)

(^d) as estimated by [Cumming et al. \(2008\)](#)

(^e) population NG76 in the NGPPS series

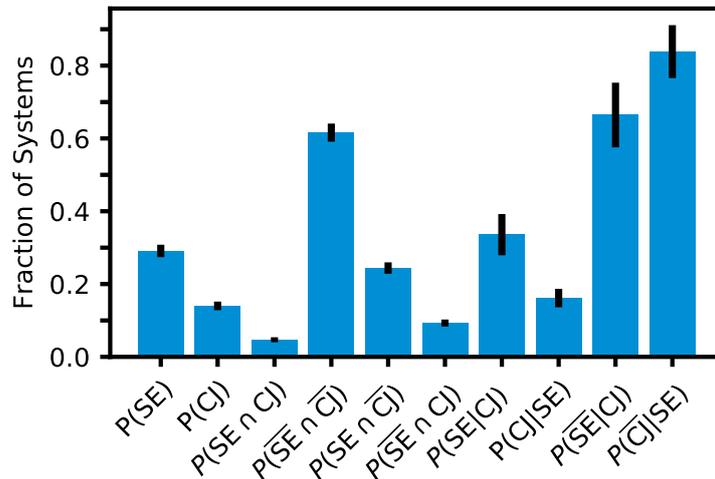


Figure 4.13: Fractions of planet hosts in the synthetic population of survived planets. The height of the bars represent the probabilities of column “Full Population” in Table 4.3. Black markers denote uncertainties assuming a Poisson statistic.

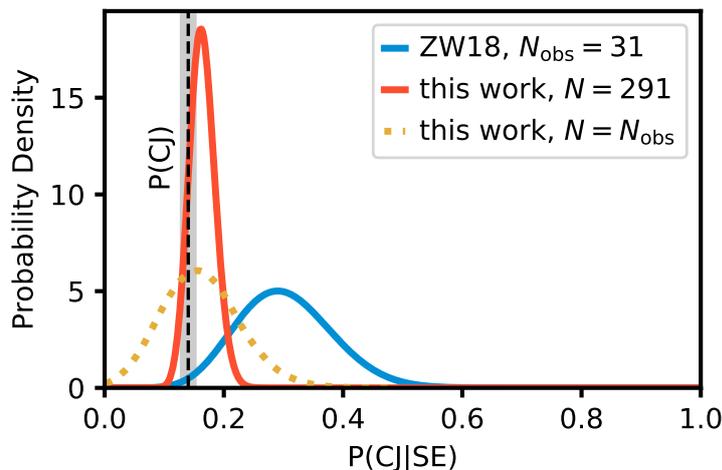


Figure 4.14: Observed and theoretical conditional probability $P(\text{CJ}|\text{SE})$. The blue curve approximates the posterior probability density to find 9 cold Jupiter systems in a sample of $N_{\text{obs}} = 31$ super-Earth systems (Zhu & Wu, 2018). Our theoretical population contains 291 systems with super-Earths, 47 of which contain cold Jupiters. The corresponding conditional probability (red curve) is enhanced compared to the overall cold Jupiter occurrence (black dashed line). We further generate a KDE of $P(\text{CJ}|\text{SE})$ from 1000 random draws of N_{obs} synthetic super-Earth systems (dotted line). While we find lower probabilities than Zhu & Wu (2018), an anti-correlation $P(\text{CJ}|\text{SE}) < P(\text{CJ})$ is unlikely.

consistent with these observables, although it contains slightly more cold Jupiters. Solar System analogs, that is, systems containing a cold Jupiter but lacking super-Earths, are rare ($P(\overline{\text{SE}} \cap \text{CJ})_{\text{syn}} = 0.09 \pm 0.01$). This quantity is difficult to constrain observationally, but Zhu & Wu (2018) give an order-of-magnitude estimate of 1%. More interesting for the relation between inner rocky planets and outer giants are the conditional probability of having at least one cold Jupiter in a super-Earth-hosting system, $P(\text{CJ}|\text{SE})$, and its inverse $P(\text{SE}|\text{CJ})$. Zhu & Wu (2018) found nine cold Jupiter-hosting systems in $N_{\text{obs}} = 31$ super-Earth systems¹ and thus report $P(\text{CJ}|\text{SE})_{\text{obs}} = 9/31 \approx 0.29$. The result is supported by Bryan et al. (2019), who find that $(39 \pm 7)\%$ of systems with inner super-Earths ($1 R_{\oplus} - 4 R_{\oplus}$, $1 M_{\oplus} - 10 M_{\oplus}$) host an outer gas giant. Both studies conclude that, compared to field stars, cold Jupiters are more prevalent around stars hosting super-Earths at short orbital distances. In our synthetic population, 47 out of 291 super-Earth systems contain at least one cold giant, which yields a rate of $P(\text{CJ}|\text{SE}) = 0.16 \pm 0.03$. In Fig. 4.14, we construct the binomial likelihood of this result. The distributions shown describe the probabilities to find k CJ systems when we randomly draw N SE systems. We compare the rate from our simulations (red curve, $k = 47$, $N = 291$) to the ones found in Zhu & Wu (2018)

¹for consistency, we consider only their sample that excludes warm Jupiters

4 Warm Super-Earths and Cold Jupiters

(blue curve, $k_{\text{obs}} = 9$, $N_{\text{obs}} = 31$). While the distributions have significant overlap, our result lies 1.7 standard deviations from the observed one. Despite these differences, the anti-correlation case $P(\text{CJ}|\text{SE}) < P(\text{CJ})$ remains unlikely.

We further estimated the variance in $P(\text{CJ}|\text{SE})$ we would expect if our sample size of super-Earth systems was the same as the observed one, $N = N_{\text{obs}}$. The dotted line in Fig. 4.14 shows the corresponding kernel density estimation (KDE) from 1000 random draws. Its standard deviation is 0.06 and the probability to find exactly nine cold Jupiter systems is 3 %.

The inverse conditional probability of finding a super-Earth in a cold Jupiter hosting system, $P(\text{SE}|\text{CJ})$, could observationally only be constrained using indirect methods. [Zhu & Wu \(2018\)](#) derived it using Bayes' law and assumptions on the individual detection probabilities $P(\text{SE})$ and $P(\text{CJ})$. They found $P(\text{SE}|\text{CJ})_{\text{obs}} = 90 \pm 20\%$. In the synthetic case, we can measure this quantity directly and obtain a much lower probability of 0.34 ± 0.06 . This result differs from the non-correlation case $P(\text{SE}|\text{CJ}) = P(\text{SE})$ by 1.1 standard deviations, suggesting that the occurrence of super-Earths is slightly enhanced in cold Jupiter hosting systems compared to field stars.

4.4.2 Removal of super-Earths

We find 93 systems that contain cold Jupiters but no super-Earths after 5×10^9 yr. This raises the question if the latter a) never existed; or b) disappeared during the formation phase.

All simulations start with 100 planet seeds of $0.01 M_{\oplus}$ each. Therefore, if the first hypothesis is true, we expect a significantly increased number of planets that had their growth stalled before reaching super-Earth mass, that is, due to the competition between solid material with other planets (particularly giants).

Thus, we confronted the fractions of planets in cold Jupiter-hosting systems without super-Earths ($\overline{\text{SE}} \cap \text{CJ}$) and with super-Earths ($\text{SE} \cap \text{CJ}$), respectively. To facilitate a comparison with the observed sample, we only used planets with $M_{\text{P}} \sin i \leq 2 M_{\oplus}$. In order to avoid biases introduced by planets that were accreted by the star or ejected from the system, we counted all planets in this mass range regardless of their ultimate fate. We find that the fraction of failed super-Earths is the same in both populations (0.80 compared to 0.81)². This shows that planetary growth to (at least) super-Earth mass was not inhibited in $\overline{\text{SE}} \cap \text{CJ}$ systems and hypothesis a) must be rejected.

To address hypothesis b), we distinguish between three scenarios that remove planets

²We caution that most of these planets retained masses close to our initial embryo mass; therefore these values should not be mistaken as occurrence rates for low-mass planets.

after they formed. They can be ejected out of the system, become accreted by the host star, or merge with another planet. Of the removed super-Earths in $\overline{\text{SE}} \cap \text{CJ}$ systems, 29 % are ejected, 11 % are accreted by the star, and 60 % are accreted by another planet.

Comparing the frequency of ejections in different populations, we find that from almost all (99 %) $\overline{\text{SE}} \cap \text{CJ}$ systems a planet in the super-Earth mass range (compare Table 4.1) was ejected. This compares to a significantly smaller fraction of 19 % for the overall population. Furthermore, 22 % of all super-Earths in systems hosting cold Jupiters were ejected, while only 2 % of super-Earths in non-cold Jupiter systems were ejected. The fraction of super-Earths that become accreted to the host star is small and comparable across the different populations, regardless of the presence of giants in the system.

An equally catastrophic and more common destiny for growing super-Earths are collisions with other protoplanets. During such events, part or all of the mass of a planet is transferred to the collision partner. In the majority of cases, this partner is a roughly terrestrial-mass body; only 20 % of events in the $\overline{\text{SE}} \cap \text{CJ}$ population correspond to a giant-mass partner. However, the “winner” of such a collision is likely to experience another planetary encounter during its lifetime, possibly with destructive consequences. We traced each accreted planet through its entire subsequent collisional history to determine which planet in the system became the final recipient of its material. In the $\overline{\text{SE}} \cap \text{CJ}$ population, only 26 % of these final accretors were one of the cold Jupiters in the system. Eventually, 34 % of accreted planets end up in super-Earth-mass planets (that might not survive), and 31 % in planets less massive than $2 M_{\oplus}$. We conclude that hypothesis b) is correct and the majority of missing super-Earths merged with another planet in their system.

In the following, we investigate the cause of these merger events. In the lower panel of Fig. 4.15, we show the eccentricity and period distributions of giant planets ($M_{\text{P}} \geq 95 M_{\oplus}$) in $\overline{\text{SE}} \cap \text{CJ}$ and $\text{SE} \cap \text{CJ}$ systems, respectively. Giants in systems with removed super-Earths have, on average, significantly higher eccentricities ($p = 5 \times 10^{-4}$). Their periods reach down to tens of days, while no giants with $P \lesssim 100$ d exist in systems with super-Earths. The planet masses follow similar distributions for both populations (Fig. 4.16), but there are differences on either extreme of the distribution: while $\text{SE} \cap \text{CJ}$ systems have a higher occurrence of “Saturns” ($M_{\text{P}} \sim 100 M_{\oplus}$), very-high-mass planets that reach into the Deuterium-burning regime occur only in the $\overline{\text{SE}} \cap \text{CJ}$ population.

Figure 4.17 shows the same distributions for planets in the super-Earth mass range. We included not only planets that survive the entire formation and evolution phase, but all planets that reached SE mass and never grew beyond. Planets that were accreted onto the stars were removed since the physical meaning of their final period and eccentricity

4 Warm Super-Earths and Cold Jupiters

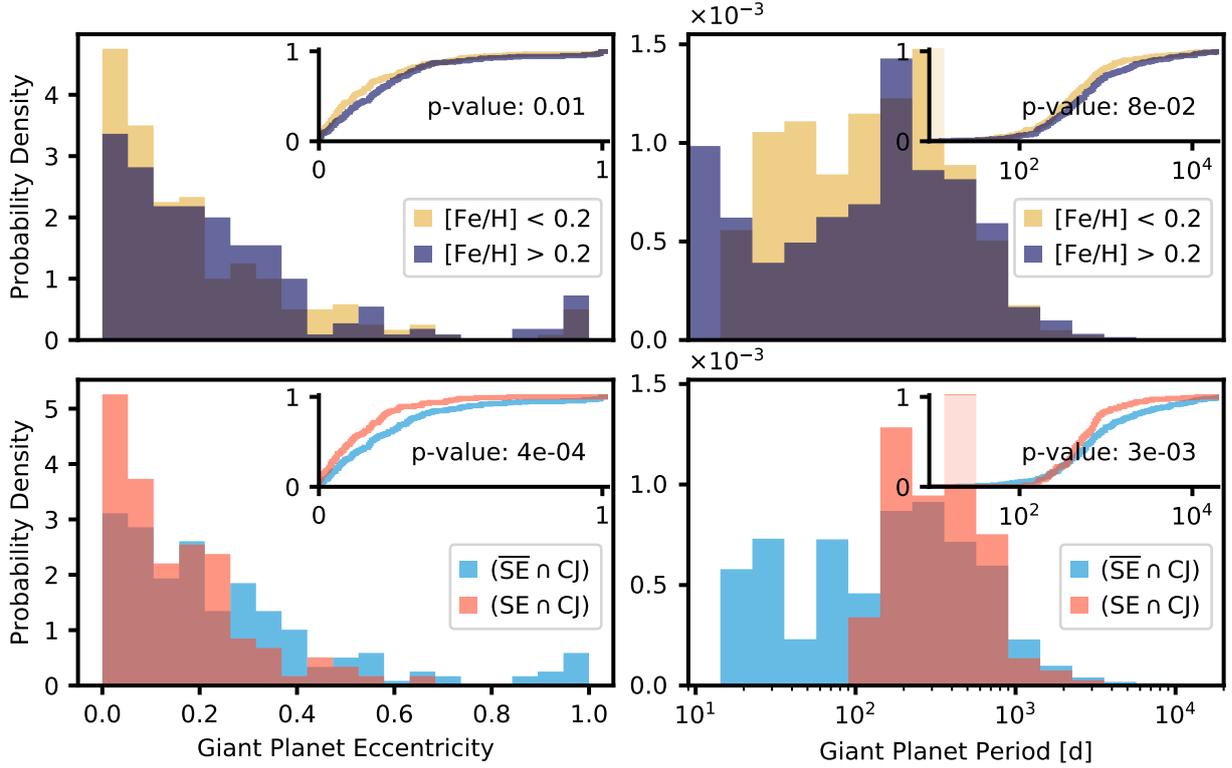


Figure 4.15: Eccentricity and period distributions of all giant planets that ever formed, regardless of their survival. This plot includes all planets with masses from $95 M_{\oplus}$, not only cold Jupiters. Insets show the corresponding empirical distribution functions.

Upper left: Eccentricity distribution for giants orbiting low-metallicity ($[\text{Fe}/\text{H}] < 0.2$, yellow) and high-metallicity ($[\text{Fe}/\text{H}] > 0.2$, blue) host stars. Planets with very low eccentricity are slightly more prevalent in low-metallicity systems.

Upper right: Period distributions of giant planets for different metallicities. A population of giants with very short periods exists only in the high-metallicity sample ($p = 8 \times 10^{-2}$).

Lower left: Eccentricity distributions of giant planets in cold Jupiter-hosting systems with and without super-Earth companions. Giants in systems without super-Earths have significantly higher eccentricities ($p = 5 \times 10^{-4}$). No super-Earths occur when a giant with $e \gtrsim 0.7$ exists.

Lower right: Period distributions of giants with and without super-Earths. The latter persist only in systems without short-period giants.

is questionable. On average, $\overline{\text{SE}} \cap \text{CJ}$ planets have higher eccentricities and larger periods than $\text{SE} \cap \text{CJ}$ planets. Ultra-short periods of less than three days are rare in $\overline{\text{SE}} \cap \text{CJ}$ systems.

4.4.3 Ice mass fractions

While detailed analyses of planetary compositions are beyond the scope of this study, we modeled the abundances of relevant chemical species and took into account the condensation of volatiles as a function of radial distance (Thiabaud et al., 2015).

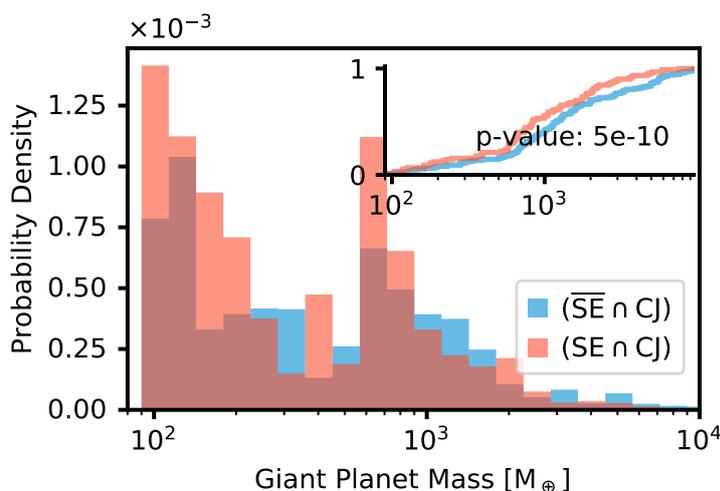


Figure 4.16: Mass distribution of giant planets in cold Jupiter-hosting systems with and without super-Earths. The mass distributions of the two populations show only minor differences.

To avoid distortions introduced by inclination effects and detection bias, we adopted for the planet classification into super-Earths and cold Jupiters an approach that reflects the architectures of our synthetic systems better than the limits in Table 4.1. This includes different mass limits for the inner planets, which are relatively abundant at higher masses than the $20 M_{\oplus}$ limit we used for the comparisons above (compare Fig. 4.7). We chose an upper limit of $47 M_{\oplus}$ (half of the lower limit for giant planets) for these planets. For each system, we:

1. checked if a giant planet exists, using our nominal mass limits. If yes, the upper period limit for super-Earths equals the period of the innermost giant. Otherwise, we used the same limit as in Table 4.1 of $P = 400$ d.
2. set mass limits for “massive” super-Earths of $1 M_{\oplus} < M_P < 47 M_{\oplus}$
3. did not impose a detection limit
4. did not draw an inclination term $\sin i$ but used the planets’ mass M_P .

Using these rules, we classified the population of systems into four distinct classes $(SE \cap CJ)$, $(\overline{SE} \cap \overline{CJ})$, $(SE \cap \overline{CJ})$, and $(\overline{SE} \cap CJ)$. For each of these system classes, we tracked the water content of individual planet cores throughout their formation history and show their ice mass fractions at $t = 5$ Gyr in Fig. 4.18.

The individual panels show balanced samples of 264 planets more massive than $0.5 M_{\oplus}$ from systems containing different combinations of super-Earths and cold Jupiters. The sample size corresponds to the number of planets in the smallest class.

4 Warm Super-Earths and Cold Jupiters

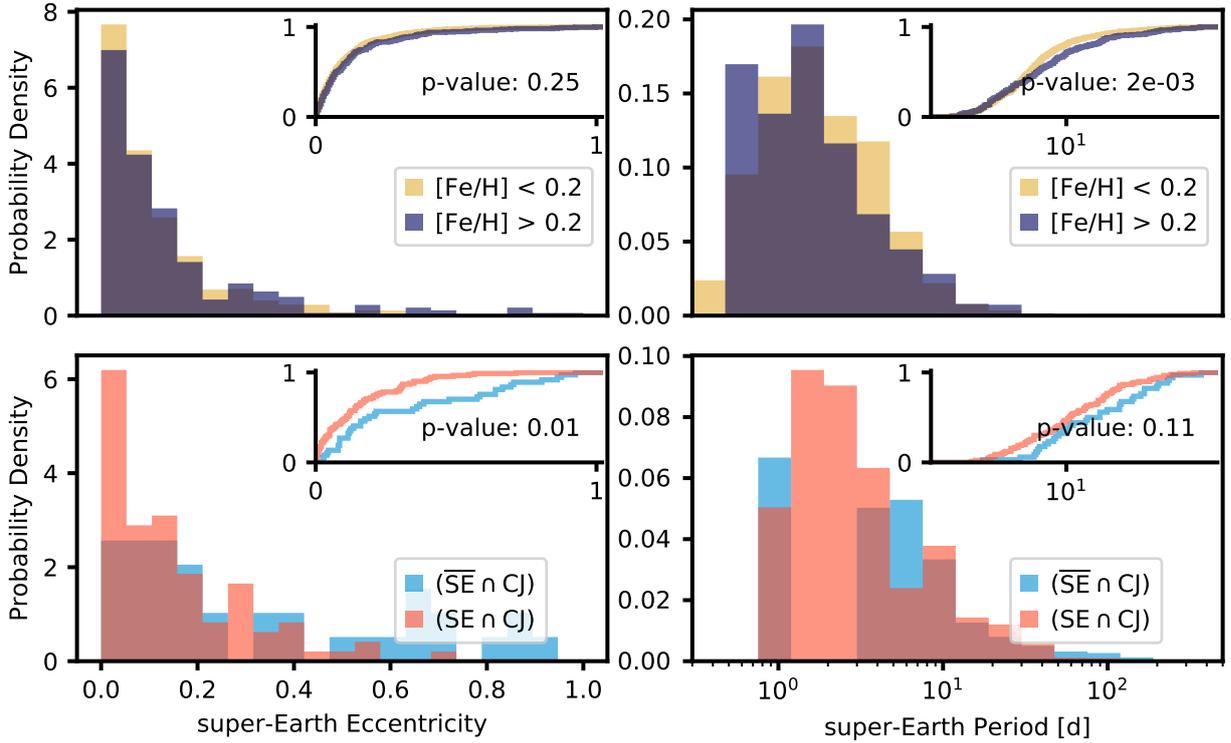


Figure 4.17: Same as Fig. 4.15, but for all super-Earths that ever formed during the systems’ histories. This includes both surviving planets and planets that we classified as super-Earths at the time of their removal (via ejection or collision events). Planets accreted to the star are not shown. With $p = 0.25$, a Kolmogorov-Smirnov test suggests a negligible statistical distance between the eccentricity distributions in high- and low-metallicity systems. On the other hand, the difference between systems with and without super-Earths is significant: where they are missing, eccentricities are strongly enhanced. The period distribution in the $SE \cap CJ$ population, which contains surviving super-Earths, is shifted toward lower values compared to $\overline{SE} \cap CJ$. These trends point to dynamical excitation of super-Earths by giant companions where they are present.

The core ice mass fractions f_{ice} of growing super-Earths are mainly determined by their initial orbital distances, which are indicated by the rug plot at the bottom of each panel. The colors of the markers correspond to the final ice mass fractions and show their strong correlations with the planet embryo positions relative to the water ice line. Cores that start outside of the ice line are strongly enhanced in water ice. Mixing of planetary compositions at later times is only possible due to different migration trajectories, collisions, and scattering events. Planets beyond ~ 1 au almost always maintain their high ice mass fraction. On the other hand, more close-in planets show varying compositions. On average, super-Earths in systems without outer giants are more ice-rich than their siblings in cold Jupiter-hosting systems. They mostly started just outside the water ice line and then migrated to their final positions, while super-Earths in cold Jupiter-hosting systems usually

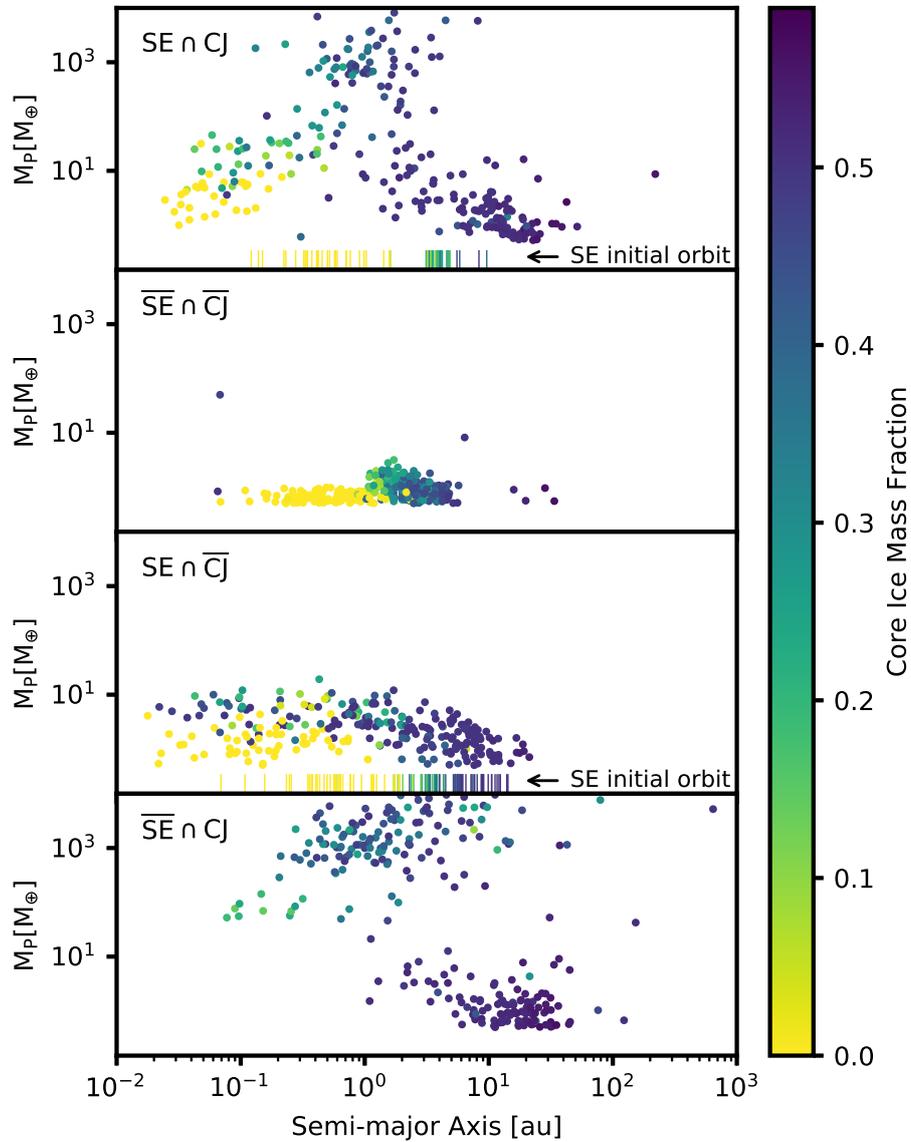


Figure 4.18: Water ice mass fractions of planet cores in the different system classes. For balanced samples of 264 planets per class, we show their position in mass-semi-major axis space, color-coded by ice mass fraction. For the systems containing super-Earths, we indicate their initial orbital distance by a rug plot with the same color-code. Planets $< 0.5 M_{\oplus}$ are not shown. The mass fractions of ice in the core are largely determined by the position of the water ice line in the protoplanetary disk, where planets beyond ~ 1 au are mostly water-rich. Super-Earths in systems without cold Jupiters have higher ice mass fractions than their siblings with giants.

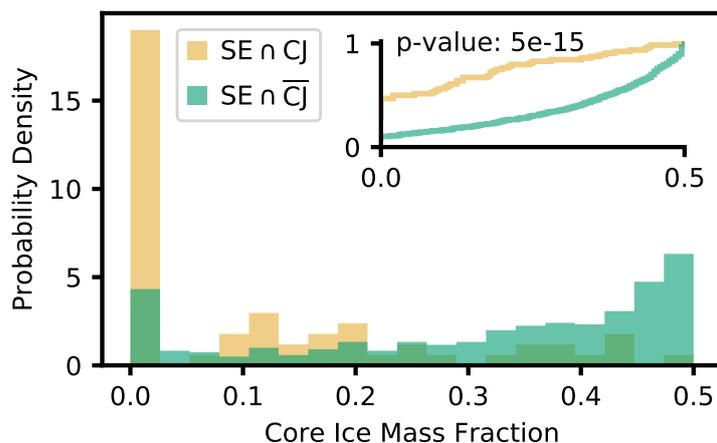


Figure 4.19: Distribution of ice mass fractions in the cores of super-Earths with and without cold Jupiter companion. The vast majority of super-Earths with giant companions is completely dry. On the other hand, super-Earths in systems without a giant planet often retain large ice mass fractions, reaching close to the maximum value of $f_{\text{ice}} \approx 0.59$. With $p \approx 10^{-15}$, the null hypothesis that both datasets are drawn from the same parent population must be rejected.

start within the ice line. The distributions of f_{ice} differ significantly between the two populations (compare Fig. 4.19). The median ice mass fractions of the two super-Earth-hosting classes amount to $f_{\text{ice}, \text{SE} \cap \overline{\text{CJ}}} = 0.23_{-0.23}^{+0.27}$ and $f_{\text{ice}, \text{SE} \cap \text{CJ}} = 0.02_{-0.02}^{+0.29}$, respectively, where upper and lower bounds denote the 84th and 16th percentiles.

These differences in composition are reflected in different bulk densities, which can be probed in a mass-radius diagram (Fig. 4.20). Here, we consider all inner ($a < 0.3$ au) planets with masses $1 M_{\oplus} - 47 M_{\oplus}$. In general, three groups of planets can be identified:

- Rocky planets without significant gaseous envelopes (lower diagonal chain of markers). These planets occupy the diagram areas with the highest densities and are the most abundant group.
- Icy planets without significant gaseous envelopes (upper diagonal chain of markers). Planets of this group have slightly lower bulk densities due to their high ice mass fractions.
- Planets that accreted and maintained significant atmospheres. These envelope-dominated planets are clearly detached toward larger radii.

As is apparent in the plot, cold Jupiter-hosting systems (top panel) are almost completely free from icy, atmosphere-less super-Earths. On the other hand, systems without cold Jupiters (center panel) are mainly populated by super-Earths with ice-rich cores. This

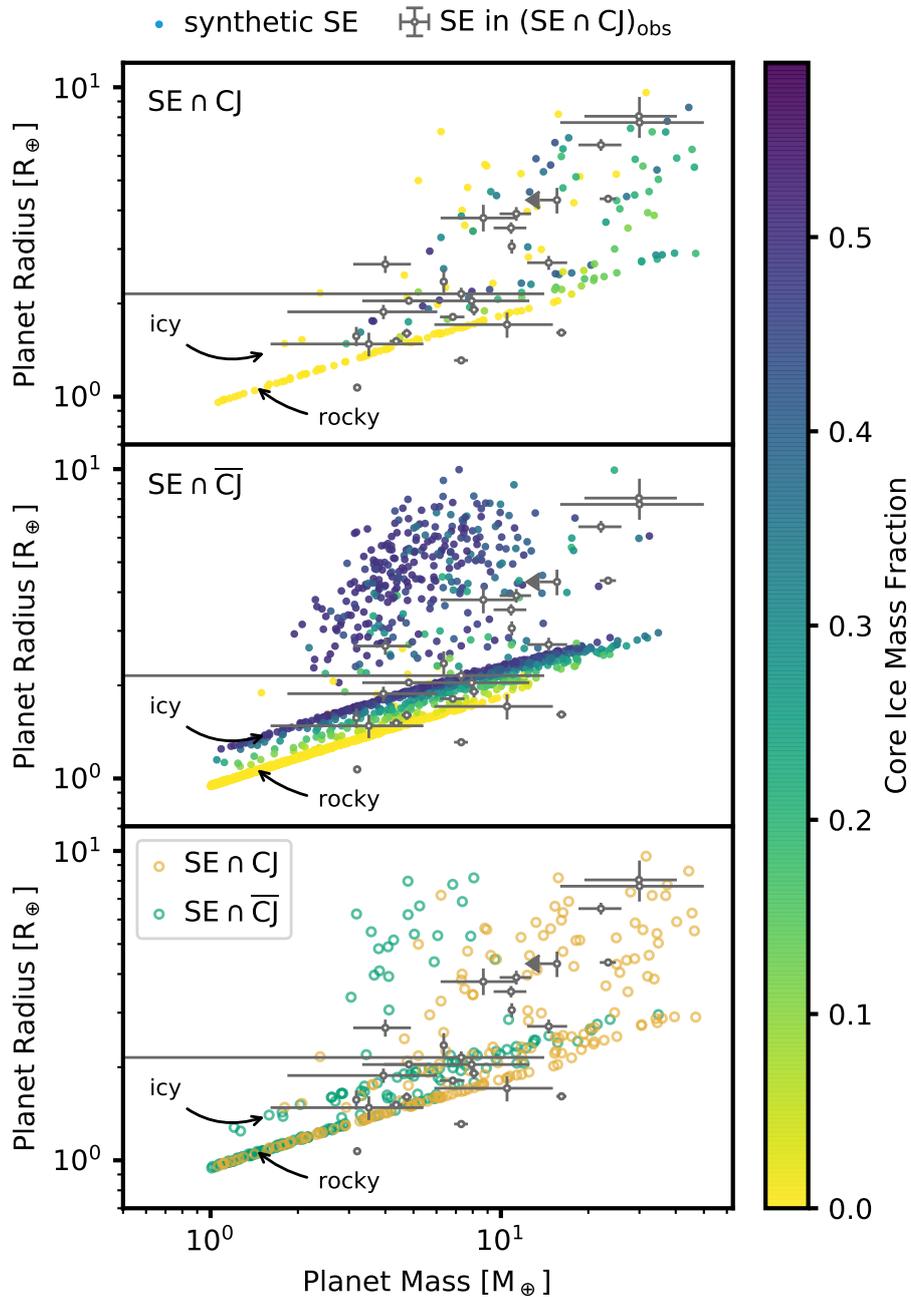


Figure 4.20: Mass-radius diagram of inner super-Earths with and without cold Jupiters. Included are all planets with masses of $1 M_{\oplus} - 47 M_{\oplus}$ and $a < 0.3$ au. The top and center panels show all such planets in $SE \cap CJ$ and $SE \cap \overline{CJ}$ systems, respectively, with their core ice mass fractions color-coded. The bottom panel shows balanced samples ($N=190$) from both system classes. On average, super-Earths in cold Jupiter-hosting systems populate regions of higher bulk density. In all panels, we overplot observed super-Earths in exoplanet systems containing both planet types, $(SE \cap CJ)_{obs}$. These planets match their synthetic counterparts well, but fail to match the $SE \cap \overline{CJ}$ population.

holds also for planets with significant H/He envelopes, which puts them to lower average masses.

In the bottom panel, $SE \cap CJ$ and $SE \cap \overline{CJ}$ systems are shown in different colors. Here, the sample sizes of the different classes are balanced. A clear difference between the two classes is visible: super-Earths in cold-Jupiter hosting systems show larger bulk densities, whereas those without a giant companion tend to populate regions of less density. The effect is even stronger for planets with large radii, that is, significant H/He envelopes, where those with masses below $\sim 10 M_{\oplus}$ exist almost only in the class without cold Jupiters.

These differences pose an interesting prediction for exoplanets with both mass and radius measurements. For a first-order comparison with real exoplanets, we constructed a sample of observed systems containing both super-Earths and outer giants, $(SE \cap CJ)_{\text{obs}}$. We did not, however, go so far as to compose an observed counterpart for the $SE \cap \overline{CJ}$ population since this sample would suffer from a strong bias: a system where no cold Jupiter was detected is not guaranteed to contain no such planets. The observed sample was constructed as follows: we obtained from the NASA exoplanet archive³ all confirmed planets and classified them in the same way we did for the synthetic population, using our flexible period limits (see above). We then kept only systems that contain both super-Earths and cold Jupiters. From these systems, we include the 26 super-Earths that have both mass and radius measured in the mass-radius diagram (Fig. 4.20). This confrontation with the theoretical sample reveals a remarkable agreement of $(SE \cap CJ)_{\text{obs}}$ with its synthetic counterpart, especially in the regime of planets with significant atmospheres. Here, $(SE \cap CJ)_{\text{obs}}$ differs substantially from the synthetic $SE \cap \overline{CJ}$ sample. In particular, $(SE \cap CJ)_{\text{obs}}$ matches the synthetic super-Earths with cold Jupiter companions $SE \cap CJ$ much better than it matches the overall super-Earth sample (compare bottom panel of Fig. 4.20).

4.4.4 Host star metallicity: effects on planet occurrences

Some planet formation models suggest that cold Jupiters and super-Earths form in different metallicity domains: host stars of low metallicity are not able to provide enough solid material for giant formation but can produce super-Earths (Ida & Lin, 2004b). On the other hand, high-metallicity hosts with larger solid reservoirs yield giants which prevent the formation of super-Earths (Izidoro et al., 2015). However, Zhu (2019) not only find the presence of super-Earths in observed high-metallicity systems but even a weak positive correlation of their occurrence.

To identify trends connected to host star metallicity, we compute all above probabilities

³queried on 2020-03-18

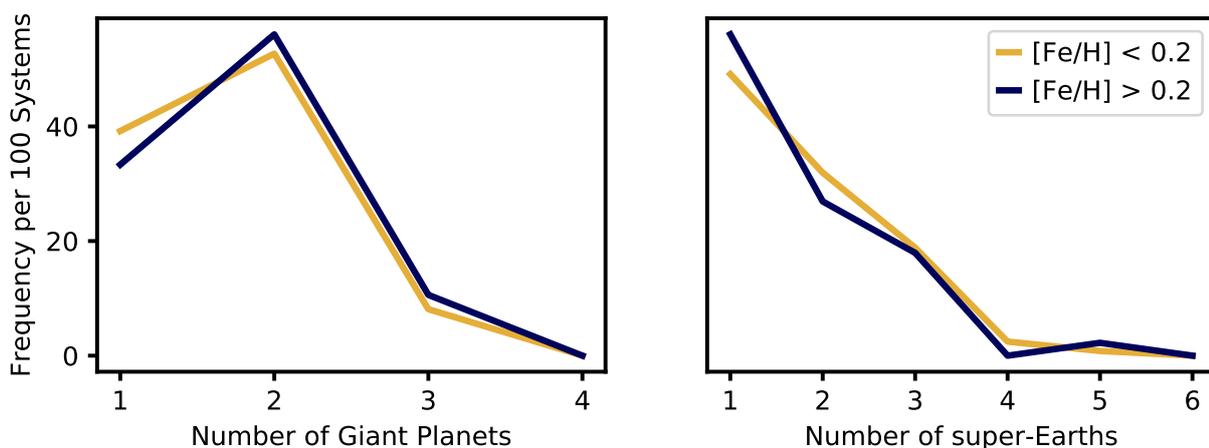


Figure 4.21: Multiplicity for different host star metallicities. *Left:* Number of giant planets per giant-hosting system. While high-metallicity systems have marginally higher multiplicity, there are no significant differences. *Right:* Number of super-Earths per super-Earth hosting system. Again, the differences are statistically indistinguishable.

not only for the full population but also for three distinct metallicity ranges: $[\text{Fe}/\text{H}] < -0.2$; $-0.2 < [\text{Fe}/\text{H}] < 0.2$; and $0.2 < [\text{Fe}/\text{H}]$. The distribution of metallicities in our synthetic population correspond to the observed values in the Solar neighborhood (see Sect. 2.5.1).

While observational studies revealed only a weak dependence of super-Earth occurrence on host star metallicity (e.g., Wang & Fischer, 2015; Zhu et al., 2016), we see a notable absence of super-Earths around low-metallicity stars with a super-Earth fraction of only 0.13 compared to 0.29 which we obtain for the full population.

Figure 4.15 shows a histogram and empirical distribution function of planet periods for two giant planet samples of low metallicity ($[\text{Fe}/\text{H}] < 0.2$) and high metallicity ($[\text{Fe}/\text{H}] > 0.2$). Host stars of high metallicity produce a rather bimodal distribution with a low-period bump. A similar feature was observed in the giant exoplanet population (Santos et al., 2006). This bimodality does not exist in the low-metallicity sample. A two-sample Kolmogorov-Smirnov test on the period distribution yields $p = 8 \times 10^{-2}$, allowing to reject the null hypothesis that the two samples stem from the same distribution.

A comparison of the number of giant planets per system between these samples (Fig. 4.21) shows no significant difference: where giant planets occur, their multiplicities follow the same distribution regardless of the metallicity. There is also no significant change in the planet mass distribution with metallicity, as has already been found in Mordasini et al. (2012a). As shown in Fig. 4.22, there is a difference in the CDF around $1000 M_{\oplus}$, but with $p = 0.06$, we cannot exclude equal source distributions.

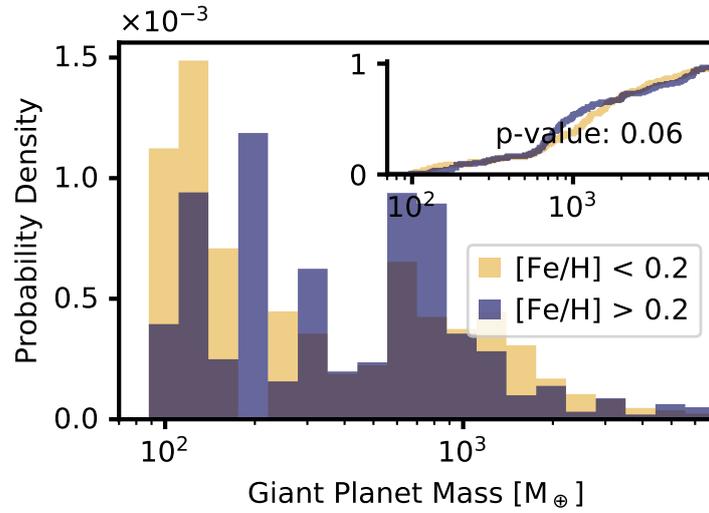


Figure 4.22: Mass distribution of giant planets in low- and high-metallicity systems. The difference between the distributions is statistically indistinguishable.

4.4.5 Possibility of reduced multiplicity in cold Jupiters

In this study, we denote as multiplicity, μ , the number of planets in a given system. Where we quote mean multiplicities across systems $\bar{\mu}$, we consider all planets that are above our detection limit of $K = 2 \text{ m s}^{-1}$ (compare Sect. 4.2.4). For the mean multiplicity of a specific planet type, we consider only systems containing at least one such planet.

Hansen (2017) suggested that dynamically hot outer giants can perturb inner terrestrial planets and decrease the multiplicities of these systems. Support for this hypothesis came from Zhu & Wu (2018), who find a deficiency of high-multiplicity systems in their cold Jupiter-hosting population, albeit with little significance due to the small sample.

In our synthetic sample, we find mean multiplicities of $\bar{\mu} = 2.27 \pm 1.08$ for the complete planet sample, $\bar{\mu}_{\text{SE}} = 1.85 \pm 0.92$ for super-Earths in super-Earth hosting systems, and $\bar{\mu}_{\text{CJ}} = 1.30 \pm 0.50$ for cold Jupiters in cold Jupiter hosting systems, quoting arithmetic mean and standard deviation. The frequency of multiplicities for different planet types is depicted in Fig. 4.23. We note that the plot shows the frequency per 100 systems containing the species (i.e. the sum of frequencies for each species equals to 100), it does not, therefore, reflect overall planet occurrences. In black, we show multiplicity frequencies for the complete sample, with orange and blue corresponding to the systems containing super-Earths and cold Jupiters, respectively. On average, the multiplicity of super-Earths is higher than for cold Jupiters. The latter show a multiplicity rate (fraction of systems with $\mu > 1$) of 28 %, consistent with the rate for observed cold Jupiters (e.g., Wright et al., 2009). Less than 4 % of all systems show a multiplicity greater than four and no systems

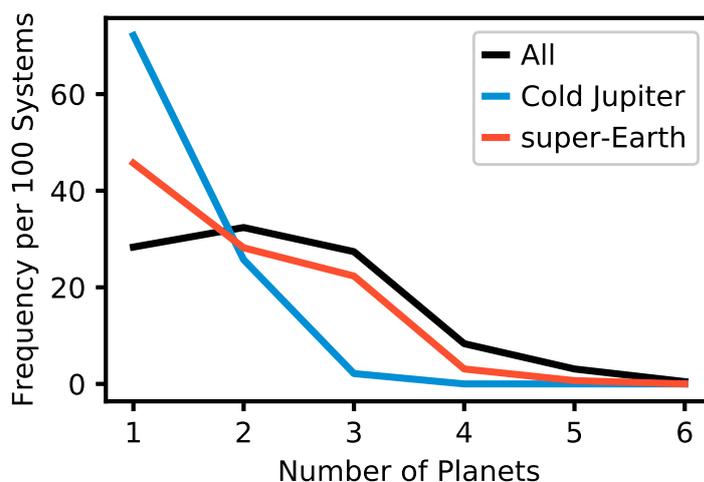


Figure 4.23: Normalized frequency of multiplicities for different planet types. The term “All” (black) includes all planets above our detection limit. For each of these multiplicities, we count the number of systems with this multiplicity and normalize it to 100 systems that host this species.

with $\mu > 5$ exist in the population.

To investigate a possible influence of the presence of an outer giant on the multiplicity of inner terrestrial planets, we examine the subpopulations of cold Jupiter systems and non-cold Jupiter systems separately. Figure 4.24 compares the frequency of super-Earth multiplicities for these two samples. Again, frequencies are normalized to 100 systems of the respective subpopulation. The mean super-Earth multiplicity is slightly enhanced in systems without outer companions (1.94 ± 0.93) compared to systems hosting a cold Jupiter (1.34 ± 0.67).

This is consistent with the observed trend of reduced multiplicity, but the effect is not significant enough to confirm it. We do not observe a difference in super-Earth multiplicities of low-metallicity systems ($[\text{Fe}/\text{H}] < 0.2$) and high-metallicity systems ($[\text{Fe}/\text{H}] > 0.2$) (see Fig. 4.21), as was suggested by Brewer et al. (2018).

4.5 Discussion

4.5.1 Observed and theoretical host star fractions

While the absolute fractions of super-Earth hosts and cold Jupiter hosts in our synthetic population are largely consistent with observations, there are considerable differences in the conditional probabilities $P(\text{CJ}|\text{SE})$ and $P(\text{SE}|\text{CJ})$. Observationally, only the former can be directly derived from counting statistics without additional assumptions on $P(\text{CJ})$ and $P(\text{SE})$.

4 Warm Super-Earths and Cold Jupiters

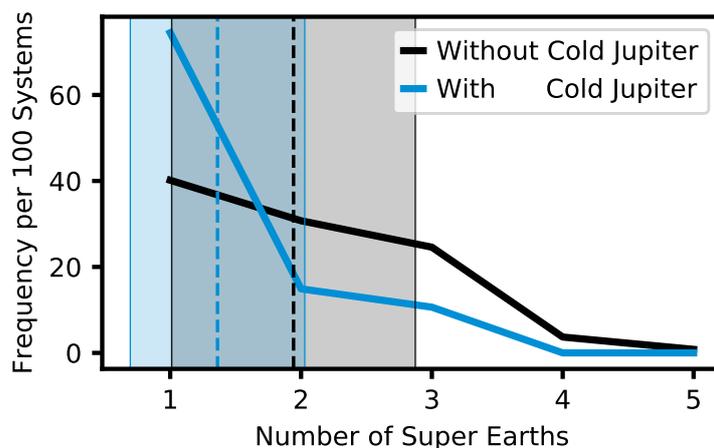


Figure 4.24: Normalized frequency of super-Earth multiplicity in systems with and without cold Jupiter. In accordance with Fig. 4.23, the frequency is normalized to 100 systems of the respective subpopulation. Dashed lines show mean multiplicities, whereas shaded regions reach to one standard deviation from the mean. super-Earth multiplicity is enhanced in systems without outer companions.

Zhu & Wu (2018) reported 9 out of 31 super-Earth systems that host a cold Jupiter, which results in a range of probabilities of $P(\text{CJ}|\text{SE})_{\text{obs}} = 0.29 \pm 0.18$ (compare Fig. 4.14). This is in line with Bryan et al. (2019), who found that super-Earth systems are enhanced in probability of also hosting a long period giant based on a sample of 65 stars. Zhu & Wu (2018) based their study on RV detections and Bryan et al. (2019) used a mixed sample of RV and transit-detected systems. Similarly, Herman et al. (2019) analyzed transiting planets and found a positive correlation between short- and long-period planets in a sample of 12 candidate systems. They conclude that outer giants occur exclusively in systems containing smaller inner planets.

With 1000 systems, our synthetic sample is significantly larger and can thus constrain the probability density tighter than the observations to $P(\text{CJ}|\text{SE})_{\text{syn}} = 0.16 \pm 0.03$. This puts our result between the non-correlation case and the strong positive correlations presented in Zhu & Wu (2018) and Bryan et al. (2019). Our figure is in agreement with the finding of Zhu & Wu (2018) within 1.7σ .

In contrast to the quantities above, the observed probability of finding a super-Earth in a system that hosts cold Jupiters, $P(\text{SE}|\text{CJ})$, could not be measured directly in these studies due to the poorly constrained detection bias for super-Earths. Zhu & Wu (2018) derived it from $P(\text{SE})$, $P(\text{CJ})$, and $P(\text{CJ}|\text{SE})$ using Bayes' theorem and report $P(\text{SE}|\text{CJ})_{\text{obs}} = 0.90 \pm 0.20$. Bryan et al. (2019) follow the same approach and, while not precisely constraining $P(\text{SE}|\text{CJ})$, come to the same conclusion. Both results suggest that nearly all cold

Jupiters are accompanied by inner super-Earths. It is thus surprising that [Barbato et al. \(2018\)](#), who conducted a survey to search for super-Earths, found zero planets in their sample of 20 cold Jupiter-hosting systems. In [Barbato et al. \(2020\)](#), it was clarified that the estimated sensitivity of the survey was inaccurate, which provides the possibility of a higher occurrence rate for inner super-Earths than the one reported. In Sect. 4.1.2, we took an assessment of their null result.

The corresponding quantity in our synthetic population is significantly lower with $P(\text{SE}|\text{CJ})_{\text{syn}} = 0.26 \pm 0.05$, indicating a lack of super-Earths in our cold Jupiter-hosting systems. In general, we do not find the reported strong positive correlation between the two planet types in our synthetic population. In the following, we discuss possible reasons for this apparent disagreement.

One explanation for the discrepancy is selection bias. We realized that the mass-period cuts chosen by [Zhu & Wu \(2018\)](#), which we adopted for our comparison, are not well reflected in the architectures of our systems but, rather, they constitute arbitrary borders in our population (compare with Fig. 4.7). In particular, our population of giant planets occupies both sides of the period limit and is located on closer orbits than the observed giant exoplanet population (e.g., [Fernandes et al., 2019](#)). The process that is responsible for their final orbit distance is the migration efficiency, which might be overestimated in our model. To investigate the relations between inner rocky planets and outer giants in a way that reflects our synthetic population better, we repeated the occurrence analysis with alternative mass/period limits: in each system, we first check if a giant planet exists according to the mass limits in Table 4.1. If it does, we set the outer period limit for our inner super-Earths to the period of the giant closest to the star, otherwise we chose a maximum period of 400 d. The planets fulfilling these period criteria are considered super-Earths if their masses obey $1 M_{\oplus} < M_{\text{P}} < 47 M_{\oplus}$. We consider the planet's actual masses instead of $M_{\text{P}} \sin i$ and no detection limit is imposed. Using these flexible limits, both $P(\text{CJ}|\text{SE})$ and $P(\text{SE}|\text{CJ})$ show a clear deficit compared to the respective unconditional probabilities. Super-Earths and cold Jupiters are anti-correlated in this case. This demonstrates that the correlations between the host star fractions are quite sensitive to the planet classification limits, which casts some doubt on the robustness of the observed trends.

At the same time, biases in the observations can falsify host star fractions, too. Exoplanet searches in general and RV surveys in particular suffer from human interventions that distort the inferences made on the underlying exoplanet demographics: to increase the significance of a candidate signal and to rule out false positive scenarios, it is very common to perform additional observations of a target star once such a signal emerges. Alarming,

this habit increases the probability of finding another planet in the same system. The “human intervention bias” thus contributes to a positive correlation in the occurrence of any two planet types and in a hardly quantifiable manner, in effect.

On a similar note, the small number of considered systems in certain observational studies (Zhu & Wu, 2018; Herman et al., 2019; Bryan et al., 2019) raises the question of whether their samples are representative of the field exoplanet population. Undoubtedly, our synthetic population is not a perfectly accurate representation of the planetary systems in nature. Hence, we do not claim that the observed trend stems merely from an unfavorable combination of selection and detection biases. However, we are concerned about the sensitivity of our results on the chosen limits. This ambiguity demonstrates yet again the importance of a thorough understanding of a sample’s selection function and of its underlying biases.

Overall, the anti-correlation in the synthetic planet population shows that giant planets on intermediate orbits can dynamically excite and ultimately destroy inner super-Earth systems (also see a discussion of this scenario in Masuda et al., 2020). We explore this mechanism in Sect. 4.5.2 in more detail. On the other hand, if the proposed positive correlation between inner super-Earths and outer giants exists, these results might indicate that inward migration of giant planets is not as efficient as hitherto assumed. More sophisticated migration models that take into account multiple interacting planets are currently not available (see, however, Masset & Snellgrove, 2001). Until they are, population syntheses with reduced migration efficiencies can possibly reconcile the observed and synthetic results. Such simulations will test if indeed an overestimation of planet migration torques is responsible for the competition between these planet types.

4.5.2 The missing super-Earths

Compared to pebble accretion models, where the formation of inner super-Earths relies on the radial drift of roughly centimeter-sized particles (Ormel & Klahr, 2010) that can be interrupted by an emerging outer giant (Morbidelli et al., 2015; Ormel et al., 2017; Bitsch et al., 2019a), the individual planet cores of a system are more independent in our model which considers only accretion of planetesimals and gas. We therefore do not expect a negative impact by massive, outer planets on the formation efficiency of cores on closer orbits.

On the other hand, such outer giant planets still block inward migration of smaller planets that formed beyond the giant’s orbit. Unlike in the model by Izidoro et al. (2015), super-Earths can form independently interior to any cold Jupiter companion in our simulations. The reason is the relatively steep radial slope of the planetesimal surface density

$\beta_s = 1.5$. The resulting plethora of solid material in the inner disk enables the formation of super-Earth-sized cores interior to the water ice line in disks massive enough to grow a giant planet.

Still, the observed occurrences in [Zhu & Wu \(2018\)](#) can overall be better matched with our final population if it contained more super-Earths. In Sect. 4.4.2 we showed that these planets are not lacking because they did not form, but because they were removed at some point in the formation and evolution phase. With smaller fractions of these failed super-Earths lost to ejections and accretion into the star, 60% disappeared in a merger event with another planet.

We consider whether the fate of the missing super-Earths was, in fact, sealed by a population of dynamically hot giants. In turning to the eccentricity distribution of giant planets (Fig. 4.15), we note that the lowest eccentricities are more prevalent for planets with a super-Earth companion in the system. There is also a population of highly eccentric giants that is missing in systems with super-Earths. The imprint of these giants can be seen in the eccentricity and period distribution of intermediate-mass planets, which differ significantly between systems with and without super-Earths (compare Fig. 4.17): whereas small eccentricities dominate in the population with super-Earths, the values are considerably higher where they are missing. Planets in $\overline{\text{SE}} \cap \text{CJ}$ systems have, in comparison with $\text{SE} \cap \text{CJ}$ systems, larger periods. Many of the super-Earths in $\text{SE} \cap \text{CJ}$ systems are on ultra-short orbits of a few days period. The reason is that such planets are safe from any destructive interaction with outer giants and can thus survive the entire formation and evolution phase.

Similar trends demonstrate instances when the sample is split in metallicity (see upper panels of Fig. 4.17), which may serve as a proxy for giant planet occurrence ([Johnson et al., 2010](#)). Our multiplicity analysis further shows that we can confirm the observed anti-correlation between cold-Jupiter occurrence and super-Earth multiplicity ([Zhu & Wu, 2018](#), see Fig. 4.24).

In summary, these findings suggest that most planetary systems produce super-Earths and where they are missing today, the stability of the system was perturbed by a giant planet. The culprit was typically not a cold Jupiter on a wide orbit, but a dynamically active and massive warm giant. A note of caution is due here since our model produces, on average, “warmer” giants than those found in the exoplanet demographics. This might lead to an overestimation of the effect of warm giants.

4.5.3 Volatile-poor super-Earths could be proxies for giant planets

For planets with both mass and radius measurements, their bulk density can be derived and the compositions of their interiors constrained. Most of these planets are expected to be accompanied by additional, often undetected planets (Zink et al., 2019; Sandford et al., 2019). Such companions, in particular a hypothetical giant planet, can place strong constraints on the formation history of the system. Our analysis of planetary compositions in Sect. 4.4.3 imply how under specific conditions the position of a planet in the mass-radius diagram could be used as a proxy for the existence of such an outer gas giant. A prerequisite for this proposal is a model that is able to produce both ice-rich and dry super-Earths. This has been proven difficult in the past, as core accretion models typically predict efficient core growth only beyond the water ice line, producing exclusively ice-rich planets (Izidoro et al., 2019). Conversely, in the population presented here, super-Earths accompanied by outer gas giants are reduced in volatile species compared to those without a giant companion. In Fig. 4.25, we illustrate schematically the reason for this dichotomy, which is rooted in disparate disk environments:

Disks that produce super-Earths but no cold giants tend to possess intermediate amounts of gas and solids. Here, most super-Earths start just outside the ice line where the additional reservoir of condensed volatiles provides the most efficient growth of a solid core. The mass surface density is however too small for the protoplanets to reach the critical masses for giant formation. They remain in the efficient Type I migration regime and experience strong inward migration, leading to the observed population of icy super-Earths in these systems.

On the other hand, disks that produce both planet types contain large solid and gas reservoirs (compare Fig. 4.10), which enables efficient growth of planetary cores to detectable sizes in a large range of orbits. In such disks, ice-poor super-Earths can form within the water ice line, while cores that accrete in regions just beyond it typically reach runaway gas accretion and grow to giant planets. They quickly enter the weaker Type II regime of planet migration and remain cold giants (Mordasini, 2018). Possible additional planets that formed further out cannot cross the giant's orbit to reach the inner system, which therefore contains only rocky planets.

These findings highlight the strong correlation between the migration history of inner super-Earths and their water content, which is largely determined by the fraction of the accretion phase spent outside of the ice line. This is also true for pebble accretion models if they assume that inward drifting pebbles lose their water ice once they cross the ice line (Bitsch et al., 2019b).

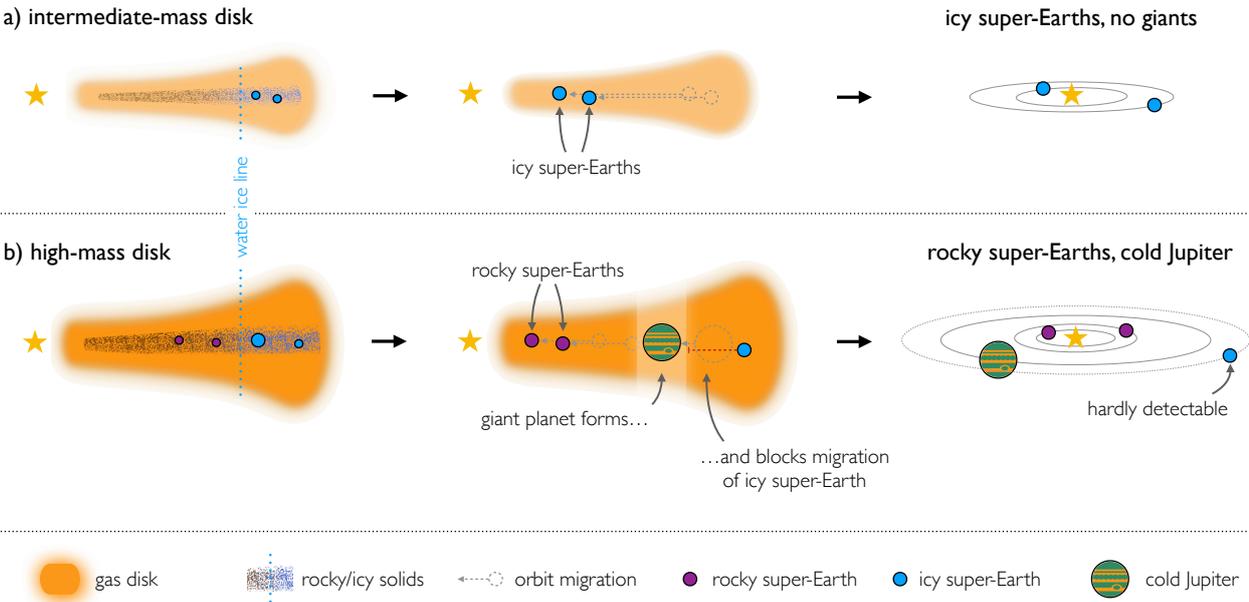


Figure 4.25: Schematic relations between solid disk mass, core ice, and system architecture:

a) a disk with just enough solid material ($M_{\text{solid}} \sim 100 M_{\oplus}$) to grow super-Earth cores will produce them preferably right behind the water ice line. The emerging ice-rich planets remain of too low mass to trigger runaway gas accretion and migrate freely to inner orbits where they can be detected.

b) a more massive disk ($M_{\text{solid}} \gtrsim 200 M_{\oplus}$) provides the conditions for giant planet formation, which again happens predominantly just beyond the ice line. The gas giant then blocks the migration of ice-rich cores that formed further out. However, the disk allows also for growth of dry super-Earths on closer orbits. In this scenario, the existence of super-Earths with high bulk densities is a proxy for giant planets in the same system.

These differences in the composition of inner super-Earths puts them into different regions in the mass-radius diagram (compare Fig. 4.20). While those planets with high ice mass fractions populate regions of larger radii and lower masses, that is, lower density, their rocky counterparts tend to occupy denser regions. Planets with significant H/He envelopes appear as a distinct group of planets with $R_{\text{P}} \gtrsim 3 R_{\oplus}$ in the diagram. In this regime, the separation is strikingly clear: while super-Earths in systems hosting cold Jupiters have typical masses of several tens of Earth masses, there are practically no gas-rich inner planets with masses below $\sim 10 M_{\oplus}$ in these systems. The reason is not, as one might suspect, a higher rate of giant impacts that can strip the envelopes of planets with low surface gravity. The frequency of such events is comparable in both populations. Instead, the higher core densities of these planets puts them to much higher masses at comparable radii. While the gaseous envelope contributes the bulk of a planet's radius,

the solid core dominates its total mass.

In conclusion, our model makes the testable prediction that volatile-poor super-Earths are more likely to host a long-period giant planet. Conversely, an inner super-Earth in a system that is also harboring a distant giant planet is likely depleted of volatiles.

Obviously, this should be interpreted in light of the assumptions we put into our model. The main simplifications that could influence our results include: the simplified disk chemistry, efficient formation of planetary embryos in the entire disk, accretion of 300 m sized planetesimals, and the generally unsatisfying constraints on planet migration.

Unfortunately, the current sample of exoplanets with known mass and radius is still too small to unequivocally test our hypothesis. However, we show in Sect. 4.4.3 that the currently available sample of 26 super-Earths with confirmed cold Jupiter-companions matches the predicted bulk densities for such systems much better than the one for the overall synthetic population. This is especially true for atmosphere-hosting planets, where the observed sample fits only the synthetic super-Earth population with giant companions and not the one without. A more rigorous benchmark would be a thorough reanalysis of observed planets using raw photometry and RV time series, and including a consistent evaluation of the underlying detection bias. This is not only beyond the scope of this study, but will require a larger sample of planets with precise photometric and spectroscopic measurements than is currently available. With the ongoing *TESS* mission (Ricker et al., 2014) and RV follow-up of its planet candidates, the number of systems for which such datasets exist is constantly increasing. Thus, statistical tests of the trends we presented here are imminent and will ultimately show if our predictions hold.

4.5.4 A negative metallicity correlation for super-Earths in cold-Jupiter hosting systems

The conditional planet host fractions as a function of host star metallicity (Sect. 4.4.4) reveal an unexpected trend: $P(\text{SE}|\text{CJ})$ correlates negatively with metallicity. Before we illustrate that this correlation is caused by an increased emergence of warm giants in high-metallicity systems, we exclude multiple alternative scenarios:

Multiple giant planets in high-metallicity systems

One factor potentially influencing the relation between metallicity and super-Earth occurrence is the formation of several massive planets per system. Such efficient formation channels may be expected in high-metallicity systems, which emerged from disks containing a large amount of solids. However, the multiplicity of giant planets does not vary with metallicity in the synthetic population (compare Fig. 4.21). While giant planets are only

present from a threshold metallicity upwards, the ratio between systems of different giant planet multiplicity remains the same with increasing $[\text{Fe}/\text{H}]$ (compare with [Emsenhuber et al. \(2020b\)](#)).

Selection bias induced through our super-Earth definition

Another possibility is that with increasing $[\text{Fe}/\text{H}]$, we are missing an increasing amount of planets in our statistics because they grow to larger bodies which we do not classify as super-Earths. Such a deficit would be apparent in the general super-Earth statistics $P(\text{SE})$, which, however, shows a positive metallicity correlation, ruling out this assumption.

Orbital properties of giant planets

A proposition we want to pursue in more detail is that high-metallicity giants are on shorter orbits and dynamically active and thus more likely to destroy a population of small planets at short orbital periods. This would make sense, especially in light of the observed close relation between super-Earth occurrence and disk solid content (compare Sect. 4.3.6) which, in our model, is tightly correlated with metallicity. With an increasing amount of solids, the occurrence rate drops just where the first giant planets emerge. To test this hypothesis, we compared orbital parameters of giants with low- and high-metallicity host stars (Fig. 4.15) and found that the high-metallicity sample extends to lower periods. Their mass distribution is inconspicuous, but their eccentricities are slightly enhanced.

We are especially interested in giants with small and intermediate orbital distances since these planets are most prone to disturbing inner super-Earths. Even though, in our model, the planets that endanger these systems are mainly warm Jupiters, in the following, we briefly discuss the planets fulfilling our criteria for a cold Jupiter (see Table 4.1). Figure 4.26 relates, for each pair of inner super-Earth and cold Jupiter, the semi-major axes of the innermost giant (yellow) and of the outermost super-Earth (red) with the metallicity of their host star. For comparison, we plot cold Jupiters that do not have inner super-Earth companions in blue. Horizontal lines denote the full orbital ranges from periapsis to apoapsis and light markers correspond to giant planets that formed but have not survived. Between the outermost super-Earths and their innermost cold Jupiter companion, a “safety gap” of $\sim 1/2$ au emerges. Almost no giants with host star metallicities greater than $[\text{Fe}/\text{H}] \approx 0.3$ have super-Earth companions and the aforementioned positive correlation between host star metallicity and cold Jupiter eccentricity is apparent. The highest eccentricities belong to giant planets in systems that lost their super-Earths.

Taken together, these results indicate that inner super-Earth systems are more likely to

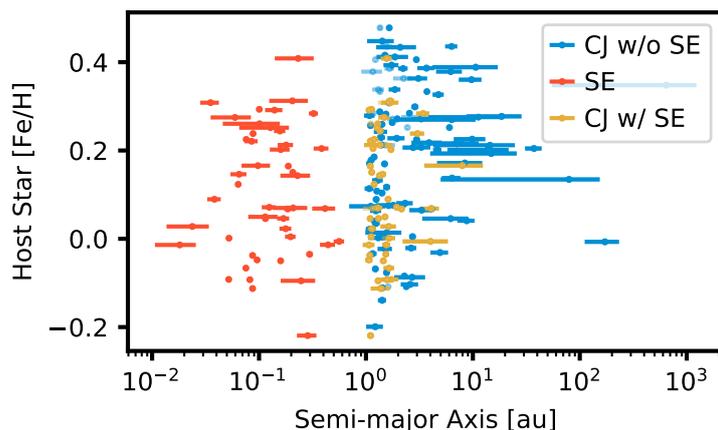


Figure 4.26: Orbital range of outermost super-Earth and innermost cold Jupiter plotted by host star metallicity. For each system, we plot the orbital range (periastron to apastron) of the outermost super-Earth (red) and innermost cold Jupiter (yellow). Cold Jupiters in systems without super-Earths are shown in blue, and light markers correspond to giant planets that did not survive. Cold Jupiter systems with host star metallicities greater than ~ 0.3 typically harbor no super-Earths, and high-metallicity giants without super-Earths are often on eccentric orbits.

be destroyed in a high-metallicity environment, where there are warmer and dynamically more active giant planets that can disrupt them.

4.6 Conclusions

In this chapter, we compared the *NG76* multi-planet population from the Bern *Generation 3* global planet formation and evolution model to observed exoplanets around Solar-type stars, focusing on the relation of close-in super-Earths and far-out giant planets (“cold Jupiters”). Our results can be summarized as follows:

1. Our synthetic planet population shows a positive intra-system correlation between the occurrences of inner super-Earths and cold Jupiters, albeit weaker than previously proposed. The reduction is attributed to warm giant planets that frequently disrupt inner systems of super-Earths. This discrepancy might hint to an overestimation of the migration efficiency of giant planets. We showed that the correlation is sensitive on the choice of mass and period limits that defines the sample of inner and outer planets.
2. We find a difference in the bulk composition of inner super-Earths with and without cold Jupiters. High-density super-Earths point to the existence of outer giant planets in the same system. Conversely, a present cold Jupiter gives rise to rocky, volatile-depleted inner super-Earths. Birth environments that produce such dry planet cores

in the inner system are also favorable for the formation of outer giants, which obstruct inward migration of icy planets that form on distant orbits. This predicted correlation can be tested observationally.

3. It is the result of a general link between the initial reservoir of solids and final system architecture: low-mass solid disks tend to produce only super-Earths but no giant planets. Intermediate-mass disks may produce both super-Earths and cold Jupiters. High-mass disks lead to the destruction of super-Earths and only giants remain.
4. Inner super-Earths initially form in nearly all systems that host an outer giant. Where they are missing today, the inner system was dynamically excited by giant planets on intermediate orbits, leading to the destruction of super-Earths.
5. The key parameter for the formation of both cold Jupiters and super-Earths is the solid content of the protoplanetary disk. With increasing initial solid mass, super-Earth occurrence rises steeply but drops for disks that are massive enough to form giant planets.
6. Outer giants reduce the multiplicity of small inner planets. In line with the tentative observational evidence ([Zhu & Wu, 2018](#)), the number of super-Earths that survive the entire formation and evolution phase is reduced where cold Jupiters occur.
7. High-metallicity giant planet hosts are less likely to harbor inner super-Earths. Planetary systems around stars with high metallicity frequently contain warm and dynamically active giant planets that can disrupt inner planetary systems.

5

A Highly Eccentric Warm Jupiter Orbiting TIC 237913194

The content of this chapter is based on the publication [Schlecker et al. \(2020a\)](#) in *The Astronomical Journal*.

Details of authorship: I am the corresponding author of this paper and led the project under supervision of Thomas Henning. I prepared the manuscript with written contributions by Felipe Rojas (Sect. 5.2.1), Rafael Brahm (Sects. 5.2.2, 5.2.4, and 5.3.1), Ludmila Carone (Sect. 5.3.3), Karan Molaverdikhani (Sect. 5.4.2), and Trifon Trifonov (Sect. 5.4.3). Ludmila Carone also prepared Fig. 5.5; all other figures in this chapter were created by me. The planet candidate was identified by Rafael Brahm, Néstor Espinoza, and Felipe Rojas, and follow-up observations of the target were coordinated and conducted by Rafael Brahm, Néstor Espinoza, Melissa J. Hobson, Paula Sarkis, and Andrés Jordán. The stellar parameters were determined by Rafael Brahm. Diana Kossakowski and myself implemented the joint modeling of the system and derived the planetary parameters. Melissa Hobson helped with interpreting the RV data and conducting the stellar activity analysis. Ludmila Carone consulted me on the approximation of the planetary equilibrium temperature. Néstor Espinoza and Joshua N. Winn provided guidance in estimating the amplitudes of phase curve modulations and of the Rossiter–McLaughlin Effect. Karan Molaverdikhani and Paul Mollière carried out the feasibility study of atmospheric characterization. Trifon Trifonov conducted the tidal evolution analysis. All authors provided comments on the paper manuscript.

5.1 Motivation

Gravitational interactions among massive planets during their formation and evolution leave an imprint on their orbital parameters. However, these imprints are frequently erased in the case of hot Jupiters, which are prone to orbital changes through tidal interactions with their host star (e.g., Eggleton et al., 1998). Planets on more distant orbits ($P \gtrsim 10$ d), although not as readily detected, are expected to retain this information and thereby provide valuable insights into the formation history of their planetary system. Unfortunately, the sample of confirmed, nearby transiting warm Jupiters is still small. The transit survey currently performed by the Transiting Exoplanet Survey Satellite (*TESS*, Ricker et al., 2014) is changing that: hundreds of giant planets on intermediate orbits are expected to be detected during the all-sky survey (Sullivan et al., 2015; Barclay et al., 2018). With this in mind, the Warm gIaNts with tEss (WINE, Brahm et al., 2019; Jordán et al., 2020) collaboration embarked on a search for such warm Jupiters. Using a network of photometric and spectroscopic facilities, we identify and follow up *TESS* planet candidates to confirm them, characterize their orbital parameters, and use them to inform planet formation theory.

Here, we report the discovery of a temperate giant planet in a highly eccentric orbit around a G3 star. By the aid of additional ground-based photometry from the *CHAT* and *LCOGT* telescopes, as well as precise radial velocity measurements from *FEROS*, we were able to tightly constrain the planet’s mass, radius, and orbital parameters. It is only the third *TESS* giant planet with $e > 0.3$ (Jordán et al., 2020; Rodriguez et al., 2019), and it has one of the most eccentric orbits reported to date for a warm Jupiter.

We show that its dynamical state is not consistent with a high-eccentricity migration scenario that would eventually result in the planet becoming a hot Jupiter. Instead, a past interaction with an undetected massive body has likely caused the planet’s extreme orbit. This valuable addition to the small sample of known warm Jupiter-hosting systems can help constrain the enigma of their origin. Through its eccentric orbit and the subsequent varying radiative forcing, the planet further holds the promise of observing potential disequilibrium processes in its atmosphere.

This chapter is organized as follows: In Section 5.2 we present the observational data used in this study. Section 5.3 covers the analysis of these data and concludes with determining properties of the planetary system and its host star. In Section 5.4 we discuss implications of our findings and put TIC 237913194b in context with the known exoplanet population. Finally, in Section 5.5 we summarize the results of our study.

We make the code used in the analysis that led to our results available in a public git

repository¹.

5.2 Observations

5.2.1 *TESS* photometry

For identifying warm Jupiter candidates, we generated light curves for all bright stars of the TICv8 catalog from the Full Frame Images (FFIs, Jenkins et al., 2016) of *TESS* using the *tesseract*² package (Rojas, in prep.). Briefly, *tesseract* receives any TIC ID or coordinate as input and performs simple aperture photometry on the FFIs via the *TESSCut* (Brasseur et al., 2019) and *lightkurve* (Barentsen et al., 2019) packages. Aperture selection was done following Lund et al. (2015). Specifically, 293253 and 479184 light curves of bright objects ($T < 14$ mag) have been generated from Sectors 1 and 2, respectively. For identifying warm Jupiter candidates, we apply a simple algorithm that goes through the light curve searching for zones that significantly deviate in the negative direction from the median flux around a given region. Then we check by visual inspection if these zones are consistent with a transit-like feature. This procedure allows us in principle to identify also single transitters in a given *TESS* Sector (e.g., Gill et al., 2020). By using this algorithm we found that the star TIC 237913194 presented transit-like periodic features in the two first Sectors of *TESS*. An initial fit to the photometric data indicated a period of $P \approx 15.17$ and a transit depth of $\delta_{\text{TESS}} \approx 0.8\%$, consistent with the properties of a warm giant candidate given the parameters of the star according to the TICv8 catalog. The *TESS* light curve of TIC 237913194 is shown in the upper panel of Fig. 5.1.

5.2.2 Photometric follow-up with CHAT

Due to the limited angular resolution of *TESS*, ground-based photometry is required to reject false positive scenarios like blended eclipsing binaries. TIC 237913194 was observed on the night of December 12, 2019 with the 0.7m Chilean-Hungarian Automated Telescope³ (CHAT) installed at the Las Campanas Observatory. Observations were performed with a Sloan i' photometric filter using a mild defocus and exposure times of 110 sec. We processed the data with a dedicated pipeline developed to produce high precision light curves using differential photometry (e.g. Espinoza et al., 2019a; Jones et al., 2019) with the LCOGT 1.0m telescopes (Brown et al., 2013). The optimal photometric precision was obtained with an aperture of 14 pixels (8.3"). We plot the obtained light curve in the

¹<https://github.com/matiscke/eccentricWarmJupiter>

²<https://github.com/astrofelipe/tesseract>

³https://www.exoplanetscience2.org/sites/default/files/submission-attachments/poster_aj.pdf

bottom right panel of Fig. 5.1. We recovered a full transit, which confirms that the transit-like features identified in the *TESS* data occur in TIC 237913194. The transit depth of $\delta_{\text{CHAT}} = 0.0087 \pm 0.0004$ is consistent with the signal identified in the *TESS* photometry.

5.2.3 Additional photometry from LCOGT

Because of a rather grazing transit configuration, the posterior probability densities from our initial fits contained a strong degeneracy between the scaled planetary radius R_{P}/R_{\star} and the impact parameter b . To lift this degeneracy and to improve the constraint on the planet radius, we obtained additional transit photometry with the Las Cumbres Observatory Global Telescope (*LCOGT*) Network on July 16, 2020 (see Fig. 5.1). The measurements were taken in the i' band and cover all phases of the transit. To maximize photometric precision, we chose an aperture of 24 pixels ($9.4''$). We recover a transit depth of $\delta_{\text{LCOGT}} = 0.0083 \pm 0.0002$. Within the uncertainties, this is consistent with the values obtained for the other instruments. Including the additional data in the fit leaves only little residual correlation between R_{P}/R_{\star} and b and strongly improved the posterior on R_{P} (see Sect. 5.3.2).

We make all our follow-up light curves available on [exoFOP](https://exofop.ipac.caltech.edu)⁴.

5.2.4 High precision spectroscopy with FEROS

We obtained high-resolution ($R = 48000$) spectra with the Fiber-fed Extended Range Optical Spectrograph (FEROS, [Kaufer et al., 1999](#)), mounted at the 2.2 m MPG telescope at La Silla Observatory. In total, 25 exposures of 1200 s were taken between June 19, 2019 and March 9, 2020. From these, we extracted radial velocities (RV) using the CERES pipeline ([Brahm et al., 2017a](#)), which performs all steps from bias, dark, and flat-field calibration to cross-correlation matching of the resulting spectrum with a G2-type binary mask. The observations were performed in the simultaneous calibration mode for tracking the instrumental velocity drift produced by changes in the spectrograph environment. This procedure involves the monitoring of a ThAr spectrum with a second fiber. The typical signal-to-noise ratio of these spectra was about 70. The time series of FEROS RV measurements are shown in Fig. 5.2 and listed in Table 5.1.

5.2.5 Contamination

We checked for possible closeby sources that could contaminate our photometric aperture with their light. Any sources within $\sim 10''$, which is the photometric aperture we used for our *LCOGT* photometry, could cause such contamination. After querying the GAIA DR2 catalog ([Gaia Collaboration et al., 2018](#)) we found the closest source to TIC 237913194 at

⁴<https://exofop.ipac.caltech.edu/tess>

Table 5.1: FEROS radial velocities and accompanying data for TIC 237913194 used in this chapter.

| BJD | RV [km/s] | σ_{RV} [km/s] | t_{exp} [s] | BIS [km/s] | σ_{BIS} [km/s] |
|-------------|-----------|-----------------------------|----------------------|------------|------------------------------|
| 2458669.798 | 29.536 | 0.009 | 1200 | -0.016 | 0.012 |
| 2458670.816 | 29.519 | 0.010 | 1200 | -0.014 | 0.013 |
| 2458672.879 | 29.533 | 0.012 | 1200 | 0.014 | 0.016 |
| 2458718.916 | 29.505 | 0.007 | 1200 | -0.006 | 0.011 |
| 2458721.768 | 29.528 | 0.007 | 1200 | -0.000 | 0.011 |
| 2458722.674 | 29.580 | 0.008 | 1200 | -0.001 | 0.011 |
| 2458723.758 | 29.595 | 0.007 | 1200 | -0.019 | 0.011 |
| 2458724.793 | 29.611 | 0.009 | 1200 | -0.019 | 0.012 |
| 2458742.740 | 29.892 | 0.009 | 1200 | -0.002 | 0.013 |
| 2458783.690 | 29.545 | 0.007 | 1200 | -0.033 | 0.011 |
| 2458785.610 | 29.646 | 0.010 | 1200 | -0.026 | 0.013 |
| 2458787.671 | 29.802 | 0.008 | 1200 | 0.012 | 0.012 |
| 2458791.610 | 29.521 | 0.009 | 1200 | -0.000 | 0.012 |
| 2458792.571 | 29.508 | 0.008 | 1200 | 0.007 | 0.011 |
| 2458800.668 | 29.614 | 0.008 | 1200 | -0.017 | 0.011 |
| 2458801.688 | 29.673 | 0.008 | 1200 | 0.014 | 0.011 |
| 2458802.619 | 29.739 | 0.007 | 1200 | -0.014 | 0.010 |
| 2458805.675 | 29.543 | 0.007 | 1200 | -0.001 | 0.010 |
| 2458847.592 | 29.701 | 0.008 | 1200 | -0.024 | 0.011 |
| 2458848.591 | 29.858 | 0.008 | 1200 | 0.006 | 0.011 |
| 2458849.592 | 29.805 | 0.008 | 1200 | 0.005 | 0.011 |
| 2458850.541 | 29.619 | 0.007 | 1200 | -0.010 | 0.011 |
| 2458911.524 | 29.560 | 0.008 | 1200 | -0.038 | 0.011 |
| 2458915.515 | 29.480 | 0.010 | 1200 | 0.023 | 0.013 |
| 2458917.514 | 29.510 | 0.010 | 1200 | 0.039 | 0.013 |

an angular separation of $\sim 46''$. We thus find no evidence for significant contamination of our photometry.

5.3 Analysis

5.3.1 Stellar Parameters

For characterizing the host star, we first determined its atmospheric parameters from the co-added FEROS spectra. Specifically, we used the ZASPE code (Brahm et al., 2017b) which compares the observed spectrum against a grid of synthetic ones generated from the ATLAS9 model atmospheres (Castelli & Kurucz, 2004). We then used the PARSEC evolutionary models (Bressan et al., 2012), as described in Brahm et al. (2019), to determine

the physical parameters of the star. Briefly, we compared the observed broad band photometric magnitudes of the star with those generated with models having different physical parameters by taking into account the distance determined from the Gaia DR2 parallax and assuming an extinction law of [Cardelli et al. \(1989\)](#) dependent on the A_V parameter. The parameter space was explored using the `emcee` ([Foreman-Mackey et al., 2013](#)) package. The obtained atmospheric and physical parameters of TIC 237913194 are listed in [Table 5.2](#). TIC 237913194 is a main sequence G-type star with a mass of $M_\star = 1.026_{-0.055}^{+0.057} M_\odot$, a radius of $R_\star = 1.088_{-0.012}^{+0.012} R_\odot$, and an age of 5.7 ± 1.7 Gyr. TIC 237913194 is slightly metal rich ($[Fe/H] = +0.14 \pm 0.05$ dex) and has an effective temperature of $T_{\text{eff}} = 5788 \pm 80$ K. We note that the quoted uncertainties do not account for possible systematic errors in the stellar evolutionary models.

5.3.2 Joint modeling

For the joint photometry and RV modeling, we employed the python package `juliet` ([Espinoza et al., 2019b](#)). This tool uses existing codes to model transit photometry (`batman`, [Kreidberg, 2015](#)) and radial velocity time series (`radvel`, [Fulton et al., 2018](#)). It further allows us to incorporate Gaussian Process Regression (GP) via the `celerite` package ([Foreman-Mackey et al., 2017](#)), which we employ to model systematic nuisance signals. To explore the parameter space, it uses the MultiNest nested sampling technique ([Feroz et al., 2009](#)), implemented in the `pyMultiNest` software package ([Buchner et al., 2014](#)). `juliet` further calculates evidences $Z_i = P(M_i(\theta)|\mathcal{D})$ for models M_i with sets of parameters θ given the data \mathcal{D} . To compare two models M_i, M_j , we compute the differences of their log-evidences,

$$\Delta \ln Z_{i,j} = \ln Z_i / Z_j = \ln [P(M_i(\theta)|\mathcal{D}) / P(M_j(\theta)|\mathcal{D})]. \quad (5.1)$$

Here, we adopted a general rule of thumb that if $\Delta \ln Z_{i,j} \geq 3$, the model with the larger log-evidence is favored. If $\Delta \ln Z_{i,j} \lesssim 3$, we consider the models to be indistinguishable and prefer the simpler one. As the MultiNest algorithm is known for showing scatter in $\ln Z$ that exceeds the reported uncertainties (e.g., [Nelson et al., 2020](#)), we always repeated the calculations several times. The variations among such runs were always smaller than one and therefore negligible for our purposes.

Model parameters

For several inferred quantities, we fitted parametrizations that allow for efficient sampling and are limited to physically plausible values:

Table 5.2: Stellar properties of TIC 237913194

| Parameter | Value | Reference |
|--|---|-----------|
| Names | TIC 237913194 | TESS |
| | 2MASS J01294694-6044238 | 2MASS |
| | UCAC4 147-001388 | UCAC 4 |
| RA (J2015.5) | 01h29m46.99s | GAIA |
| DEC (J2015.5) | -60d44m23.67s | GAIA |
| pm ^{RA} (mas yr ⁻¹) | 18.053 ± 0.036 | GAIA |
| pm ^{DEC} (mas yr ⁻¹) | 10.523 ± 0.034 | GAIA |
| π (mas) | 3.23 ± 0.02 | GAIA |
| T (mag) | 11.486 ± 0.006 | TESS |
| B (mag) | 12.746 ± 0.015 | APASS |
| V (mag) | 12.144 ± 0.069 | APASS |
| J (mag) | 10.858 ± 0.023 | 2MASS |
| H (mag) | 10.571 ± 0.024 | 2MASS |
| K _s (mag) | 10.485 ± 0.021 | 2MASS |
| WISE1 (mag) | 10.463 ± 0.023 | WISE |
| WISE2 (mag) | 10.518 ± 0.021 | WISE |
| WISE3 (mag) | 10.408 ± 0.059 | WISE |
| T_{eff} (K) | 5788 ± 80 | this work |
| log g (dex) | 4.376 ± 0.021 | this work |
| [Fe/H] (dex) | +0.14 ± 0.05 | this work |
| $v \sin i$ (km s ⁻¹) | 2.18 ± 0.41 | this work |
| M_{\star} (M_{\odot}) | 1.026 ^{+0.057} _{-0.055} | this work |
| R_{\star} (R_{\odot}) | 1.088 ^{+0.012} _{-0.012} | this work |
| L_{\star} (L_{\odot}) | 1.196 ± 0.050 | this work |
| Age (Gyr) | 5.7 ± 1.7 | this work |
| A_V (mag) | 0.117 ^{+0.068} _{-0.063} | this work |
| ρ_{\star} (g cm ⁻³) | 1.12 ± 0.11 | this work |

- **Limb darkening coefficients:** to ensure uniform sampling of only physical solutions, we used a triangular sampling scheme. As outlined in [Kipping \(2013\)](#), we transformed the quadratic limb darkening coefficients u_1, u_2 to $q_1 = (u_1 + u_2)^2$ and $q_2 = 0.5u_1(u_1 + u_2)^{-1}$. For ground-based photometry, we assumed a linear limb darkening profile and $q_1 = u_1$.
- **Prior for the stellar density ρ_{\star} :** from our stellar modeling (Sect. 5.3.1), we obtained a distribution for the stellar density ρ_{\star} which we used as a prior for our joint fit instead of the scaled semi-major axis of the planetary orbit.
- **Eccentricity and argument of periastron:** we parameterized the orbital eccentricity

5 Discovery of A Highly Eccentric Warm Jupiter

e and the argument of periastron ω as $\mathcal{S}_1 = \sqrt{e} \sin \omega$ and $\mathcal{S}_2 = \sqrt{e} \cos \omega$ and ensure at each iteration that $e = \mathcal{S}_1^2 + \mathcal{S}_2^2 \leq 1$.

Given the observed RV variations and empirical mass-radius relationships (e.g., [Chen & Kipping, 2016](#); [Neil & Rogers, 2020](#)), it is justified to neglect extreme radius ratios. We thus constrained the sampling to $R_P/R_\star < 0.5$.

Limb darkening

The limb darkening profile of TIC 237913194 is poorly constrained by our available data; we therefore simultaneously fit for it in the joint fit. An optimal choice of a limb darkening law is not straight-forward: there is a trade-off between accuracy and computational cost, and the performances of different laws depend on the noise level of the light curve (see [Schlecker \(2016\)](#) for a more detailed discussion). To account for the different noise levels in space-based and ground-based photometry ([Espinoza & Jordán, 2016](#)), we decided to use a quadratic limb darkening law for *TESS* photometry and a linear law for the *CHAT* and *LCOGT* light curves.

RV analysis

The RV time series show a strong signal with a period corresponding to the candidate transiting planet ($P = 15.17$ d, see [Fig. 5.2](#), [Fig. 5.3](#)). To assess the evidence of this signal being of planetary origin, we compared models with and without a planet based on only the FEROS RV dataset. We further evaluated models including more than one planet and compared the log-evidences of the different cases:

1. **No planet:** we assumed that all RV variations are due to astronomical and instrumental “jitter”. The only free parameters were μ_{FEROS} and σ_{FEROS} using the same prior distributions as in [Table 5.4](#). This “flat” model resulted in a log-evidence $\ln(Z) = -161$.
2. **Single planet:** we assumed there is a planetary signal in the RV data and widened the orbital period prior to a uniform distribution of 1 d – 30 d. The RV semi-amplitude K was free to vary between zero and 1000 m s^{-1} (uniform prior). For T_0 , we chose a uniform prior ranging from the first photometric observation to 124 d later, which corresponds to half the RV baseline. All other free parameters had the same priors as in our nominal model. This fit converged to a similar solution as our final model with a period distribution consistent with the intervals between the observed transit events. $\ln(Z) = -135$.

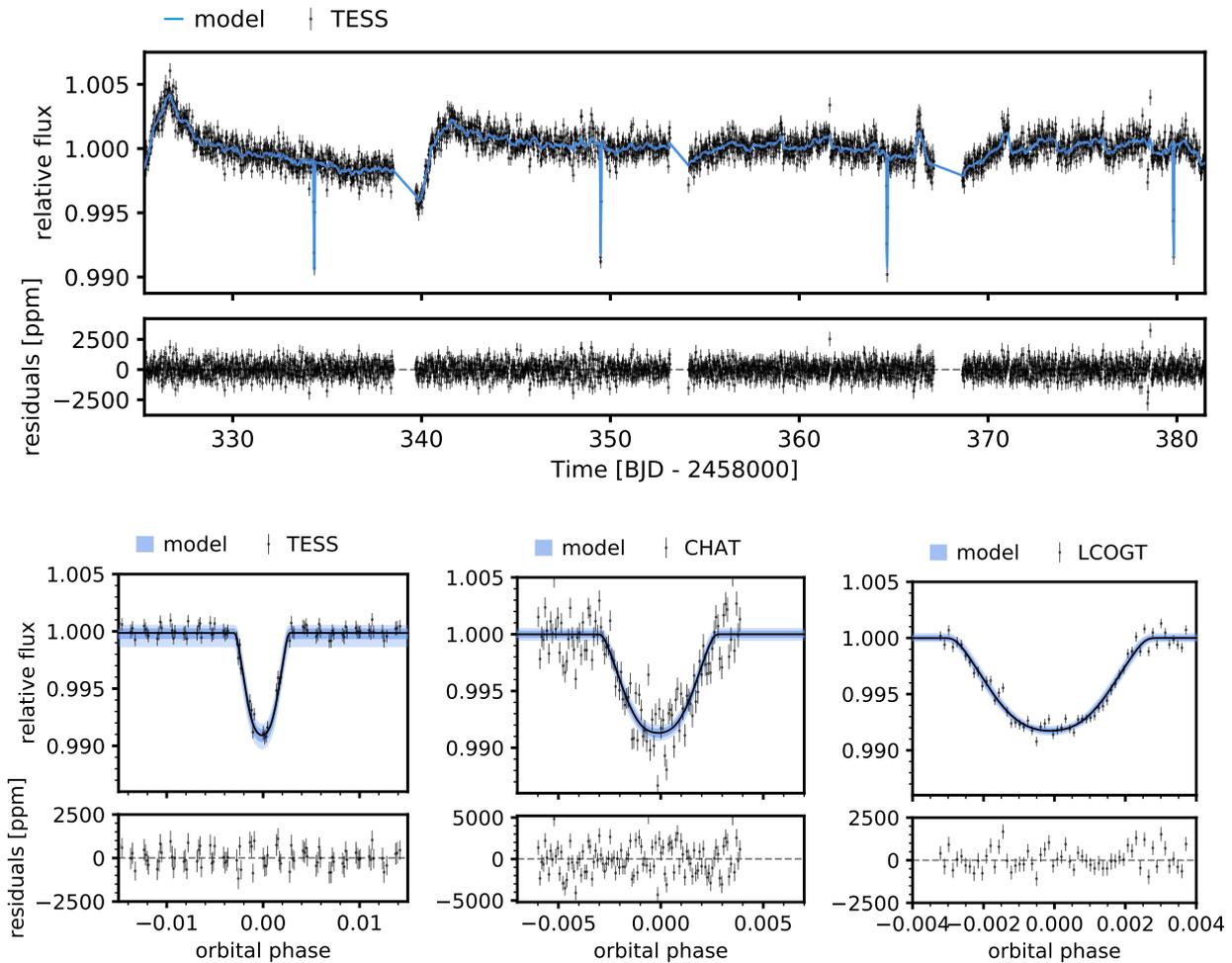


Figure 5.1: Photometry for TIC 237913194. Gray points represent the relative flux and errors. Solid lines show the theoretical light curve using the best-fit parameters derived in the joint modeling including GP. Blue shaded regions denote the 68 % and 95 % credibility bands of the model. Residuals are shown below each light curve. *Top:* Full TESS light curve generated from 30-minute-cadence photometry of Sectors 1 and 2. *Bottom left:* Phase-folded TESS photometry around the transit events. *Bottom center:* Follow-up photometry of a single transit obtained with CHAT in the i' band. *Bottom right:* LCOGT photometry of a single transit (i' band). This additional transit photometry lifted the $R_P/R_\star - b$ degeneracy and strongly improved our constraint on the planet radius.

5 Discovery of A Highly Eccentric Warm Jupiter

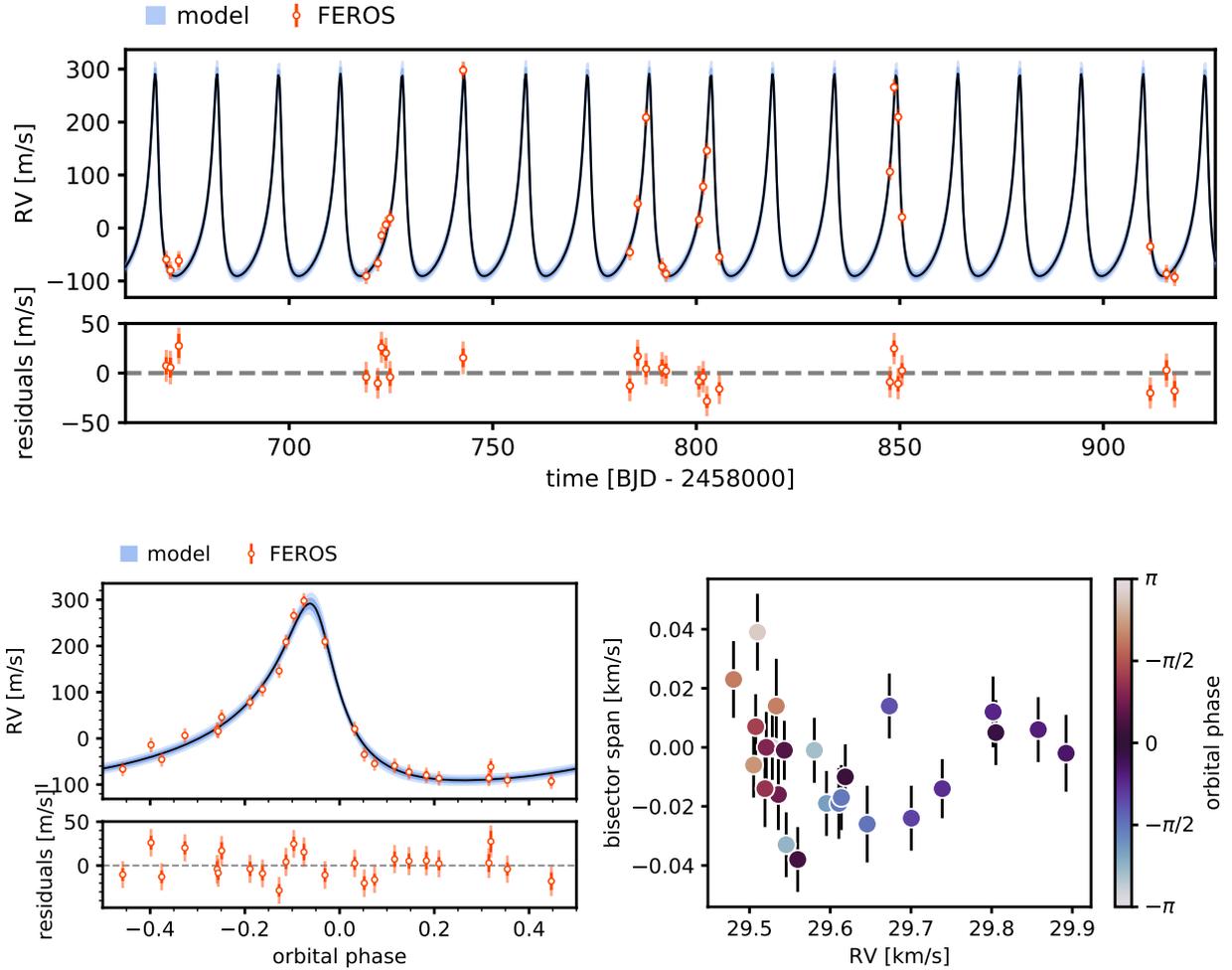


Figure 5.2: Radial velocity time series for TIC 237913194. Light error bars reveal the best-fit jitter term, which we added in quadrature to the RV errors. The model generated with the derived parameters of our joint modeling is plotted with a black line, and blue bands denote its 68% and 95% posterior credibility intervals. Residuals are obtained by subtracting the median posterior model from the data. *Top:* RV time series measured with the FEROS spectrograph. *Left:* phase-folded RV measurements obtained with FEROS. *Right:* bisector span as a function of radial velocity. The color of each measurement represents the orbital phase at which it was taken, assuming our best-fit period.

Table 5.3: Model evidences from RV fits for different models

| Model | log-evidence $\ln Z$ | $\Delta \ln Z$ |
|--------------------------------------|----------------------|----------------|
| 0 planets | -160.94 ± 0.14 | -26.41 |
| 1 planet, eccentric orbit | -134.53 ± 0.02 | 0 |
| 1 planet, circular orbits | -160.46 ± 0.04 | -25.93 |
| 2 planets, circular orbits | -164.34 ± 0.02 | -29.81 |
| 2 planets, circular&eccentric orbits | -150.76 ± 0.01 | -16.23 |
| 2 planets, eccentric orbits | -144.46 ± 0.01 | -9.93 |

Notes. $\Delta \ln Z$ states the difference in log-evidence compared to our best model “1 planet, eccentric orbit”.

3. **Single planet, circular orbit:** same as 2., but fixing the eccentricity to zero (and ω to an arbitrary 90°). The fit converged to a solution with P similar to the period distribution in our final joint fit, but the jitter term is strongly increased to account for the large RV variations. With $\ln(Z) = -160$, the evidence of this model is similar to the one belonging to the no-planet hypothesis.
4. **Two planets, circular orbits:** same as 3., but assuming a second planet in the system. For this hypothetical additional planet, we let the orbital period vary within $1 \text{ d} - 30 \text{ d}$ and used the same uniform prior $\mathcal{U}(0, 1000) \text{ m s}^{-1}$ for its RV semi-amplitude K_2 . The $\sim 15 \text{ d}$ periodicity is recovered, but no stable solution in favor of a two-planet-scenario is evident. $\ln(Z) = -164$
5. **Two planets, circular and eccentric orbits:** same as 4., but one planet with freely varying eccentricity. The eccentric, 15 d candidate signal is recovered. The period and RV semi-amplitude of the second planet are poorly constrained. $\ln(Z) = -151$
6. **Two planets, both on eccentric orbits:** same as 4., but with free eccentricity for both planets. Again, the 15 d signal is strongly recovered, while the weak signal of an additional planet is poorly constrained. $\ln(Z) = -144$

We list all model evidences in Table 5.3. The log-evidence difference between the preferred model (2. Single planet) and the runner-up (6. Two planets, both on eccentric orbits) $\Delta \ln Z_{6,2} \approx 9$, which corresponds to a Bayes factor of $\sim 10^4$. The difference to the flat model is as large as $\Delta \ln Z_{1,2} \approx 26$, implying a Bayes factor of $\sim 10^{11}$. Thus, the planetary model is strongly favored above the flat model and an eccentric single-planet solution is preferred.

To test if the candidate signal could potentially be associated with stellar activity, we produced generalized Lomb-Scargle periodograms (Zechmeister & Kürster, 2009) for the

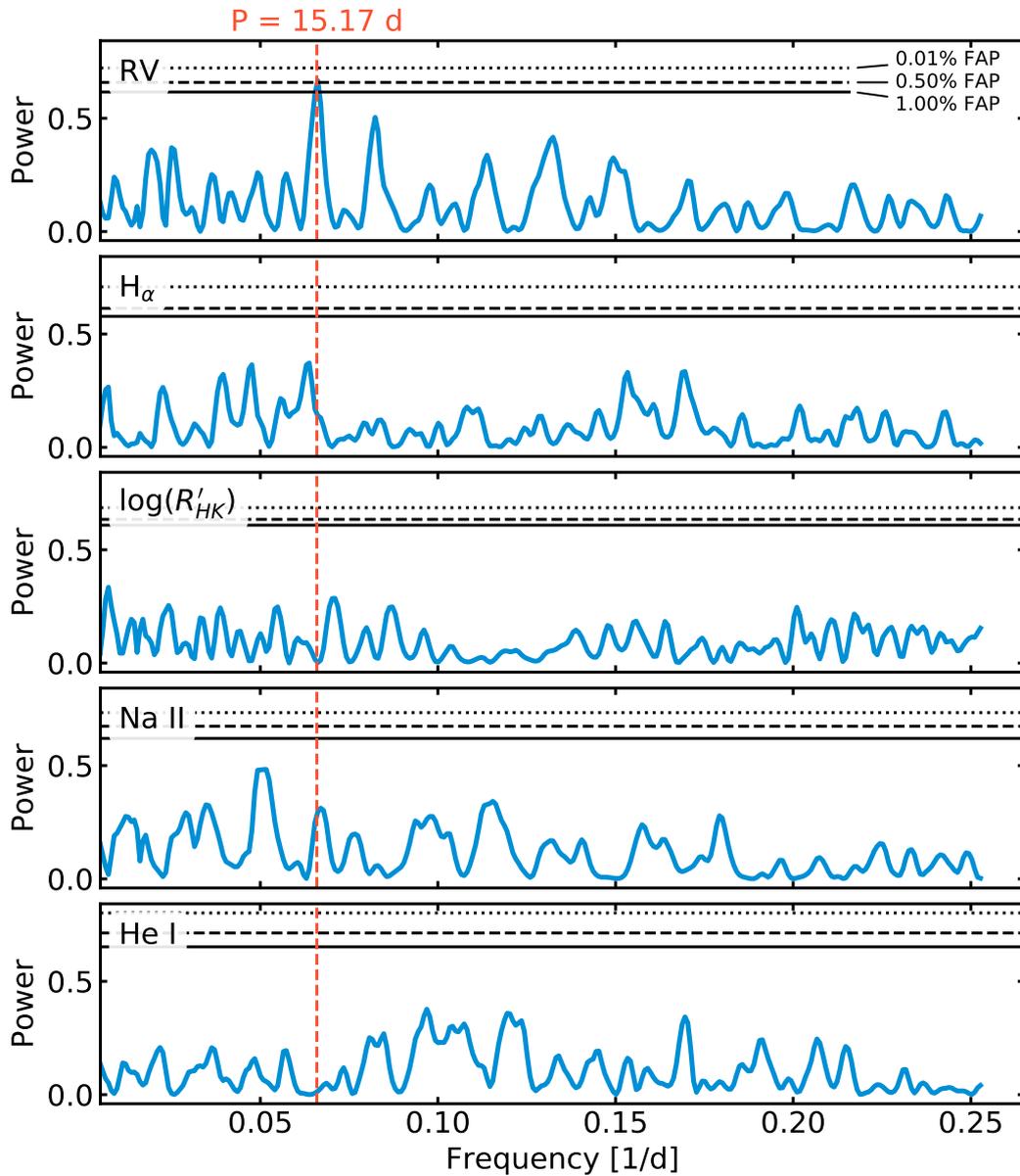


Figure 5.3: Generalized Lomb-Scargle periodograms of radial velocity time series and common activity indicators. Solid, dashed, and dotted lines mark 1%, 0.5%, and 0.01% false alarm probabilities, which we computed via bootstrap resamplings. The orange line marks the orbital period of TIC 237913194b.

radial velocity time series, as well as for common activity indicators based on FEROS data (Fig. 5.3). In particular, we obtained the H_α , $\log(R'_{HK})$, Na II, and He I activity indices, which trace chromospheric activity. We computed H_α following Boisse et al. (2009). As TIC 237913194 is a G-type star, we used the regions defined by Duncan et al. (1991) and the calibrations of Noyes et al. (1984) for $\log(R'_{HK})$. For Na II and He I we followed Gomes da Silva et al. (2011). For each time series (RV, H_α , $\log(R'_{HK})$, Na II, and He I), we computed the power levels for 1 %, 0.5 %, and 0.01 % false alarm probabilities (FAP) by a bootstrap method and plot them as solid, dashed, and dotted lines, respectively. There is a strong signal in the periodogram of the radial velocities at the 15.17d period, below 0.5 % FAP. Meanwhile, there are no significant signals visible in the periodograms of any of the activity indices, indicating that the radial velocity signal is unlikely to come from quasi-periodic stellar activity.

To further show beyond doubt that the measured radial velocity variations represent reflex motions of the star, we also tested if they can be caused by variations in the stellar photosphere. A well-established method to do this is the inspection of atmospheric line profiles, which should be constant in time for actual stellar velocity changes. Specifically, the bisector span (BIS) can serve as a diagnostic to search for possible false positive scenarios (e.g., Queloz et al., 2001). We are interested in its correlation with the RV time series and orbital phase, and confront these variables in the right panel of Fig. 5.2. Here, we plot the FEROS RV measurements against their bisector spans. The points are color-coded by the orbital phase of the measurements, where zero phase is at $T_0 + nP$ using median values from our nominal fit. While bisector span and RV show no evidence for correlation, this may not be true for bisector span and orbital phase. A Spearman's rank coefficient of $0.45^{+0.31}_{-0.39}$, where we quote 95 % confidence intervals from a bootstrapped sampling method⁵, permits the suspicion of a positive correlation. However, due to the small number of data points we cannot reject the null hypothesis that there is no monotonic association between the two variables. In addition, the bisector span variations are on the order of 10 m s^{-1} and cannot account for the observed RV semi-amplitude of $K \approx 191 \text{ m s}^{-1}$. The line profiles thus provide further evidence that the observed RV variations are indeed due to velocity changes of the target star and not caused by atmospheric variations.

In addition to the tests described above, the non-sinusoidal pattern of the RVs is a strong indication for orbital motion as opposed to stellar activity. We conclude that the radial velocity time series independently confirms the planet hypothesis. In the following, we refer to the confirmed exoplanet as TIC 237913194b.⁶

⁵pingouin.compute_bootci from the Pingouin python package (Vallat, 2018)

⁶We submitted the object to exoFOP as a community TESS Object of Interest (CTOI); it is now listed as

Joint photometry and RV fit

We performed a simultaneous fit on photometric and spectroscopic datasets (*TESS*, *CHAT*, *LCOGT*, and *FEROS*) to jointly constrain all planetary parameters of TIC 237913194b. To account for the long cadence in the *TESS* light curve, we modeled the transits with 20-fold supersampling using the exposure time of the actual observations. The initial photometric fit (see Sect. 5.2.1) provided narrow constraints on the orbital period P and time of mid-transit T_0 . We used the median values obtained there to construct Gaussian priors for these parameters, but we enlarged the dispersions (compare Table 5.4). For the instrument-specific flux offsets M_{TESS} , M_{CHAT} , and M_{LCOGT} we assumed Gaussian priors based on our photometric fit. A Gaussian prior on the stellar density was motivated by the analysis of the stellar parameters presented in Sect. 5.3.1. Additional confidence for this prior stems from a separate joint fit where we used an uninformative prior ($\mathcal{J}(10^2, (10^4)^2)$) on ρ_* . The resulting posterior probability, $\rho_* = 996_{-421}^{+257}$, and the result from our stellar analysis agree within the uncertainties. For all other parameters, we chose uninformative priors to sample the whole physically plausible parameter space.

There are potentially time-correlated processes such as instrumental red noise, stellar variability, or blended sources that are not covered by our astrophysical model. We account for this red noise by adding a Gaussian Process (GP) component to the *TESS* photometry with an exponential kernel as implemented in the *celerite* software package (Foreman-Mackey et al., 2017). This adds two additional hyperparameters: an amplitude $\sigma_{\text{TESS}}^{\text{GP}}$ and a timescale $\tau_{\text{TESS}}^{\text{GP}}$. For comparison, we performed the same fit with and without a GP component. The variant including GP performed significantly better than the white-noise model and we thus consider it our nominal model.

In the same manner, we tested adding a GP component to the *LCOGT* photometry, which shows possible systematic effects in the residuals (compare Fig. 5.1). Here, we chose a Matérn 3/2 kernel, which again adds two hyperparameters for timescale and amplitude to the model. The model including the GP component consistently performed comparable ($|\Delta \ln Z| < 1$) or worse than the model without, which is why we chose to continue with the less complex noise model without GP.

Figures 5.1 and 5.2 show the photometric and radial velocity time series resulting from this model using its median parameters (solid lines). Dark (light) blue bands show the 68 % (95 %) credibility bands of the model. The residuals below the time series show the measured data with the median posterior model subtracted; both RV and photometry residuals appear inconspicuous. Figure 5.4 shows the posterior distributions for the

TOI 2179.01.

Table 5.4: Prior parameter distributions

| Parameter name | Prior | Units | Description |
|---|----------------------------------|-------------------|--|
| Stellar Parameters | | | |
| ρ_* | $\mathcal{N}(1120, 110^2)$ | kg/m ³ | Stellar density |
| Planetary parameters | | | |
| P | $\mathcal{N}(15.16, 0.2^2)$ | d | Period |
| t_0 | $\mathcal{N}(2458319.17, 0.2^2)$ | d | Time of transit center |
| R_P/R_* | $\mathcal{U}(0.0, 1.5)$ | — | Impact factor |
| $b = (a/R_*) \cos(i)$ | $\mathcal{U}(0.0, 0.5)$ | — | Planet-to-star ratio |
| K | $\mathcal{U}(140.0, 260.0)$ | m/s | Radial velocity semi-amplitude |
| $S_1 = \sqrt{e} \sin \omega$ | $\mathcal{U}(-1, 1)$ | — | Parametrization for e and ω |
| $S_2 = \sqrt{e} \cos \omega$ | $\mathcal{U}(-1, 1)$ | — | Parametrization for e and ω |
| RV instrumental parameters | | | |
| μ_{FEROS} | $\mathcal{U}(-30, 30)$ | m/s | Systemic velocity for FEROS |
| σ_{FEROS} | $\mathcal{J}(1.0, 100.0^2)$ | ppm | Extra jitter term for FEROS |
| $\text{RV}_{\text{linear}}$ | $\mathcal{N}(0.0, 1.0^2)$ | m/s/d | Linear term for the RVs ^a |
| $\text{RV}_{\text{intercept}}$ | $\mathcal{N}(0.0, 10000^2)$ | m/s | Intercept term for the RVs ^a |
| Photometry instrumental parameters | | | |
| D_{TESS} | 1.0 (fixed) | — | Dilution factor for <i>TESS</i> |
| M_{TESS} | $\mathcal{N}(0.0, 0.1^2)$ | ppm | Relative flux offset for <i>TESS</i> |
| σ_{TESS} | $\mathcal{J}(10^{-5}, (10^5)^2)$ | ppm | Extra jitter term for <i>TESS</i> |
| $q_{1,\text{TESS}}$ | $\mathcal{U}(0.0, 1.0)$ | — | Linear limb-darkening parametrization |
| $q_{2,\text{TESS}}$ | $\mathcal{U}(0.0, 1.0)$ | — | Quadratic limb-darkening parametrization |
| D_{CHAT} | 1.0 (fixed) | — | Dilution factor for <i>CHAT</i> |
| M_{CHAT} | $\mathcal{N}(0.0, 0.1^2)$ | ppm | Relative flux offset for <i>CHAT</i> |
| σ_{CHAT} | $\mathcal{J}(10^{-5}, (10^5)^2)$ | ppm | Extra jitter term for <i>CHAT</i> |
| $q_{1,\text{CHAT}}$ | $\mathcal{U}(0.0, 1.0)$ | — | Linear limb-darkening parametrization |
| D_{LCOGT} | 1.0 (fixed) | — | Dilution factor for <i>LCOGT</i> |
| M_{LCOGT} | $\mathcal{N}(0.0, 0.1^2)$ | ppm | Relative flux offset for <i>LCOGT</i> |
| σ_{LCOGT} | $\mathcal{J}(10^{-5}, (10^5)^2)$ | ppm | Extra jitter term for <i>LCOGT</i> |
| $q_{1,\text{LCOGT}}$ | $\mathcal{U}(0.0, 1.0)$ | — | Linear limb-darkening parametrization |
| Additional parameters | | | |
| $\sigma_{\text{TESS}}^{\text{GP}}$ | $\mathcal{J}(10^{-8}, 0.0005^2)$ | — | Amplitude of the GP component |
| $\tau_{\text{TESS}}^{\text{GP}}$ | $\mathcal{J}(0.0001, 2^2)$ | — | Timescale of the GP component |

Notes. $N(\mu, \sigma)$ stands for a normal distribution with mean μ and standard deviation σ , $U(a, b)$ stands for a uniform distribution between a and b , and $J(a, b)$ stands for a Jeffrey's prior (that is, a log-uniform distribution) defined between a and b .

^(a) These parameters were only used to search for an additional linear RV trend and are not included in our nominal joint fit.

5 Discovery of A Highly Eccentric Warm Jupiter

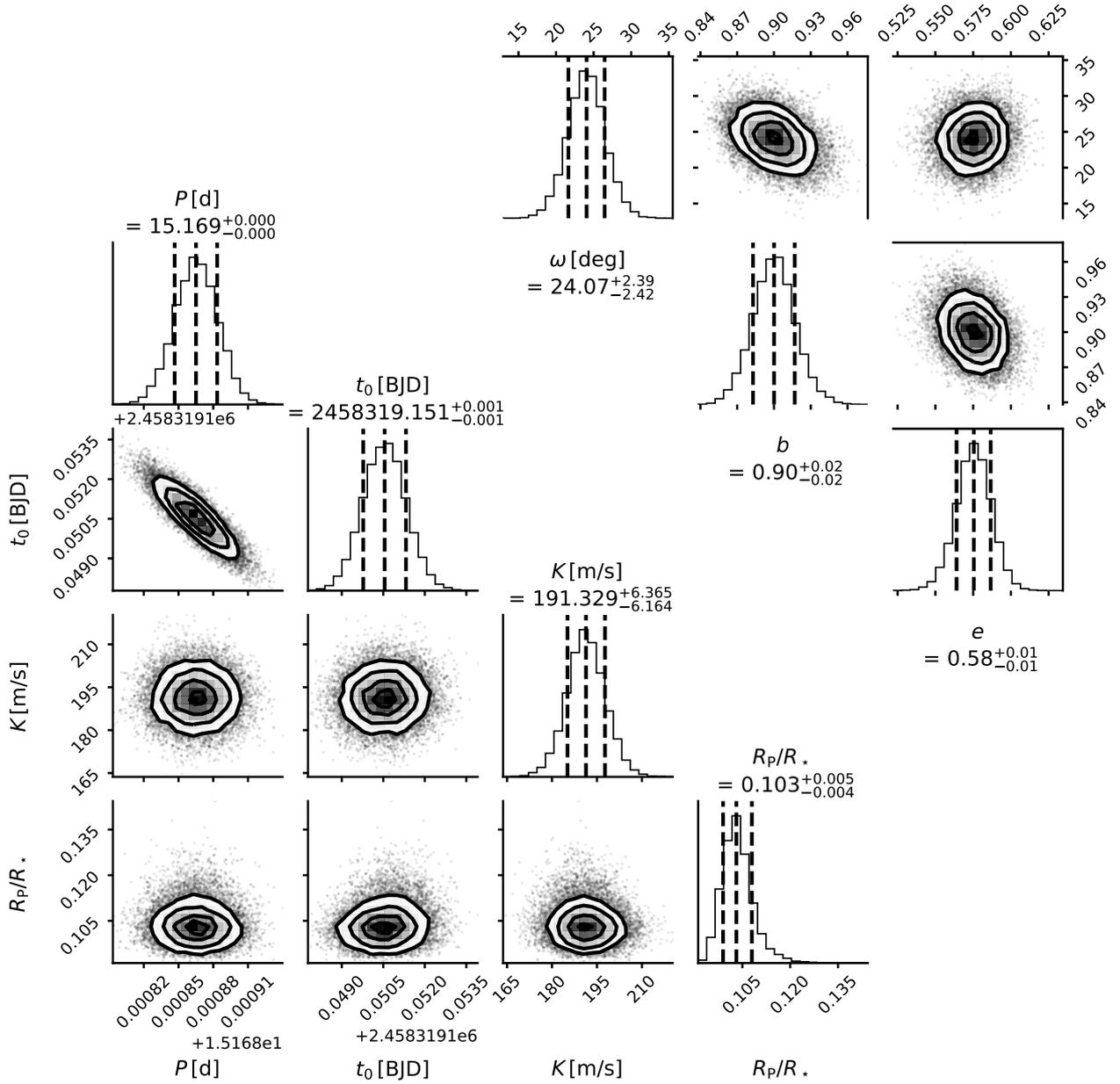


Figure 5.4: Posterior distributions of planetary parameters. The lower left triangle shows fitting parameters and the upper right triangle shows derived parameters of TIC 237913194b’s orbit. The stated values represent 16th, 50th, and 84th percentiles. See Appendix B.1 for the full sample and a discussion about correlated parameters.

planet’s main parameters as sampled in our nominal fit. All distributions are approximately Gaussian and barely correlated, except for the planet-to-star ratio R_P/R_* (see Appendix B.1 for a discussion). We present the posterior distributions of the model parameters alongside the derived physical parameters in Table 5.5, where we state for each parameter distribution the 16th, 50th, and 84th percentile. Notably, with a planetary bulk density $\rho_P \approx 1847 \text{ kg m}^{-3}$, TIC 237913194b’s average density is comparable to Neptune’s. By sheer coincidence, the planet’s period and eccentricity resemble that of the *TESS* spacecraft (13.7 d, $e = 0.55$) (Ricker et al., 2014).

Search for additional planets

We repeated the joint fit with an additional linear RV term to search for any long-period companions that would locally cause a linear trend in the RVs. To this end, we include intercept and slope parameters with wide, normal priors for another joint fit. The result is consistent with an RV slope of zero and the log-likelihood of the model including the linear trend is suppressed with $\Delta \ln Z \approx 7$. We conclude that the data at hand does not support additional outer companions in the system.

There is also no evidence of interior planets, which is expected since the deep intrusion of TIC 237913194b into the inner system leaves only limited room for stable inner orbits (e.g., Gladman, 1993). In fact, planets like TIC 237913194b have been suggested to be a main cause for the destruction of inner systems of low-mass planets (Schlecker et al., 2020a).

5.3.3 Approximation of the planetary equilibrium temperature

The equilibrium temperature T_{eq} that the planet maintains if it is in energy balance with the radiation input from the host star is a determining factor for the physical properties of its atmosphere. Due to TIC 237913194b’s considerable eccentricity, this input is not constant over its orbit. To give a first-order estimate on the temperature range that the planet can assume, we investigated two extreme cases of planetary heat adjustment:

1. instantaneous heat adjustment ($T_{\text{eq,inst}}$). Here, we assumed that the planetary atmosphere adjusts to the changing irradiation without any time delay. For this situation, we used the approximation in Kaltenegger & Sasselov (2011, Equation 3).
2. orbitally averaged heat adjustment ($T_{\text{eq,avg}}$). In this case, the planetary temperature remains in equilibrium with the incoming stellar energy, i.e. $T_{\text{eq}} = \text{const.}$ over one orbit. To approximate this temperature, we used a temporal average for elliptic orbits (Méndez & Rivera-Valentín, 2017, Equation 16).

Table 5.5: Posterior parameters

| Parameter | Value |
|---|--|
| Stellar Parameters | |
| ρ_* (kg/m ³) | 1076 ⁺⁹⁵ ₋₉₃ |
| Planetary parameters | |
| P (d) | 15.168865 ^{+0.000018} _{-0.000018} |
| t_0 (BJD UTC) | 2458319.15055 ^{+0.00077} _{-0.00077} |
| a/R_* | 23.85 ^{+0.67} _{-0.69} |
| $b = (a/R_*) \cos(i)$ | 0.900 ^{+0.017} _{-0.017} |
| R_P/R_* | 0.1031 ^{+0.0048} _{-0.0042} |
| K (m/s) | 191.3 ^{+6.4} _{-6.2} |
| e | 0.575 ^{+0.011} _{-0.011} |
| ω | 24.1 ^{+2.4} _{-2.4} |
| $S_1 = \sqrt{e} \sin \omega$ | 0.309 ^{+0.029} _{-0.030} |
| $S_2 = \sqrt{e} \cos \omega$ | 0.692 ^{+0.014} _{-0.015} |
| RV instrumental parameters | |
| μ_{FEROS} (m/s) | 14.0 ^{+3.4} _{-3.3} |
| σ_{FEROS} (m/s) | 13.6 ^{+3.2} _{-2.6} |
| Photometry instrumental parameters | |
| M_{TESS} (ppm) | 0.00017 ^{+0.00041} _{-0.00039} |
| σ_{TESS} (ppm) | 221 ⁺²⁸ ₋₃₁ |
| $q_{1,\text{TESS}}$ | 0.33 ^{+0.39} _{-0.24} |
| $q_{2,\text{TESS}}$ | 0.27 ^{+0.35} _{-0.20} |
| $\sigma_{\text{TESS}}^{GP}$ | 0.00000152 ^{+0.00000110} _{-0.00000041} |
| τ_{TESS}^{GP} | 0.40 ^{+0.16} _{-0.18} |
| M_{CHAT} (ppm) | 0.00149 ^{+0.00025} _{-0.00025} |
| σ_{CHAT} (ppm) | 1625 ⁺¹⁵⁰ ₋₁₄₀ |
| $q_{1,\text{CHAT}}$ | 0.55 ^{+0.20} _{-0.24} |
| M_{LCOGT} (ppm) | -0.00038 ^{+0.00016} _{-0.00016} |
| σ_{LCOGT} (ppm) | 480 ⁺⁸⁵ ₋₈₀ |
| $q_{1,\text{LCOGT}}$ | 0.67 ^{+0.17} _{-0.20} |
| Derived parameters | |
| i (deg) | 87.0 ^{+1.5} _{-1.7} |
| R_P [R _{Jup}] | 1.117 ^{+0.054} _{-0.047} |
| M_P [M _{Jup}] | 1.942 ^{+0.092} _{-0.091} |
| a [au] | 0.1207 ^{+0.0037} _{-0.0037} |
| ρ_P [kg m ⁻³] | 1847 ⁺²⁸⁰ ₋₂₆₀ |
| T_{eq} [K] ^a | 974 |

Notes. ^(a) Time-averaged equilibrium temperature computed according to equation 16 of Méndez & Rivera-Valentín (2017). We assumed zero albedo, a unity broadband thermal emissivity, and $\beta = 0.5$, i.e. only half of the planetary surface re-radiates the absorbed flux.

For both extremes, we assumed that the heat flux from the planet’s interior is negligible compared to the stellar irradiation and ignored any internal energy sources. The infrared emissivity ε_{IR} was fixed to unity. We further assumed two cases for how atmospheric flow distributes the incoming stellar energy over the planetary surface and parametrized this property with the fraction of planetary surface that re-radiates stellar flux β . We distinguished between $\beta = 0.5$, i.e. emission only from one hot hemisphere, and $\beta = 1$ where the whole globe emits (Seager et al., 2005; Kaltenegger & Sasselov, 2011; Carone et al., 2014; Méndez & Rivera-Valentín, 2017).

Planets colder than 1000 K are expected to be relatively cloudy (e.g., Stevenson, 2016; Parmentier et al., 2016). Here, we parametrized different cloudiness with albedos $\alpha = 0, 0.3, \text{ and } 0.6$ following Parmentier et al. (2016). With the above assumptions and in the case of instantaneous heat adjustment, we derived a range of $T_{\text{eq,inst}} \approx 900 \text{ K} - 1300 \text{ K}$ at secondary eclipse and $T_{\text{eq,inst}} \approx 700 \text{ K} - 1100 \text{ K}$ at transit (Fig. 5.5). We list $T_{\text{eq,inst}}$ at critical times in Table 5.6 together with the values for orbitally averaged heat adjustment $T_{\text{eq,avg}}$. The latter is constant over one orbit and covers a temperature range of $T_{\text{eq,avg}} \approx 650 \text{ K} - 975 \text{ K}$.

Due to the orientation of the orbit (compare Fig. 5.6), the temperature $T_{\text{eq,inst}}$ during transit is assumed to be about 200 K colder compared to the temperature at secondary eclipse. In reality, however, some delay in heat adjustment based on radiative and dynamical timescales is expected. Therefore, the temperature during a secondary eclipse, which would occur before passage of periastron, could be colder than in our estimate. Likewise, the temperature at transit, occurring after periastron passage, would be warmer than expected (see, e.g., Lewis et al., 2013, for a qualitative discussion of the thermal evolution of an exoplanet on an eccentric orbit). We emphasize again that our goal is to estimate to first order possible temperature ranges for TIC 237913194b for which these simplified assumptions are sufficient.

5.3.4 Secondary eclipses, phase curve modulations, and Rossiter-MacLaughlin effect

No secondary eclipses or phase curve signals are evident in the photometric time series. We used the starry software package (Luger et al., 2019) to estimate the planetary phase curve based on our derived orbital parameters and a simple toy model for the planetary brightness distribution. This model neglects any heat redistribution between the hot and cold hemispheres of the planet, resulting in a ‘dipole’ brightness map where the hot side points to the substellar point at periastron. In this scenario, the emission of the cold hemisphere and planetary limb darkening are negligible. The total luminosity of the planet

5 Discovery of A Highly Eccentric Warm Jupiter

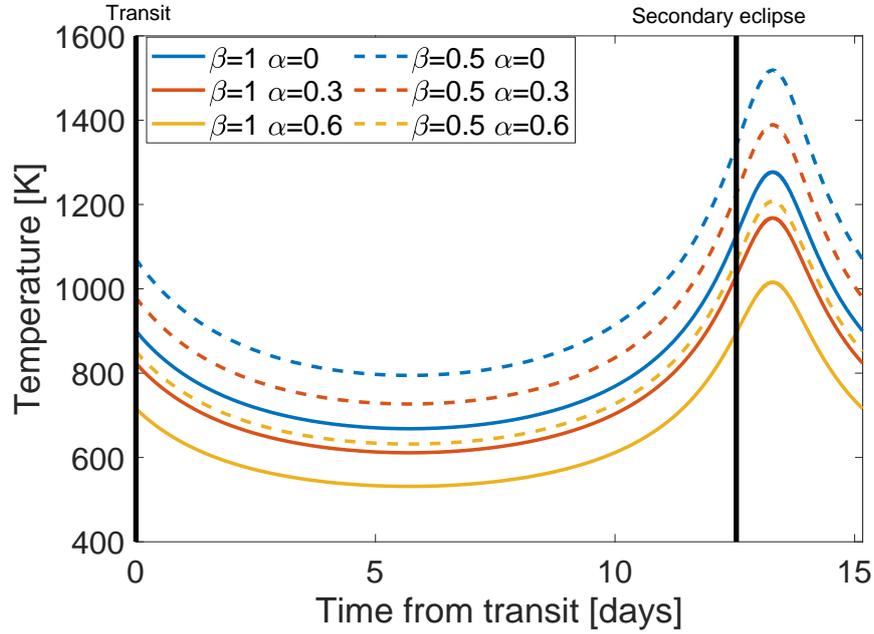


Figure 5.5: Evolution of the planetary equilibrium temperature in the case of instantaneous heat adjustment for different albedos α and re-radiation factors β . We assumed unity infrared emissivity ε_{IR} . Black lines denote the time of transit and secondary eclipse, respectively. Due to the high eccentricity, $T_{\text{eq,inst}}$ varies by several hundred Kelvin within one orbit. It stays below 1000 K for most of the orbit.

Table 5.6: Theoretical temperature constraints of TIC 237913194b in the course of one orbit.

| Time | orbital distance [au] | $T_{\text{eq,inst}}$ [K] | | | $T_{\text{eq,avg}}$ [K] | | |
|-------------------|--------------------------|--------------------------|----------------|----------------|-------------------------|----------------|----------------|
| | | $\alpha = 0$ | $\alpha = 0.3$ | $\alpha = 0.6$ | $\alpha = 0$ | $\alpha = 0.3$ | $\alpha = 0.6$ |
| $\beta = 1$ | | | | | | | |
| apoastron | 0.1900 | 668 | 611 | 531 | | | |
| transit | 0.1047 | 900 | 823 | 716 | | | |
| periastron | 0.0520 | 1277 | 1168 | 1016 | 819 | 749 | 651 |
| secondary eclipse | 0.0668 | 1127 | 1031 | 896 | | | |
| $\beta = 0.5$ | | | | | | | |
| apoastron | 0.1900 | 795 | 727 | 632 | | | |
| transit | 0.1047 | 1070 | 979 | 851 | | | |
| periastron | 0.0520 | 1519 | 1389 | 1208 | 974 | 891 | 775 |
| secondary eclipse | 0.0668 | 1340 | 1226 | 1066 | | | |

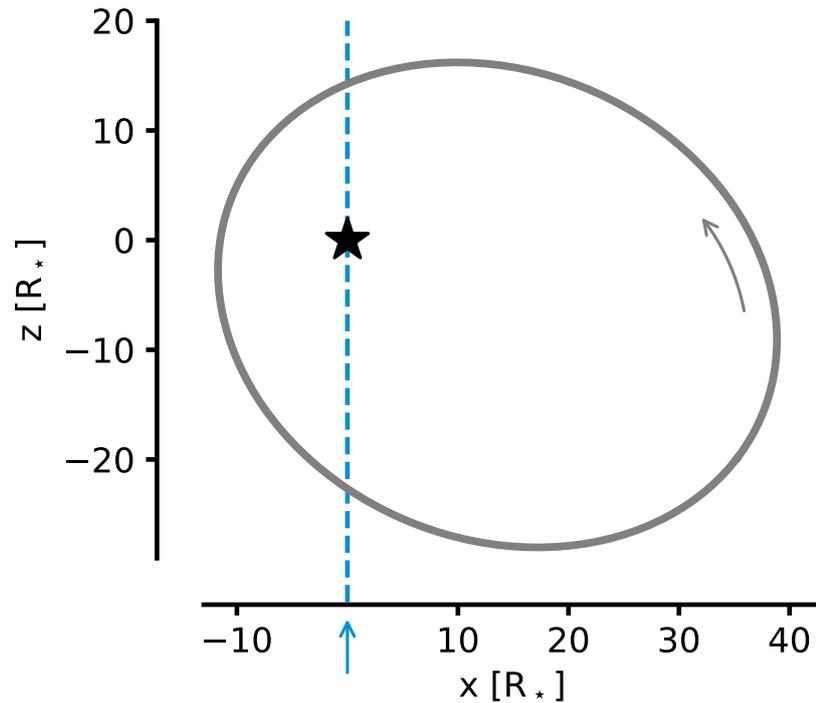


Figure 5.6: Orbit aspect ratio and orientation. The orbit of TIC 237913194b is plotted with stellar radii as length units. The dashed blue line shows our line of sight with respect to the orbit.

is then that of a half-sphere black body with radius R_P and temperature T_{hot} . For the temperature of the hot hemisphere, we adopt two cases: firstly, our estimate of the equilibrium temperature for the case of orbitally averaged heat adjustment and $\beta = 0.5$, hence $T_{\text{hot}} = T_{\text{eq,avg}} = 974$ K. Secondly, we consider the hottest temperature in Table 5.6 and assume $T_{\text{hot}} = 1519$ K. With a resulting peak-to-peak phase curve amplitude of ~ 6 ppm in the cool case and ~ 32 ppm in the hot case, a future detection of the phase curve or secondary eclipse might be challenging despite the expected precision of the *JWST*.

Measuring the Rossiter-McLaughlin effect (RM effect, [Rossiter, 1924](#); [McLaughlin, 1924](#)) has proved a useful tool to measure the alignment of planetary orbits with the spin axis of host stars. The different proposed scenarios for the formation and migration theory of warm Jupiters differ in their predicted impact on the spin-orbit alignment (e.g., [Triaud, 2018](#)). A detection of the RM effect could thus shed light on TIC 237913194b’s enigmatic formation history. Analytical formulas exist to estimate the amplitude of its RV signature (e.g., [Gaudi & Winn, 2007](#), equation 6), however, the large impact parameter in TIC 237913194b’s transit geometry would entail a large error. Instead, we modeled the RV anomaly due to the Rossiter-McLaughlin effect with a velocity-weighted brightness map in *starry* using the median posterior values of the system’s physical parameters (see Ta-

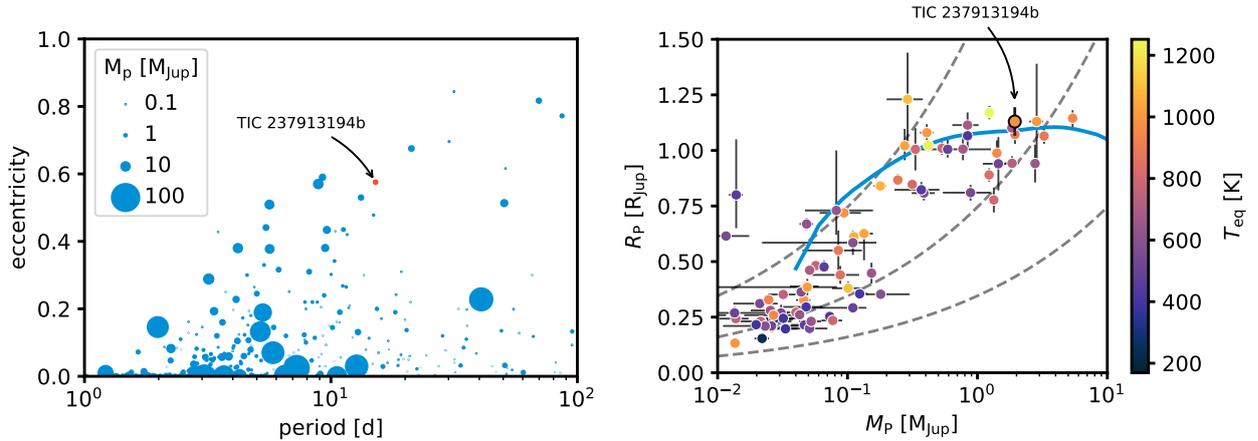


Figure 5.7: Comparison to other well-characterized warm Jupiters. *Left:* period-eccentricity plot of transiting exoplanets with periods of 1 d – 100 d and measured eccentricity from the TEPcat catalog (Southworth, 2011). Marker sizes scale with planet mass. With $e = 0.58$, TIC 237913194b occupies the 98th percentile in this population and contributes to a sparse sample of planets with very high eccentricities.

Right: mass-radius diagram of *warm* ($P = 10$ d – 100 d) planets from the same catalog. The color of the markers represent the equilibrium temperatures of the planets, and dashed gray lines are isodensity curves of 0.3, 3, and 30 g cm⁻³, respectively. The solid blue line marks the predicted mass-radius relation for giant planets with a 10 M_⊕ core (Fortney et al., 2007). TIC 237913194b lies very close to this line. The error bars for its mass are too small to be seen.

bles 5.2 and 5.5).

The resulting amplitude of the signal $K_{\text{RM}} \approx 10 \text{ m s}^{-1}$, which is just in the range of current state-of-the-art spectroscopic facilities.

5.4 Discussion

5.4.1 TIC 237913194b’s place in the exoplanet population

In Fig. 5.7 we compare TIC 237913194b with well-studied transiting exoplanets (Southworth, 2011)⁷. The left panel shows the periods and eccentricities of these planets (blue markers); our discovery is marked in red. We included planets with both mass and radius measurements that have constrained eccentricities (not only upper limits) and show those in the period range 1 d – 100 d. Marker sizes in the plot correspond to planet masses. Warm Jupiters with high eccentricities appear to be rare: 98% of this population have smaller eccentricities than TIC 237913194b, making it one of the most eccentric planets

⁷TEPCat catalog, queried on 2020-08-11.

in this period range. It lies at the edge of a demographic feature that we discuss in the following.

On close orbits, the planet occurrence rate dn/de shows a rapid drop with increasing eccentricity. With increasing period, the position of this ridge shifts to larger eccentricities. Through this, a triangular under-density of planets with high eccentricity on very short orbits emerges (upper left corner in Fig. 5.7). While exoplanet detection sensitivities are expected to have a dependency on eccentricity, the effect is too small to account for the observed dearth of planets (Burke, 2008). A plausible physical explanation would be tidal circularization (Adams & Laughlin, 2006; Dawson & Johnson, 2018). As discussed in Sect. 5.4.3, the strength of this mechanism is expected to scale inversely with orbital distance, which would explain the period-dependence of the distribution. The detection of planets close to this ridge can thus help constrain theories of tidal interaction and giant planet migration, which are crucial components for explaining planetary systems with close-in giant planets. Our discovery of TIC 237913194b adds to the small current sample of such planets.

In the right panel of Fig. 5.7, we put our planet into context of warm Jupiters with mass and radius measurements. Here, we include only planets from TEPcat with periods of 10 d – 100 d, and color-code them by equilibrium temperature. We further plot a theoretically predicted mass-radius relation for planets with a $10 M_{\oplus}$ core (Fortney et al., 2007, blue line). TIC 237913194b is located close to this curve, which indicates that its bulk density is consistent with established structural models.

5.4.2 The atmosphere of TIC 237913194b

The large eccentricity of TIC 237913194b makes it a promising test bed to study the response of its atmosphere to external forcing (e.g., compare Carone et al., 2020). Atmospheres at the inferred temperature ranges are susceptible to a variety of chemical disequilibrium processes such as photochemistry and chemical quenching (e.g. Molaverdikhani et al., 2019a; Moses et al., 2013; Venot et al., 2012, 2020; Tsai et al., 2018; Kawashima & Ikoma, 2019). While a thorough inspection of these processes is beyond the scope of this study, we used the physical parameters constrained here to demonstrate the feasibility of atmospheric characterization. To this end, we used self-consistent models for cloud-free (Molaverdikhani et al., 2019b) and cloudy (Molaverdikhani et al., 2020) irradiated planetary atmospheres and calculated synthetic transmission and emission spectra using petitCODE (Mollière et al., 2015; Mollière et al., 2017). For this fiducial model, we assumed solar composition, zero bond albedo, instantaneous thermal equilibrium, and $\beta = 0.5$. This resulted in an equilibrium temperature of the planet during transit and

secondary eclipse of 1070 K and 1340 K, respectively (see Table 5.6). Using the Pandexo package (Batalha et al., 2017), we predicted uncertainties for *JWST* observations of a single transit or secondary eclipse for the three observing modes MIRI-LRS, NIRISS-SOSS, and NIRSpec-G395M.

We find relative magnitudes of the largest synthetic spectral features of ~ 100 ppm in transmission and ~ 1000 ppm in emission. While some emission features are well above the predicted noise floor, the largest transmission features are on the order of the predicted uncertainties for a single transit observation. However, since these two techniques probe different regions of the atmosphere and at different orbital phases, a joint analysis of the transmission and emission spectra may provide important clues on atmospheric dynamics and heat redistribution.

5.4.3 TIC 237913194b’s large eccentricity

The peculiarly large eccentricity of TIC 237913194b could be an important lead in not only understanding the dynamical origin of the system, but also planet evolution theories in general. Possible origins of large warm Jupiter eccentricities include interaction with a massive companion through either scattering events (e.g., Rasio & Ford, 1996), secular interactions (e.g., Petrovich & Tremaine, 2016; Kozai, 1962; Lidov, 1962), or giant impacts (Freikh et al., 2019); planet-disk interactions (e.g., Lubow, 1991; Petrovich et al., 2019); or a combination of processes. The absence of evidence for an additional perturber that might sustainably excite TIC 237913194b’s eccentricity or that could be the counterpart in a recent scattering event makes it challenging to distinguish between these scenarios.

However, TIC 237913194b is subject to tidal dissipation through secular interaction with the host star (Goldreich & Soter, 1966) and the rates of semi-major axis and eccentricity decay, given the current orbital parameters, can be estimated (Yoder & Peale, 1981). A short orbit circularization timescale compared to the lifetime of the star would exclude the planet-disk interaction scenario and could provide an upper limit on the time that has passed since a hypothetical perturbation event. In this case, we would observe the system during high-eccentricity migration and TIC 237913194b would eventually become a hot Jupiter in a circular orbit.

Several caveats have to be considered when trying to trace the system back in time close to its primordial orbital configuration. First, the tidal evolution of a and e are strongly coupled, which may result in ambiguities. In addition, the tidal evolution strongly depends on the stellar and planetary tidal dissipation rates, typically parameterized by the dimensionless “reduced tidal quality factors” Q'_p and Q'_* . Here, $Q' = 1.5 Q/k_2$, where k_2 is the

Love number of second order. Estimates of the planetary tidal quality factor range from $Q'_p = 10^4$ to $Q'_p = 10^7$ (e.g. [Lainey et al., 2009](#); [Lainey et al., 2020](#); [Hansen, 2012](#)). The stellar tidal dissipation factor is even less well constrained but theoretical and observational works suggest $Q'_* \gtrsim 10^7$ (e.g. [Carone & Pätzold, 2007](#); [Hansen, 2012](#); [Damiani & Lanza, 2015](#)).

We studied the star-planet tides of TIC 237913194b using the EqTide⁸ code ([Barnes, 2017](#)), which calculates the tidal evolution of two bodies based on models by [Ferraz-Mello et al. \(2008\)](#) and a "constant-phase-lag" (CPL) model ([Goldreich & Soter, 1966](#); [Cheng et al., 2014](#)). For our tidal-torque test, we adopted a Stellar rotational period of 30 d and an initial planetary rotational period of 0.5 d (i.e., similar to the Solar and Jupiter rotational periods). We adopted tidal quality factors Q'_p in the range of $3 \times 10^4 - 10^6$. For the primary, we chose a fixed value of $Q'_* = 10^8$ (e.g., [Hansen, 2010](#); [Penev et al., 2012](#); [Bonomo et al., 2017](#)). We tested a large set of increased initial semi-major axes and eccentricities and integrated with EqTide. The results agree with the observed eccentricity and semi-major axis only for those samples that started a few percent above the observed values. This suggests that the orbital period of ~ 15 d is too large for significant tidal circularization within the age of the system (~ 5.5 Gyr, see [Table 5.2](#)) and TIC 237913194b's orbit has only slightly evolved from its primordial configuration. These results are in line with [Barnes \(2015\)](#), who showed that Jovians with periods longer than ~ 8 days and a typical eccentricity of 0.3 do not experience significant orbital and eccentricity decay. While we cannot determine the origin of the high orbital eccentricity, we conclude that the planet we observe today is not a credible progenitor of a future hot Jupiter.

5.5 Conclusions

We have presented the discovery of TIC 237913194b (TOI 2179b), a transiting warm Jupiter orbiting its G-type host in a very eccentric ($e \approx 0.58$) 15-day orbit. Its transit signal was detected using *TESS* full frame images from Sectors 1 and 2. We confirmed the planetary nature of the signal using ground-based photometry (*CHAT*, *LCOGT*) and high-precision spectroscopy (*FEROS*). Our main results are:

- a planetary mass $M_p = 1.942^{+0.091}_{-0.091} M_J$ and radius $R_p = 1.117^{+0.054}_{-0.047} R_J$, yielding a bulk density similar to Neptune's.
- with $e \approx 0.58$ one of the highest eccentricities among all currently known warm giants.

⁸<https://github.com/RoryBarnes/EqTide>

5 *Discovery of A Highly Eccentric Warm Jupiter*

- slow tidal evolution, prohibitive of a hot Jupiter progenitor scenario.
- an attractive opportunity for future observations of the planet’s atmosphere, which might harbor observable chemical disequilibrium processes due to the greatly varying external forcing.
- good prospects for detecting the Rossiter-MacLaughlin effect, which would be a valuable contribution to the still small sample of warm Jupiters with constrained spin-orbit obliquities.

This study made use of the following software packages: `astropy` (Robitaille et al., 2013), `juliet` (Espinoza et al., 2019b), `CERES` (Brahm et al., 2017a; Jordán et al., 2014), `ZASPE` (Brahm et al., 2017b), `tesseract` (Rojas, in prep.), `TESSCut` (Brasseur et al., 2019), `lightkurve` (Barentsen et al., 2019), `radvel` (Fulton et al., 2018), `emcee` (Foreman-Mackey et al., 2013), `corner.py` (Foreman-Mackey, 2016), `MultiNest` (Feroz et al., 2009), `PyMultiNest` (Buchner et al., 2014), `batman` (Kreidberg, 2015), `starry` (Luger et al., 2019), `celerite` (Foreman-Mackey et al., 2017), `petitCODE` (Mollière et al., 2015; Mollière et al., 2017), `EqTide` (Barnes, 2017), `Pingouin` (Vallat, 2018).

6

Planetary Systems Around Low-Mass Stars

This chapter presents work in progress, and the bulk of its scientific content will be the subject of a forthcoming paper with me as the main author. The project builds up on the results of Burn et al. (in press), where I am the second author. My contribution to that publication was to help with preparing and running the simulations, defining the new boundary conditions for the adapted model, and providing comments to the manuscript. I further wrote code that was used in the statistical analysis of the study and created Fig. 2 in the paper.

This chapter further makes use of a survey completeness analysis and planet occurrence rates derived in Sabotta et al. (subm.), which I contributed to as second author. There, I assisted in developing the approach and code to compute the survey completeness, provided advice during the project, and helped compiling and revising the manuscript.

Details of authorship: The present content of this chapter was created by me; I received advice by Remo Burn, Thomas Henning, Christoph Mordasini, Hubert Klahr, and Silvia Sabotta.

6.1 Motivation

In the previous chapters, I have established origins of and relations between different planet types in planetary systems around solar-type stars. I have kept the stellar mass fixed to $1 M_{\odot}$ to permit comparability with the most extensively studied planetary system, the Solar System. This choice also enabled me to use the Bern model of planet formation and evolution “as is”, since it has been developed for planets around solar-type stars. However, the most abundant planet hosts in the solar neighborhood are M dwarfs (Hsu et al., 2020). There has been a growing awareness that these stars are rewarding study objects regarding some of the most pressing questions on planet formation and habitability. It has been established that low-mass stars host more small, potentially habitable planets (Mulders et al., 2015a; Dressing & Charbonneau, 2015), and much has been learned about their properties from recent demographic studies: mass, radius, occurrence rates, also as a function of host star mass and metallicity have been in the focus of recent studies (Bonfils et al., 2013; Mulders et al., 2015b; Fulton & Petigura, 2018; Pinamonti et al., 2019; Hardegree-Ullman et al., 2019; Hsu et al., 2020). This preliminary work has provided important foundations, but it has relied on either observational data or on pure theoretical arguments (e.g., Burn et al., in press). It is thus very timely to confront the observational evidence regarding exoplanets around M dwarfs with planet formation theory. In this chapter, I take a first step toward this goal by relating the results of theoretical simulations with a well-defined sample from an extensive observation program.

Key tools to enable this comparison come from recent work by Burn et al. (in press) and Sabotta et al. (subm.), each of which I have significantly contributed to. In Burn et al. (in press), we have adapted the model introduced in Sect. 2.5.1 to lower stellar masses and also adapted the distributions of the Monte Carlo parameters to this stellar mass regime (compare Sect. 2.5.2). Using the modified model, we have computed a synthetic population of planetary systems with host star masses ranging between $0.1 M_{\odot} - 1 M_{\odot}$. This population termed “NGM” features a discrete grid of stellar masses, and each system initially starts with 50 planetary embryos. A total of 4996 systems were simulated. Several predictions from our investigation in Burn et al. (in press) can be tested observationally. I present the properties of the NGM population in Sect. 2.5.2.

The sample of observed exoplanets investigated here comes from the CARMENES M dwarf survey (Quirrenbach et al., 2010, 2013; Reiners et al., 2018b). To confront its planet detections with the results from Burn et al. (in press), it is crucial to take into account the unavoidable biases that occur in the observed datasets. In Sabotta et al. (subm.), we quantified CARMENES’ detection bias by injecting artificial RV signals into the ob-

served RV time series and trying to retrieve them using a method similar to the discovery method in the survey. By taking this bias into account, we also derived occurrence rates of planets in the parameter space spanned by the minimum planet mass $M \sin i$, the orbital period P , and the stellar mass M_* . With both the synthetic population and the bias-corrected occurrence rates in place, I set out to statistically compare the observed and synthetic planetary systems. This comparison serves two purposes: firstly, it shows which of CARMENES' planet detections our model can explain and which not. Secondly, it may reveal populations of simulated planets that would have had to be detected in the survey.

6.2 The CARM₁₂₅ observational sample

One of the most comprehensive searches for exoplanets around M dwarfs is the CARMENES high-precision RV survey. The CARMENES *instrument* consists of two independent Echelle spectrographs, one for visual wavelengths $0.55 \mu\text{m} - 1.05 \mu\text{m}$ and one for NIR wavelengths $0.95 \mu\text{m} - 1.7 \mu\text{m}$ (Quirrenbach et al., 2013). Both channels are fibered from the Calar Alto 3.5 m telescope. In its Guaranteed Time Observations (GTO), CARMENES targets a sample of ~ 350 stars whose spectral type distribution peaks at M4V (Reiners et al., 2018b). This survey started beginning of 2018 and has since produced more than 18,000 spectra (Sabotta et al., *subm.*) and led to various exoplanet discoveries (e.g., Sarkis et al., 2018; Ribas et al., 2018; Luque et al., 2018; Morales et al., 2019; Zechmeister et al., 2019; Stock et al., 2020; Nowak et al., 2020). The survey has already been defined with the goal to perform a population-level, statistical analysis on the datasets it produces. For a subset of the GTO stars, observations have already been completed. This subset, presented in Sabotta et al. (*subm.*) and here termed CARM₁₂₅, is used as the observational sample in this chapter. For each of the 125 targets in the sample, we computed detection probabilities on a grid in planet mass and orbital period via injection-and-retrieval tests. 27 planets that are hosted in 22 planetary systems have been discovered in CARM₁₂₅.

6.3 Observing the synthetic population

In order to make a comparison to observed data meaningful, we have to take into account selection effects and detection biases of the surveys and instrument we consider. The NGM population consists of 4996 simulated systems in the stellar mass range $0.1 M_\odot - 1.0 M_\odot$. Notably, the stellar mass is not a continuous variable but was fixed to the discrete masses $0.1 M_\odot$, $0.3 M_\odot$, $0.5 M_\odot$, $0.7 M_\odot$, and $1.0 M_\odot$, which were nearly uniformly sampled. Table 6.1 lists the simulation runs for each host star mass with the corresponding stellar

Table 6.1: Synthetic planet population NGM. Each of the simulated systems started with $N_{\text{emb,ini}} = 50$ planetary embryos. Five stellar masses and corresponding effective temperatures were sampled with different weights to match the observed distribution, and a total of 100,000 systems were drawn. Each system was assigned a random, isotropic inclination angle i . The population NG75 extends NGM to solar-type stars (see [Emsenhuber et al., 2020b](#), Burn et al. (in press)) but did not have to be sampled to match CARM₁₂₅'s M_{\star} distribution. Table adapted from Burn et al. (in press).

| Name | M_{\star} | $T_{\text{eff},5\text{Gyr}}^{(a)}$ | $N_{\text{emb,ini}}$ | Simulated Systems | Sampling Weights | Resampled Systems |
|---------------------|-----------------|------------------------------------|----------------------|-------------------|------------------|-------------------|
| NGM10 | 0.1 M_{\odot} | 2811 K | 50 | 1000 | 0.232 | 23282 |
| NGM14 | 0.3 M_{\odot} | 3416 K | 50 | 997 | 0.272 | 27234 |
| NGM11 | 0.5 M_{\odot} | 3682 K | 50 | 1000 | 0.304 | 30420 |
| NGM12 | 0.7 M_{\odot} | 4430 K | 50 | 999 | 0.192 | 19064 |
| NG75 ^(b) | 1.0 M_{\odot} | 5731 K | 50 | 1000 | 0 | 0 |

^(a) following [Baraffe et al. \(2015\)](#)

^(b) population also discussed in [Emsenhuber et al. \(2020b\)](#).

effective temperature at a simulation time of 5 Gyr, $T_{\text{eff},5\text{Gyr}}$. The initial number of planetary embryos per system $N_{\text{emb,ini}}$ was always 50. On the other hand, the CARM₁₂₅ sample has a characteristic distribution of stellar masses, as shown in Fig. 6.1. We approximated this distribution by weighted resampling of our synthetic population. For this purpose, we first computed a histogram of the observed stellar mass sample with the bin edges defined as the center between the discrete NGM masses. The normalized histogram counts then served as sampling weights, which for CARM₁₂₅ amount to 0.232, 0.272, 0.304, 0.192, and 0 for NGM's host star masses 0.1 M_{\odot} – 1.0 M_{\odot} (see Tab. 6.1); that is, the 1.0 M_{\odot} population has no contribution. In total, we sampled 100,000 systems with replacement. As the synthetic planets will be compared to RV-detected exoplanets, for which only minimum masses $M \sin(i)$ are known, we assigned them random orbital inclination angles i . Here, we assumed an isotropic distribution of orbit orientations and, for each system, drew a $\sin(i)$ term following Eqn 4.2. Hence, despite our oversampling of the NGM population, no planet occurs more than once with the exact same properties. The resulting M_{\star} distribution approximates the one of the CARM₁₂₅ sample (compare Fig. 6.1). While the oversampled population contains 5×10^6 planets, in the following we consider only the 2,421,124 planets that survived the formation and evolution phase until an assumed observation time $t_{\text{obs}} = 5$ Gyr.

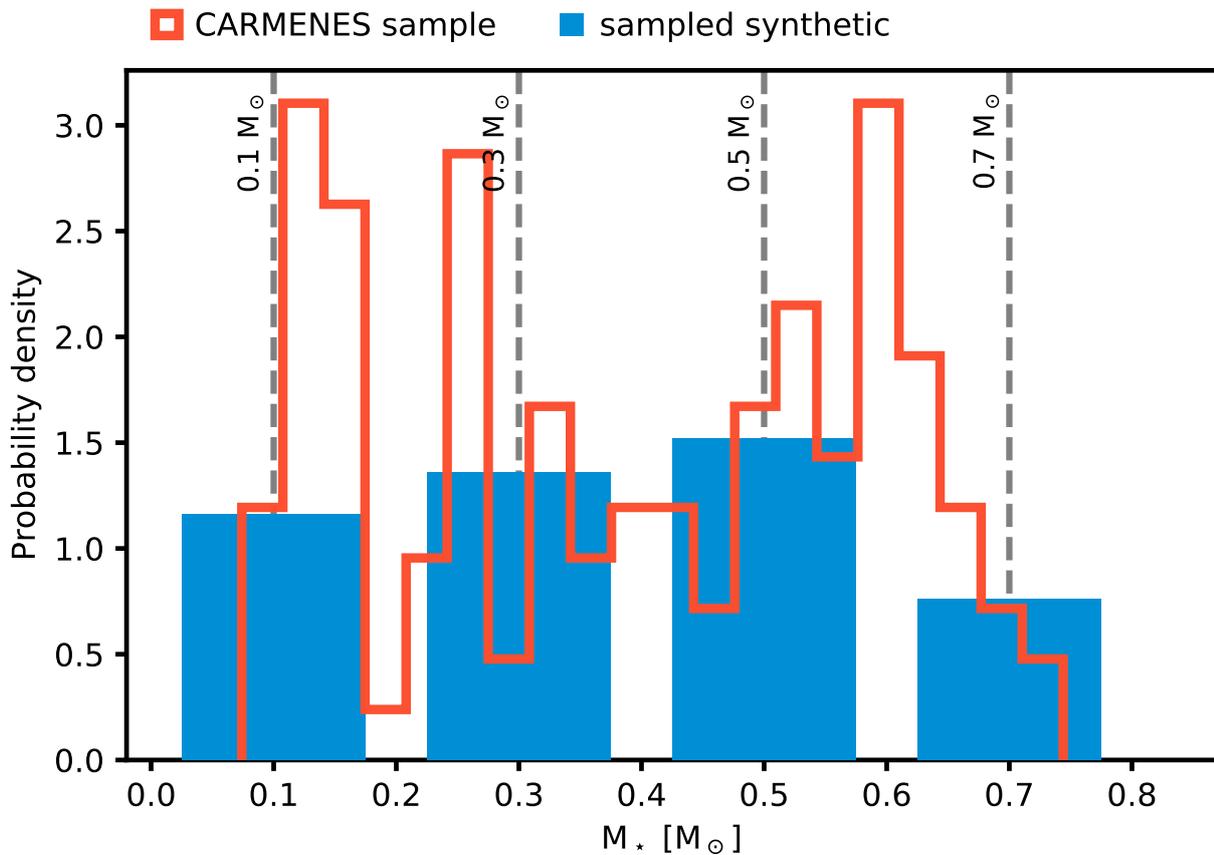


Figure 6.1: Stellar mass distributions of the observed and simulated populations. The host star masses in the synthetic NGM population are discrete (dashed lines). By weighted resampling of its systems according to the CARM₁₂₅ sample (red), we obtain a new distribution that approximates it (blue).

6.4 Observed and theoretical occurrence rates

With the simulated planet population we have generated, we can now derive synthetic occurrence rates. It is important to note the difference to the *fraction of stars* hosting a certain planet, which we have investigated in Chapter 4. In contrast, an occurrence rate n refers to the number of objects – in our case simulated or observed planets – per finite interval in some property they have (e.g., their mass) (Winn, 2018). In order to facilitate comparison with the sample of RV-detected exoplanets, we will express planetary occurrence rates in the parameter space spanned by the minimum mass and the orbital period. To be independent of the choice of interval, we can express an occurrence rate *density* corresponding

6 Planetary Systems Around Low-Mass Stars

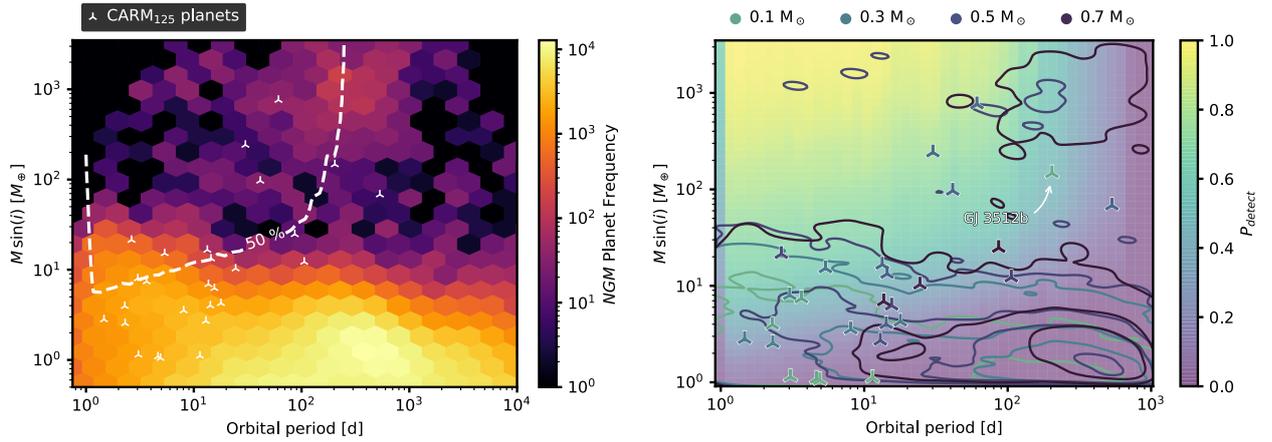


Figure 6.2: Planet occurrence of the *NGM* population and *CARM*₁₂₅ planets in minimum mass-period space. *Left:* Overall *NGM* M dwarf planet population regardless of specific host star mass. All synthetic planets surviving until $t = 5$ Gyr are shown. The population is strongly dominated by low-mass planets of less than $10 M_{\oplus}$, and the highest occurrence density lies at orbital periods of a few hundred days. For comparison, the white markers show positions of detected planets in the *CARM*₁₂₅ sample, and the white dashed line indicates the location where the survey completeness is 50%. As in the synthetic population, the majority of the observed planets populates low planet masses. *Right:* Survey completeness and dependency on host star mass. The background colors indicate detection probabilities in the *CARM*₁₂₅ sample, and the markers denoting detected planets are color-coded by host star mass. Likewise, the different colors for the synthetic population (contours) correspond to the different stellar mass bins. Low-mass planets ($M_P \lesssim 20 M_{\oplus}$) of all host star masses are well covered by the *NGM* population. Simulated giant planets are scarce and occur only around stars of earlier spectral types. In particular, GJ 3512b, a giant planet hosted by a late (M5.5) M dwarf, can not be explained by our model. There are also no synthetic counterparts for planets in the mass range $20 M_{\oplus} - 200 M_{\oplus}$, where a deep valley exists in the population.

to the occurrence rate n as

$$\Gamma_{M,P} = \frac{\partial^2 n}{\partial \log M \sin(i) \partial \log P}. \quad (6.1)$$

The left panel of Fig. 6.2 approximates this occurrence rate density by counting synthetic planets in a binned minimum mass-period space. It is obvious that low-mass planets represent the bulk of the population with the highest occurrence density for roughly Earth-mass planets at orbital periods of a few hundred days. Giant planets occur mostly at similar periods and are much rarer. They clearly separate from the rest of the population, leaving very few planets in the $20 M_{\oplus} - 200 M_{\oplus}$ mass range. Such a trend is not obvious in the detected planets from the *CARM*₁₂₅ sample. Although the observed planets are subject to detection biases, it is noticeable that the mass distribution here is rather continuous and does not reveal multiple modes.

As has been shown before, the efficiency and result of planet formation is a strong function of the host star mass (Mulders et al., 2015a; Raymond et al., 2007; Hsu et al., 2020, Burn et al., in press). Despite the small number statistics, I now factor in this mass dependency and discuss planet occurrences as a function of the stellar mass. In the right panel of Fig. 6.2, CARM₁₂₅ planets are color-coded by host star mass and overlaid on the synthetic population, which I divided according to the stellar mass bins defined above. The survey completeness of the CARM₁₂₅ sample is illustrated by the 2D histogram in the background. In Sabotta et al. (subm.), we computed this “sensitivity map” via injection and retrieval experiments, where we injected for each cell on the grid artificial RV signals into the observed time series of targets in the CARM₁₂₅ sample. For each pair of $M \sin(i)$ and P , this was done multiple times with random orbital phases. We then counted the fraction of successful recoveries of such “mock planets” by our detection pipeline, resulting in a unique completeness map for each target star. To obtain the *survey* completeness shown in Fig. 6.2, we took for each grid cell its arithmetic mean across the sample.

While it is apparent that the least massive planets are found exclusively around the latest M dwarfs, it is difficult to disentangle this result from the greatly increased detection probability for low-mass stars (compare Eqn 2.25). Despite the high survey completeness in the respective parameter domain, no hot Jupiters (planets with $M \sin i > 100 M_{\oplus}$ and orbital period of less than 10 d) are present in the CARMENES sample (Sabotta et al., subm.). The synthetic population agrees with this result.

I caution that with 27 planets detected in a sample of 125 stars, one can not make strong statements about stellar mass trends in this two-dimensional planet parameter space. Nevertheless, given the differences identified above it is worthwhile to explore the stellar mass-dependent planet occurrences in more detail. In Fig. 6.3, I jointly subdivide the synthetic and observed populations into the four stellar mass bins defined above. The synthetic population shows a clear trend of increasing giant planet occurrence with host star mass. In general, giant planets are rare and occur only in the stellar mass bins $0.5 M_{\odot}$ and $0.7 M_{\odot}$. None of the CARM₁₂₅ giants can be reconciled with this trend. The observed giant planet closest to a synthetic counterpart is GJ 876b, which orbits its early M dwarf host with a period of 61 d and has a projected mass of $M \sin(i) \approx 761 M_{\oplus}$ (Marcy et al., 2001; Rivera et al., 2005; Trifonov et al., 2018). With a stellar mass of $0.37 M_{\odot}$, the system just barely ended up in the stellar mass bin with zero occurrence density in the giant domain and is in fact relatively close to the few synthetic giant planets in the $0.5 M_{\odot}$ population. There are four discovered planets on intermediate and large orbits ($P = 10 \text{ d} - 1000 \text{ d}$) with projected masses $M \sin(i) = 20 M_{\oplus} - 200 M_{\oplus}$, where NGM shows a deep valley in the occurrences:

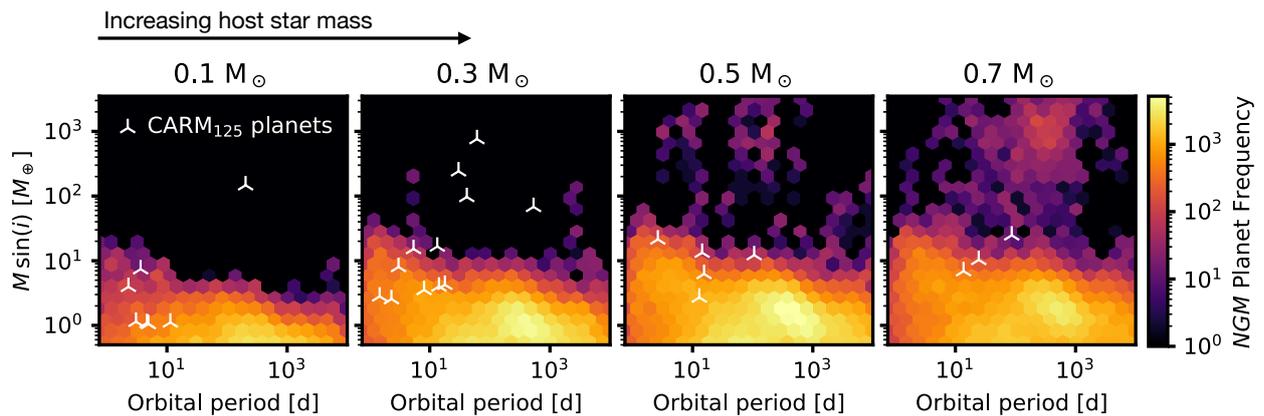


Figure 6.3: As Fig. 6.2, but partitioned according to host star mass. In the NGM population, the domination by low-mass planets on intermediate orbits is similar for all M_* , but giant planets occur only around stars with masses $0.5 M_\odot$ and higher. In contrast, the sub-giant and giant planets in CARM₁₂₅ occur most frequently around less massive stars. GJ 3512b, the giant planet in the $0.1 M_\odot$ panel, is particularly at odds with theoretical predictions. No simulated planets with $M \sin(i) \gtrsim 10 M_\oplus$ occur in this stellar mass bin.

GJ 1148b,c (Haghighipour et al., 2010; Trifonov et al., 2018), HD 147379b (Reiners et al., 2018a), and GJ 3512b (Morales et al., 2019). While none of them would have been expected based on our simulations, the $147 M_\oplus$ giant GJ 3512b is particularly difficult to reconcile with theoretical predictions: it orbits a late (M5.5) M dwarf with a very low-mass of $(0.123 \pm 0.009) M_\odot$ (Morales et al., 2019). Our theoretical model produces no giant planets in this stellar mass regime, and its existence remains a challenge for planet formation theories based on core accretion (Liu et al., 2020; Schib et al., 2021, but also see Kurtovic et al., 2021).

On the other hand, low-mass planets ($M \sin(i) \lesssim 20 M_\oplus$) are well represented by our synthetic systems. Populations from all stellar mass bins contribute to the synthetic planet occurrence in this domain and the maximum of the occurrence rate density is invariably at orbits ~ 100 d – 1000 d and terrestrial planet masses of a few M_\oplus . All but one CARM₁₂₅ planets less massive than $20 M_\oplus$ lie in the domain where significant occurrence density exists. The exception is GJ 4276b with $P \approx 13.4$ d and $M \sin(i) \approx 16.6 M_\oplus$ (Nagel et al., 2019), which sits in a small gap of planet occurrence in the $0.3 M_\odot$ bin that might be explained by sampling error.

6.5 Planetary mass function

By marginalizing over the orbital period axis, a minimum mass distribution of the samples can be obtained. Figure 6.4 shows a histogram of the uncorrected planet counts of

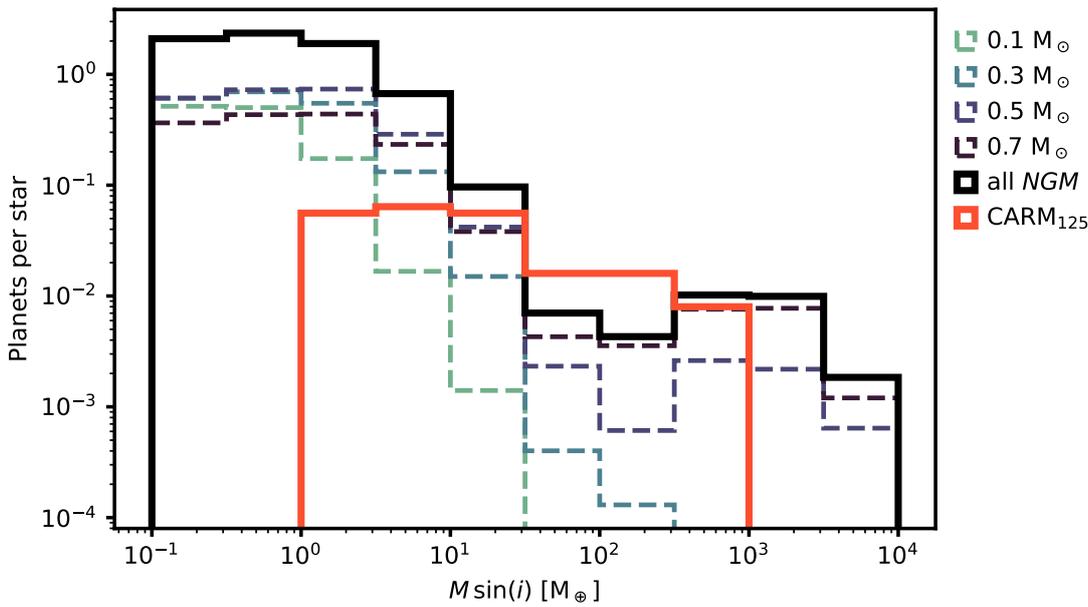


Figure 6.4: Minimum mass distribution of short-orbit ($P < 600$ d) NGM planets and of the detected planets in the CARM₁₂₅ sample, normalized to planets per star. NGM shows a steep slope and a bimodal distribution for larger stellar masses. No giant planets $\gtrsim 100 M_{\oplus}$ occur at stellar masses below $0.5 M_{\odot}$. CARM₁₂₅ shows a rather continuous distribution from terrestrial-mass planets to giant planets with small planets dominating. Note that no correction for observational biases is included.

CARM₁₂₅, as well as histograms for the planets in NGM, where I included all planets with periods < 600 d. This corresponds roughly to half of the observational baseline of the CARMENES GTO survey. The histograms were normalized to planets per star in the total CARM₁₂₅ or NGM sample, respectively. For the synthetic population, I subdivided the histogram by host star mass to highlight differences. It can be seen that for the more massive host stars, the synthetic distribution is bimodal. At stellar masses below $0.5 M_{\odot}$, no giant planets occur and the histogram features a single slope. The CARM₁₂₅ histogram is overall more flat, which is certainly partly due to the planetary mass-dependent detection bias. Noteworthy is, however, that no valley at intermediate masses and generally no obvious structure in the distribution is apparent. Instead, the CARM₁₂₅ planet sample shows a continuous negative slope in $\log M \sin(i)$.

For a direct comparison between the observed and theoretical planetary mass function, the inherent observational biases have to be taken into account. By conducting injection-and-retrieval experiments on the RV data for each of the stars in the CARM₁₂₅ sample, Sabotta et al. (subm.) derived sensitivities in minimum mass-period space for each individual target while accounting for instrumental and stellar noise. They averaged these

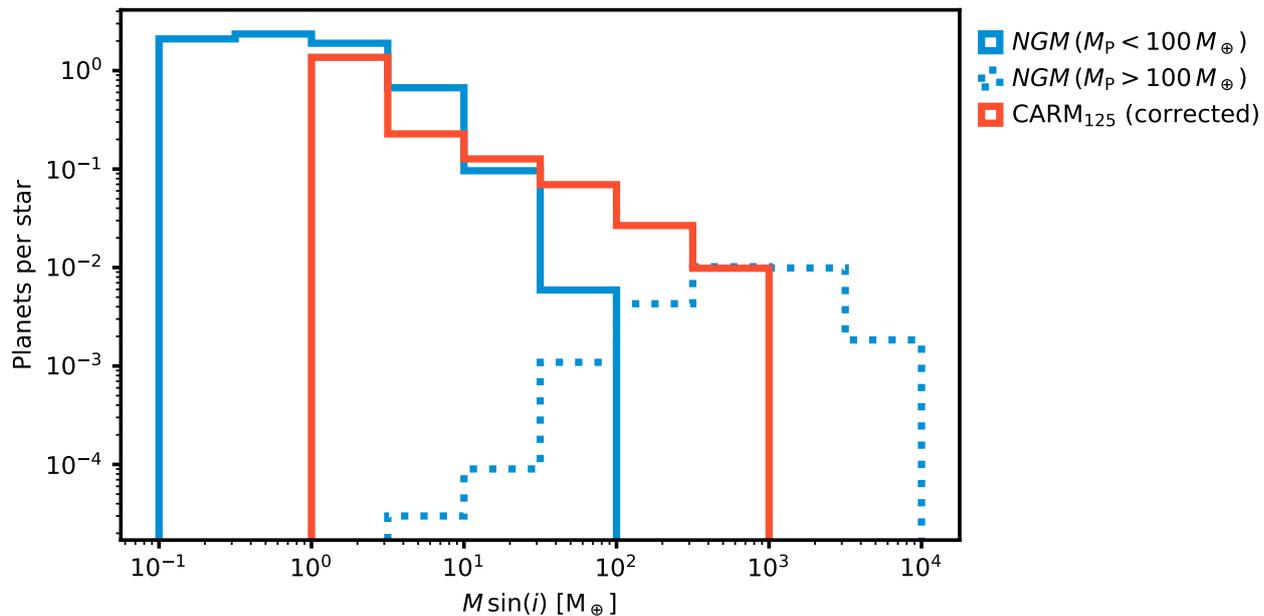


Figure 6.5: Minimum mass distribution of simulated planets and of the sensitivity-weighted CARM₁₂₅ planets. The NGM sample is divided into giant planets ($M_P > 100 M_\oplus$) and the rest; this classification is based on the actual planet mass M_P . Giant planets left aside, the NGM occurrences are steeper in $M \sin(i)$.

sensitivity maps to obtain the survey completeness illustrated in the right panel of Fig. 6.2. Weighting the 27 planet detections currently known in the CARM₁₂₅ subsample with this completeness yields bias-corrected occurrence rates, which are shown marginalized over the period axis in Fig. 6.5. As in the previous figure, the histogram was normalized to planets per star. The bias-corrected planetary mass function resembles a smooth power law. Compared to the uncorrected histogram, its slope is somewhat steeper in $M \sin(i)$, which is owed to the decreasing detection probability with decreasing planet mass. Again, the bimodal distribution of the simulated planets is contradictory. Excluding giant planets ($M \sin(i) > 100 M_\oplus$) removes the bimodality but results in a steeper planetary mass function than the observed one.

6.6 Discussion

6.6.1 Giant planets around M dwarfs

The stellar mass-dependent occurrences of giant planets in the NGM and CARM₁₂₅ samples are at odds with each other. While the Bern model generally produces only giant planets around host stars more massive than $\sim 0.3 M_\odot$, the observed CARMENES giants orbit only host stars with masses lower than $0.5 M_\odot$. This is despite the higher survey sensi-

tivity around earlier stars (Sabotta et al., *subm.*). In fact, the average detection probability of a Saturn-mass planet is nearly twice as high in the subsample of early M dwarfs with $M_* > 0.34 M_\odot$ as in the late M dwarf ($M_* < 0.34 M_\odot$) sample (~ 0.8 vs. ~ 0.5). This further substantiates the puzzling disagreement with theoretical predictions regarding the giant planet occurrence as a function of stellar mass. However, the small observed sample size should be noted: CARM₁₂₅ contains only four giant planets, and GJ 3512b might not be representative as it was hypothesized to have formed via disk instability (Morales et al., 2019). An extension of the analysis presented here with additional data will help to validate or refuse this trend, and we are currently in the process of investigating additional datasets.

A brief outline of the existing literature on this topic is worthwhile here. From the point of view of disk and exoplanet observations, Manara et al. (2018) found that the combined solid masses of planetary systems frequently exceed those of the most massive dust disks ($\sim 30 M_\oplus \approx 10^{-4} M_\odot$). They also discuss different populations for single and multiple systems and speculate about a dedicated formation pathway for massive giant planets around very low-mass stars.

Miguel et al. (2020) performed a population synthesis of planetary systems around (very) low-mass stars ($M_* = 0.05 M_\odot - 0.25 M_\odot$) based on a semi-analytical model assuming classical planetesimal accretion. Their model was originally designed to study circumplanetary disks and features a low-viscosity gas disk model as well as Type I and Type II migration (Miguel & Ida, 2016). They find efficient planet formation only in sufficiently massive disks ($\gtrsim 10^{-2} M_\odot$). Even under these conditions, their model fails to form any planets more massive than $5 M_\oplus$.

Zawadzki et al. (2021) use N -body calculations to simulate the formation of planets around $0.2 M_\odot$ stars. Under the assumptions of early planetesimal formation (Lenz et al., 2019) and including Type I migration, they find efficient growth of planetary cores through early collisions of planetary embryos. Their setup corresponds to a solid disk mass of $\sim 2 \times 10^{-2} M_\odot$ at the start of their simulations. While they do not model gas accretion onto planets, many of these cores grow to super-Earths, some of which in the mass range where runaway accretion could be triggered.

Burn et al. (*in press*) discuss some modifications to the model used here that would enable the formation of giant planets around very low-mass stars. If Type I migration is artificially suppressed, planetary cores are able to reach masses beyond $10 M_\oplus$ without rapidly migrating into the star, and some giant planets form. Even though there is no obvious evidence justifying such tuning of the migration scheme, it does to some degree

mimic “planet traps” due to inverted gas pressure gradients, the existence of which is indicated by the numerous observed disk substructures (e.g., [Andrews et al., 2018a](#)). This way out of the giant planet conundrum still requires relatively high initial solid disk masses $M_{\text{solid},0} \gtrsim 66 M_{\oplus} \approx 2 \times 10^{-4} M_{\odot}$. The efficiency of giant planet formation can be further enhanced if the planetesimals are more concentrated toward the inner disk.

The present data situation does not allow a definite conclusion whether the observed population of gas giants around very low-mass stars can be explained by core accretion, and under which specific conditions this is possible. Further research on the subject, both theoretical and observational, is urgently needed.

6.6.2 Planetary mass function

Confronting the observed CARM₁₂₅ planets with the synthetic NGM population revealed a discrepancy between their planetary mass functions: while the distribution recovered from the discovered planets follows a smooth power law, the simulated planets show a significant bimodality with a gap between about $20 M_{\oplus}$ and $200 M_{\oplus}$. This valley was already previously predicted for planets around sunlike stars ([Mordasini et al., 2009b](#)) and persists in the current version of the Bern model ([Emsenhuber et al., 2020b](#), [Emsenhuber et al., in prep.](#)). It separates all planets that attained solid cores massive enough to enter runaway gas accretion and became giant planets from those that did not. As detailed in [Sect. 2.4](#), the reason for the low occurrence of such intermediate-mass planets is the short timescale on which runaway gas accretion happens. Once a planet reaches the threshold mass, it will quickly grow into a giant planet. Only the unlikely event of gas disk dispersal exactly during this growth phase would fix the planet’s mass to intermediate values. Preliminary observational evidence for this sub-Saturn desert has been provided by [Mayor et al. \(2011\)](#), who computed bias-corrected occurrence rates and the mass distribution for the HARPS RV survey, albeit for mostly solar-type stars. They reported a decrease of their bias-corrected mass distribution “between a few Earth masses and $\sim 40 M_{\oplus}$ ”. The significance of this dip has recently been questioned ([Bennett et al., 2021](#)).

Testing the existence or non-existence of this demographic feature in planetary systems around low-mass stars is very interesting from the perspective of planet formation, as it might provide clues about possible different formation pathways for different stellar types. The present observational sample of 27 confirmed planets is not large enough for any statistically sound claim. A more quantitative analysis of the observed and theoretical planetary mass functions and the inclusion of additional datasets will shed light on the statistical significance and physical origin of the discrepancy.

7

Conclusions and Future Work

In this chapter, I summarize the work presented in Chapters 3–6, highlight key results that represent a progress with respect to the questions posed at the beginning, and outline some directions for possible future projects based on these results.

7.1 Summary

Recent advancements in understanding the architectures of planetary systems have been based on either observational data alone (e.g., [Weiss et al., 2018](#); [He et al., 2020b](#)) or on purely theoretical methods (e.g., [Lambrechts et al., 2019](#)). In this thesis, I connected planet formation theory with results of extensive observation programs and drew connections between properties of protoplanetary disks and the architectures of the emerging planetary systems. Using global planet formation models, I have shown that intra-system relations between planets can be linked to properties of the protoplanetary disk they emerged from.

Chapter 3 takes a data-driven approach to investigate the relations between properties of simulated planets and their formation histories. The analysis is based on synthetic planet populations calculated with the Generation III Bern model of planet formation, which follows the core accretion paradigm and assumes solar-type host stars. As part of a cluster analysis, I trained an unsupervised machine learning algorithm on typical exoplanet observables to identify groups of similar planets. Without presupposing any known planet types, four classes of planets are identified. They roughly correspond to the observed exoplanet types of (super-)Earths, (sub-)Neptunes, and giant planets, plus an additional unobserved class of small planets on distant orbits. I showed that the future cluster affiliation of the planets can be reliably predicted from properties of their parental disks. The most reliable predictors are the initial orbital distance at which the planetary embryos are injected, followed by the total planetesimal mass available. When N -body interactions

between planets are taken into account, as in the nominal *NG76* population, the predetermination of planet types is slightly weakened. This particularly concerns sub-Neptunes, which frequently undergo giant collisions and turn into super-Earths. My results highlight the impact of gravitational interactions for systems of low-mass planets, but also the importance of currently missing self-consistent models of planetary embryo formation.

In Chapter 4, I present a study of the relations between inner super-Earths and outer giant planets in the same system. The main goal of the study is to shed light on a suggested mutual enhancement of the occurrences of these planet types. I confront the multi-planet population analyzed in Chapter 3 with an observed exoplanet sample while taking into account its observational biases. This shows indeed an enhanced formation of one planet type whenever the other is present, albeit weaker than previously proposed. I attribute the discrepancy to dynamically active giant planets on intermediate orbits, which frequently disrupt existing inner systems of smaller planets. A comparison of super-Earths with and without a cold Jupiter companion results in two predictions: 1. A joint occurrence of both planet types requires intermediate-mass planetesimal disks, and 2. Inner super-Earths with higher density have a higher likelihood of having a giant planet companion.

Chapter 5 presents the discovery of a highly eccentric warm Jupiter of the type I suggest being responsible for the disruption of compact low-mass planetary systems. After initial detection in Transiting Exoplanet Survey Satellite (*TESS*, [Ricker et al., 2014](#)) full-frame images, it was followed up with ground-based photometry and radial velocity time series. By jointly analyzing these datasets in a Bayesian scheme, I confirm the planetary nature of the signals. I precisely constrain the planet's mass to $M_P = 1.942^{+0.091}_{-0.091} M_J$ and its radius to $R_P = 1.117^{+0.054}_{-0.047} R_J$, implying a bulk density similar to Neptune's. Being on one of the most eccentric orbits of all known warm Jupiters ($e \approx 0.58$), the planet likely had a past encounter with another massive body in the system. I present a tidal evolution analysis that shows a large dissipation timescale, suggesting that the planet is *not* a hot Jupiter progenitor caught during its high-eccentricity migration. The newly discovered planet further represents an opportunity to study energy redistribution in the atmosphere of a warm Jupiter with high eccentricity.

Chapter 6 focuses on planetary systems around low-mass stars by confronting a dedicated M dwarf population synthesis to detected exoplanets from the CARMENES ([Quirrenbach et al., 2010](#)) survey. I examine both samples as a function of stellar host mass and find a strong dependence of planet occurrences on this parameter. In particular, CARMENES' discoveries suggest a population of giant planets around stars with $M_\star < 0.5 M_\odot$, which the planet formation model does not reproduce. There is further a differ-

ence in the planetary mass function, which appears smooth for the observed planets and bimodal for the simulated planets. The continuation of this work in progress will shed light on the origins of these discrepancies.

7.2 Conclusions

The results presented in these chapters can be put into a larger context by summarizing their consequence for the questions stated at the beginning of this thesis (Sect. 1.1):

1. Are the increasingly complex multi-planet models still deterministic?

While interplanetary N -body interactions introduce an element of chaos - in the mathematical sense - the links between the initial conditions of the model and the observables of the planets it produces are largely preserved (Chapter 3).

2. Is the formation of one planet type or another already set by primordial properties of the disk and host star?

The type of planet formed can be reliably predicted from its initial conditions (Chapter 3). Initial orbital positions of planetary embryos are the most important predictors (although this does not reflect true “ $t = 0$ ”). The high predictability of core accretion models should not be misunderstood as a legitimation for the single-planet approximation: gravitational interactions are vital for realistic simulations of low-mass planets and their system architectures. Different types of architectures can be traced back to differently sized solid reservoirs in the originating protoplanetary disks, which is a proxy for stellar mass and metallicity. Primordial conditions thus not only determine the properties of individual planets, but also the architectures of planetary systems (Chapter 4).

3. What are the relations between inner systems of rocky planets and outer giant planets, and can the core accretion scenario explain these relations?

The formation of warm super-Earths and cold Jupiters in the same system is enhanced compared to the individual appearances; although the trend is weaker than what has been proposed through observations (Chapter 4). Under the core accretion paradigm and in the range of assumptions our model is operating, this mutual increase of occurrence is not due to direct interactions of the planets or indirect influences that are expected in the pebble accretion picture. Instead, the effect depends on disk properties, above all the solid content, which determine if, where, and how efficiently planetary cores are formed.

4. Are giant planets on wide orbits favorable or detrimental for systems of inner rocky

planets?

The main reason for the attenuated theoretical super-Earth-cold Jupiter relation is the occurrence of warm and dynamically active giant planets that frequently disrupt inner systems. These planets form preferably in high-metallicity environments, which is also a prerequisite for a system including both inner super-Earths and outer giant planets (Chapter 4).

5. Is there observational evidence for interactions among these planet types?

The highly eccentric warm Jupiter whose discovery I presented in Chapter 5 is one possible culprit for the violent destruction of systems of inner rocky planets. Its extreme dynamical state points to a past interaction with an additional, undetected massive companion. Hypothetical planets interior to TIC 237913194b's orbit could not maintain long-term stability of their orbits.

6. Are the bulk properties of planets connected to the architectures of their planetary systems?

Planets should have different volatile content depending on the architectures of their system. From this follows the testable prediction that systems with high-density super-Earths are more likely to host an outer giant (Chapter 4). This link between the bulk densities and system architectures is again due to different birth environments: depending on the disk mass, either icy super-Earths without giant planets or dry super-Earths with giant companions form.

7. Can population synthesis models reproduce the planetary systems around low-mass stars?

Chapter 6 addresses this question by comparing a dedicated synthetic planet population with the exoplanets discovered by the CARMENES survey. While a conclusive answer requires further research, two differences relevant for planet formation processes stand out: 1. Compared to the rather smooth, continuous mass distribution of the observed planets, a bimodality in the simulated population is noticeable, and 2. The simulations produce giant planets only around higher-mass stars, whereas all of CARMENES' giant planets orbit stars less massive than $0.4 M_{\odot}$.

7.3 Outlook

Exoplanet demographics is a rapidly evolving field, and this is particularly true for its fledgling sub-discipline of planetary systems exploration. There are many future oppor-

tunities for expanding on the research I presented here, both from a theoretical and an observational point of view.

7.3.1 Extensions of the formation model

The global model of planet formation and evolution used in this thesis (Emsenhuber et al., 2020a, Sect. 2.5.1) is being continuously developed and is making progress in addressing wider regions of the planetary and stellar parameter spaces (e.g., Alibert et al., 2005; Morasini et al., 2009a; Alibert et al., 2013; Burn et al., 2021, Voelkel et al., in prep.). In my view, the following current shortcomings of the model are particularly important to address.

Disk model

The disk model still uses as initial conditions derived quantities, for instance the total integrated flux of unresolved disks to infer their masses. The corresponding measurements rely on assumptions on the opacities and disk temperature profiles. Furthermore, smooth power law radial density profiles are assumed for all simulated protoplanetary disks, which is in conflict with numerous observations of disks with rich sub-structures (Andrews et al., 2018b). In an era where more and more highly resolved disk observations are available, this approach is outdated. Future available mm/submm data will allow pinning down the disk parameters at the time of planet formation. Constraints on disk sub-structures and dynamical traces of embedded planets (e.g., Teague et al., 2018) should be taken into account. Introducing structured disks might help resolve discrepancies due to too efficient planet migration (see, e.g., Chapter 4, Burn et al., in press).

Pebble accretion

Planetary embryos grow by the accretion of both planetesimals and pebbles. The accretion of planetesimals is of special importance for the early stages of growing from 100 km to the level of several hundreds of kilometers required to start pebble accretion. As discussed in Sect. 2.2.7, the accretion of small grains drifting through the disk has been shown to efficiently grow planetary cores, facilitating subsequent gas accretion. This paradigm has been solidified over the past years (Ormel & Klahr, 2010; Lambrechts & Johansen, 2012; Birnstiel et al., 2012; Schoonenberg et al., 2019; Bitsch et al., 2019a; Lambrechts et al., 2019). It is thus very timely to develop combined models of pebble and planetesimal accretion.

Planetary embryo formation

The current ad-hoc injection of planetary embryos into the planet formation model is a major shortcoming of our framework (and in fact of all current population synthesis models). Voelkel et al. (2021a,b) took the first steps in developing a consistent treatment of embryo formation. The next stage should be to further develop and include their embryo formation model into the Bern model, replacing the current Monte Carlo embryo insertion.

Evaporation and re-condensation beyond the ice line

Due to the evaporation, outward diffusion and subsequent re-condensation of water, the region just beyond the water ice line should be significantly enriched in pebbles and planetesimals. This enrichment is expected to have a large impact on embryo and core formation (Stevenson & Lunine, 1988; Cuzzi & Zahnle, 2004; Drążkowska & Alibert, 2017). The importance of these effects and consequence for the populations should be studied in future versions of our planet formation model.

7.3.2 Fitting for system architectures

On the level of planet parameters, it is relatively straightforward to compare two samples of planets, regardless of whether they are observed or simulated. The typical use case is the comparison of an observed parameter distribution with the corresponding one from a synthetic sample (or several synthetic samples stemming from different model assumptions). This way, it can be measured – for example with a Kolmogorov-Smirnov test – how well the simulations reproduce exoplanet observables and which changes in the model would improve the match.

It is less obvious how one should compare the *architectures of catalogs* of systems. I suggest to compute, for both observed and simulated planets, quantitative statistical measures that consider all planets in a system simultaneously. Suitable measures were for instance presented by Gilbert & Fabrycky (2020), who used them to quantify system similarities and distinguish between real demographic features and observational biases that may mimic them. Similarly, Alibert (2019) defined a metric that reduces the dimensionality of multi-planet systems in period-radius space. These or similar measures could also serve as a statistical target for simulated multi-planet systems.

While a Markov Chain Monte Carlo (MCMC) approach is certainly not feasible for elaborate simulations that take weeks to compute on a supercomputer, a re-sampling scheme that aims to match target values of the architectural measures could be implemented. An-

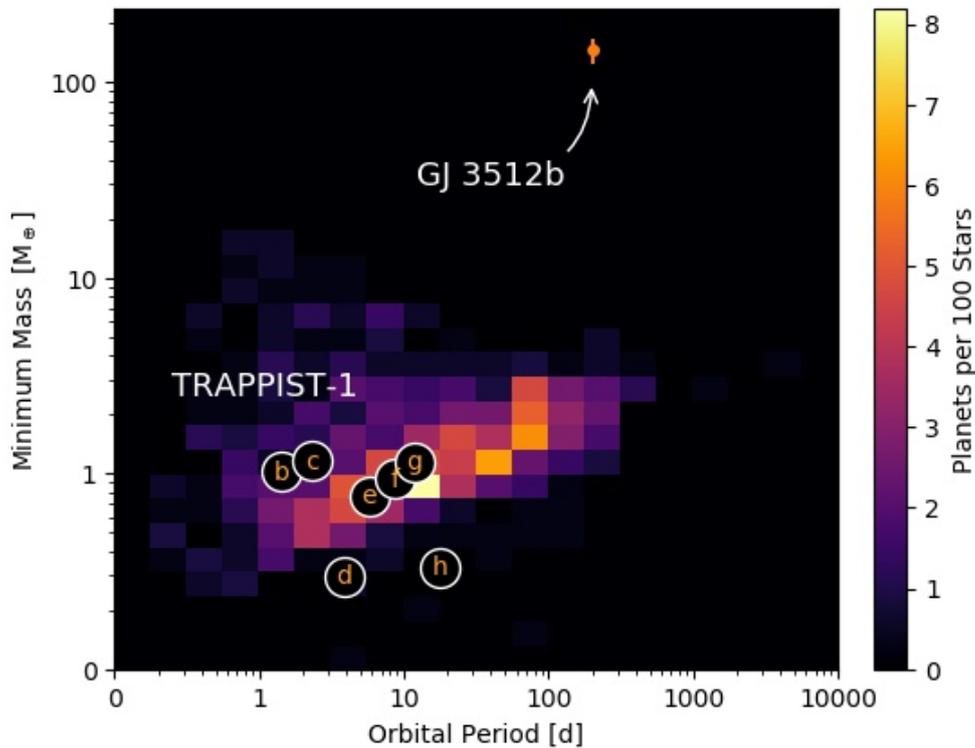


Figure 7.1: Observed and simulated M dwarf planets. The plot shows occurrence rates of synthetic planets from a $0.1 M_{\odot}$ host star population in the minimum mass-period plane. A simplified calculation of a detection bias was applied to remove all planets with an Radial Velocity (RV) semi-amplitude below 1 m s^{-1} . While planets similar to those in the TRAPPIST-1 system (Gillon et al., 2017; Luger et al., 2017) are the norm, the model does not form any giant planets like GJ 3512b (Morales et al., 2019). This incongruity between model and observation is a major motivation to further develop the planet formation physics in our model. Required extensions include the consistent modelling of planetesimal formation, embryo formation, and pebble accretion.

other potentially insightful approach would be to compute such measures a posteriori for a range of different models: the effect of particular “model switches” or tuning parameters on the system architectures could be analyzed in a quantitative manner.

7.3.3 The architectures of M dwarf planetary systems

Low-mass stars have been shown to be the most abundant planet hosts (Hsu et al., 2020), and their small mass and size facilitates the detection and characterization of small planets in the habitable zone. It is thus a very promising avenue for exoplanet demographics to take the stellar mass into the equation. Building on the foundations laid out in Chapter 6, selection effects and detection biases need to be taken into account before more detailed statistical analyses can provide meaningful results. Since the completeness of the CARMENES survey has been determined for each target star (Sabotta et al., *subm.*), it is

possible to assign each synthetic planet a detection probability depending on its period and mass. An NGM population *as observed by CARMENES* can thus be produced and directly compared to the uncorrected observed population. Efforts to implement this are underway (Schlecker et al., in prep.).

To increase the statistical significance of the analysis, it would be rewarding to expand the observed CARMENES sample with data of the complete Guaranteed Time Observations (GTO) sample, and beyond that with data from other surveys such as the HARPS M dwarf sample (Bonfils et al., 2013). A joint analysis of combined datasets may be more informative, but special care is required in taking into account the different selection functions. For example, the HARPS M dwarf survey covered a slightly different stellar mass distribution compared to CARMENES, although with a similar median mass close to $0.3 M_{\odot}$.

It would further be interesting to test if the super-Earth-cold Jupiter relation and the composition-architecture link presented in Chapter 4 also exist in the M dwarf regime. Current and future surveys targeting M dwarfs such as TESS (Ricker et al., 2014), SPECULOOS (Sebastian et al., 2021), EDEN (Gibbs et al., 2020), and PLATO (Rauer et al., 2014) will increase the available sample of planets around low-mass stars and facilitate the exploration of their properties.

These ideas are closely related to the questions that have been addressed in this thesis. Beyond that, the architectures of multi-planetary systems provide plenty of broader questions for the years to come, some of which I would like to point out below.

- The spectral type M covers a large bandwidth of stellar masses, ranging from stars not too different from the Sun to extremely low-mass stars with only $\sim 0.1 M_{\odot}$. *Which stellar mass bin is most favorable for temperate, potentially habitable planets, and how does this compare to the prediction from population syntheses?*
- A particularly exciting and unexpected discovery is the existence of giant planets around very low-mass stars, for example the ~ 0.5 Jupiter masses cold giant around the $\sim 0.12 M_{\odot}$ star GJ 3512b (see Sect. 6.4, Morales et al., 2019). *Can these objects be explained by the core accretion paradigm of planet formation and, if so, what are the conditions needed to form them?* The preliminary comparison of observed planets around ultra-cool M dwarfs and their synthetic counterparts in Fig. 7.1 shows already qualitatively which planets our current model can explain and which not.
- Planetary searches targeting low-mass stars allow the detection of planets down

to the terrestrial regime. The continuous push toward lower detection limits lets population-wide studies of these systems appear imminent. But *which observations are needed to test the predicted trends in the M dwarf population?* Using data from the *TESS* mission and ground-based M dwarf surveys such as CARMENES, one could isolate the most interesting targets for follow-up with *James Webb Space Telescope* (*JWST*, [Beichman et al., 2014](#)), *The Roman Space Telescope*, and future space-based missions. For *JWST*, the focus should be on maximizing the utility of the inherent small sample the mission will provide to distinguish between competing theories.

By continuous improvement of the theoretical modeling, and with the current and expected advances in observational instrumentation, I am optimistic that some of the above questions will be answered within this decade.

To my successor(s), if you are still tapped into the bitstream: your future is bright. As it seems, you will have the chance to participate in a major revision of the current narrative of planet formation. Enjoy the ride.

A

Appendix to Chapter 3

A.1 The choice of a clustering algorithm

A.1.1 Clustering algorithms

For the cluster analysis in Sect. 3.3, we examined several other clustering algorithms in addition to Gaussian Mixture Model (GMM, [McLachlan, 1988](#))¹ and explored their behavior on our data set. For each method, we used its implementation in `scikit-learn` ([Pedregosa et al., 2011](#)) and, where applicable, chose the default Euclidean distance metric. The algorithms considered are centroid, density, or hierarchical-based. A centroid-based method we explored was K-means ([MacQueen, 1967](#); [Lloyd, 1982](#)). In the density-based group, we tested DBSCAN and OPTICS ([Ester et al., 1996](#); [Ankerst et al., 1999](#)). For hierarchical clustering, we examined Agglomerative clustering ([Ward, 1963](#)) besides GMMs.

K-Means² ([MacQueen, 1967](#); [Lloyd, 1982](#)) is a centroid-based clustering algorithm: it randomly initializes k centroids and associates each data point to the centroid that is closest to it, then shifts the centroids to the mean of their cluster. These steps are repeated until no changes occur. The algorithm requires only a single hyperparameter k , which is the number of clusters.

Agglomerative clustering³ ([Ward, 1963](#)) is a bottom-up hierarchical clustering algorithm: each data point begins as its own cluster and incrementally merges similar pairs of clusters into a new cluster. This process is repeated until there are k clusters left, where k is the hyperparameter for the number of clusters. When testing this algorithm, we used a hyperparameter called linkage to quantify ‘similarity’ between pairs of clusters (e.g., [Ward, 1963](#); [Szekely & Rizzo, 2005](#)). Empirically, we found that the “Ward” linkage is optimal.

¹`sklearn.mixture.GaussianMixture`

²`sklearn.cluster.KMeans`

³`sklearn.cluster.AgglomerativeClustering`

DBSCAN⁴ (Ester et al., 1996) is a density-based clustering algorithm classifying each data point as either a core point (with at least `minPts` neighboring points within a distance ϵ), a reachable point (that is within distance ϵ of the core point), or an outlier (that is not reachable by any core point). All core points and their reachable points form a cluster, but outliers do not. The method we tested is an advancement of DBSCAN with improved performance on datasets of varying density. This method called OPTICS⁵ (Ankerst et al., 1999) has one hyperparameter: `minPts` – the minimum number of points nearby to make a core point.

A.1.2 Validation metrics and choice of method

Each of these methods has hyperparameters, that is, parameters that are not derived during model training but that control the learning process itself. We used a number of validation metrics to quantify the clustering performance for each method and specific choice of hyperparameters. Some of these metrics are method-specific and can only be used with a specific algorithm. These are the elbow method (e.g., Thorndike, 1953; Ketchen & Shook, 1996), the Bayesian and Aikake Information Criteria (BIC and AIC, e.g., Akaike, 1973; Schwarz, 1978; Cavanaugh & Neath, 2019), and the dendrogram method (e.g., Nielsen, 2016). The elbow method is used to evaluate the performance of the K-Means algorithm. By plotting the within-cluster sum-of-squares against k , an ‘elbow’-shaped curve emerges. The ideal k will be one close to the ‘elbow’. The reasoning for this is that we aim to find the first k that minimizes the within-cluster sum-of-squares. BIC and AIC are used for GMM. Both are based on information theory and are used to prevent overfitting and underfitting to choose the most optimized model. The dendrogram method is used to judge the bottom-up process of Agglomerative clustering. It shows the clustering at each hierarchy, where the y-axis is the distance between clusters and the x-axis shows the clusters. Therefore, the goal is to perform a horizontal cut such that the vertical distance is maximized. As one traverses up the hierarchy, the vertical distance naturally increases.

In addition to these scores, we used the following scalar-valued metrics that can be used for any method: the Silhouette score (Rousseeuw, 1987), the Caliński-Harabasz score (CH, Caliński & Harabasz, 1974), and the Davies-Bouldin score (DB, Davies & Bouldin, 1979). The Silhouette score is computed from the mean intra-cluster distance and the mean nearest-cluster distance. Silhouette scores range between -1 and 1 with 1 being the best and -1 being the worst, and values near 0 implying overlapping clusters. We aimed to maximize this score. The Caliński-Harabasz score is the ratio of the within-

⁴`sklearn.cluster.DBSCAN`

⁵`sklearn.cluster.OPTICS`

cluster dispersion and the between-cluster dispersion, where dispersion is the sum of the squared distances. Again, we aimed to maximize this score. The Davies-Bouldin score determines the clustering performance by using the ratio of the within-cluster distances to the between-cluster distances. As a result, compact clusters that are far apart give better scores. The minimum score is 0, and we aimed to minimize this score.

A.1.3 Model selection

Our approach in selecting the best clustering method was as follows: first, we applied each method to the $\{a, M_P, R_P\}$ subspace of the *NG73* planet population for a wide range of hyperparameters. We then compared the validation metrics computed for the resulting clusterings. The scores did not always agree unanimously, which is expected, as the structures in our multidimensional data set are rather complex and the scores consider different goals regarding an optimal clustering. The next step was thus to produce, for each combination of method and hyperparameters, scatter plots that showed the clustering results in different projections of $\{a, M_P, R_P\}$ space. Using these plots, we could compare the different partitionings and determine the most sensible model. Figure A.1 shows these diagnostic plots for k-means, OPTICS, and Agglomerative clustering, using the choice of hyperparameters considered most appropriate. The diagnostic plots for GMMs are shown in Fig. A.2. Based on this selection procedure, GMMs showed the best performance and we considered it our nominal method for clustering.

A free parameter of GMMs is the number of components N , which we chose using the same two-step approach as in the method selection. After the validation metrics suggested $N = 4, N = 6$ for *NG73* and $N = 3, N = 5$ for *NG76* (see Fig. 3.1), we assessed the diagnostic plots shown in Fig. A.2. For *NG73*, we found that the GMM with $N = 6$ reaches similar scores than $N = 4$ but traces less reliably the underdensities in the domain and partly draws cluster borders through rather arbitrary regions. We thus chose the GMM with $N = 4$ as our nominal model for the single-planet case. For *NG76*, the model with more components reliably detects visible overdensities and outperforms the less complex model. Hence, we adopted the GMM with $N = 5$ as the nominal model for the multi-planet case.

A.2 Boundary conditions for giant planet formation

In Sect. 3.5.3, I characterize the cluster of giant planets in $a_{\text{start}} - M_{\text{solid},0}$ space, where it occupies a distinct triangular region. Two quantities were used for this characterization: the

NG73: single-planet

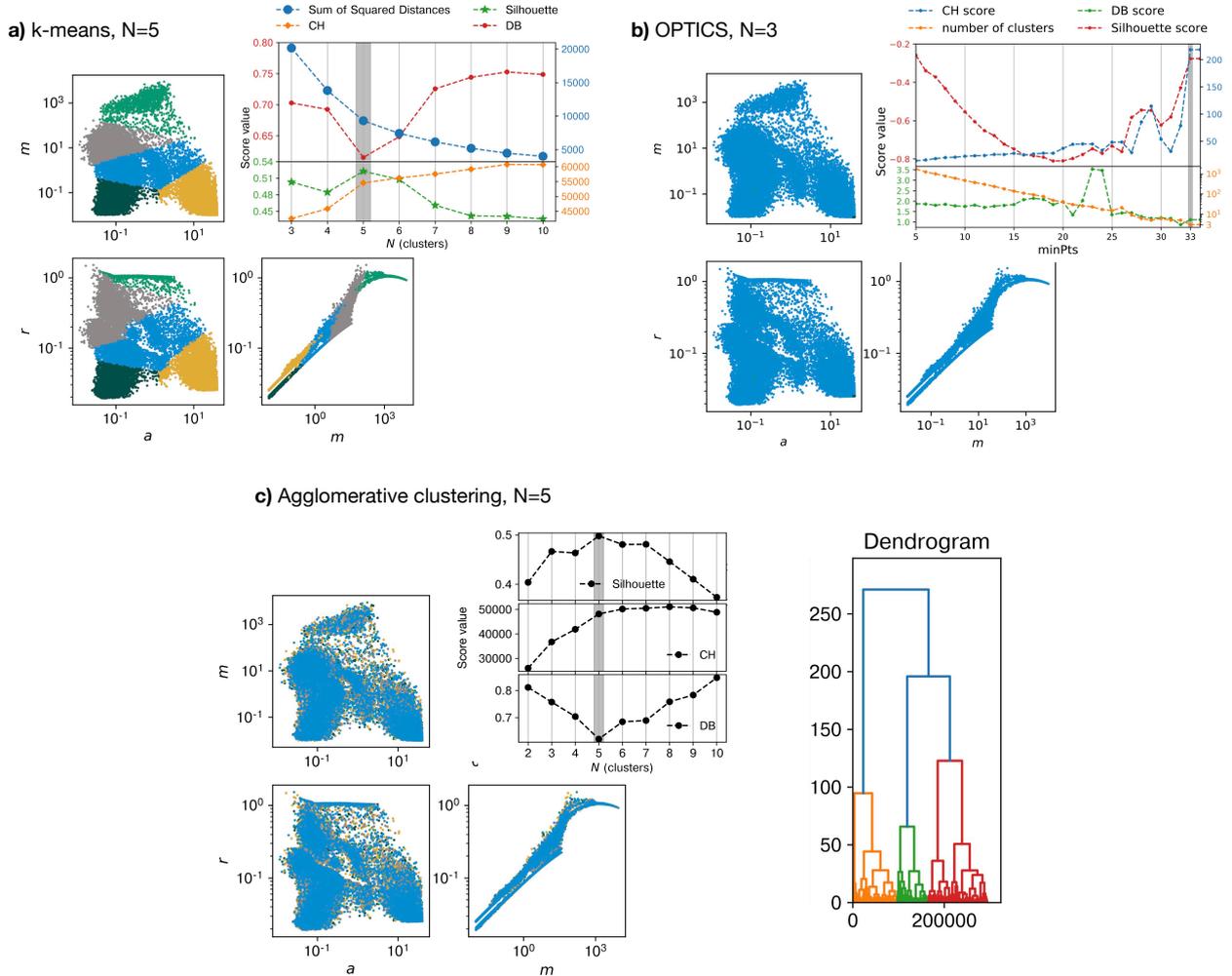


Figure A.1: Diagnostic plots for clustering method selection. For each alternative clustering algorithm we explored, we show the validation metrics we used to choose hyperparameters. Based on these metrics, we show the resulting clustering for the most promising choices in the corner plots. a) Even in the best case ($N = 5$), k-means' approach to draw cluster borders is too simplistic to account for the structure in our data. b) For the numerically best choice of $minPts$, OPTICS finds three clusters of extremely different sizes. Most of the data belong to a single cluster that covers the whole domain, and no sensible relation to the data point density is apparent. c) Agglomerative clustering suggests the existence of five clusters. Again, no reasonable partitioning is visible. The lower right panel shows the dendrogram corresponding to this clustering.

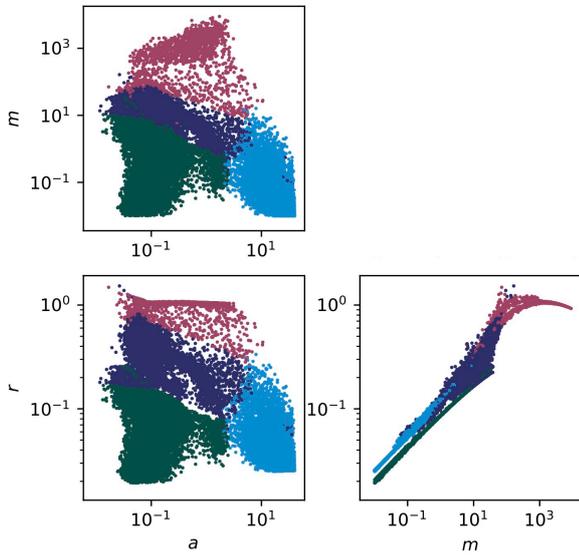
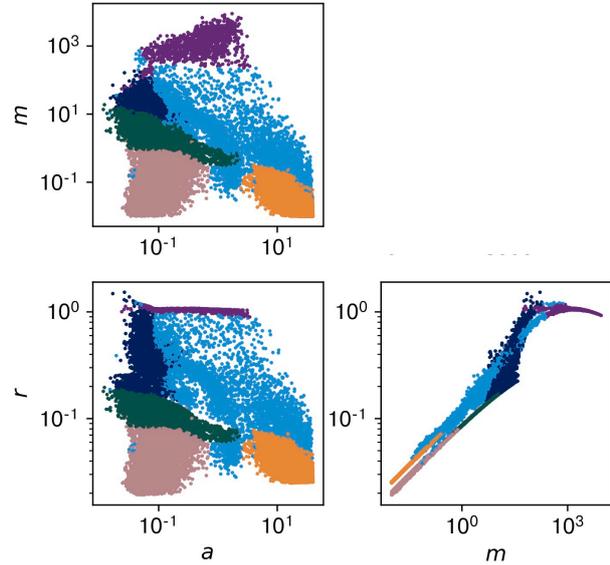
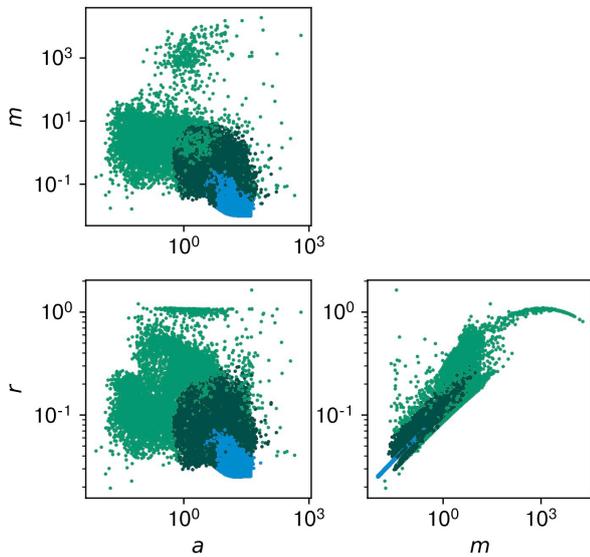
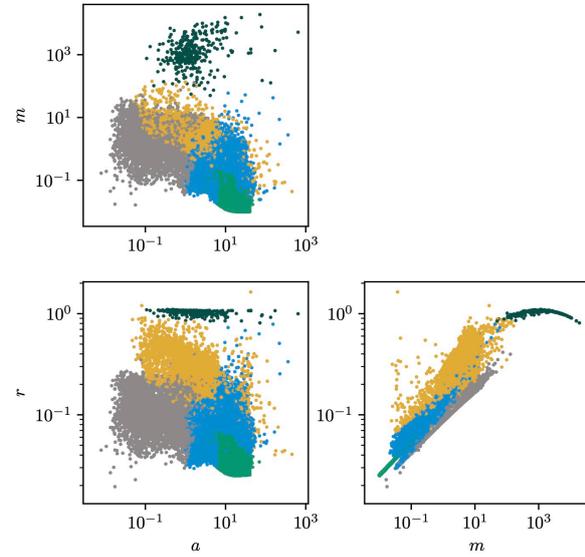
NG73: single-planet**a) GMM, $N=4$** **b) GMM, $N=6$** **NG76: multi-planet****c) GMM, $N=3$** **d) GMM, $N=5$** 

Figure A.2: Diagnostic plots for GMM clustering model selection. According to our validation metrics, the best candidate number of clusters are $N = 4$, $N = 6$ for NG73 and $N = 3$, $N = 5$ for NG76 (compare Fig. 3.1). The panels a)–d) show the clustering results of these choices. The models in a) ($N = 4$) and d) ($N = 5$) trace the over- and underdensities in the domain best and we consider them our nominal models.

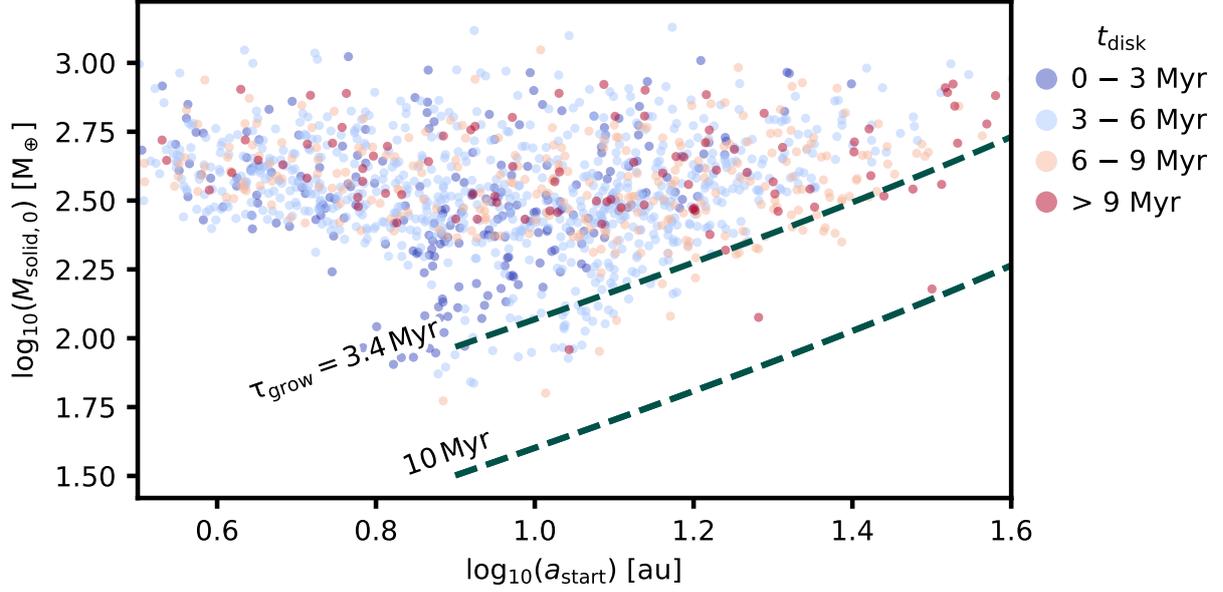


Figure A.3: Planetesimal disk mass and initial planet core position of giant planets forming in disks of different lifetimes. Markers show the distribution of all planets classified as “cluster 3: giant planet” in $a_{\text{start}} - M_{\text{solid},0}$ space, color-coded by the lifetime of their nascent disk. We overplot isolines of planetesimal masses corresponding to specific growth timescales τ_{grow} for reaching a core mass of $10 M_{\oplus}$. Giant planet growth is limited by the disk lifetime, and the formation of giant planets far out requires high planetesimal masses and long lifetimes.

total solid disk mass as a function of orbital distance 1. for different planetesimal isolation masses and 2. for different growth timescales. I derived these properties in Sect. 2.2.

Figure A.3 shows the cluster of giant planets in the space spanned by two important initial disk properties, a_{start} and $M_{\text{solid},0}$. The colors correspond to different lifetimes of the protoplanetary disk in which they formed. Most giants grow (and survive) in disks with lifetimes 3 Myr – 6 Myr. Only long-living disks enable formation of giant planets at low solid disk masses and large orbital distances. In short-lived disks, there is only a narrow region of embryo starting positions where giant planets grow at low planetesimal surface densities.

B

Appendix to Chapter 5

B.1 Joint fit posteriors

Figure B.1 shows all combinations of one- and two-dimensional projections of the posterior space from our joint fit in Sect. 5.3.2. On the diagonal, we state for each parameter the median value and its 16th and 84th percentile.

There is some residual degeneracy in the $b - R_P/R_*$ plane. In previous fits that did not include *LCOGT* data, the distribution extended far into the range of large impact parameters and planet-to-star ratios, marking a degenerate solution. This effect is physically plausible: a larger planet with higher impact parameter can to some degree mimic a smaller one with lower impact parameter. In the joint fit shown in Fig. B.1, this degeneracy is lifted and both b and R_P/R_* are well constrained.

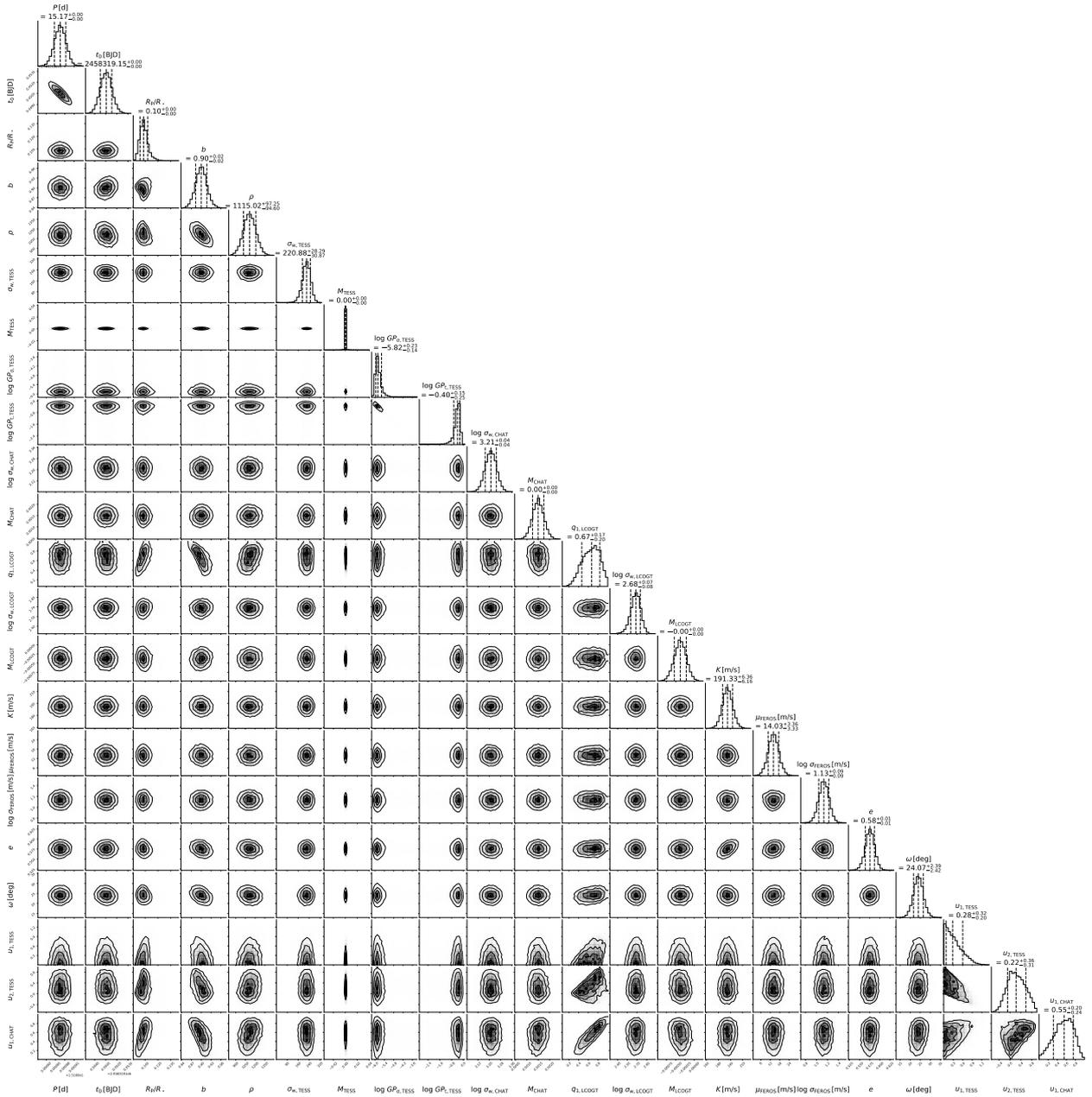


Figure B.1: Corner plot showing the posteriors of all parameters in our joint fit. The stated values represent 16th, 50th, and 84th percentiles, and we consider the median our 'best fit'. Individual sample points are not drawn to limit file size.

List of Abbreviations

| | |
|-----------------|---|
| ALMA | Atacama Large Millimeter/Submillimeter Array. 9 , 17 |
| CARMENES | Calar Alto high-Resolution search for M dwarfs with exo-Earths with Near-infrared and optical Echelle Spectrographs. 6 , 41 , 162 , 163 , 167 , 169 , 170 , 174 , 176 , 179–181 |
| CO | Carbon monoxide. 12 |
| EPOS | Exoplanet Population Observation Simulator. 40 |
| EUV | extreme-Ultraviolet. 15 |
| FIR | Far-Infrared. 7 |
| FUV | far-Ultraviolet. 15 |
| GLS | Generalized Lomb-Scargle periodogram. 31 |
| GMM | Gaussian Mixture Model. 62–67 , 76 , 183–185 , 187 |
| GTO | Guaranteed Time Observations. 163 , 169 , 180 |
| HD | Hydrogen deuteride. 12 |
| HST | Hubble Space Telescope. 15 |
| ISM | Interstellar medium. 8 |
| JWST | James Webb Space Telescope. 35 , 155 , 158 , 181 |
| KBO | Kuiper belt object. 21 |
| LS | Lomb-Scargle periodogram. 31 |
| MCMC | Markov Chain Monte Carlo. 34 , 178 |
| MIR | Mid-Infrared. 7 |
| MMSN | Minimum mass solar nebula. 42 |
| MOA-II | Microlensing Observations in Astrophysics II. 39 |
| MPIA | Max Planck Institute for Astronomy. 57 |
| NGPPS | The New Generation Planetary Population Synthesis. 57 , 59 |
| NIR | Near-Infrared. 7 , 163 |

| | |
|--------------|--|
| PLATO | PLAnetary Transits and Oscillations of stars. 180 |
| PMS | Pre-main-sequence. 8 |
| ppm | parts per million. 32, 155 |
| RV | Radial Velocity. 6, 29–31, 34–38, 162–165, 167, 169, 172, 179 |
| SED | Spectral energy distribution. 7, 16 |
| SFD | Size frequency distribution. 21 |
| SMA | Submillimeter Array. 17 |
| TESS | Transiting Exoplanet Survey Satellite. 6, 34, 35, 137, 147–149, 174, 180, 181, 193 |
| TIC | <i>TESS</i> Input Catalog. 34 |
| YSO | Young Stellar Object. 7, 8, 12, 16, 193 |
| ZAMS | Zero age main sequence. 8 |

List of Figures

| | | |
|------|--|----|
| 2.1 | Evolutionary stages of young stellar objects (YSO) | 8 |
| 2.2 | Schematic diagram of a protoplanetary disk | 9 |
| 2.3 | Schematic summary of the steps from dust to planets. | 18 |
| 2.4 | Accretion phases of giant planets in the core accretion paradigm | 25 |
| 2.5 | Light curve profile of a transiting planet | 32 |
| 2.6 | Field of view and scanning strategy of the <i>TESS</i> primary survey | 35 |
| 2.7 | Distributions of currently confirmed exoplanets | 38 |
| 2.8 | Valley in the radius distribution of small planets | 40 |
| 2.9 | Comparison of different population synthesis models and the physical pre- scriptions they include | 43 |
| 2.10 | Cumulative distribution of Monte Carlo parameters | 52 |
| 2.11 | Distribution of host star metallicities | 52 |
| 3.1 | Validation scores for Gaussian Mixture Models with different numbers of components N | 63 |
| 3.2 | Single-planet population clustering. | 64 |
| 3.3 | Multi-planet population clustering. | 65 |
| 3.4 | Model validation via generative models | 67 |
| 3.5 | Early time evolution of the clusters identified by the Gaussian Mixture Model | 68 |
| 3.6 | Confusion matrix of the planet classification. | 71 |
| 3.7 | Pairwise relationships between all disk parameters, sorted by cluster affi- liation | 72 |
| 3.8 | Relation between disk features and planet species | 73 |
| 3.9 | Radius and period distributions of Neptunes and (super-)Earths | 75 |
| 3.10 | Confusion of classifications for N -body population. | 76 |

| | | |
|------|---|-----|
| 3.12 | Relation between disk features and planet species | 78 |
| 3.13 | The four clusters of planets in $a_{\text{start}} - M_{\text{solid},0}$ space of their nascent proto-planetary disk | 82 |
| 4.1 | Probability of finding different numbers of super-Earths in a survey of 20 target stars | 91 |
| 4.2 | Time evolution of systems with cold Jupiters and super-Earths | 96 |
| 4.3 | Same as Fig. 4.2, but for systems containing neither super-Earths nor cold Jupiters. | 97 |
| 4.4 | Same as Fig. 4.2, but for systems containing super-Earths and no cold Jupiters | 98 |
| 4.5 | Same as Fig. 4.2, but for systems containing cold Jupiters and no super-Earths | 99 |
| 4.6 | Occurrence map of the synthetic population | 100 |
| 4.7 | Occurrences of the synthetic population in the mass-period plane | 101 |
| 4.8 | Distribution of planet radii in the synthetic population | 102 |
| 4.9 | Dependence of host star metallicity on planet radius | 104 |
| 4.10 | Initial conditions for different populations | 107 |
| 4.11 | Correlation Map of disk properties and planet occurrence rates | 108 |
| 4.12 | Mean planet occurrence per system as a function of initial solid disk mass . | 109 |
| 4.13 | Fractions of planet hosts in the synthetic population of survived planets . . | 110 |
| 4.14 | Observed and theoretical conditional probability $P(\text{CJ} \text{SE})$ | 111 |
| 4.15 | Eccentricity and period distributions of all giant planets that ever formed, regardless of their survival | 114 |
| 4.16 | Mass distribution of giant planets in cold Jupiter-hosting systems with and without super-Earths | 115 |
| 4.17 | Same as Fig. 4.15, but for all super-Earths that ever formed during the systems' histories | 116 |
| 4.18 | Water ice mass fractions of planet cores in the different system classes . . . | 117 |
| 4.19 | Distribution of ice mass fractions in the cores of super-Earths with and without cold Jupiter companion | 118 |
| 4.20 | Mass-radius diagram of inner super-Earths with and without cold Jupiters . | 119 |
| 4.21 | Multiplicity for different host star metallicities | 121 |
| 4.22 | Mass distribution of giant planets in low- and high-metallicity systems . . . | 122 |
| 4.23 | Normalized frequency of multiplicities for different planet types | 123 |
| 4.24 | Normalized frequency of super-Earth multiplicity in systems with and without cold Jupiter | 124 |
| 4.25 | Schematic relations between solid disk mass, core ice, and system architecture | 129 |

| | | |
|------|---|-----|
| 4.26 | Orbital range of outermost super-Earth and innermost cold Jupiter plotted by host star metallicity | 132 |
| 5.1 | Photometry for TIC 237913194 | 143 |
| 5.2 | Radial velocity time series for TIC 237913194 | 144 |
| 5.3 | Generalized Lomb-Scargle periodograms of radial velocity time series and common activity indicators | 146 |
| 5.4 | Posterior distributions of planetary parameters | 150 |
| 5.5 | Evolution of the planetary equilibrium temperature in the case of instantaneous heat adjustment for different albedos α and re-radiation factors β | 154 |
| 5.6 | Orbit aspect ratio and orientation | 155 |
| 5.7 | Comparison to other well-characterized warm Jupiters | 156 |
| 6.1 | Stellar mass distributions of the observed and simulated populations | 165 |
| 6.2 | Planet occurrence of the <i>NGM</i> population and <i>CARM</i> ₁₂₅ planets in minimum mass-period space | 166 |
| 6.3 | As Fig. 6.2, but partitioned according to host star mass. | 168 |
| 6.4 | Minimum mass distribution of short-orbit ($P < 600$ d) <i>NGM</i> planets and of the detected planets in the <i>CARM</i> ₁₂₅ sample, normalized to planets per star | 169 |
| 6.5 | Minimum mass distribution of simulated planets and of the sensitivity-weighted <i>CARM</i> ₁₂₅ planets | 170 |
| 7.1 | Observed and simulated M dwarf planets | 179 |
| A.1 | Diagnostic plots for clustering method selection | 186 |
| A.2 | Diagnostic plots for GMM clustering model selection. | 187 |
| A.3 | Planetesimal disk mass and initial planet core position of giant planets forming in disks of different lifetimes | 188 |
| B.1 | Corner plot showing the posteriors of all parameters in our joint fit | 190 |

List of Tables

| | | |
|-----|--|-----|
| 2.1 | Model parameters for the <i>NG76</i> multi-planet population | 53 |
| 2.2 | Fraction of systems harboring different planet types in the <i>NGM</i> M dwarf planet population | 55 |
| 3.1 | Feature importances of disk properties | 72 |
| 3.2 | Best-fit parameters for the broken power law in Equation 3.3 | 81 |
| 4.1 | Planet Classifications | 93 |
| 4.2 | Initial conditions of different populations | 105 |
| 4.3 | Fractions of Stars hosting super-Earths and cold Jupiters | 110 |
| 5.1 | FEROS radial velocities and accompanying data for TIC 237913194 used in this chapter | 139 |
| 5.2 | Stellar properties of TIC 237913194 | 141 |
| 5.3 | Model evidences from RV fits for different models | 145 |
| 5.4 | Prior parameter distributions | 149 |
| 5.5 | Posterior parameters | 152 |
| 5.6 | Theoretical temperature constraints of TIC 237913194b in the course of one orbit | 154 |
| 6.1 | Synthetic planet population <i>NGM</i> | 164 |

List of Own Publications

Publications as first author

- **Schlecker, M.**; Pham, D.; Burn, R.; Alibert, Y.; Mordasini, C.; Emsenhuber, A.; Klahr, H.; Henning, T.; Mishra, L., *The New Generation Planetary Population Synthesis (NGPPS). V. Predetermination of planet types in global core accretion models*, ArXiv, 2021 (arXiv:2104.11750)
- **Schlecker, M.**; Mordasini, C.; Emsenhuber, A.; Klahr, H.; Henning, T.; Burn, R.; Alibert, Y.; Benz, W., *The New Generation Planetary Population Synthesis (NGPPS). III. Warm super-Earths and cold Jupiters: A weak occurrence correlation, but with a strong architecture-composition link*, ArXiv, 2020 (arXiv:2007.05563)
- **Schlecker, M.**; Kossakowski, D.; Brahm, R.; Espinoza, N.; Henning, T.; Carone, L.; Molaverdikhani, K.; Trifonov, T.; Mollière, P.; Hobson, M. J.; Jordán, A.; Rojas, F. I.; Klahr, H.; Sarkis, P.; Bakos, G.; Bhatti, W.; Osip, D.; Suc, V.; Ricker, G.; Vanderspek, R.; Latham, D. W.; Seager, S.; Winn, J. N.; Jenkins, J. M.; Vezie, M.; Villaseñor, J. N.; Rose, M. E.; Rodriguez, D. R.; Rodriguez, J. E.; Quinn, S. N.; Shporer, A., doi:10.3847/1538-3881/abbe03, *A Highly Eccentric Warm Jupiter Orbiting TIC 237913194*, AJ, **160**, 275, 2020 (arXiv:2010.03570)

Publications as contributing author cited in this thesis

- Gibbs, A.; Bixel, A.; Rackham, B. V.; Apai, D.; **Schlecker, M.**; Espinoza, N.; Mancini, L.; Chen, W.; Henning, T.; Gabor, P.; Boyle, R.; Perez Chavez, J.; Mousseau, A.; Dietrich, J.; Jay Socia, Q.; Ip, W.; Ngeow, C.; Tsai, A.; Bhandare, A.; Marian, V.; Baehr, H.; Brown, S.; Häberle, M.; Keppler, M.; Molaverdikhani, K.; Sarkis, P., doi:10.3847/1538-3881/ab7926, *EDEN: Sensitivity Analysis and Transiting Planet Detection Limits for Nearby Late Red Dwarfs*, AJ, **159**, 169, 2020 (arXiv:2002.10017)

- Stock, S.; Nagel, E.; Kemmer, J.; Passegger, V. M.; Reffert, S.; Quirrenbach, A.; Caballero, J. A.; Czesla, S.; Béjar, V. J. S.; Cardona, C.; Díez-Alonso, E.; Herrero, E.; Lalitha, S.; **Schlecker, M.**; Tal-Or, L.; Rodríguez, E.; Rodríguez-López, C.; Ribas, I.; Reiners, A.; Amado, P. J.; Bauer, F. F.; Bluhm, P.; Cortés-Contreras, M.; González-Cuesta, L.; Dreizler, S.; Hatzes, A. P.; Henning, T.; Jeffers, S. V.; Kaminski, A.; Kürster, M.; Lafarga, M.; López-González, M. J.; Montes, D.; Morales, J. C.; Pedraz, S.; Schöfer, P.; Schweitzer, A.; Trifonov, T.; Zapatero Osorio, M. R.; Zechmeister, M., doi:10.1051/0004-6361/202038820, *The CARMENES search for exoplanets around M dwarfs. Three temperate-to-warm super-Earths*, A&A, **643**, 2020 (arXiv:2010.00474)
- Nowak, G.; Luque, R.; Parviainen, H.; Pallé, E.; Molaverdikhani, K.; Béjar, V. J. S.; Lillo-Box, J.; Rodríguez-López, C.; Caballero, J. A.; Zechmeister, M.; Passegger, V. M.; Cifuentes, C.; Schweitzer, A.; Narita, N.; Cale, B.; Espinoza, N.; Murgas, F.; Hidalgo, D.; Zapatero Osorio, M. R.; Pozuelos, F. J.; Aceituno, F. J.; Amado, P. J.; Barkaoui, K.; Barrado, D.; Bauer, F. F.; Benkhaldoun, Z.; Caldwell, D. A.; Casasayas Barris, N.; Chaturvedi, P.; Chen, G.; Collins, K. A.; Collins, K. I.; Cortés-Contreras, M.; Crossfield, I. J. M.; de León, J. P.; Díez Alonso, E.; Dreizler, S.; El Mufti, M.; Esparza-Borges, E.; Essack, Z.; Fukui, A.; Gaidos, E.; Gillon, M.; Gonzales, E. J.; Guerra, P.; Hatzes, A.; Henning, T.; Herrero, E.; Hesse, K.; Hirano, T.; Howell, S. B.; Jeffers, S. V.; Jehin, E.; Jenkins, J. M.; Kaminski, A.; Kemmer, J.; Kielkopf, J. F.; Kossakowski, D.; Kotani, T.; Kürster, M.; Lafarga, M.; Latham, D. W.; Law, N.; Lissauer, J. J.; Lodieu, N.; Madrigal-Aguado, A.; Mann, A. W.; Massey, B.; Matson, R. A.; Matthews, E.; Montañés-Rodríguez, P.; Montes, D.; Morales, J. C.; Mori, M.; Nagel, E.; Oshagh, M.; Pedraz, S.; Plavchan, P.; Pollacco, D.; Quirrenbach, A.; Reffert, S.; Reiners, A.; Ribas, I.; Ricker, G. R.; Rose, M. E.; **Schlecker, M.**; Schlieder, J. E.; Seager, S.; Stangret, M.; Stock, S.; Tamura, M.; Tanner, A.; Teske, J.; Trifonov, T.; Twicken, J. D.; Vanderspek, R.; Watanabe, D.; Wittrock, J.; Ziegler, C.; Zohrabi, F., doi:10.1051/0004-6361/202037867, *The CARMENES search for exoplanets around M dwarfs. Two planets on opposite sides of the radius gap transiting the nearby M dwarf LTT 3780*, A&A, **642**, 2020 (arXiv:2003.01140)
- Kossakowski, D.; Espinoza, N.; Brahm, R.; Jordán, A.; Henning, T.; Rojas, F.; Kürster, M.; Sarkis, P.; **Schlecker, M.**; Pozuelos, F. J.; Barkaoui, K.; Jehin, E.; Gillon, M.; Matthews, E.; Horch, E. P.; Ciardi, D. R.; Crossfield, I. J. M.; Gonzales, E.; Howell, S. B.; Matson, R.; Schlieder, J.; Jenkins, J.; Ricker, G.; Seager, S.; Winn, J. N.; Li, J.; Rose, M. E.; Smith, J. C.; Dynes, S.; Morgan, E.; Villasenor, J. N.; Charbonneau, D.; Jaffe, T.; Yu, L.; Bakos, G.; Bhatti, W.; Bouchy, F.; Collins, K. A.; Collins, K. I.; Csubry,

Z.; Evans, P.; Jensen, E. L. N.; Lovis, C.; Marmier, M.; Nielsen, L. D.; Osip, D.; Pepe, F.; Relles, H. M.; Ségransan, D.; Shporer, A.; Stockdale, C.; Suc, V.; Turner, O.; Udry, S., doi:10.1093/mnras/stz2433, *TOI-150b and TOI-163b: two transiting hot Jupiters, one eccentric and one inflated, revealed by TESS near and at the edge of the JWST CVZ*, MNRAS, **490**, 1094, 2019 (arXiv:1906.09866)

- Morales, J. C.; Mustill, A. J.; Ribas, I.; Davies, M. B.; Reiners, A.; Bauer, F. F.; Kossakowski, D.; Herrero, E.; Rodríguez, E.; López-González, M. J.; Rodríguez-López, C.; Béjar, V. J. S.; González-Cuesta, L.; Luque, R.; Pallé, E.; Perger, M.; Baroch, D.; Johansen, A.; Klahr, H.; Mordasini, C.; Anglada-Escudé, G.; Caballero, J. A.; Cortés-Contreras, M.; Dreizler, S.; Lafarga, M.; Nagel, E.; Passegger, V. M.; Reffert, S.; Rosich, A.; Schweitzer, A.; Tal-Or, L.; Trifonov, T.; Zechmeister, M.; Quirrenbach, A.; Amado, P. J.; Guenther, E. W.; Hagen, H. -.; Henning, T.; Jeffers, S. V.; Kaminski, A.; Kürster, M.; Montes, D.; Seifert, W.; Abellán, F. J.; Abril, M.; Aceituno, J.; Aceituno, F. J.; Alonso-Floriano, F. J.; Ammler-von Eiff, M.; Antona, R.; Arroyo-Torres, B.; Az-zaro, M.; Barrado, D.; Becerril-Jarque, S.; Benítez, D.; Berdiñas, Z. M.; Bergond, G.; Brinkmüller, M.; del Burgo, C.; Burn, R.; Calvo-Ortega, R.; Cano, J.; Cárdenas, M. C.; Cardona Guillén, C.; Carro, J.; Casal, E.; Casanova, V.; Casasayas-Barris, N.; Chaturvedi, P.; Cifuentes, C.; Claret, A.; Colomé, J.; Czesla, S.; Díez-Alonso, E.; Dorda, R.; Emsenhuber, A.; Fernández, M.; Fernández-Martín, A.; Ferro, I. M.; Fuhrmeister, B.; Galadí-Enríquez, D.; Gallardo Cava, I.; García Vargas, M. L.; Garcia-Piquer, A.; Gesa, L.; González-Álvarez, E.; González Hernández, J. I.; González-Peinado, R.; Guàrdia, J.; Guijarro, A.; de Guindos, E.; Hatzes, A. P.; Hauschildt, P. H.; Hedrosa, R. P.; Hermelo, I.; Hernández Arabi, R.; Hernández Otero, F.; Hintz, D.; Holgado, G.; Huber, A.; Huke, P.; Johnson, E. N.; de Juan, E.; Kehr, M.; Kemmer, J.; Kim, M.; Klüter, J.; Klutsch, A.; Labarga, F.; Labiche, N.; Lalitha, S.; Lampón, M.; Lara, L. M.; Launhardt, R.; Lázaro, F. J.; Lizon, J. -.; Llamas, M.; Lodieu, N.; López del Fresno, M.; López Salas, J. F.; López-Santiago, J.; Magán Madinabeitia, H.; Mall, U.; Mancini, L.; Mandel, H.; Marfil, E.; Marín Molina, J. A.; Martín, E. L.; Martín-Fernández, P.; Martín-Ruiz, S.; Martínez-Rodríguez, H.; Marvin, C. J.; Mirabet, E.; Moya, A.; Naranjo, V.; Nelson, R. P.; Nortmann, L.; Nowak, G.; Ofir, A.; Pascual, J.; Pavlov, A.; Pedraz, S.; Pérez Medialdea, D.; Pérez-Calpena, A.; Perryman, M. A. C.; Rabaza, O.; Ramón Ballesta, A.; Rebolo, R.; Redondo, P.; Rix, H. -.; Rodler, F.; Rodríguez Trinidad, A.; Sabotta, S.; Sadegi, S.; Salz, M.; Sánchez-Blanco, E.; Sánchez Carrasco, M. A.; Sánchez-López, A.; Sanz-Forcada, J.; Sarkis, P.; Sarmiento, L. F.; Schäfer, S.; **Schlecker, M.**; Schmitt, J. H. M. M.; Schöfer, P.; Solano, E.; Sota, A.;

Stahl, O.; Stock, S.; Stuber, T.; Stürmer, J.; Suárez, J. C.; Taberner, H. M.; Tulloch, S. M.; Veredas, G.; Vico-Linares, J. I.; Vilardell, F.; Wagner, K.; Winkler, J.; Wolthoff, V.; Yan, F.; Zapatero Osorio, M. R., doi:10.1126/science.aax3198, *A giant exoplanet orbiting a very-low-mass star challenges planet formation models*, *Science*, **365**, 1441, 2019 (arXiv:1909.12174)

- Luque, R.; Pallé, E.; Kossakowski, D.; Dreizler, S.; Kemmer, J.; Espinoza, N.; Burt, J.; Anglada-Escudé, G.; Béjar, V. J. S.; Caballero, J. A.; Collins, K. A.; Collins, K. I.; Cortés-Contreras, M.; Díez-Alonso, E.; Feng, F.; Hatzes, A.; Hellier, C.; Henning, T.; Jeffers, S. V.; Kaltenegger, L.; Kürster, M.; Madden, J.; Molaverdikhani, K.; Montes, D.; Narita, N.; Nowak, G.; Ofir, A.; Oshagh, M.; Parviainen, H.; Quirrenbach, A.; Reffert, S.; Reiners, A.; Rodríguez-López, C.; **Schlecker, M.**; Stock, S.; Trifonov, T.; Winn, J. N.; Zapatero Osorio, M. R.; Zechmeister, M.; Amado, P. J.; Anderson, D. R.; Batalha, N. E.; Bauer, F. F.; Bluhm, P.; Burke, C. J.; Butler, R. P.; Caldwell, D. A.; Chen, G.; Crane, J. D.; Dragomir, D.; Dressing, C. D.; Dynes, S.; Jenkins, J. M.; Kaminski, A.; Klahr, H.; Kotani, T.; Lafarga, M.; Latham, D. W.; Lewin, P.; McDermott, S.; Montañés-Rodríguez, P.; Morales, J. C.; Murgas, F.; Nagel, E.; Pedraz, S.; Ribas, I.; Ricker, G. R.; Rowden, P.; Seager, S.; Shectman, S. A.; Tamura, M.; Teske, J.; Twicken, J. D.; Vanderspeck, R.; Wang, S. X.; Wohler, B., doi:10.1051/0004-6361/201935801, *Planetary system around the nearby M dwarf GJ 357 including a transiting, hot, Earth-sized planet optimal for atmospheric characterization*, *A&A*, **628**, 2019 (arXiv:1904.12818)

Publications as contributing author not cited in this thesis

- Hobson, M. J.; Brahm, R.; Jordán, A.; Espinoza, N.; Kossakowski, D.; Henning, T.; Rojas, F.; **Schlecker, M.**; Sarkis, P.; Trifonov, T.; Thorngren, D.; Binnenfeld, A.; Shahaf, S.; Zucker, S.; Ricker, G. R.; Latham, D. W.; Seager, S.; Winn, J. N.; Jenkins, J. M.; Addison, B.; Bouchy, F.; Bowler, B. P.; Briegal, J. T.; Bryant, E. M.; Collins, K. A.; Daylan, T.; Grieves, N.; Horner, J.; Huang, C.; Kane, S. R.; Kielkopf, J.; McLean, B.; Mengel, M. W.; Nielsen, L. D.; Okumura, J.; Jones, M.; Plavchan, P.; Shporer, A.; Smith, A. M. S.; Tilbrook, R.; Tinney, C. G.; Twicken, J. D.; Udry, S.; Unger, N.; West, R.; Wittenmyer, R. A.; Wohler, B.; Torres, P.; Wright, D. J., doi:10.3847/1538-3881/abeaa1, *A Transiting Warm Giant Planet around the Young Active Star TOI-201*, *AJ*, **161**, 235, 2021 (arXiv:2103.02685)
- Addison, B. C.; Wright, D. J.; Nicholson, B. A.; Cale, B.; Mocnik, T.; Huber, D.; Plavchan, P.; Wittenmyer, R. A.; Vanderburg, A.; Chaplin, W. J.; Chontos, A.; Clark, J. T.; Eastman, J. D.; Ziegler, C.; Brahm, R.; Carter, B. D.; Clerte, M.; Espinoza, N.;

Horner, J.; Bentley, J.; Jordán, A.; Kane, S. R.; Kielkopf, J. F.; Laychock, E.; Mengel, M. W.; Okumura, J.; Stassun, K. G.; Bedding, T. R.; Bowler, B. P.; Burnelis, A.; Blanco-Cuaresma, S.; Collins, M.; Crossfield, I.; Davis, A. B.; Evensberget, D.; Heitzmann, A.; Howell, S. B.; Law, N.; Mann, A. W.; Marsden, S. C.; Matson, R. A.; O'Connor, J. H.; Shporer, A.; Stevens, C.; Tinney, C. G.; Tylor, C.; Wang, S.; Zhang, H.; Henning, T.; Kossakowski, D.; Ricker, G.; Sarkis, P.; **Schlecker, M.**; Torres, P.; Vanderpek, R.; Latham, D. W.; Seager, S.; Winn, J. N.; Jenkins, J. M.; Mireles, I.; Rowden, P.; Pepper, J.; Daylan, T.; Schlieder, J. E.; Collins, K. A.; Collins, K. I.; Tan, T.; Ball, W. H.; Basu, S.; Buzasi, D. L.; Campante, T. L.; Corsaro, E.; González-Cuesta, L.; Davies, G. R.; de Almeida, L.; do Nascimento, Jose-Dias, Jr.; García, R. A.; Guo, Z.; Handberg, R.; Hekker, S.; Hey, D. R.; Kallinger, T.; Kawaler, S. D.; Kayhan, C.; Kuzlewicz, J. S.; Lund, M. N.; Lyttle, A.; Mathur, S.; Miglio, A.; Mosser, B.; Nielsen, M. B.; Serenelli, A. M.; Aguirre, V. S.; Themeßl, N., doi:10.1093/mnras/staa3960, *TOI-257b (HD 19916b): a warm sub-saturn orbiting an evolved F-type star*, MNRAS, **502**, 3704, 2021 (arXiv:2001.07345)

- Dreizler, S.; Crossfield, I. J. M.; Kossakowski, D.; Plavchan, P.; Jeffers, S. V.; Kemmer, J.; Luque, R.; Espinoza, N.; Pallé, E.; Stassun, K.; Matthews, E.; Cale, B.; Caballero, J. A.; **Schlecker, M.**; Lillo-Box, J.; Zechmeister, M.; Lalitha, S.; Reiners, A.; Soubkiou, A.; Bitsch, B.; Zapatero Osorio, M. R.; Chaturvedi, P.; Hatzes, A. P.; Ricker, G.; Vanderpek, R.; Latham, D. W.; Seager, S.; Winn, J.; Jenkins, J. M.; Aceituno, J.; Amado, P. J.; Barkaoui, K.; Barbieri, M.; Batalha, N. M.; Bauer, F. F.; Benneke, B.; Benkhaldoun, Z.; Beichman, C.; Berberian, J.; Burt, J.; Butler, R. P.; Caldwell, D. A.; Chintada, A.; Chontos, A.; Christiansen, J. L.; Ciardi, D. R.; Cifuentes, C.; Collins, K. A.; Collins, K. I.; Combs, D.; Cortés-Contreras, M.; Crane, J. D.; Daylan, T.; Dragomir, D.; Esparza-Borges, E.; Evans, P.; Feng, F.; Flowers, E. E.; Fukui, A.; Fulton, B.; Furlan, E.; Gaidos, E.; Geneser, C.; Giacalone, S.; Gillon, M.; Gonzales, E.; Gorjian, V.; Hellier, C.; Hidalgo, D.; Howard, A. W.; Howell, S.; Huber, D.; Isaacson, H.; Jehin, E.; Jensen, E. L. N.; Kaminski, A.; Kane, S. R.; Kawauchi, K.; Kielkopf, J. F.; Klahr, H.; Kosiarek, M. R.; Kreidberg, L.; Kürster, M.; Lafarga, M.; Livingston, J.; Louie, D.; Mann, A.; Madrigal-Aguado, A.; Matson, R. A.; Mocnik, T.; Morales, J. C.; Muirhead, P. S.; Murgas, F.; Nandakumar, S.; Narita, N.; Nowak, G.; Oshagh, M.; Parviainen, H.; Passegger, V. M.; Pollacco, D.; Pozuelos, F. J.; Quirrenbach, A.; Reefe, M.; Ribas, I.; Robertson, P.; Rodríguez-López, C.; Rose, M. E.; Roy, A.; Schweitzer, A.; Schlieder, J.; Shectman, S.; Tanner, A.; Şenavcı, H. V.; Teske, J.; Twicken, J. D.; Villaseñor, J.; Wang, S. X.; Weiss, L. M.; Wittrock, J.; Yilmaz, M.; Zohrabi, F., doi:10.1051/0004-6361/202038016, *The*

CARMENES search for exoplanets around M dwarfs. LP 714-47 b (TOI 442.01): populating the Neptune desert, *A&A*, **644**, 2020 (arXiv:2011.01716)

- Brahm, R.; Nielsen, L. D.; Wittenmyer, R. A.; Wang, S.; Rodriguez, J. E.; Espinoza, N.; Jones, M. I.; Jordán, A.; Henning, T.; Hobson, M.; Kossakowski, D.; Rojas, F.; Sarkis, P.; **Schlecker, M.**; Trifonov, T.; Shahaf, S.; Ricker, G.; Vanderspek, R.; Latham, D. W.; Seager, S.; Winn, J. N.; Jenkins, J. M.; Addison, B. C.; Bakos, G.; Bhatti, W.; Bayliss, D.; Berlind, P.; Bieryla, A.; Bouchy, F.; Bowler, B. P.; Briceño, C.; Brown, T. M.; Bryant, E. M.; Caldwell, D. A.; Charbonneau, D.; Collins, K. A.; Davis, A. B.; Esquerdo, G. A.; Fulton, B. J.; Guerrero, N. M.; Henze, C. E.; Hogan, A.; Horner, J.; Huang, C. X.; Irwin, J.; Kane, S. R.; Kielkopf, J.; Mann, A. W.; Mazeh, T.; McCormac, J.; McCully, C.; Mengel, M. W.; Mireles, I.; Okumura, J.; Plavchan, P.; Quinn, S. N.; Rabus, M.; Saesen, S.; Schlieder, J. E.; Segransan, D.; Shiao, B.; Shporer, A.; Siverd, R. J.; Stassun, K. G.; Suc, V.; Tan, T.; Torres, P.; Tinney, C. G.; Udry, S.; Vanzi, L.; Vezie, M.; Vines, J. I.; Vuckovic, M.; Wright, D. J.; Yahalomi, D. A.; Zapata, A.; Zhang, H.; Ziegler, C., doi:10.3847/1538-3881/abba3b, *TOI-481 b and TOI-892 b: Two Long-period Hot Jupiters from the Transiting Exoplanet Survey Satellite*, *AJ*, **160**, 235, 2020 (arXiv:2009.08881)
- Kemmer, J.; Stock, S.; Kossakowski, D.; Kaminski, A.; Molaverdikhani, K.; **Schlecker, M.**; Caballero, J. A.; Amado, P. J.; Astudillo-Defru, N.; Bonfils, X.; Ciardi, D.; Collins, K. A.; Espinoza, N.; Fukui, A.; Hirano, T.; Jenkins, J. M.; Latham, D. W.; Matthews, E. C.; Narita, N.; Pallé, E.; Parviainen, H.; Quirrenbach, A.; Reiners, A.; Ribas, I.; Ricker, G.; Schlieder, J. E.; Seager, S.; Vanderspek, R.; Winn, J. N.; Almenara, J. M.; Béjar, V. J. S.; Bluhm, P.; Bouchy, F.; Boyd, P.; Christiansen, J. L.; Cifuentes, C.; Cloutier, R.; Collins, K. I.; Cortés-Contreras, M.; Crossfield, I. J. M.; Crouzet, N.; de Leon, J. P.; Della-Rose, D. D.; Delfosse, X.; Dreizler, S.; Esparza-Borges, E.; Essack, Z.; Forveille, T.; Figueira, P.; Galadí-Enríquez, D.; Gan, T.; Glidden, A.; Gonzales, E. J.; Guerra, P.; Harakawa, H.; Hatzes, A. P.; Henning, T.; Herrero, E.; Hodapp, K.; Hori, Y.; Howell, S. B.; Ikoma, M.; Isogai, K.; Jeffers, S. V.; Kürster, M.; Kawauchi, K.; Kimura, T.; Klagyivik, P.; Kotani, T.; Kurokawa, T.; Kusakabe, N.; Kuzuhara, M.; Lafarga, M.; Livingston, J. H.; Luque, R.; Matson, R.; Morales, J. C.; Mori, M.; Muirhead, P. S.; Murgas, F.; Nishikawa, J.; Nishiumi, T.; Omiya, M.; Reffert, S.; Rodríguez López, C.; Santos, N. C.; Schöfer, P.; Schwarz, R. P.; Shiao, B.; Tamura, M.; Terada, Y.; Twicken, J. D.; Ueda, A.; Vievard, S.; Watanabe, N.; Zechmeister, M., doi:10.1051/0004-6361/202038967, *Discovery of a hot, transiting, Earth-sized planet and a second temperate, non-transiting planet around the M4 dwarf GJ 3473 (TOI-488)*, *A&A*, **642**, 2020 (arXiv:2009.10432)

- Jahnke, K.; Fendt, C.; Fouesneau, M.; Georgiev, I.; Herbst, T.; Kaasinen, M.; Kosakowski, D.; Rybizki, J.; **Schlecker, M.**; Seidel, G.; Henning, T.; Kreidberg, L.; Rix, H., doi:10.1038/s41550-020-1202-4, *An astronomical institute's perspective on meeting the challenges of the climate crisis*, *Nature Astronomy*, **4**, 812, 2020 (arXiv:2009.11307)
- Bluhm, P.; Luque, R.; Espinoza, N.; Pallé, E.; Caballero, J. A.; Dreizler, S.; Livingston, J. H.; Mathur, S.; Quirrenbach, A.; Stock, S.; Van Eylen, V.; Nowak, G.; López, E. D.; Csizmadia, S.; Zapatero Osorio, M. R.; Schöfer, P.; Lillo-Box, J.; Oshagh, M.; González-Álvarez, E.; Amado, P. J.; Barrado, D.; Béjar, V. J. S.; Cale, B.; Chaturvedi, P.; Cifuentes, C.; Cochran, W. D.; Collins, K. A.; Collins, K. I.; Cortés-Contreras, M.; Díez Alonso, E.; El Mufti, M.; Ercolino, A.; Fridlund, M.; Gaidos, E.; García, R. A.; Georgieva, I.; González-Cuesta, L.; Guerra, P.; Hatzes, A. P.; Henning, T.; Herrero, E.; Hidalgo, D.; Isopi, G.; Jeffers, S. V.; Jenkins, J. M.; Jensen, E. L. N.; Ká bath, P.; Kaminski, A.; Kemmer, J.; Korth, J.; Kosakowski, D.; Kürster, M.; Lafarga, M.; Mallia, F.; Montes, D.; Morales, J. C.; Morales-Calderón, M.; Murgas, F.; Narita, N.; Passegger, V. M.; Pedraz, S.; Persson, C. M.; Plavchan, P.; Rauer, H.; Redfield, S.; Reffert, S.; Reiners, A.; Ribas, I.; Ricker, G. R.; Rodríguez-López, C.; Santos, A. R. G.; Seager, S.; **Schlecker, M.**; Schweitzer, A.; Shan, Y.; Soto, M. G.; Subjak, J.; Tal-Or, L.; Trifonov, T.; Vanaverbeke, S.; Vanderspek, R.; Wittrock, J.; Zechmeister, M.; Zohrabi, F., doi:10.1051/0004-6361/202038160, *Precise mass and radius of a transiting super-Earth planet orbiting the M dwarf TOI-1235: a planet in the radius gap?*, *A&A*, **639**, 2020 (arXiv:2004.06218)
- Espinoza, N.; Brahm, R.; Henning, T.; Jordán, A.; Dorn, C.; Rojas, F.; Sarkis, P.; Kosakowski, D.; **Schlecker, M.**; Díaz, M. R.; Jenkins, J. S.; Aguilera-Gomez, C.; Jenkins, J. M.; Twicken, J. D.; Collins, K. A.; Lissauer, J.; Armstrong, D. J.; Adibekyan, V.; Barrado, D.; Barros, S. C. C.; Battley, M.; Bayliss, D.; Bouchy, F.; Bryant, E. M.; Cooke, B. F.; Demangeon, O. D. S.; Dumusque, X.; Figueira, P.; Giles, H.; Lillo-Box, J.; Lovis, C.; Nielsen, L. D.; Pepe, F.; Pollacco, D.; Santos, N. C.; Sousa, S. G.; Udry, S.; Wheatley, P. J.; Turner, O.; Marmier, M.; Ségransan, D.; Ricker, G.; Latham, D.; Seager, S.; Winn, J. N.; Kielkopf, J. F.; Hart, R.; Wingham, G.; Jensen, E. L. N.; Helminiak, K. G.; Tokovinin, A.; Briceño, C.; Ziegler, C.; Law, N. M.; Mann, A. W.; Daylan, T.; Doty, J. P.; Guerrero, N.; Boyd, P.; Crossfield, I.; Morris, R. L.; Henze, C. E.; Chacon, A. D., doi:10.1093/mnras/stz3150, *HD 213885b: a transiting 1-d-period super-Earth with an Earth-like composition around a bright ($V = 7.9$) star unveiled by TESS*, *MNRAS*, **491**, 2982, 2020 (arXiv:1903.07694)

- Tiedemann, L.; Breunig, E.; Burwitz, V.; Fürmetz, M.; Hartner, G.; Kink, W.; Menz, B.; Predehl, P.; Röser, H.; **Schlecker, M.**; Valsecchi, G., doi:10.1117/12.2023973, *The development of the μ ROSI X-ray telescope*, SPIE, **8859**, 885905, 2013

Other contributions

- **Schlecker, M.**, *lcps: Light curve pre-selection*, Astrophysics Source Code Library, 2018
- **Schlecker, M.**, doi:10.5281/zenodo.221659, *Irregular Variability in Kepler Photometry*, Master's Thesis, 2016

Bibliography

- Adams, F. C. & Laughlin, G. 2006, *The Astrophysical Journal*, 649, 1004
- Adams, F. C., Meyer, M. R., & Adams, A. D. 2021, 1
- Akaike, H. 1973, *Proceedings of the 2nd International Symposium on Information Theory*, 267–281
- Alcalá, J. M., Natta, A., Manara, C. F., et al. 2014, *Astronomy and Astrophysics*, 561
- Alessi, M., Pudritz, R. E., & Cridland, A. J. 2017, *Monthly Notices of the Royal Astronomical Society*, 464, 428
- Alessi, M., Pudritz, R. E., & Cridland, A. J. 2020, *Monthly Notices of the Royal Astronomical Society*, 493, 1013
- Alexander, R., Pascucci, I., Andrews, S., Armitage, P., & Cieza, L. 2014, *Protostars and Planets VI*
- Alibert, Y. 2019, A new metric to quantify the similarity between planetary systems - application to dimensionality reduction using T-SNE.
- Alibert, Y., Carron, F., Fortier, A., et al. 2013, *Astronomy and Astrophysics*, 558, 1
- Alibert, Y., Mordasini, C., & Benz, W. 2011, 63, 1
- Alibert, Y., Mordasini, C., Benz, W., & Winisdoerffer, C. 2005, *Astronomy & Astrophysics*, 434, 343
- Alibert, Y. & Venturini, J. 2019, *Astronomy & Astrophysics*, 626, A21
- Andre, P., Ward-Thompson, D., & Barsony, M. 1993, *Astrophysical Journal*, 406, 122

- Andrews, S. M. 2020, *Annual Review of Astronomy and Astrophysics*, 58, 483
- Andrews, S. M., Huang, J., Pérez, L. M., et al. 2018a, *The Astrophysical Journal*, 869, L41
- Andrews, S. M., Rosenfeld, K. A., Kraus, A. L., & Wilner, D. J. 2013, *Astrophysical Journal*, 771, 129
- Andrews, S. M., Terrell, M., Tripathi, A., et al. 2018b, *The Astrophysical Journal*, 865, 157
- Andrews, S. M. & Williams, J. P. 2005, *The Astrophysical Journal*, 631, 1134
- Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2009, *The Astrophysical Journal*, 700, 1502
- Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2010, *Astrophysical Journal*, 723, 1241
- Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, *Nature*, 536, 437
- Ankerst, M., Breunig, M. M., peter Kriegel, H., & Sander, J. 1999, in (ACM Press), 49–60
- Ansdell, M., Williams, J. P., Manara, C. F., et al. 2017, *The Astronomical Journal*, 153, 240
- Ansdell, M., Williams, J. P., Trapman, L., et al. 2018, *Alma survey of lupus protoplanetary disks II: Gas disk radii*
- Ansdell, M., Williams, J. P., van der Marel, N., et al. 2016, *The Astrophysical Journal*, 828, 46
- Armitage, P. J. 2007, arXiv e-prints
- Armitage, P. J. 2018, in *Handbook of Exoplanets* (Cham: Springer International Publishing), 2185–2203
- Armitage, P. J. 2019, *Saas-Fee Advanced Course*, 45, 1
- Asplund, M., Grevesse, N., & Sauval, A. 2005, in *Astronomical Society of the Pacific Conference Series*, Vol. 336, *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis*, ed. I. I. I. Barnes Thomas G. & F. N. Bash, 25
- Ataiee, S., Baruteau, C., Alibert, Y., & Benz, W. 2018, *Astronomy and Astrophysics*, 615, 1
- Avduevskij, V. S., Avduevsky, V. S., Marov, M. Y., et al. 1971, *Journal of Atmospheric Sciences*, 28, 263

- Bakos, G. A., Kovacs, G., Torres, G., et al. 2007, *The Astrophysical Journal*, 670, 826
- Balbus, S. A. & Hawley, J. F. 1991, *The Astrophysical Journal*, 376, 214
- Ballard, S. & Johnson, J. A. 2016, *The Astrophysical Journal*, 816, 66
- Bally, J., O'Dell, C. R., & McCaughrean, M. J. 2000, *The Astronomical Journal*, 119, 2919
- Baraffe, I., Homeier, D., Allard, F., & Chabrier, G. 2015, *A&A*, 577, A42
- Barbato, D., Pinamonti, M., Sozzetti, A., et al. 2020, 1
- Barbato, D., Sozzetti, A., Desidera, S., et al. 2018, *Astronomy & Astrophysics*, 615, A175
- Barclay, T., Pepper, J., & Quintana, E. V. 2018, *The Astrophysical Journal Supplement Series*, 239, 2
- Barenfeld, S. A., Carpenter, J. M., Ricci, L., & Isella, A. 2016, *The Astrophysical Journal*, 827, 142
- Barentsen, G., Hedges, C. L., De Miranda Cardoso, J. V., et al. 2019, in *American Astronomical Society Meeting Abstracts*, Vol. 233, *American Astronomical Society Meeting Abstracts #233*, 109.08
- Barnes, J. W. 2007, *Publications of the Astronomical Society of the Pacific*, 119, 986
- Barnes, R. 2015, *International Journal of Astrobiology*, 14, 321
- Barnes, R. 2017, *Celestial Mechanics and Dynamical Astronomy*, 129, 509
- Barranco, J. A., Pei, S., & Marcus, P. S. 2018
- Batalha, N. E., Mandell, A., Pontoppidan, K., et al. 2017, *Publications of the Astronomical Society of the Pacific*, 129, 064501
- Beatty, T. G., Pepper, J., Siverd, R. J., et al. 2012, *Astrophysical Journal Letters*, 756
- Beichman, C., Benneke, B., Knutson, H., et al. 2014, *Publications of the Astronomical Society of the Pacific*, 126, 1134
- Bell, K. R. & Lin, D. N. C. 1994, *The Astrophysical Journal*, 427, 987
- Benítez-Llambay, P., Masset, F., & Beaugé, C. 2011, *Astronomy and Astrophysics*, 528, 2

- Bennett, D. P., Akeson, R., Alibert, Y., et al. 2019, arXiv, 1
- Bennett, D. P., Ranc, C., & Fernandes, R. B. 2021, arXiv e-prints
- Benz, W., Ida, S., Alibert, Y., Lin, D., & Mordasini, C. 2014, in *Protostars and Planets VI*
- Bergin, E. A., Cleeves, L. I., Gorti, U., et al. 2013, *Nature*, 493, 644
- Birnstiel, T. & Andrews, S. M. 2014, *Astrophysical Journal*, 780
- Birnstiel, T., Dullemond, C. P., & Brauer, F. 2010, *Astronomy and Astrophysics*, 513, 1
- Birnstiel, T., Klahr, H., & Ercolano, B. 2012, *Astronomy and Astrophysics*, 539
- Bitsch, B., Izidoro, A., Johansen, A., et al. 2019a, *Astronomy and Astrophysics*, 623
- Bitsch, B., Lambrechts, M., & Johansen, A. 2015, *Astronomy and Astrophysics*, 582, C2
- Bitsch, B., Lambrechts, M., & Johansen, A. 2018a, *Astronomy and Astrophysics*, 609, 4
- Bitsch, B., Morbidelli, A., Johansen, A., et al. 2018b
- Bitsch, B., Raymond, S. N., & Izidoro, A. 2019b, *Astronomy and Astrophysics*, 624
- Bodenheimer, P., Hubickyj, O., & Lissauer, J. J. 2000, *Icarus*, 143, 2
- Bodenheimer, P. & Pollack, J. B. 1986, *Icarus*, 67, 391
- Bohlin, R., Savage, B., & Drake, J. 1978, *Astrophysical Journal*, 224, 132
- Boisse, I., Moutou, C., Vidal-Madjar, A., et al. 2009, *A&A*, 495, 959
- Bolton, S., Adriani, A., Adumitroaie, V., et al. 2017, *Science*, 356, 821
- Bonfils, X., Delfosse, X., Udry, S., et al. 2013, *Astronomy & Astrophysics*, 549, A109
- Bonomo, A. S., Desidera, S., Benatti, S., et al. 2017, *Astronomy & Astrophysics*, 602, A107
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, 327, 977
- Borucki, W. J. & Summers, A. L. 1984, *Icarus*, 58, 121
- Boss, A. P. 1995, *Science*, 267, 360
- Bouvier, J., Alencar, S. H. P., Harries, T. J., Johns-Krull, C. M., & Romanova, M. M. 2007, in *Protostars and Planets V*, ed. B. Reipurth, D. Jewitt, & K. Keil, 479

- Box, G. E. P. & Cox, D. R. 1964, *Journal of the Royal Statistical Society: Series B (Methodological)*, 26, 211
- Brahm, R., Espinoza, N., Jordán, A., et al. 2019, *AJ*, 158, 45
- Brahm, R., Jordán, A., & Espinoza, N. 2017a, *Publications of the Astronomical Society of the Pacific*, 129, 1
- Brahm, R., Jordán, A., Hartman, J., & Bakos, G. 2017b, *Monthly Notices of the Royal Astronomical Society*, 467, 971
- Brasseur, C. E., Phillip, C., Fleming, S. W., Mullally, S. E., & White, R. L. 2019, *Astrocut: Tools for creating cutouts of TESS images*
- Breiman, L. 1996, *Machine Learning*, 24, 123
- Breiman, L. 2001, *Machine Learning*, 45, 5
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. 1984, *Classification and regression trees*, The Wadsworth statistics/probability series (Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software)
- Bressan, A., Marigo, P., Girardi, L., et al. 2012, *MNRAS*, 427, 127
- Brewer, J. M., Wang, S., Fischer, D. A., & Foreman-Mackey, D. 2018, *The Astrophysical Journal*, 867, L3
- Brown, T. M., Baliber, N., Bianco, F. B., et al. 2013, *Publications of the Astronomical Society of the Pacific*, 125, 1031
- Brügger, N., Burn, R., Coleman, G. A., Alibert, Y., & Benz, W. 2020, *Astronomy and Astrophysics*, 640
- Bruno, G. 1584, *De l'infinito universo et mundi*
- Bryan, M. L., Knutson, H. A., Howard, A. W., et al. 2016, *The Astrophysical Journal*, 821, 89
- Bryan, M. L., Knutson, H. A., Lee, E. J., et al. 2019, *The Astronomical Journal*, 157, 52
- Bryson, S., Coughlin, J., Batalha, N. M., et al. 2020, *The Astronomical Journal*, 159, 279
- Buchhave, L. A., Bitsch, B., Johansen, A., et al. 2018, *The Astrophysical Journal*, 856, 37

- Buchhave, L. A., Bizzarro, M., Latham, D. W., et al. 2014, *Nature*, 509, 593
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, *Astronomy and Astrophysics*, 564, 1
- Burke, C. J. 2008, *The Astrophysical Journal*, 679, 1566
- Burn, R. 2020, PhD thesis, University of Bern
- Burn, R., Marboeuf, U., Alibert, Y., & Benz, W. 2019, *Astronomy & Astrophysics*, 629, A64
- Burn, R., Schlecker, M., Mordasini, C., et al. 2021
- Caliński, T. & Harabasz, J. 1974, *Communications in Statistics*, 3, 1
- Cameron, A. 1978, *Moon and Planets*, 18, 5
- Cañas, C. I., Stefansson, G., Monson, A. J., et al. 2019, *The Astrophysical Journal*, 877, L29
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245
- Carmona, A., Van Der Plas, G., Van Den Ancker, M. E., et al. 2011, *Astronomy and Astrophysics*, 533, 1
- Carone, L., Keppens, R., & Decin, L. 2014, *MNRAS*, 445, 930
- Carone, L., Mollière, P., Zhou, Y., et al. 2020, arXiv e-prints, 1
- Carone, L. & Pätzold, M. 2007, *Planet. Space Sci.*, 55, 643
- Castelli, F. & Kurucz, R. L. 2004, ArXiv e-prints, astro
- Cavanaugh, J. E. & Neath, A. A. 2019, *Wiley Interdisciplinary Reviews: Computational Statistics*, 11, e1460
- Chabrier, G. 2003, *Publications of the Astronomical Society of the Pacific*, 115, 763
- Chabrier, G., Baraffe, I., Leconte, J., Gallardo, J., & Barman, T. 2009in (*AIP*), 102–111
- Chambers, J. 2006, *Icarus*, 180, 496
- Chambers, J. 2018, *The Astrophysical Journal*, 865, 30
- Chambers, J. E. 1999, *MNRAS*, 304, 793
- Chambers, J. E. 2009, *Astrophysical Journal*, 705, 1206

- Chambers, J. E. 2012, Mercury: A software package for orbital dynamics, *Astrophysics Source Code Library*
- Chambers, J. E., Wetherill, G. W., & Boss, A. P. 1996, *Icarus*, 119, 261
- Chaparro Molano, G., Bautista, F., Miguel, Y., et al. 2019, *Proceedings of the International Astronomical Union*, 152
- Charbonneau, D. & Deming, D. 2007, *Exoplanet Task Force (AAAC)*, 1
- Chauvin, G., Desidera, S., Lagrange, A. M., et al. 2017, *Astronomy and Astrophysics*, 605, L9
- Chen, C., Liaw, A., & Breiman, L. 2004
- Chen, J. & Kipping, D. 2016, *The Astrophysical Journal*, 834, 17
- Cheng, W. H., Lee, M. H., & Peale, S. J. 2014, *Icarus*, 233, 242
- Chiang, E. & Laughlin, G. 2013, *Monthly Notices of the Royal Astronomical Society*, 431, 3444
- Chiang, E. I. & Goldreich, P. 1997, *The Astrophysical Journal*, 490, 368
- Cimerman, N. P., Kuiper, R., & Ormel, C. W. 2017, *Monthly Notices of the Royal Astronomical Society*, 471, 4662
- Claret, A. 2000, *Astronomy and Astrophysics*, 363
- Clarke, C. J., Gendrin, A., Sotomayor, M., et al. 2001, *Monthly Notices of the Royal Astronomical Society*, 328, 485
- Cloutier, R., Artigau, É., Delfosse, X., et al. 2018, *The Astronomical Journal*, 155, 93
- Coleman, G. A., Leleu, A., Alibert, Y., & Benz, W. 2019, *Astronomy and Astrophysics*, 631
- Coleman, G. A. L. & Nelson, R. P. 2014, *MNRAS*, 445, 479
- Cortes, C. & Vapnik, V. 1995, *Machine learning*, 20, 273
- Crida, A., Morbidelli, A., & Masset, F. 2006, *Icarus*, 181, 587
- Cridland, A. J., Pudritz, R. E., & Alessi, M. 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 3274

- Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, *Publications of the Astronomical Society of the Pacific*, 120, 531
- Cumming, A., Marcy, G. W., & Butler, R. P. 1999, *The Astrophysical Journal*, 526, 890
- Cuzzi, J. N. & Zahnle, K. J. 2004, *The Astrophysical Journal*, 614, 490
- D'Alessio, P., Calvet, N., Hartmann, L., Lizano, S., & Canto, J. 1999, *The Astrophysical Journal*, 527, 893
- D'Alessio, P., Canto, J., Calvet, N., & Lizano, S. 1998, *The Astrophysical Journal*, 500, 411
- Damiani, C. & Lanza, A. F. 2015, *A&A*, 574, A39
- D'Angelo, G., Kley, W., & Henning, T. 2003, *The Astrophysical Journal*, 586, 540
- David, T. J., Contardo, G., Sandoval, A., et al. 2020
- Davies, D. L. & Bouldin, D. W. 1979, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1, 224
- Dawson, R. I. & Johnson, J. A. 2018, *Annual Review of Astronomy and Astrophysics*, 56, 175
- Diaz-Cordoves, J. & Gimenez, A. 1992, *Astronomy and Astrophysics*, 259
- Dittkrist, K. M., Mordasini, C., Klahr, H., Alibert, Y., & Henning, T. 2014, *Astronomy and Astrophysics*, 567
- Draine, B. T. 2003, in *Annual Review of Astronomy and Astrophysics*, Vol. 41, 241–289
- Drążkowska, J. & Alibert, Y. 2017, *Astronomy & Astrophysics*, 608, A92
- Drążkowska, J., Alibert, Y., & Moore, B. 2016, *Astronomy and Astrophysics*, 594, 1
- Dreizler, S., Jeffers, S. V., Rodríguez, E., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 493, 536
- Dressing, C. D. & Charbonneau, D. 2013, *The Astrophysical Journal*, 767, 95+
- Dressing, C. D. & Charbonneau, D. 2015, *Astrophysical Journal*, 807, 45
- Dullemond, C. & Monnier, J. 2010, *Annual Review of Astronomy and Astrophysics*, 48, 205

- Dullemond, C. P., van Zadelhoff, G., & Natta, A. 2002, *Astronomy and Astrophysics*, 389
- Dumusque, X., Santos, N. C., Udry, S., Lovis, C., & Bonfils, X. 2011, *Astronomy and Astrophysics*, 527
- Duncan, D. K., Vaughan, A. H., Wilson, O. C., et al. 1991, *ApJS*, 76, 383
- Dürmann, C. & Kley, W. 2017, *Astronomy and Astrophysics*, 598, 1
- Eastman, J., Gaudi, B. S., & Agol, E. 2013, *Publications of the Astronomical Society of the Pacific*, 125, 83
- Eggleton, P. P., Kiseleva, L. G., & Hut, P. 1998, *The Astrophysical Journal*, 499, 853
- Eisner, J. A., Graham, J. R., Akeson, R. L., & Najita, J. 2009, *Astrophysical Journal*, 692, 309
- Eisner, J. A., Hillenbrand, L. A., White, R. J., Akeson, R. L., & Sargent, A. I. 2005, *The Astrophysical Journal*, 623, 952
- Eisner, J. A., Monnier, J. D., Woillez, J., et al. 2010, *Astrophysical Journal*, 718, 774
- Elbakyan, V., Johansen, A., Lambrechts, M., Akimkin, V., & Vorobyov, E. 2020, *Astronomy & Astrophysics*, 1
- Emsenhuber, A., Mordasini, C., Burn, R., et al. 2020a, arXiv e-prints
- Emsenhuber, A., Mordasini, C., Burn, R., et al. 2020b, arXiv e-prints
- Espinoza, N., Brahm, R., Henning, T., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 491, 2982
- Espinoza, N., Brahm, R., Jordán, A., et al. 2016, *The Astrophysical Journal*, 830, 43
- Espinoza, N., Hartman, J. D., Bakos, G. Á., et al. 2019a, *AJ*, 158, 63
- Espinoza, N. & Jordán, A. 2016, *Monthly Notices of the Royal Astronomical Society*, 457, 3573
- Espinoza, N., Kossakowski, D., & Brahm, R. 2019b, *MNRAS*, 490, 2262
- Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in (AAAI Press), 226–231
- Facchini, S., Birnstiel, T., Bruderer, S., & Van Dishoeck, E. F. 2017, *Astronomy and Astrophysics*, 605, 1

- Fedele, D., van den Ancker, M. E., Henning, T., Jayawardhana, R., & Oliveira, J. M. 2010, *Astronomy and Astrophysics*, 510, A72
- Fernandes, R. B., Mulders, G. D., Pascucci, I., Mordasini, C., & Emsenhuber, A. 2019, *The Astrophysical Journal*, 874, 81
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, 398, 1601
- Ferraz-Mello, S., Rodríguez, A., & Hussmann, H. 2008, *Celestial Mechanics and Dynamical Astronomy*, 101, 171
- Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, *Publications of the Astronomical Society of the Pacific*, 128, 1
- Fischer, D. A. & Valenti, J. 2005, *The Astrophysical Journal*, 622, 1102
- Fjeldbo, G., Fjeldbo, W. C., & Eshleman, V. R. 1966, *J. Geophys. Res.*, 71, 2307
- Flock, M., Fromang, S., Turner, N. J., & Benisty, M. 2016, *The Astrophysical Journal*, 827, 144
- Flock, M., Fromang, S., Turner, N. J., & Benisty, M. 2017, *The Astrophysical Journal*, 835, 230
- Flock, M., Turner, N. J., Mulders, G. D., et al. 2019a, *Astronomy and Astrophysics*, 630, 1
- Flock, M., Turner, N. J., Mulders, G. D., et al. 2019b, *Astronomy & Astrophysics*
- Ford, E. B., Quinn, S. N., & Veras, D. 2008, *The Astrophysical Journal*, 678, 1407
- Foreman-Mackey, D. 2016, *The Journal of Open Source Software*, 24
- Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *The Astronomical Journal*, 154, 220
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306
- Foreman-Mackey, D., Hogg, D. W., & Morton, T. D. 2014, *The Astrophysical Journal*, 795, 64
- Fortier, A., Alibert, Y., Carron, F., Benz, W., & Dittkrist, K. M. 2013, *Astronomy and Astrophysics*, 549, 1

Fortier, A., Benvenuto, O. G., & Brunini, A. 2009, *Astronomy and Astrophysics*, 500, 1249

Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, *The Astrophysical Journal*, 659, 1661

Fortney, J. J. & Nettelmann, N. 2010, *Space Science Reviews*, 152, 423

Fouchet, L., Alibert, Y., Mordasini, C., & Benz, W. 2012, *Astronomy & Astrophysics*, 540, A107

France, K., Duvvuri, G., Egan, H., et al. 2020, arXiv

Freikh, R., Jang, H., Murray-Clay, R. A., & Petrovich, C. 2019, *The Astrophysical Journal*, 884, L47

Fressin, F., Torres, G., Charbonneau, D., et al. 2013, *Astrophysical Journal*, 766

Frink, S., Mitchell, D. S., Quirrenbach, A., et al. 2002, *The Astrophysical Journal*, 576, 478

Fulton, B. J. & Petigura, E. A. 2018, *The Astronomical Journal*, 156, 264

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, *The Astronomical Journal*, 154, 109

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *Astronomy and Astrophysics*, 1, 1

Gaidos, E. & Mann, A. W. 2013, *Astrophysical Journal*, 762

Gail, H. P. 2004, *Astronomy and Astrophysics*, 413, 571

Galiano, A., Palomba, E., D'Amore, M., et al. 2020, *Icarus*, 351, 113959

Gaudi, B. S. & Winn, J. N. 2007, *The Astrophysical Journal*, 655, 550

Gerbig, K., Murray-Clay, R. A., Klahr, H., & Baehr, H. 2020, *The Astrophysical Journal*, 895, 91

Gibbs, A., Bixel, A., Rackham, B. V., et al. 2020, *The Astronomical Journal*, 159, 169

Gilbert, G. J. & Fabrycky, D. C. 2020, *The Astronomical Journal*, 159

Gill, S., Wheatley, P. J., Cooke, B. F., et al. 2020, *Astrophysical Journal Letters*, 898, L11

Gillon, M., Jehin, E., Lederer, S. M., et al. 2016, *Nature*, 533, 221

- Gillon, M., Jehin, E., Queloz, D., et al. 2018, in *Ground-based and Airborne Telescopes VII*, ed. R. Gilmozzi, H. K. Marshall, & J. Spyromilio (SPIE), 49
- Gillon, M., Triaud, A. H., Demory, B. O., et al. 2017, *Nature*, 542, 456
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, *The Astrophysical Journal*, 825, 29
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, *Monthly Notices of the Royal Astronomical Society*, 476, 759
- Gladman, B. 1993, *Icarus*, 106, 247
- Goldreich, P., Lithwick, Y., & Sari, R. 2004, *Annual Review of Astronomy and Astrophysics*, 42, 549
- Goldreich, P. & Soter, S. 1966, *Icarus*, 5, 375
- Goldreich, P. & Tremaine, S. 1979, *The Astrophysical Journal*, 233, 857
- Goldreich, P. & Tremaine, S. 1980, *ApJ*, 241, 425
- Gomes da Silva, J., Santos, N. C., Bonfils, X., et al. 2011, *A&A*, 534, A30
- Gonzalez, G. 1997, *Monthly Notices of the Royal Astronomical Society*, 285, 403
- Gorti, U. & Hollenbach, D. 2009, *Astrophysical Journal*, 690, 1539
- Greenberg, R., Bottke, W. F., Carusi, A., & Valsecchi, G. B. 1991, *Icarus*, 94, 98
- Guerrero, N. M., Seager, S., Huang, C. X., et al. 2021, arXiv e-prints
- Guillot, T. 2005, *Annual Review of Earth and Planetary Sciences*, 33, 493
- Guilloteau, S., Dutrey, A., Piétu, V., & Boehler, Y. 2011, *Astronomy and Astrophysics*, 529, 1
- Gullbring, E., Hartmann, L., Briceno, C., & Calvet, N. 1998, *The Astrophysical Journal*, 492, 323
- Günther, H. M. 2013, *Astronomische Nachrichten*, 334, 67
- Günther, M. N. & Daylan, T. 2021, *The Astrophysical Journal Supplement Series*, 254, 13
- Gupta, A. & Schlichting, H. E. 2019, *Monthly Notices of the Royal Astronomical Society*, 487, 24

Haghighipour, N., Vogt, S. S., Paul Butler, R., et al. 2010, *Astrophysical Journal*, 715, 271

Haisch, Jr., K. E., Lada, E. A., & Lada, C. J. 2001, *The Astrophysical Journal*, 553, L153

Hansen, B. M. 2010, *Astrophysical Journal*, 723, 285

Hansen, B. M. 2012, *Astrophysical Journal*, 757

Hansen, B. M. 2017, *Monthly Notices of the Royal Astronomical Society*, 467, 1531

Hardegree-Ullman, K. K., Cushing, M. C., Muirhead, P. S., & Christiansen, J. L. 2019

Hart, M. H. 1979, *Icarus*, 37, 351

Hartley, H. O. 1958, *Biometrics*, 14, 174

Hartman, J. D., Bakos, G. Á., Bayliss, D., et al. 2019, *Astronomical Journal*, 157, 55

Hartmann, L., Calvet, N., Gullbring, E., & D'Alessio, P. 1998, *The Astrophysical Journal*, 495, 385

Hasegawa, Y. & Pudritz, R. E. 2011, *Monthly Notices of the Royal Astronomical Society*, 417, 1236

Hayashi, C. 1981, *Progress of Theoretical Physics Supplement*, 70, 35

He, M. Y., Ford, E. B., & Ragozzine, D. 2020a, 17, 1

He, M. Y., Ford, E. B., Ragozzine, D., & Carrera, D. 2020b

Henning, T. 2010, *Annual Review of Astronomy and Astrophysics*, 48, 21

Henning, T. & Semenov, D. 2013, *Chemical Reviews*, 113, 9016

Herbst, E. & Klemperer, W. 1973, *The Astrophysical Journal*, 185, 505

Herman, M. K., Zhu, W., & Wu, Y. 2019, *The Astronomical Journal*, 157, 248

Hillenbrand, L. A., Strom, S. E., Vrba, F. J., & Keene, J. 1992, *The Astrophysical Journal*, 397, 613

Ho, S. & Turner, E. L. 2011, *Astrophysical Journal*, 739

Ho, T. 1998, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 832

- Hobson, M. J. & Gomez, M. 2017, *New Astronomy*, 55, 1
- Hollenbach, D., Johnstone, D., Lizano, S., & Shu, F. 1994, *The Astrophysical Journal*, 428, 654
- Horner, J. & Jones, B. W. 2010, *Astronomy & Geophysics*, 51, 6.16
- Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, *The Astrophysical Journal Supplement Series*, 201, 15
- Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010, *Science*, 330, 653
- Hsu, D. C., Ford, E. B., Ragozzine, D., & Ashby, K. 2019, *The Astronomical Journal*, 158, 109
- Hsu, D. C., Ford, E. B., Ragozzine, D., & Morehead, R. C. 2018, *The Astronomical Journal*, 155, 205
- Hsu, D. C., Ford, E. B., & Terrien, R. 2020, *Monthly Notices of the Royal Astronomical Society*, 498, 2249
- Huang, C., Wu, Y., & Triaud, A. H. M. J. 2016, *The Astrophysical Journal*, 825, 98
- Hueso, R. & Guillot, T. 2005, *Astronomy and Astrophysics*, 442, 703
- Hyodo, R., Ida, S., & Charnoz, S. 2019, *Astronomy and Astrophysics*, 629
- Ida, S., Guillot, T., Hyodo, R., Okuzumi, S., & Youdin, A. N. 2020
- Ida, S. & Lin, D. N. C. 2004a, *The Astrophysical Journal*, 604, 388
- Ida, S. & Lin, D. N. C. 2004b, *The Astrophysical Journal*, 616, 567
- Ida, S. & Lin, D. N. C. 2005, *The Astrophysical Journal*, 626, 1045
- Ida, S. & Lin, D. N. C. 2008, *The Astrophysical Journal*, 673, 487
- Ida, S. & Lin, D. N. C. 2010, *Astrophysical Journal*, 719, 810
- Ida, S. & Makino, J. 1993, *Icarus*, 106, 210
- Ikoma, M., Nakazawa, K., & Emori, H. 2000, *The Astrophysical Journal*, 537, 1013
- Inaba, S., Tanaka, H., Nakazawa, K., Wetherill, G. W., & Kokubo, E. 2001, *Icarus*, 149, 235

- Irwin, J., Hodgkin, S., Aigrain, S., et al. 2008, *Monthly Notices of the Royal Astronomical Society*, 384, 675
- Isella, A., Carpenter, J. M., & Sargent, A. I. 2009, *Astrophysical Journal*, 701, 260
- Isella, A., Testi, L., Natta, A., et al. 2007, *Astronomy & Astrophysics*, 469, 213
- Izidoro, A., Bitsch, B., Raymond, S. N., et al. 2019, arXiv, 1
- Izidoro, A., Raymond, S. N., Morbidelli, A., Hersant, F., & Pierens, A. 2015, *Astrophysical Journal Letters*, 800, L22
- Jain, A. K. & Dubes, R. C. 1988, *Algorithms for clustering data*
- J Jeans, J. H. 1902, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 199, 1
- Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, in *Proc. SPIE, Vol. 9913, Software and Cyberinfrastructure for Astronomy IV*, 99133E
- Jin, S. & Mordasini, C. 2018, *The Astrophysical Journal*, 853, 163
- Jin, S., Mordasini, C., Parmentier, V., et al. 2014, *Astrophysical Journal*, 795
- Johansen, A., Ida, S., & Brasser, R. 2019, *Astronomy and Astrophysics*, 622, 1
- Johansen, A., Oishi, J. S., Low, M. M. M., et al. 2007, *Nature*, 448, 1022
- Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, *Publications of the Astronomical Society of the Pacific*, 122, 905
- Jolliff, B. L. & Robinson, M. S. 2019, *Physics Today*, 72, 44
- Jones, M. I., Brahm, R., Espinoza, N., et al. 2019, *A&A*, 625, A16
- Jordán, A., Brahm, R., Bakos, G. Á., et al. 2014, *AJ*, 148, 29
- Jordán, A., Brahm, R., Espinoza, N., et al. 2020, *AJ*, 159, 145
- Jordán, A., Brahm, R., Espinoza, N., et al. 2020, *The Astronomical Journal*, 159, 145
- Kaltenegger, L. & Sasselov, D. 2011, *Astrophysical Journal Letters*, 736, 2
- Kanagawa, K. D., Tanaka, H., & Szuszkiewicz, E. 2018, *The Astrophysical Journal*, 861, 140

Kaufer, A., Stahl, O., Tubbesing, S., et al. 1999, *The Messenger*, 95, 8

Kawahara, H. & Masuda, K. 2019, *The Astronomical Journal*, 157, 218

Kawashima, Y. & Ikoma, M. 2019, *ApJ*, 877, 109

Kemmer, J., Stock, S., Kossakowski, D., et al. 2020, *Astronomy and Astrophysics*, 642, 1

Kenyon, S. J. & Hartmann, L. 1987, *The Astrophysical Journal*, 323, 714

Keppeler, M., Benisty, M., Müller, A., et al. 2018, *Astronomy and Astrophysics*, 617, 1

Ketchen, D. J. & Shook, C. L. 1996, *Strategic Management Journal*, 17, 441

King, G. W. & Wheatley, P. J. 2021, *Monthly Notices of the Royal Astronomical Society: Letters*, 501, L28

Kipping, D. M. 2013, *Monthly Notices of the Royal Astronomical Society*, 435, 2152

Klahr, H. & Bodenheimer, P. 2003, *Astrophys. J.*, 582

Klahr, H. & Bodenheimer, P. 2006, *The Astrophysical Journal*, 639, 432

Klahr, H. & Schreiber, A. 2020, *The Astrophysical Journal*, 901, 54

Klahr, H. & Schreiber, A. 2021, *The Astrophysical Journal*, 911, 9

Kley, W. & Nelson, R. P. 2012, *Annual Review of Astronomy and Astrophysics*, 50, 211

Klinglesmith, D. A. & Sobieski, S. 1970, *The Astronomical Journal*, 75, 175

Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, *The Astrophysical Journal*, 713, L79

Kokubo, E. & Ida, S. 1995, *Icarus*, 114, 247

Kokubo, E. & Ida, S. 1996, *Icarus*, 123, 180

Kokubo, E. & Ida, S. 1998, *Icarus*, 131, 171

Kokubo, E. & Ida, S. 2000, *Icarus*, 143, 15

Kokubo, E. & Ida, S. 2002, *The Astrophysical Journal*, 581, 666

Kopal, Z. 1950, *Harvard College Observatory Circular*, 454

- Kossakowski, D., Espinoza, N., Brahm, R., et al. 2019, *Monthly Notices of the Royal Astronomical Society*, 19, 1
- Kozai, Y. 1962, *The Astronomical Journal*, 67, 591
- Kreidberg, L. 2015, *PASP*, 127, 1161
- Krivov, A. V. & Wyatt, M. C. 2021, *Monthly Notices of the Royal Astronomical Society*, 500, 718
- Kruijssen, J. M. D., Longmore, S. N., & Chevance, M. 2020, *The Astrophysical Journal*, 905, L18
- kumar Kopparapu, R., Wolf, E. T., Haqq-Misra, J., et al. 2016, *The Astrophysical Journal*, 819, 84
- Kürster, M., Endl, M., Rouesnel, F., et al. 2003, *Astronomy & Astrophysics*, 403, 1077
- Kurtovic, N. T., Pinilla, P., Long, F., et al. 2021, *Astronomy & Astrophysics*, 645, A139
- Lada, C. J. 1987, in *Star Forming Regions*, ed. M. Peimbert & J. Jugaku, Vol. 115, 1
- Lada, C. J. & Lada, E. A. 2003, *Annual Review of Astronomy and Astrophysics*, 41, 57
- Lada, C. J. & Wilking, B. A. 1984, *Astrophysical Journal*, 287, 610
- Lagrange, A. M., Bonnefoy, M., Chauvin, G., et al. 2010, *Science*, 329, 57
- Lainey, V., Arlot, J.-E., Karatekin, Ö., & Van Hoolst, T. 2009, *Nature*, 459, 957
- Lainey, V., Casajus, L. G., Fuller, J., et al. 2020, *Nature Astronomy*
- Lambrechts, M. & Johansen, A. 2012, *Astronomy and Astrophysics*, 544, 1
- Lambrechts, M. & Johansen, A. 2014, *Astronomy and Astrophysics*, 572
- Lambrechts, M., Johansen, A., & Morbidelli, A. 2014, *Astronomy and Astrophysics*, 572, 1
- Lambrechts, M., Morbidelli, A., Jacobson, S. A., et al. 2019, *Astronomy & Astrophysics*, 627, A83
- Lammer, H., Selsis, F., Ribas, I., et al. 2003, *The Astrophysical Journal*, 598, L121
- Latham, D. W., Rowe, J. F., Quinn, S. N., et al. 2011, *Astrophysical Journal Letters*, 732, 2

- Lee, E. J. & Chiang, E. 2015, *Astrophysical Journal*, 811, 41
- Lemaître, G., Nogueira, F., & Aridas, C. K. 2017, *Journal of Machine Learning Research*, 18, 1
- Lenz, C. T. 2020, PhD thesis, University of Heidelberg
- Lenz, C. T., Klahr, H., & Birnstiel, T. 2019, *The Astrophysical Journal*, 874, 36
- Lewis, N. K., Knutson, H. A., Showman, A. P., et al. 2013, *Astrophysical Journal*, 766
- Lidov, M. L. 1962, *Planetary and Space Science*, 9, 719
- Lin, D. & Papaloizou, J. 1980, *Monthly Notices of the Royal Astronomical Society*, 191, 37
- Lin, D. N. C. & Papaloizou, J. 1979, *Monthly Notices of the Royal Astronomical Society*, 186, 799
- Lin, D. N. C. & Papaloizou, J. 1986, *The Astrophysical Journal*, 309, 846
- Lin, J. W., Lee, E. J., & Chiang, E. 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 4338
- Lissauer, J. J. 1987, *Icarus*, 69, 249
- Lissauer, J. J. 1993, *Annual Review of Astronomy and Astrophysics*, 31
- Lissauer, J. J., Ragozzine, D., Fabrycky, D. C., et al. 2011, *Astrophysical Journal, Supplement Series*, 197
- Liu, B., Lambrechts, M., Johansen, A., & Liu, F. 2019, 1265
- Liu, B., Lambrechts, M., Johansen, A., Pascucci, I., & Henning, T. 2020
- Liu, S. F., Hori, Y., Lin, D. N., & Asphaug, E. 2015, *Astrophysical Journal*, 812
- Lloyd, S. P. 1982, *IEEE Transactions on Information Theory*, 28, 129
- Lodders, K. 2003, *The Astrophysical Journal*, 591, 1220
- Lomb, N. R. 1976, *Astrophysics and Space Science*, 39, 447
- Long, F., Herczeg, G. J., Pascucci, I., et al. 2017, *The Astrophysical Journal*, 844, 99
- Lopez, E. D. & Fortney, J. J. 2013, *Astrophysical Journal*, 776

- Lopez, E. D. & Fortney, J. J. 2014, *Astrophysical Journal*, 792
- Lopez, E. D., Fortney, J. J., & Miller, N. 2012, *Astrophysical Journal*, 761
- Lubow, S. H. 1991, *The Astrophysical Journal*, 381, 259
- Luger, R., Agol, E., Foreman-Mackey, D., et al. 2019, *The Astronomical Journal*, 157, 64
- Luger, R., Sestovic, M., Kruse, E., et al. 2017, *Nature Astronomy*, 1, 1
- Lund, M. N., Handberg, R., Davies, G. R., Chaplin, W. J., & Jones, C. D. 2015, *Astrophysical Journal*, 806, 30
- Luque, R., Nowak, G., Pallé, E., et al. 2018, *Astronomy and Astrophysics*, 620, 1
- Luque, R., Trifonov, T., Reffert, S., et al. 2019, *Astronomy and Astrophysics*, 631, 4
- Lüst, R. 1952, *Zeitschrift Naturforschung Teil A*, 7, 87
- Lynden-Bell, D. & Pringle, J. E. 1974, *Monthly Notices of the Royal Astronomical Society*, 168, 603
- Macintosh, B., Graham, J. R., Barman, T., et al. 2015, *Science*, 350, 64
- MacQueen, J. 1967, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics* (Berkeley, Calif.: University of California Press), 281–297
- Mamajek, E. E., Usuda, T., Tamura, M., & Ishii, M. 2009, in *AIP Conference Proceedings* (AIP), 3–10
- Manara, C. F., Morbidelli, A., & Guillot, T. 2018, *Astronomy and Astrophysics*, 618, 1
- Manara, C. F., Mordasini, C., Testi, L., et al. 2019, *Astronomy & Astrophysics*, 631, L2
- Manara, C. F., Natta, A., Rosotti, G. P., et al. 2020, *Astronomy & Astrophysics*, 639, A58
- Manara, C. F., Rosotti, G., Testi, L., et al. 2016, *Astronomy & Astrophysics*, 3, 3
- Maoz, D. 2016, *Astrophysics in a Nutshell* (Princeton University Press)
- Marcy, G. W., Butler, R. P., Fischer, D., et al. 2001, *The Astrophysical Journal*, 556, 296
- Marleau, G.-D., Klahr, H., Kuiper, R., & Mordasini, C. 2017, *The Astrophysical Journal*, 836, 221

- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Science*, 322, 1348
- Masset, F. & Snellgrove, M. 2001, *Monthly Notices of the Royal Astronomical Society*, 320, 0
- Masuda, K., Winn, J. N., & Kawahara, H. 2020, *The Astronomical Journal*, 159, 38
- Matsakos, T. & Königl, A. 2016, *The Astrophysical Journal*, 820, L8
- Matsuyama, I., Johnstone, D., & Hartmann, L. 2003, *The Astrophysical Journal*, 582, 893
- Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv e-prints, 1
- Mayor, M. & Queloz, D. 1995, *Nature*, 378, 355
- Mazeh, T., Holczer, T., & Faigler, S. 2016, *Astronomy and Astrophysics*, 589
- McKee, C. F. & Ostriker, E. C. 2007, *Annual Review of Astronomy and Astrophysics*, 45, 565
- McLachlan, G. 1988, *Mixture models : inference and applications to clustering* (New York, N.Y: M. Dekker)
- McLaughlin, D. B. 1924, *ApJ*, 60, 22
- Méndez, A. & Rivera-Valentín, E. G. 2017, *ApJL*, 837, L1
- Méndez, A. & Rivera-Valentín, E. G. 2017, *The Astrophysical Journal*, 837, L1
- Miguel, Y., Cridland, A., Ormel, C. W., Fortney, J. J., & Ida, S. 2020, *Monthly Notices of the Royal Astronomical Society*, 491, 1998
- Miguel, Y. & Ida, S. 2016, *Icarus*, 266, 1
- Millholland, S., Wang, S., & Laughlin, G. 2017, 1
- Miotello, A., van Dishoeck, E. F., Williams, J. P., et al. 2016, *Astronomy and Astrophysics*, 599, 1
- Mizuno, H. 1980, *Progress of Theoretical Physics*, 64, 544
- Mizuno, H., Nakazawa, K., & Hayashi, C. 1978, *Progress of Theoretical Physics*, 60, 699
- Molaverdikhani, K., Henning, T., & Mollière, P. 2019a, *The Astrophysical Journal*, 883, 194

- Molaverdikhani, K., Henning, T., & Mollière, P. 2019b, *The Astrophysical Journal*, 873, 32
- Molaverdikhani, K., Henning, T., & Mollière, P. 2020, *The Astrophysical Journal*, 899, 53
- Moldenhauer, T. W., Kuiper, R., Kley, W., & Ormel, C. W. 2021, *Astronomy and Astrophysics*, 646, 1
- Mollière, P., Van Boekel, R., Bouwman, J., et al. 2017, *Astronomy and Astrophysics*, 600, 1
- Mollière, P., van Boekel, R., Dullemond, C., Henning, T., & Mordasini, C. 2015, *The Astrophysical Journal*, 813, 47
- Molyarova, T., Akimkin, V., Semenov, D., et al. 2017, *The Astrophysical Journal*, 849, 130
- Morales, J. C., Mustill, A. J., Ribas, I., et al. 2019, *Science*, 365, 1441
- Morbidelli, A., Bottke, W. F., Nesvorný, D., & Levison, H. F. 2009, *Icarus*, 204, 558
- Morbidelli, A., Lambrechts, M., Jacobson, S., & Bitsch, B. 2015, *Icarus*, 258, 418
- Morbidelli, A. & Nesvorny, D. 2012, *Astronomy and Astrophysics*, 546, 1
- Morbidelli, A., Nesvorny, D., Bottke, W. F., & Marchi, S. 2021, *Icarus*, 356, 1
- Mordasini, C. 2008, PhD thesis, Universitaet Bern
- Mordasini, C. 2018, in *Handbook of Exoplanets* (Cham: Springer International Publishing), 2425–2474
- Mordasini, C. 2020, *Astronomy and Astrophysics*, 638, 1
- Mordasini, C., Alibert, Y., & Benz, W. 2009a, *Astronomy and Astrophysics*, 501, 1139
- Mordasini, C., Alibert, Y., Benz, W., Klahr, H., & Henning, T. 2012a, *Astronomy and Astrophysics*, 541, 1
- Mordasini, C., Alibert, Y., Benz, W., & Naef, D. 2009b, *Astronomy and Astrophysics*, 501, 1161
- Mordasini, C., Alibert, Y., Georgy, C., et al. 2012b, *Astronomy and Astrophysics*, 547, A112
- Mordasini, C., Mollière, P., Dittkrist, K. M., Jin, S., & Alibert, Y. 2015, *International Journal of Astrobiology*, 14, 201

- Mordasini, C. A., Alibert, Y., Klahr, H. H., & Henning, T. 2012c, *Astronomy and Astrophysics*, 547, A111
- Moses, J. I., Line, M. R., Visscher, C., et al. 2013, *ApJ*, 777, 34
- Mulders, G. D. 2018, in *Handbook of Exoplanets* (Cham: Springer International Publishing), 2009–2034
- Mulders, G. D., O'Brien, D. P., Ciesla, F. J., Apai, D., & Pascucci, I. 2020, *The Astrophysical Journal*, 897, 72
- Mulders, G. D., Pascucci, I., & Apai, D. 2015a, *Astrophysical Journal*, 798
- Mulders, G. D., Pascucci, I., & Apai, D. 2015b, *Astrophysical Journal*, 814, 130
- Mulders, G. D., Pascucci, I., Apai, D., & Ciesla, F. J. 2018, *The Astronomical Journal*, 156, 24
- Mulders, G. D., Pascucci, I., Manara, C. F., et al. 2017, *ApJ*, 847, 31
- Murchikova, L. & Tremaine, S. 2020, *The Astronomical Journal*, 160, 160
- Nagel, E., Czesla, S., Schmitt, J. H., et al. 2019, *Astronomy and Astrophysics*, 622
- Nakagawa, Y., Sekiya, M., & Hayashi, C. 1986, *Icarus*, 67, 375
- Nakamoto, T. & Nakagawa, Y. 1994, *The Astrophysical Journal*, 421, 640
- Narang, M., Manoj, P., Furlan, E., et al. 2018, *The Astronomical Journal*, 156, 221
- Ndugu, N., Bitsch, B., & Jurua, E. 2017, arXiv, 1
- Ndugu, N., Bitsch, B., & Jurua, E. 2018, *Monthly Notices of the Royal Astronomical Society*, 474, 886
- Neil, A. R. & Rogers, L. A. 2020, *The Astrophysical Journal*, 891, 12
- Nelson, B. E., Ford, E. B., Buchner, J., et al. 2020, *The Astronomical Journal*, 159, 73
- Neugebauer, M. & Snyder, C. W. 1966, *J. Geophys. Res.*, 71, 4469
- Nielsen, F. 2016, in *Introduction to HPC with MPI for Data Science* (Springer International Publishing), 195–211

- Nowak, G., Luque, R., Parviainen, H., et al. 2020, *Astronomy & Astrophysics*, 642, A173
- Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., & Vaughan, A. H. 1984, *ApJ*, 279, 763
- Nutzman, P. & Charbonneau, D. 2008, *PASP*, 120, 317
- Öberg, K. I. & Bergin, E. A. 2021, *Physics Reports*, 893, 1
- Ohtsuki, K., Stewart, G. R., & Ida, S. 2002, *Icarus*, 155, 436
- Ormel, C. W. 2017, in *Formation, Evolution, and Dynamics of Young Solar Systems*, ed. M. Pessah & O. Gressel (Springer International Publishing), 197–228
- Ormel, C. W. & Klahr, H. H. 2010, *Astronomy and Astrophysics*, 520, A43
- Ormel, C. W., Liu, B., & Schoonenberg, D. 2017, *Astronomy and Astrophysics*, 604, 1
- Owen, J. E., Clarke, C. J., & Ercolano, B. 2012, *Monthly Notices of the Royal Astronomical Society*, 422, 1880
- Owen, J. E. & Jackson, A. P. 2012, *Monthly Notices of the Royal Astronomical Society*, 425, 2931
- Owen, J. E. & Murray-Clay, R. 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 2206
- Owen, J. E. & Wu, Y. 2013, *Astrophysical Journal*, 775, 1
- Paardekooper, S.-J., Baruteau, C., & Kley, W. 2011, *MNRAS*, 410, 293
- Paardekooper, S. J. & Mellema, G. 2006, *Astronomy and Astrophysics*, 453, 1129
- Papaloizou, J. & Lin, D. N. C. 1984, *The Astrophysical Journal*, 285, 818
- Papanastassiou, D. A., Wasserburg, G. J., & Burnett, D. S. 1970, *Earth and Planetary Science Letters*, 8, 1
- Parker, R. J. 2020, *Royal Society Open Science*, 7
- Parmentier, V., Fortney, J. J., Showman, A. P., Morley, C., & Marley, M. S. 2016, *ApJ*, 828, 22
- Pascucci, I., Testi, L., Herczeg, G. J., et al. 2016, *The Astrophysical Journal*, 831, 125

- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *Journal of Machine Learning Research*, 12, 2825
- Penev, K., Jackson, B., Spada, F., & Thom, N. 2012, *Astrophysical Journal*, 751
- Perri, F. & Cameron, A. 1974, *Icarus*, 22, 416
- Perri, F. & Cameron, A. G. W. 1974, *Icarus*, 22, 416
- Perryman, M. 2018, *The Exoplanet Handbook* (Cambridge University Press)
- Perryman, M., Hartman, J., Bakos, G., & Lindegren, L. 2014, *Astrophysical Journal*, 797
- Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013a, *Proceedings of the National Academy of Sciences of the United States of America*, 110, 19651
- Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013b, *Proceedings of the National Academy of Sciences of the United States of America*, 110, 19273
- Petigura, E. A., Marcy, G. W., Winn, J. N., et al. 2018, *The Astronomical Journal*, 155, 89
- Petrovich, C. & Tremaine, S. 2016, *The Astrophysical Journal*, 829, 132
- Petrovich, C., Wu, Y., & Ali-Dib, M. 2019, *The Astronomical Journal*, 157, 5
- Pinamonti, M., Sozzetti, A., Bonomo, A. S., & Damasso, M. 2017, *Monthly Notices of the Royal Astronomical Society*, 468, 3775
- Pinamonti, M., Sozzetti, A., Giacobbe, P., et al. 2019, *Astronomy and Astrophysics*, 625
- Piso, A. M. A. & Youdin, A. N. 2014, *Astrophysical Journal*, 786
- Podolak, M. 2003, *Icarus*, 165, 428
- Pohl, A. 2018, PhD thesis
- Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, *Icarus*, 124, 62
- Pringle, J. E. 1981, *Annual Review of Astronomy and Astrophysics*, 19, 137
- Queloz, D., Henry, G. W., Sivan, J. P., et al. 2001, *Astronomy and Astrophysics*, 379, 279
- Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2013, *Proceedings of the International Astronomical Union*, 8, 395

- Quirrenbach, A., Amado, P. J., Mandel, H., et al. 2010, in *Ground-based and Airborne Instrumentation for Astronomy III*, ed. I. S. McLean, S. K. Ramsay, & H. Takami, Vol. 7735, International Society for Optics and Photonics (SPIE), 455–468
- Rafikov, R. R. 2004, *The Astronomical Journal*, 128, 1348
- Rasio, F. A. & Ford, E. B. 1996, *Science*, 274, 954
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Experimental Astronomy*, 38, 249
- Raymond, S. N., Armitage, P. J., & Gorelick, N. 2009, *Astrophysical Journal*, 699, 88
- Raymond, S. N. & Cossou, C. 2014, *Monthly Notices of the Royal Astronomical Society: Letters*, 440, 11
- Raymond, S. N. & Izidoro, A. 2017, *Icarus*, 297, 134
- Raymond, S. N., Kokubo, E., Morbidelli, A., Morishima, R., & Walsh, K. J. 2014, *Protostars and Planets VI*
- Raymond, S. N., Mandell, A. M., & Sigurdsson, S. 2006, *Science*, 313, 1413
- Raymond, S. N. & Morbidelli, A. 2020, arXiv, 1
- Raymond, S. N., Scalo, J., & Meadows, V. S. 2007, *The Astrophysical Journal*, 669, 606
- Reiners, A., Ribas, I., Zechmeister, M., et al. 2018a, *Astronomy & Astrophysics*, 609, L5
- Reiners, A., Zechmeister, M., Caballero, J. A., et al. 2018b, *Astronomy & Astrophysics*, 612, A49
- Ribas, Á., Bouy, H., & Merín, B. 2015, *Astronomy & Astrophysics*, 576, A52
- Ribas, Á., Merín, B., Bouy, H., et al. 2014, *Astronomy & Astrophysics*, 561, 12
- Ribas, I. & Miralda-Escudé, J. 2007, *Astronomy and Astrophysics*, 464, 779
- Ribas, I., Tuomi, M., Reiners, A., et al. 2018, *Nature*, 563, 365
- Richert, A. J., Getman, K. V., Feigelson, E. D., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 477, 5191
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 014003

- Rivera, E. J., Lissauer, J. J., Butler, R. P., et al. 2005, *The Astrophysical Journal*, 634, 625
- Robitaille, T. P., Tollerud, E. J., Greenfield, P., et al. 2013, *Astronomy and Astrophysics*, 558, A33
- Rodriguez, J. E., Quinn, S. N., Huang, C. X., et al. 2019, *The Astronomical Journal*, 157, 191
- Ronco, M. P., Guilera, O. M., & de Elía, G. C. 2017, *Monthly Notices of the Royal Astronomical Society*, 471, 2753
- Rosenfeld, K. A., Andrews, S. M., Hughes, A. M., Wilner, D. J., & Qi, C. 2013, *Astrophysical Journal*, 774
- Rossiter, R. A. 1924, *ApJ*, 60, 15
- Rousseuw, P. J. 1987, *Journal of Computational and Applied Mathematics*, 20, 53
- Salpeter, E. E. 1955, *The Astrophysical Journal*, 121, 161
- Sandford, E., Kipping, D., & Collins, M. 2019, *Monthly Notices of the Royal Astronomical Society*, 489, 3162
- Santos, N. C., Israelian, G., & Mayor, M. 2001, *Astronomy and Astrophysics*, 373, 1019
- Santos, N. C., Israelian, G., & Mayor, M. 2004, *Astronomy and Astrophysics*, 415, 1153
- Santos, N. C., Israelian, G., Mayor, M., et al. 2005, *Astronomy & Astrophysics*, 437, 1127
- Santos, N. C., Israelian, G., Mayor, M., Rebolo, R., & Udry, S. 2003, *Astronomy and Astrophysics*, 398, 363
- Santos, N. C., Pont, F., Melo, C., et al. 2006, *Astronomy and Astrophysics*, 450, 825
- Sarkis, P., Henning, T., Kürster, M., et al. 2018, *The Astronomical Journal*, 155, 257
- Scargle, J. D. 1982, *Astrophysical Journal*, 263, 835
- Schäfer, U., Johansen, A., & Banerjee, R. 2020, *Astronomy & Astrophysics*, 635, A190
- Schib, O., Mordasini, C., Wenger, N., Marleau, G.-D., & Helled, R. 2021, *Astronomy & Astrophysics*, 645, A43
- Schlaufman, K. C. & Winn, J. N. 2016, *The Astrophysical Journal*, 825, 62

Schlecker, M. 2016, Irregular Variability in Kepler Photometry

Schlecker, M., Kossakowski, D., Brahm, R., et al. 2020a, *The Astronomical Journal*, 160, 275

Schlecker, M., Mordasini, C., Emsenhuber, A., et al. 2020b, arXiv, 1

Schlecker, M., Pham, D., Burn, R., et al. 2021, ArXiv e-prints, arXiv:2104.11750

Schlichting, H. E., Fuentes, C. I., & Trilling, D. E. 2013, *Astronomical Journal*, 146

Schoonenberg, D., Liu, B., Ormel, C. W., & Dorn, C. 2019, *Astronomy & Astrophysics*, 627, A149

Schwarz, G. 1978, *Annals of Statistics*, 6, 461

Schwarzschild, K. & Villiger, W. 1906, *The Astrophysical Journal*, 23, 284

Scott Gaudi, B., Christiansen, J. L., & Meyer, M. R. 2020, arXiv, 1

Seager, S., Richardson, L. J., Hansen, B. M. S., et al. 2005, *ApJ*, 632, 1122

Sebastian, D., Gillon, M., Ducrot, E., et al. 2021, *Astronomy and Astrophysics*, 645, 1

Shakura, N. I. & Sunyaev, R. a. 1973, *Astronomy and Astrophysics*, 24, 337

Shields, A. L., Ballard, S., & Johnson, J. A. 2016, *Physics Reports*, 663, 1

Shields, A. L., Meadows, V. S., Bitz, C. M., et al. 2013, *Astrobiology*, 13, 715

Sierks, H., Barbieri, C., Lamy, P. L., et al. 2015, *Science*, 347, 1

Singer, K. N., McKinnon, W. B., Gladman, B., et al. 2019, *Science*, 363, 955

Soderblom, D., Hillenbrand, L., Jeffries, R., Mamajek, E., & Naylor, T. 2014, in *Protostars and Planets VI*, ed. H. Beuther, R. S. Klessen, C. P. Dullemond, & T. Henning, 219

Solomon, P. M. & Werner, M. W. 1971, *The Astrophysical Journal*, 165, 41

Sousa, S. G., Adibekyan, V., Delgado-Mena, E., et al. 2018, *Astronomy and Astrophysics*, 620, 1

Southworth, J. 2011, *Monthly Notices of the Royal Astronomical Society*, 417, 2166

Spearman, C. 1904, *The American Journal of Psychology*, 15, 72

- Spencer, J. R., Stern, S. A., Moore, J. M., et al. 2020, *Science*, 367
- Stassun, K. G., Oelkers, R. J., Pepper, J., et al. 2018, *The Astronomical Journal*, 156, 102
- Stevenson, D. J. & Lunine, J. I. 1988, *Icarus*, 75, 146
- Stevenson, K. B. 2016, *ApJL*, 817, L16
- Stewart, G. R. & Wetherill, G. W. 1988, *Icarus*, 74, 542
- Stock, S., Kemmer, J., Reffert, S., et al. 2020, *Astronomy and Astrophysics*, 636
- Struve, O. 1952, *The Observatory*, 72, 199
- Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, *Astrophysical Journal*, 809, 77
- Suzuki, D., Bennett, D. P., Ida, S., et al. 2018, *The Astrophysical Journal*, 869, L34
- Suzuki, D., Bennett, D. P., Sumi, T., et al. 2016, *The Astrophysical Journal*, 833, 145
- Szabó, G. M. & Kiss, L. L. 2011, *Astrophysical Journal Letters*, 727, 2
- Szekely, G. J. & Rizzo, M. L. 2005, *Journal of Classification*, 22, 151
- Szulágyi, J., Morbidelli, A., Crida, A., & Masset, F. 2014, *Astrophysical Journal*, 782
- Takeuchi, T. & Lin, D. 2002, *ApJ*, 1344
- Tanaka, H., Takeuchi, T., & Ward, W. R. 2002, *The Astrophysical Journal*, 565, 1257
- Teague, R., Bae, J., Bergin, E. A., Birnstiel, T., & Foreman-Mackey, D. 2018, *The Astrophysical Journal*, 860, L12
- Thiabaud, A., Marboeuf, U., Alibert, Y., et al. 2014, *Astronomy & Astrophysics*, 562, A27
- Thiabaud, A., Marboeuf, U., Alibert, Y., Leya, I., & Mezger, K. 2015, *Astronomy & Astrophysics*, 574, A138
- Thommes, E. W., Duncan, M. J., & Levison, H. F. 2003, *Icarus*, 161, 431
- Thommes, E. W., Matsumura, S., & Rasio, F. A. 2008, *Science*, 321, 814
- Thorndike, R. L. 1953, *Psychometrika*, 18, 267

- Tin Kam Ho. 1995, in Proceedings of 3rd International Conference on Document Analysis and Recognition, Vol. 1, 278–282 vol.1
- Torres, G., Winn, J. N., & Holman, M. J. 2008, *The Astrophysical Journal*, 677, 1324
- Toulmin, Priestley, I., Baird, A. K., Clark, B. C., et al. 1977, *J. Geophys. Res.*, 82, 4625
- Trapman, L., Facchini, S., Hogerheijde, M. R., van Dishoeck, E. F., & Bruderer, S. 2019, *Astronomy & Astrophysics*, 629, A79
- Triaud, A. H. M. J. 2018, *Handbook of Exoplanets*, 1375
- Trifonov, T., Kürster, M., Zechmeister, M., et al. 2018, *Astronomy and Astrophysics*, 609, 1
- Trifonov, T., Tal-Or, L., Zechmeister, M., et al. 2020, *Astronomy and Astrophysics*, 636, 1
- Tripathi, A., Andrews, S. M., Birnstiel, T., & Wilner, D. J. 2017, *The Astrophysical Journal*, 845, 44
- Tsai, S.-M., Kitzmann, D., Lyons, J. R., et al. 2018, *ApJ*, 862, 31
- Tsiganis, K., Gomes, R., Morbidelli, A., & Levison, H. F. 2005, *Nature*, 435, 459
- Turbet, M. & Selsis, F. 2021, arXiv e-prints
- Tychoniec, Ł., Tobin, J. J., Karska, A., et al. 2018, *The Astrophysical Journal Supplement Series*, 238, 19
- Vallat, R. 2018, *The Journal of Open Source Software*, 3, 1026
- Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 479, 4786
- Vanderburg, A., Rowden, P., Bryson, S., et al. 2020, *The Astrophysical Journal*, 893, L27
- Venot, O., Hébrard, E., Agúndez, M., et al. 2012, *A&A*, 546, A43
- Venot, O., Parmentier, V., Blečić, J., et al. 2020, *ApJ*, 890, 176
- Venturini, J., Guilera, O. M., Haldemann, J., Ronco, M. P., & Mordasini, C. 2020a, *Astronomy & Astrophysics*, 643, L1
- Venturini, J., Ronco, M. P., & Guilera, O. M. 2020b, *Space Science Reviews*, 216

- Venuti, L., Bouvier, J., Cody, A. M., et al. 2017, *Astronomy & Astrophysics*, 599, A23
- Voelkel, O., Deienno, R., Kretke, K., & Klahr, H. 2021a, *Astronomy & Astrophysics*, 645, A131
- Voelkel, O., Deienno, R., Kretke, K., & Klahr, H. 2021b, *Astronomy & Astrophysics*, 645, A132
- Voelkel, O., Klahr, H., Mordasini, C., Emsenhuber, A., & Lenz, C. 2020, *Astronomy & Astrophysics*, 642, A75
- Walsh, K. J., Morbidelli, A., Raymond, S. N., O'Brien, D. P., & Mandell, A. M. 2011, *Nature*, 475, 206
- Wang, J. & Fischer, D. A. 2015, *The Astronomical Journal*, 149, 14
- Wang, S. 2017, *Research Notes of the AAS*, 1, 26
- Ward, J. H. 1963, *Journal of the American Statistical Association*, 58, 236
- Ward, W. R. 1997, *Icarus*, 126, 261
- Weidenschilling, S. J. 1977, *Monthly Notices of the Royal Astronomical Society*, 180, 57
- Weidenschilling, S. J. 2011, *Icarus*, 214, 671
- Weiss, L. M., Isaacson, H. T., Marcy, G. W., et al. 2018, *The Astronomical Journal*, 156, 254
- Weizsäcker, C. F. 1948, *Zeitschrift für Naturforschung - Section A Journal of Physical Sciences*, 3, 524
- Whipple, F. 1972, in *From Plasma to Planet*, ed. A. Elvius, 211
- Williams, J. P. & Cieza, L. A. 2011, *Annual Review of Astronomy and Astrophysics*, 49, 67
- Winn, J. N. 2010, eprint arXiv, 1001, 2010
- Winn, J. N. 2018, in *Handbook of Exoplanets No. 1963* (Cham: Springer International Publishing), 1–18
- Winter, A. J., Kruijssen, J. M., Longmore, S. N., & Chevance, M. 2020, *Nature*, 586, 528
- Wittenmyer, R. A., Butler, R. P., Tinney, C. G., et al. 2016, *The Astrophysical Journal*, 819, 28

- Wittenmyer, R. A., Wang, S., Butler, J. H. R. P., et al. 2020, *Monthly Notices of the Royal Astronomical Society*, 8, 1
- Wolszczan, A. & Frail, D. A. 1992, *Nature*, 355, 145
- Wright, J. T., Upadhyay, S., Marcy, G. W., et al. 2009, *The Astrophysical Journal*, 693, 1084
- Wuchterl, G. 2007, *Extrasolar Planets: XVI Canary Islands Winter School of Astrophysics*, 89
- Wyatt, M. C., Kral, Q., & Sinclair, C. A. 2019, *Monthly Notices of the Royal Astronomical Society*, 802, 782
- Yoder, C. F. & Peale, S. J. 1981, *Icarus*, 47, 1
- Zawadzki, B., Carrera, D., & Ford, E. B. 2021, *Monthly Notices of the Royal Astronomical Society*, 503, 1390
- Zechmeister, M., Dreizler, S., Ribas, I., et al. 2019, *Astronomy & Astrophysics*, 627, A49
- Zechmeister, M. & Kürster, M. 2009, *Astronomy and Astrophysics*, 496, 577
- Zechmeister, M., Reffert, S., Hatzes, A. P., Endl, M., & Quirrenbach, A. 2008, *Astronomy and Astrophysics*, 491, 531
- Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, *Proceedings of the National Academy of Sciences of the United States of America*, 116, 9723
- Zhang, K., Bergin, E. A., Blake, G. A., Cleeves, L. I., & Schwarz, K. R. 2017, *Nature Astronomy*, 1, 1
- Zhao, B., Tomida, K., Hennebelle, P., et al. 2020, *Space Science Reviews*, 216
- Zhu, W. 2019, *The Astrophysical Journal*, 873, 8
- Zhu, W. 2020, *The Astronomical Journal*, 159, 188
- Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, *The Astrophysical Journal*, 860, 101
- Zhu, W., Wang, J., & Huang, C. 2016, *The Astrophysical Journal*, 832, 1
- Zhu, W. & Wu, Y. 2018, *The Astronomical Journal*, 156, 92
- Zink, J. K., Christiansen, J. L., & Hansen, B. M. 2019, *Monthly Notices of the Royal Astronomical Society*, 483, 4479

Acknowledgements

Teamwork makes the dream work.

A doctoral project is a team effort, and it takes a lot of help to get the job done. I owe my advisors Thomas Henning and Hubert Klahr the deepest gratitude for their support and guidance over the past three and a half years. Thomas, it has been a privilege to work with you, and I am deeply indebted to you for rewarding challenges, the trust in me to accomplish them, and your support to make it happen. Thank you for your unwavering advise, both scientifically and beyond. Hubert, thank you for giving me the freedom to pursue my own ideas. You understand how to create a relaxed work environment where nobody feels afraid to ask questions. It was fun to discuss science and random other things with the whole “Klahr group” within this laid back setting. I would like to thank my examination committee Thomas Henning, Sabine Reffert, Kees Dullemond, and Stephanie Hansmann-Menzemer for agreeing to evaluate my work, and it is a pleasure to acknowledge the members of my PhD Advisory Committee (PAC) Wolfgang Brandner, Tom Herbst, Ralf Launhardt, Karin Lind, and Fabian Walter.

In the course of my PhD, I had the opportunity to collaborate with more amazing scientists than I could possibly list here. I feel fortunate to have been able to work with Yann Alibert, Daniel Apai, Remo Burn, Néstor Espinoza, Paul Mollière, Christoph Mordasini, Dang Pham, Silvia Sabotta, and many others. I learned a lot from interesting discussions with Robin Baeyens, Gesa Bertrang, Til Birnstiel, Francesco Biscani, Bertram Bitsch, Faustine Cantalloube, Neige Frankel, Julien Milli, Gijs Mulders, Gabriele Pichierri, Trifon Trifonov, and Oliver Voelkel. A huge thanks goes to the EDEN-HD team, in particular the initial bunch who got the stone rolling with me: Asmita Bhandare, Miriam Keppler, Karan Molaverdikhani, and Paula Sarkis. Nicolas Kurtovic, great job pushing the project forward as a new team lead!

This work relied heavily on all kinds of non-scientific support, and I am grateful for all the help I got by the MPIA family, from the canteen team to the technical service to the cleaning personnel. I owe a special thanks to MPIA's incredibly dedicated IT support team, who never got tired of fixing the things I broke, be it hardware or software. I would also like to thank my peers at the International Max Planck Research School for Astronomy and Cosmic Physics (IMPRS) for lively seminars, fun pub nights, and for not nagging too much about the degree of housing comfort we chose for the first-year retreat. Special thanks go to my co-student representative Diana Kossakowski for the great team work in organizing this and other events. Thank you to Christian Fendt and Huong Witte-Nguy for greatly improving the PhD experience by managing the IMPRS program in an amazingly warm and down-to-Earth way.

I am extremely grateful to Michael Becker, Julia Bodensteiner, Remo Burn, Robert Glas, and Miriam Keppler for proof-reading various parts of my thesis. As always, when technical difficulties drove me into despair, Robert was ready to help with his inexhaustible troubleshooting skills. Julia, excellent time syncing! It was fabulous to be able to rely on your camaraderie and friendship during the most intense phases of both our PhDs: the application phase, the next (postdoc) application phase, and the writing phase. The cover image is courtesy of Michael Becker who, through technical talent and an unlimited appetite to learn something new, was able to turn his balcony in the midst of the city of Erlangen into a formidable observatory. A big shout-out goes to my colleagues in Office 217 for maintaining an island of colorful joviality at MPIA: Hans Baehr, Tobias Buck, Samantha Brown, Sven Buder, Timmy Delage, Lizxandra Flores, Hector Hiss, Diana Kossakowski, and Kristoffer Nelson. To the "mountain crew" Miriam, Paula, and Gabriele: I will miss our adventures!

Every journey starts with a chance. Without my inspiring teacher Walter Boehm, I would never have studied physics. Thank you for sparking my scientific curiosity. To my parents Maria and Ludwig: thank you for your constant love and support, no matter what unusual turns I decided to take in my life! I have been fortunate to have someone special by my side on this journey, someone who understood the hardships and knew how to pull me up again when I was struggling. It is hard to express in words what a crucial role she has played in the success of my doctorate, in particular in the final months leading up to finishing this thesis. Miriam, I am incredibly grateful that I could hike this trail with you. More adventures are waiting for us to dare!

What can never be adequately acknowledged is the nutrition that every scientist needs: the everlasting supply of enigmas and challenges that nature provides.