Imperial College London

Department of Physics

The Atmospheric Evolution of Small, Close-In Exoplanets

James G. Rogers

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy at Imperial College London, July 2022.

Statement of Originality

This thesis is my own work, except where specifically indicated in the text. Some parts of this thesis are based on work completed in collaboration with others, and some have also been published as journal articles:

- Chapter 1 (specifically Section 1.6) is based on work as part of the *California Kepler Survey* collaboration, which is published as:

  Petigura, E. A., ***Rogers, J. G.****, Isaacson, H., Owen, J. E., Kraus, A. L., Winn, J. N., MacDougall, M. G., Howard, A. W., Fulton, B., Kosiarek, M. R., Weiss, L. M., Behmard, A., and Blunt, S. The California-Kepler Survey. X. The Radius Gap as a Function of Stellar Mass, Metallicity, and Age. Astronomical Journal, 163, 179 (2022).*

  In this work I was responsible for developing tools for demographic analysis, specifically quantifying the location and slope of the observed radius gap.

- Chapter 3 is based on work that has been completed in collaboration with A. Gupta, J. E. Owen and H. E. Schlichting, and is published as:

  Rogers, J. G.*, Gupta, A., Owen, J. E., and Schlichting, H. E. Photoevaporation versus core-powered mass-loss: model comparison with the 3D radius gap. Monthly Notices of the Royal Astronomical Society, 508, 5886–5902 (2021).*

  In this work I performed models of photoevaporation as well as all statistical analysis. A. Gupta provided models for core-powered mass-loss. Discussion and advice were provided by J. E. Owen and H. E. Schlichting.

- Chapter 4 is based on work that has been completed in collaboration with J. E. Owen and is published as:

  Rogers, J. G. *and Owen, J. E. Unveiling the planet population at birth. Monthly Notices of the Royal Astronomical Society, 503, 1526–1542 (2021).*

  In this work I ran all models and statistical analysis. J. E. Owen provided discussion and advice.

- Chapters 5 is based on work that has been completed in collaboration with J. E. Owen.

  In this work I ran all models and statistical analysis. J. E. Owen provided discussion and advice.

- Chapters 6 is based on work that has been completed in collaboration with C. Janó Muñoz, J. E. Owen and R. A. Booth.

  In this work, I ran all models, trained all machine learning algorithms and performed all statistical analysis. C. Janó Muñoz performed proof of concept work on machine learning whilst J. E. Owen and R. A. Booth provided discussion and advice.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its con- tents are licensed under a Creative Commons Attribution-Non Commercial-No Derivatives 4.0 International Licence (CC BY-NC-ND). Under this licence, you may copy and redistribute the material in any medium or format on the condition that; you credit the author, do not use it for commercial purposes and do not distribute modified versions of the work. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

I owe sincere thanks to so many people for their support during my time at Imperial College London. Firstly, my supervisor, James, who's wisdom and approach to science have been an immense inspiration, not to mention his endless patience and support through the metaphorical marathon of a PhD. Perhaps most importantly though, he has pushed me to achievements I did know I was capable of. Although not scientific, this also includes 95% of an actual marathon (which in fact ended in catastrophe, hopefully my PhD won't end the same way).

My family have been there for me at every stage, and I owe them all so much. My parents: Graham and Louise, who instilled a passion and curiosity in the natural world from a very young age, that still drives me today. My older sister and brother: Helen and Michael, who have been the most amazing role models since day one. My grandparents, for their unshakeable faith in me. Finally my nephews: Billy and Seb, who decided to sneak rude words into my paper drafts, only for me to find later.

My friends: Matt, James and Jamie, have been by my side ever since our first week at Imperial, 8 years ago. They make me laugh more than anyone. As engineer graduates, their disbelief that anything can be modelled in 1D amuses me.

Thanks also go to other students in my year: Dan, Joe, Ian, Lena and Richard. Despite attempting a PhD during a global pandemic, we've had great fun and I will miss the time we have spent in the office and, more importantly, the pub. Finally, I would like to thank other scientists who have given me fantastic advice and guidance along the way: Yvonne, Subu, Hilke, Erik, Akash, Richard, Matthäus, Marija, Chris and Christina.

Abstract

The results from the Kepler space observatory have revealed a wealth of planets with small sizes and small orbital separations, referred to as "super-Earths" or "sub-Neptunes". This thesis will focus on such planets.

Atmospheres of super-Earths/sub-Neptunes are vulnerable to hydrodynamic escape due to the proximity to their host stars. I begin by comparing two mechanisms by which this can occur: photoevaporation, which is caused by high energy radiation from the host star; and core-powered mass-loss, which relies on the stellar bolometric luminosity in conjunction with the cooling luminosity from the planetary core. I construct a robust statistical test that can be performed on future exoplanet surveys in order to determine which of these mechanisms is dominant.

I then go on to show that if one takes the population of exoplanets observed today, one can rewind the clock on atmospheric evolution to infer the planet properties at birth. I perform this with photoevaporation models to place constraints on planetary properties that are currently unobservable with standard techniques. I find that a typical super-Earth/sub-Neptune has a core mass of $\sim 4M_{\oplus}$, a core composition consistent with that of Earth and an initial atmospheric mass fraction after protoplanetary disc dispersal of $\sim 2\%$.

This latter result highlights a discrepancy between atmospheric mass-loss and gas accretion models, the latter of which predicts initial atmospheric masses $\gtrsim 10\%$. I perform sophisticated evolution models of planets immersed in protoplanetary discs that self-consistently model gas accretion and mass-loss induced by rapid disc dispersal. I find that this mechanism efficiently removes atmospheric mass and provides a robust solution to the aforementioned discrepancy. This mass-loss phase induces rapid contraction and advection of high entropy material into the escaping atmosphere. This process, which I find is strongly controlled by the disc dispersal rate, produces a planet that has prematurely cooled, given its age.

Contents

Statement of Originality **3**

Copyright Declaration **5**

Acknowledgements **7**

Abstract **9**

1 Introduction **19**
1.1 Exoplanet Detection . . . 20
1.2 Demographic Surveys . . . 23
1.3 Mass-Radius Diagrams . . . 25
1.4 Escaping Exoplanet Atmospheres . . . 28
1.5 The Radius Gap . . . 32
1.6 CKS-Cool . . . 35
1.7 Physics-Based Inference . . . 36

2 The Theory of Atmospheric Mass-Loss and the Origin of the Radius Gap **41**
2.1 Collisional Fluids . . . 41
2.2 Fluid Equations and Hydrostatic Equilibrium . . . 42
2.3 Isothermal Parker Wind . . . 44
2.4 Radiation Transport . . . 46
2.5 Radiative-Convective Equilibrium . . . 48
2.6 An analytic model for atmospheric structure . . . 56
2.7 An analytic model for atmospheric mass-loss . . . 59
2.8 Carving the Radius Gap . . . 60
2.9 Model Limitations . . . 63
2.10 Core-Powered Mass-Loss . . . 63

3 Photoevaporation Vs. Core-Powered Mass-Loss: Model Comparison with the 3D Radius Gap **65**
3.1 Introduction . . . 65
3.2 Method . . . 66
3.2.1 The Radius Gap in 3D . . . 66
3.2.2 The Models . . . 70
3.2.3 Model Predictions . . . 72
3.2.4 Synthetic Transit Surveys . . . 74
3.2.5 Gradient Calculation . . . 75
3.3 Results . . . 77
3.3.1 Measurement Uncertainty . . . 80
3.3.2 False Positives . . . 80

3.3.3 Changing the Stellar Mass Distribution . . . 82
3.3.4 Changing Model Parameters . . . 82
3.3.5 Results for Existing Surveys . . . 86
3.4 Discussion . . . 86
3.4.1 The Perfect Survey . . . 86
3.4.2 The Effect of Survey Completeness . . . 87
3.4.3 Limitations of the Statistical Analysis . . . 89
3.4.4 Independence of Core Mass Scaling . . . 89
3.4.5 A Route Forward . . . 89
3.5 Conclusions . . . 90

4 Unveiling the Planet Population at Birth 93
4.1 Introduction . . . 93
4.2 Method . . . 93
4.2.1 Evolving the Planets . . . 97
4.2.2 Detecting the Planets . . . 97
4.2.3 Likelihood Calculation . . . 99
4.3 Results . . . 104
4.3.1 Model I . . . 104
4.3.2 Model II . . . 105
4.3.3 Model III . . . 105
4.4 Convergence and full posteriors . . . 108
4.5 Discussion . . . 108
4.5.1 Atmospheric Mass Fraction Distribution . . . 115
4.5.2 Core Mass Distribution . . . 116
4.5.3 Core Composition . . . 120
4.5.4 Core Composition Spread . . . 121
4.5.5 Planet Distribution Beyond 100 day Orbital Period . . . 123
4.5.6 Changing the Evolution Model . . . 123
4.5.7 Improving the Model . . . 125
4.6 Conclusion . . . 125

5 Under the Light of a New Star: Evolution of Planets through Disc Dispersal and Boil-Off 127
5.1 Introduction . . . 127
5.1.1 Protoplanetary Discs . . . 127
5.1.2 Atmospheric Accretion . . . 129
5.1.3 Disc Dispersal . . . 130
5.1.4 Boil-Off . . . 133
5.2 Method . . . 134
5.2.1 MESA . . . 134
5.2.2 Outer Boundary Condition . . . 135
5.2.3 Accretion and Mass-Loss . . . 136
5.2.4 Building Planets . . . 141
5.3 Results . . . 143
5.3.1 Formation time vs. disc dispersal timescale . . . 144
5.3.2 Orbital separation vs. Core mass . . . 147
5.4 Discussion . . . 148
5.4.1 Boil-off sets the initial conditions for photoevaporation . . . 150
5.4.2 A Gentle Breeze . . . 150
5.4.3 Planet formation history is removed . . . 151
5.4.4 Model Improvements . . . 151

5.4.5 Future direction . . . 153
5.5 Conclusion . . . 153

6 Summary and Outlook **155**
6.1 Summary . . . 155
6.2 Outlook . . . 156
6.2.1 Rapid Evaluation of Exoplanet Evolution Models . . . 157
6.2.2 Variable Stellar X-ray Luminosity . . . 160
6.2.3 X-ray vs. EUV dominated mass-loss . . . 162
6.2.4 Closing Remarks . . . 166

Appendices **191**

List of Figures

1.1 Kepler injection recovery . . . 24
1.2 Kepler planets . . . 25
1.3 Mass-radius diagram . . . 26
1.4 Lyman-α transit spectroscopy . . . 29
1.5 Stellar XUV evolution . . . 32
1.6 Radius gap predictions . . . 33
1.7 Radius gap observations . . . 34
1.8 CKS-Cool . . . 37
1.9 Photoevaporation vs. core accretion . . . 38
1.10 Core-powered mass-loss radius gap . . . 40

2.1 Isothermal Parker wind . . . 45
2.2 Planet schematic . . . 50
2.3 Radiative convective equilibrium . . . 53
2.4 Analytic model integrals . . . 57
2.5 Atmospheric mass fraction evolution . . . 61
2.6 Radius gap formation . . . 62

3.1 Photoevaporation vs. core-powered mass-loss . . . 67
3.2 Stellar and planet mass scalings . . . 69
3.3 Core mass distributions . . . 71
3.4 Method limitation . . . 76
3.5 Posteriors summary . . . 78
3.6 Overlap integral I . . . 79
3.7 Overlap integral II . . . 81
3.8 Overlap integral III . . . 83
3.9 Overlap integral IV . . . 84
3.10 Survey results . . . 85
3.11 Occurrence vs. detection . . . 88

4.1 BHM flowchart . . . 94
4.2 Completeness . . . 98
4.3 Inferred period distribution . . . 102
4.4 Inferred stellar mass distribution . . . 103
4.5 Likelihood convergence . . . 104
4.6 Best fit model . . . 106
4.7 Model II and III . . . 107
4.8 Core composition . . . 107
4.9 Model I corner . . . 109
4.10 Model I auto-correlation . . . 110
4.11 Model II corner . . . 111

4.12 Model II auto-correlation . . . 112
4.13 Model III corner . . . 113
4.14 Model III auto-correlation . . . 114
4.15 Atmospheric mass fraction evolution . . . 115
4.16 Photoevaporation vs. core accretion II . . . 117
4.17 Mass-radius predicted . . . 118
4.18 Ice-mass fractions . . . 120
4.19 Clean vs. noisy gap . . . 122
4.20 Long orbital periods . . . 123
4.21 Radius gap evolution . . . 124

5.1 Disc dispersal observations . . . 131
5.2 Disc dispersal schematic . . . 132
5.3 MESA boundary condition . . . 136
5.4 Accretion vs. mass-loss . . . 138
5.5 Breeze threshold . . . 139
5.6 Cooling timescale . . . 143
5.7 Boil-off I . . . 144
5.8 Boil-off II . . . 146
5.9 Boil-off III . . . 147
5.10 Boil-off IV . . . 148
5.11 Boil-off V . . . 149

6.1 Emulator flowchart . . . 158
6.2 Emulator results . . . 159
6.3 X-ray intrinsic scatter I . . . 160
6.4 X-ray intrinsic scatter II . . . 161
6.5 X-ray vs. EUV evolution . . . 163
6.6 X-ray vs. EUV dominated mass-loss . . . 165

List of Tables

4.1 Comparison of Model I with Wu (2019) 105
4.2 Core composition comparison 106

Chapter 1

Introduction

Until recently, planet formation has been poorly understood. For centuries astronomers could only build a picture of planet formation by looking at the other celestial bodies in our own Solar System. At the time, it would not have been inconceivable to think that the Solar system and its abundance of planets was unique. However, in 1995, this restricted view was opened up with the first detection of a planet orbiting another Sun-like star, referred to as an exoplanet.

Exoplanets allow us to understand planet formation and evolution on a population level, instead of just relying on the planets in our own Solar System. Nowadays, exoplanets are found around stars of different ages, masses, metallicities and in different regions of the Galaxy, allowing us to piece together the story of planet formation on a broader scale. The general picture that has built up since the first exoplanet detection is that our Solar System is not unique in its abundance of planets. However, it is the differences between these systems and our own that are perhaps the most striking. Indeed, in early planet formation models, most systems were predicted to mimic that of our own: with small, terrestrial planets; Mercury, Venus, Earth and Mars interior to the giant planets; Jupiter and Saturn, followed by the ice giants; Uranus and Neptune (Boss, 1995). To our surprise however, some of the first planets to be detected were giant and found closer to their star than Mercury is to the Sun. In light of this, formation models were constructed which relied on rapid growth of giant planets and orbital migration to bring them close to their host stars (e.g. Ida and Lin, 2004, 2005). Although these models were able to explain the newly observed population, they also predicted a scarcity of close-in, intermediate mass planets i.e. between the mass of Earth and Jupiter since planet growth was expected to be such an efficient process. However, the observed planet population subverted expectations again, with a large sample of planets detected in precisely this intermediate mass range.

Surprising results that contradict theoretical prejudice have been a common theme in the field of exoplanets thus far. However, our understanding is still, broadly speaking, in its infancy. Although great leaps in knowledge have been made as to the diversity and architecture of planetary systems, the connection between formation and the population of planets that we observe (including our own) still remains a mystery. As with the first formation models, it is still true today that the observations (almost always) drive theoretical development and understanding. Therefore, the key to gaining a more complete picture of planet formation and where our Solar System fits in, is to perform large-scale *demographic* surveys which deeply probe all of parameter space. Of course, there is always a limit as to how effective an instrument can be, meaning that another crucial aspect of demographic surveys is to understand the completeness of these surveys i.e. which type of planets are physically detectable. This then allows one to determine the underlying population of planets, given the observed sample.

In this thesis, I use recent demographic surveys as a tool to understand planet formation and evolution. One of the essential themes I emphasise is that the key to understanding the origin of exoplanets is to understand their evolution. If one can 'rewind the clock' on the processes that have controlled them, then it allows us to understand the conditions in which they formed. To begin, I discuss the various exoplanet detection methods that have revolutionised our view of planets in our galaxy, before discussing demographic surveys and the processes which govern exoplanet evolution.

1.1 Exoplanet Detection

There are multiple ways of detecting exoplanets, each with its own advantages and limitations. In general, they work by measuring observables from stellar system such as flux, spectra etc. which are then used to infer planetary properties. Among these methods are radial velocity, transit, microlensing and direct imaging. This latter technique, being the easiest to conceptually understand, aims to directly image a planet by resolving it from its host star. These are typically young (and hence self-luminous in infra-red) massive planets that are found at large orbital separation (e.g. Chauvin et al., 2004; Kalas et al., 2008; Marois et al., 2008; Chauvin et al., 2017; Currie et al., 2022). The benefit of this method is the ability to look at the emission spectrum of such planets, which will give insight into their chemical inventories. However, resolving a planet from its star is an extremely challenging task, meaning that this method has only yielded ~ 60 planets to-date.

A different approach to direct imaging is *gravitational microlensing*, which exploits the lensing of light from a background source as it passes through the warped spacetime around a massive foreground object, such as a planet-hosting star. As this foreground object passes through the line of sight of an observer with the background source, the image of the background source will be magnified in brightness. If the foreground source does indeed host a planet, a characteristic spike in brightness will be detected on top of this magnification. From this, the mass ratio between host star and planet can be calculated. With suitable follow-up observations of the host star, planet mass and orbital separation can be calculated (e.g. Bond et al., 2004; Bennett et al., 2006; Beaulieu et al., 2006; Gaudi et al., 2008; Mróz et al., 2020). The benefit of this technique is that it can probe exoplanets at large orbital separations and low masses. The upcoming *Nancy Grace Roman Space Telescope* will utilise microlensing events to detect exoplanets and is expected to be one of first surveys to successfully probe planets smaller than Earth at large orbital distances $\gtrsim 1$ AU (Green et al., 2012).

Chronologically speaking, the majority of the first exoplanets were discovered using the *radial velocity* method, which seeks to detect the Doppler shift in a host star's spectrum due the barycentric movement caused by an orbiting planet. For a given orbit inclination i, the projected stellar radial velocity can be determined which in-turn allows one to place an upper limit on planet mass i.e. $M_{\rm p} \sin i$, where $M_{\rm p}$ is planet mass[1]. Indeed the first exoplanet to be discovered around a solar-type star, 51 Pegasi b, was detected in 1995 with this method (Mayor and Queloz, 1995), which was found to be $\sim 0.5 M_{\rm J}$ at an orbital distance of ~ 0.05AU. This discovery came as a surprise since, clearly, we have no analogue for such a planet in our own Solar System. Nonetheless, in the following years, high-precision radial velocity surveys were established, such as HARPS (Mayor et al., 2003), which found many planets similar to 51 Pegasi b, which were labelled as 'hot-Jupiters'.

The proximity of such planets led many to the conclusion that another observational method, known as the *transit* method may be employed to corroborate the existence of such planets. In

[1]Note that if a transit observation is also detected, then the inclination is inherently known, meaning that the upper limit is indeed the actual planet mass.

the transit method, one performs photometry of a potential planet-hosting star for given period of time. If a planet orbits the star and also transits i.e. blocks some of the stellar light received by the observer, then one can infer its size relative to its host. In the following toy-model, a planet is referred to as 'in-transit' as it transits the star and 'out-of-transit' when not. For a flux received out-of-transit, $F_{\rm out}$, and flux received in-transit, $F_{\rm in}$, the *transit depth* δ scales as the fraction of the stellar disc area that is covered by the planet:

$$\delta \equiv \frac{F_{\rm in} - F_{\rm out}}{F_{\rm out}} \propto \left(\frac{R_{\rm p}}{R_*}\right)^2. \tag{1.1.1}$$

Hence, in order to calculate the planet's size, one requires knowledge of the stellar radius. In practise, this is calculated from photometric observations of the host star, which allow one to calculate its effective temperature; the stellar radius then comes from the Stefan-Boltzmann law.

One of the core themes of this thesis is the importance of large-scale demographic surveys. As discussed, the goal for such a survey is not simply to find as many exoplanets as possible, but to also understand the observational biases. It is only then that one can understand the underlying population of exoplanets. In the case of transit observations, there are two important sources of bias: firstly a simple geometric argument states that there is a small probability that a planet will transit i.e. its orbital plane will be such that it blocks stellar light incident to an observer. If one assumes that orbital plane inclinations are randomly orientated and also make a small-angle approximation, the probability of transit, $P_{\rm tr}$, is:

$$P_{\rm tr} = \frac{R_*}{a}, \tag{1.1.2}$$

where a is the planet's semi-major axis. This result shows that planets around small stars and with large orbital distances will have a very low transit probability. For reference, the transit probability for Jupiter at its current orbital distance $\sim 0.1\%$, whereas a typical hot-Jupiter will have a transit probability of $\sim 10\%$. Since the early radial velocity measurements revealed a population of massive close-in planets, the transit method thus became viable. The other source of bias is from a signal-to-noise ratio (SNR) argument. Following Wright et al. (2011); Kipping and Sandford (2016), assuming the transit can be modelled as a rectangular box function, the incident intensity in and out-of-transit can be modelled as:

$$I = \begin{cases} I_{\rm out} = \Gamma & \text{out-of-transit} \\ I_{\rm in} = \Gamma(1-\delta) & \text{in-transit,} \end{cases} \tag{1.1.3}$$

where Γ is the normalised average photon collection rate of the host star. The signal, S, is the fractional change in photon count, which is equivalent to the transit depth:

$$S = \delta = \frac{I_{\rm in} - I_{\rm out}}{I_{\rm out}}. \tag{1.1.4}$$

The uncertainty in the signal i.e. the noise $\sigma_{\rm S}$, is calculated in quadrature:

$$\sigma_{\rm S}^2 = \frac{I_{\rm out}^2 \sigma_{I_{\rm in}}^2 + I_{\rm in}^2 \sigma_{I_{\rm out}}^2}{I_{\rm out}^4}. \tag{1.1.5}$$

For a perfect detector, the photon-count can be modelled as a Poisson process, the uncertainty in $I_{\rm out}$ and $I_{\rm in}$ is an integrated average:

$$\sigma_I = \frac{\sqrt{\int_{t=0}^{t_{\rm f}} I\, dt}}{\int_{t=0}^{t_{\rm f}} dt}. \tag{1.1.6}$$

In this simple model, we observe a star for a given amount of time $T_{\rm obs}$ and the transit duration is $T_{\rm tr}$. We therefore integrate to $t_{\rm f} = T_{\rm tr}$ in-transit and $t_{\rm f} = T_{\rm obs}$ out-of-transit. Substituting Eq. 1.1.3 into the above, one finds from Eq. 1.1.5 that:

$$\sigma_{\rm S} \approx \frac{1}{\sqrt{\Gamma T_{\rm tr}}}, \tag{1.1.7}$$

where I have assumed that $T_{\rm obs} \gg T_{\rm tr}$ and $\delta \ll 1$. As such, the SNR for a single transit is:

$$\text{SNR} = \frac{S}{\sigma_{\rm S}} \approx \delta\sqrt{\Gamma T_{\rm tr}}. \tag{1.1.8}$$

Naturally, detecting multiple transits of the same planet will increase the signal-to-noise, which we can account for with the following:

$$\text{SNR} = \left(\frac{R_{\rm p}}{R_*}\right)^2 \sqrt{\Gamma\, T_{\rm obs} \frac{T_{\rm tr}}{P}}, \tag{1.1.9}$$

where P is the orbital period of the planet. Note that in order to use the transit method to detect a planet and measure its orbital period, one typically requires ≥ 3 transits to make a statistically significant detection. Furthermore, any instrument that performs photometry will have an SNR threshold below which transits cannot be detected. As such, this additional bias is introduced which preferentially favours planets with small orbital periods and large transit depths. Note that this means that planets of a given size are easier to detect around smaller mass, and therefore smaller size stars. Note also that in this toy-model, I have assumed a box function for the transit. In reality, transits are more accurately modelled with a trapezoidal function (e.g. Mandel and Agol, 2002; Kipping and Sandford, 2016) which accounts for ingress and egress of the planet onto the stellar disc, which will be effected by limb-darkening. In addition, one should also account for the impact parameter b, which is a measure of how close the transit trajectory is to the stellar disc mid-plane. These effects will change the shape of the transit and are therefore taken into account for actual planet detection pipelines (e.g. Christiansen et al., 2012; Fulton et al., 2017).

The first planet to be detected transiting a Sun-like star was HD 209458b from Charbonneau et al. (2000), which was found to be $1.3 R_{\rm Jup}$ at an orbital period of 3.5 days. In the proceeding years, various transit surveys such as HATnet (Bakos et al., 2004), SuperWasp (Pollacco et al., 2006), KELT (Pepper et al., 2007) and CoRoT (Auvergne et al., 2009) were established to perform photometry on wide-field searches of large populations of prospective host stars. Whilst the first results from these surveys were typically follow-up observations of planets detected with radial velocity measurements, they quickly began detecting new exoplanets, many of which were also hot-Jupiters. However, it wasn't long before a new class of exoplanet was also detected by radial velocity and transit surveys, namely *super-Earths* or *sub-Neptunes*. The first of such planets discovered around a main sequence star, Gliese 876d, had a mass of $\sim 8.5 M_\oplus$ and orbital period of ~ 3 days (Rivera et al., 2005). Other detections confirmed that such planets had radii and masses between Earth and Neptune, and typical orbital periods between 1 and 100 days (e.g. Mayor et al., 2009, 2011). Like with hot-Jupiters, there are no analogues of such planets in our own Solar System and hence the emerging picture was that our Solar System may not be the ordinary. However, to answer this question rigorously requires a more complete picture of the exoplanet population, often referred to as *demographic* survey, which must take into account the biases of the adopted detection method. One of the key ingredients for a demographic survey is a well-defined sample of stars e.g. volume-limited, magnitude-limited etc., that are homogeneously observed by the instrument. This in-turn allows one to quantify the biases of the survey i.e. the probability of detecting a planet given its size, mass, orbital parameters which also include instrumental bias. Largely speaking, however, this is more observational challenging for a radial velocity survey versus a transit survey.

1.2 Demographic Surveys

As discussed, one of the ultimate goals in exoplanet detection is a demographic survey. The key question one can ask for a demographic survey is, given the sample of detected planets, what is the underlying distribution? The simplest example of such a demographic calculation is the occurrence rate i.e how many planets of a certain type exist given the observed sample?

With regards to demographic surveys, the issue with radial velocity measurements is that, for a statistically significant detection for a single star, long observations with very stable spectrometers are required. For reference, the radial velocity of the Sun due to Earth's orbit is ~ 10 cm s^{-1}, which is only now a possibility in the near-future with next generation high-precision radial velocity instruments. As a result, observational targets are typically chosen *a priori* such that they are high signal-to-noise and likely to yield radial velocity signatures. Note that one cannot simply combine all historic radial velocity measurements performed by a given instrument to create an *ad hoc* demographic survey, since then one would need to quantify the probability of a team being granted observational time for a given planet, which is of course not possible. Even so, successful demographic RV surveys were performed in a hand full of cases (e.g. Howard et al., 2010; Mayor et al., 2011) which suggested that, although more hot-Jupiters had been detected, they were in-fact far less common when compared to smaller mass planets. Nonetheless, it was the transit survey performed by *Kepler* that would truly shed light on this topic, owing to the benefits of a dedicated exoplanet transit survey.

Kepler was the first (and only, to-date) large-scale demographic survey of its kind with the precision capable of detecting an Earth-like planet around a Sun-like star (Borucki et al., 2010). The space-based telescope stared at a 115 square-degree field-of-view and collected continuous optical photometry for $\sim 150,000$ stars simultaneously (Borucki et al., 2011). The stars were picked with a well defined magnitude-limited selection-function implying that biases were quantifiable (e.g. Christiansen et al., 2012; Fulton et al., 2017). In Howard et al. (2012), the first calculation of occurrence rate from the first 16 months of data (from Batalha et al., 2013) confirmed that super-Earths and sub-Neptunes were in-fact more populous that hot-Jupiters. Further demographics analysis by Fressin et al. (2013); Petigura et al. (2013) would refine these calculations to account for *false-positive events*. Transit surveys are, by nature, searching for transient astrophysical events and are therefore prone to incorrectly identify objects such as eclipsing binaries as transiting exoplanets. To remedy this, works by Fressin et al. (2013); Petigura et al. (2013) performed simulations of *Kepler* targets with injected false positives and fed those simulations into the *Kepler* detection pipeline to determine if this could mimic the transit that was actually detected. As such, they determined that the overall false-positive rate of *Kepler* detections was $\sim 10\%$. A similar technique was also used to accurately quantify survey biases. By injecting modelled transit signatures into the data, one can determine if the *Kepler* pipeline would be able to detect the given planet around a given star. This technique, often referred to as *pipeline injection recovery*, is hence a way to measure the signal-to-noise threshold (see Eq. 1.1.9) and overall detection efficiency. The injection recovery fraction for the *Kepler* pipeline is shown in Figure 1.1, which shows that a SNR ~ 10 was required for a probable detection. Over the course of the *Kepler* mission, pipeline injection recovery was performed rigorously in the works of Christiansen et al. (2012, 2015, 2016), allowing the early demographic results of Fressin et al. (2013); Petigura et al. (2013) to be performed. In addition, later demographic analyses have considered the effect of planet multiplicity on occurrence rates (e.g. Mulders et al., 2015; Zhu et al., 2018; Mulders et al., 2018; Zink et al., 2019). These works have shown that the occurrence rates of giant planets and super-Earth/sub-Neptunes are $\sim 5\%$ and $\sim 40\%$ respectively. In particular, it is expected that at least $\sim 40\%$ of Sun-like stars have nearly co-planar planetary systems with seven or more exoplanets.

Throughout its life, *Kepler* discovered over 2500 transiting exoplanets, leading to an unprece-

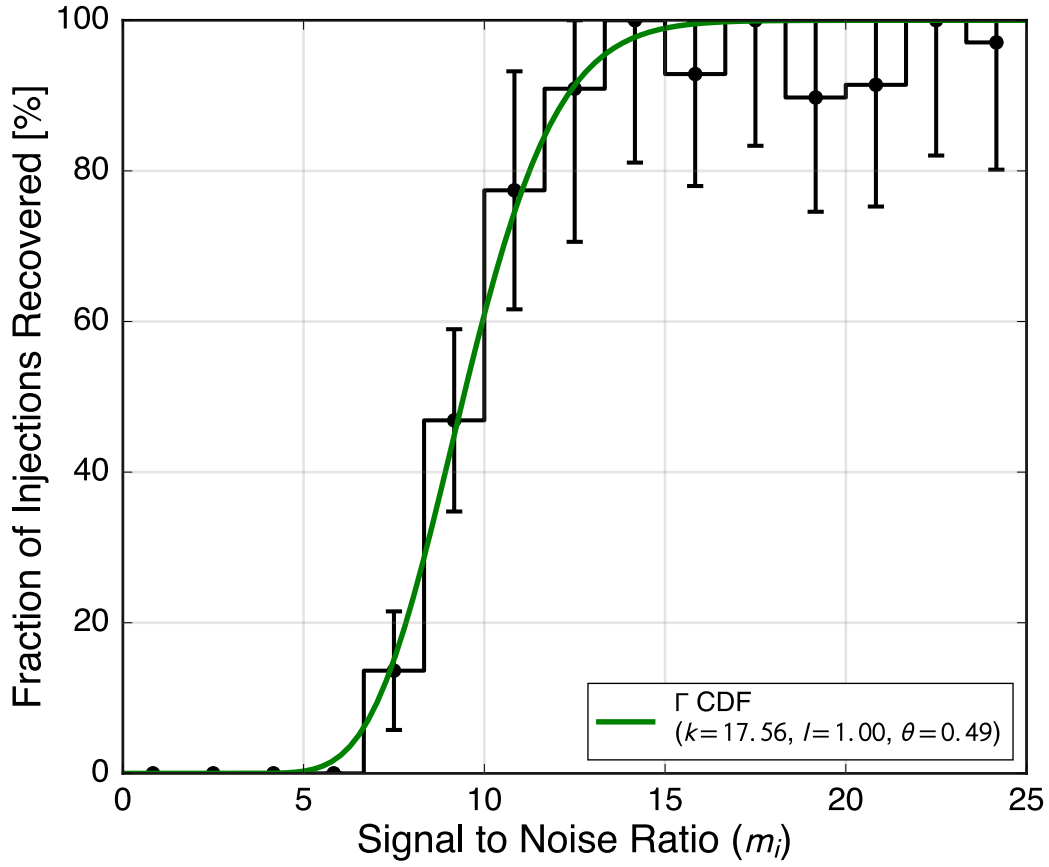


Figure 1.1: The injection recovery for the Kepler pipeline is shown as function of signal-to-noise ratio (SNR), which represents the probability of a planet being detected. One can see that planets have a higher probability of being detected for an SNR $\gtrsim 10$. In green is an analytic fit for injection recovery from Christiansen et al. (2015, 2016). Figure taken from Fulton et al. (2017).

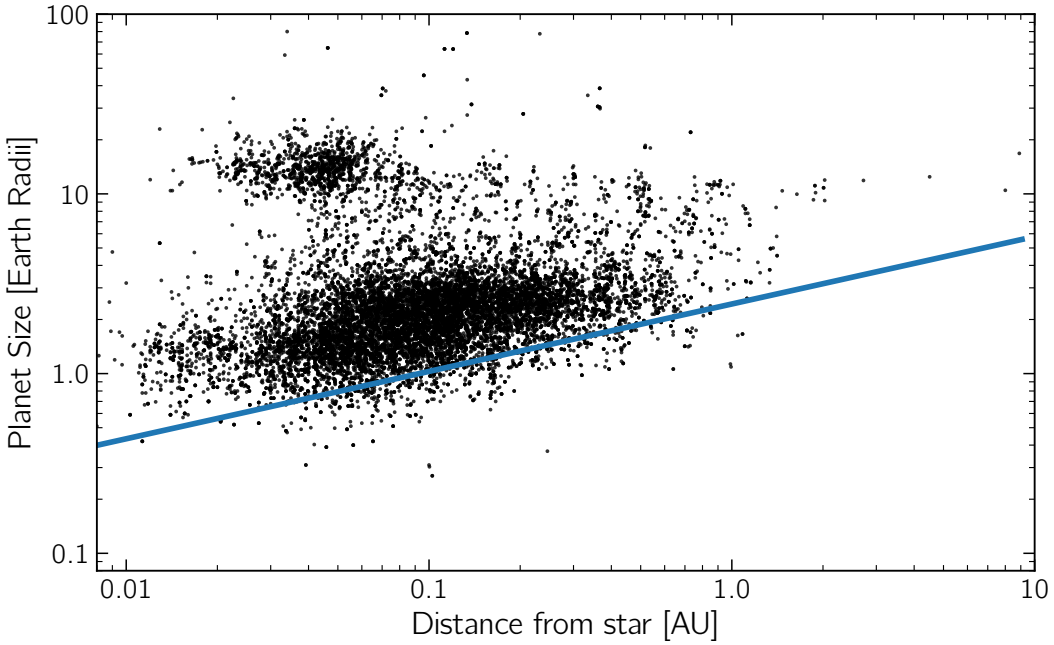


Figure 1.2: Orbital separations and size are shown for exoplanets detected by Kepler space-mission. The majority have sizes between $1-4R_\oplus$ *and orbits of* $\sim 0.1AU$*, providing a stark difference to the planets in our own Solar System. Note however that is an incomplete view of the underlying population i.e. transit survey such as Kepler preferentially detect large planets at short orbital periods. Kepler required typical signal-to-noise ratio of* ~ 10 *which is shown in blue from Eq. 1.1.9 for a typical photometric precision and observation times. Data taken from NASA exoplanet archive, accessed 25/05/22.*

dented view of planets in our galaxy, which is shown in Figure 1.2. One can clearly see two populations of planets: the first being the hot-Jupiters, which are $\sim 10R_\oplus$ in size and within 0.1 AU (e.g. Fischer and Valenti, 2005; Winn et al., 2010; Wright et al., 2012), and the other being the population of super-Earths and sub-Neptunes (within $0.01-1$AU and $\lesssim 4R_\oplus$). It is this latter population of planets which are the focus of this thesis, which is inspired by one of the key questions to come from the *Kepler* survey: how did these planets form, and why were they so common?

1.3 Mass-Radius Diagrams

As discussed, transit observations provide planet sizes. If one is to move forward in understanding the nature of super-Earths and sub-Neptunes, one must obtain mass measurements. Although the radial velocity method was originally used to detect planets, more recently it is predominantly used as a tool for follow-up observations i.e. finding the mass of a transiting exoplanet detected by missions such as *Kepler* or *TESS* (Ricker et al., 2015). In addition, one can use transit-timing variations i.e. small changes in transit periodicity in multi-planet systems to also infer planet masses (e.g. Agol et al., 2005; Holman and Murray, 2005). From mass and radius, one can calculate the bulk density.

It is worth noting that during a transit observation, the orbiting planet travels across the face of the star in chord-like trajectory, with the width (and hence the transit depth) according to the optical depth of the planet. For a naked core[2] a transit simply probes the core radius, but for a planet with a sufficiently large atmosphere, it measures the radius at which the atmosphere

[2]For the entirety of this thesis, 'core' refers to the non-gaseous portion of a planet, as opposed to the geological interpretation which refers to the iron core at the centre of terrestrial planets.

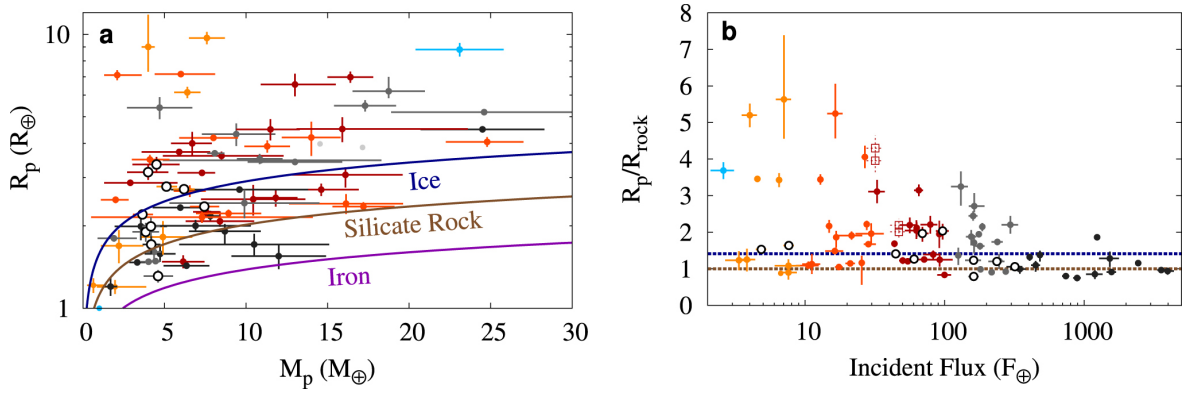


Figure 1.3: In the left-hand panel, measured masses and radii are shown for small close-in exoplanets. Composition curves from Fortney et al. (2007) are shown for bodies of pure iron, silicate and ice compositions. In the right-hand panel, planets sizes are shown relative to how large they would be, given their mass, if made of pure silicate rock. This is shown as a function of incident flux, which also colour-codes all points in both panels. In the right-hand panel, lines of constant composition are shown for silicates and ice in brown and blue respectively. Planet above this latter line must host significant H/He atmospheres to explain their underdensities. Figure taken with permission from Jontof-Hutter et al. (2016).

becomes optically thick - typically near the photosphere of the planet[3] (Rogers and Seager, 2010, see also Chapter 2). One can confirm this with a simple calculation: assuming Rayleigh scattering, the photon cross-section for molecular hydrogen at a given wavelength, λ, is (e.g. Vardya, 1962):

$$\sigma_{H_2} = 6.6 \times 10^{-5} \text{ cm}^{-2} \left(\frac{\lambda}{600 \text{ nm}} \right)^{-4}. \tag{1.3.1}$$

In order for an isothermal hydrogen-dominated atmosphere to be optically thick within a single scale height, the number density of molecular hydrogen required is:

$$n_{H_2} \approx 3.2 \times 10^4 \text{ cm}^{-3} \left(\frac{M_p}{5M_\oplus} \right) \left(\frac{R_c}{1.5R_\oplus} \right)^{-2} \left(\frac{T_{eq}}{500\text{K}} \right)^{-1} \left(\frac{\lambda}{600 \text{ nm}} \right)^{4}, \tag{1.3.2}$$

where R_c is the core radius and T_{eq} is the planetary equilibrium temperature. This shows that a very small amount of molecular hydrogen is required in order to be optically thick. It also suggests that a very small amount of mass is needed to increase the size of a planet significantly. Note that this is explored in Chapter 2. As a consequence, the measured radii in transit observations probe the atmosphere (if one is present) and not just the core radius. In this case, one would expect the bulk density of the planet to be reduced compared to the case of a bare core.

A mass-radius relation seeks to quantify how the size of a planet varies with mass and is therefore a function of planet composition i.e. bulk density. In early work from Wu and Lithwick (2013) and Weiss and Marcy (2014), the mass and radii for samples of exoplanets were investigated to reveal that the bulk density of super-Earths/sub-Neptunes decreases with increasing radii. They found a joint relation, which was consistent with an Earth-like density ~ 5.5 g cm^{-3} for planets $< 1.5R_\oplus$ and significantly less dense for larger planets. In addition, they found that the bulk density decreases for planets with a higher incident stellar flux. Since then, many works have corroborated this finding (e.g. Hadden and Lithwick, 2014; Rogers, 2015; Dressing et al., 2015; Wolfgang et al., 2016; Chen and Kipping, 2017). A summary of these findings is shown in Figure 1.3, in which planets with mass and radii measurements are shown in the left-hand mass-radius plane. In addition, the right-hand panel shows the size of these planets relative to

[3]This is the case for an atmosphere in which the scale height is far smaller than the photospheric radius.

the size of body made of pure rock given their mass. In other words, planets with a composition consistent with pure silicate rock will have a value of 1, whilst over-dense/under-dense planets will sit below/above this line respectively. This metric is shown as a function of incident flux. One can clearly see that a large population of planets are under-dense, which suggests that the planets host volatiles which reduce their bulk density. One possibility for this is the presence of water-ice, a so-called 'water-world' (Valencia et al., 2007; Zeng et al., 2019). For reference, Figure 1.3 also shows a horizontal line which is corresponds to bodies made of pure water-ice. One can see that even this scenario is not consistent with many of the planets. One finds that the only way to construct a planet with such a reduced bulk-density is with the presence of an extended H/He atmosphere (e.g. Adams et al., 2008; Wu and Lithwick, 2013; Rogers and Seager, 2010).

As discussed, a mass-radius relation allows one to begin to place constraints on planet formation and evolution. In general, planets will form whilst immersed in a protoplanetary disc which surrounds its young protostar and provides the solid material to form a solid core. The fact that many detected planets required large H/He dominated atmospheres fits with the core-nucleated gas accretion model (e.g. Perri and Cameron, 1974; Mizuno, 1980; Pollack et al., 1996). In this scenario, planets that have grown a solid core of sufficient mass will start to accrete gaseous material from the local protoplanetary disc. As this atmosphere cools, it contracts, which allows more material to in-fall onto the planet, increasing the mass of the atmosphere. Since the molecular cloud that collapsed to form the protostar and disc is H/He dominated, it follows that the primordial atmospheres of such planets should be the same. Note that the physics of gas accretion is explored in detail in Chapter 5. Another proponent in favour of large extended atmospheres comes from the incident flux trend that can also be seen in the right-hand panel of Figure 1.3. One can see that planets at high-incident flux tend to be consistent with a rocky body, which fits the hypothesis that the atmospheres of such planets are not stable to the high levels of irradiation, potentially inducing atmospheric escape. In addition, it is worth noting the work of Ciardi et al. (2013), in which it was noted that in many *Kepler* systems with multiple super-Earth/sub-Neptunes, planets at shorter orbital periods were smaller than counterparts further out from the host star. This ordering was not present for pairs of planets $\lesssim 3R_{\oplus}$. Again, this is suggestive of particularly close-in planets being unable to host large atmospheres.

Although insightful, mass-radius relations are limited by degeneracy. For a planet with a given mass and radius, there are many internal compositions e.g. rock mass-fractions, ice mass-fractions, atmospheric mass-fractions that are consistent with these observations (e.g. Rogers and Seager, 2010). Ultimately, gaining an understanding in the composition of these planets is crucial as it will place constraints on planet-formation models. One approach is the *migration model* (e.g. Ida and Lin, 2005, 2010; Bodenheimer and Lissauer, 2014; Raymond and Cossou, 2014; Bitsch et al., 2018; Raymond et al., 2018), in which planets form in the outer disc whilst immersed in a solid-enhanced region beyond the point where water exists as ice, instead of vapour. This point is often referred to as the ice-line and is typically at ~ 5AU for a protoplanetary disc around a solar-mass star. The planets that form beyond the ice-line then migrate inwards to the orbital period we observe them at today and will therefore have core compositions consistent with one rich in water/ice. Another popular approach however, that was developed in response to the ubiquity of close-in exoplanets, is the *in-situ model* (e.g. Hansen and Murray, 2012; Chiang and Laughlin, 2013), in which planetary embryos form in the inner disc, close to the location we observe them at today. As a result, the constituents that accreted to build up their cores are the silicate materials that drift into the inner-disc (Chatterjee and Tan, 2014; Jankovic et al., 2019). One would thus predict a core composition of such planets to be similar to that of Earth. In reality, core formation may draw upon physics from both schemes as it is well established that solids must migrate towards the star in the form of pebbles or embryos (e.g. Weidenschilling, 1977; Panić et al., 2009; Andrews et al., 2012; Isella et al., 2012; Powell

et al., 2017). What is not clear, even from mass-radius relations, is which of the models is the main driver of core formation.

The composition degeneracies are deepened when considering gas accretion. As discussed in Lee et al. (2014); Lee and Chiang (2015a), gas accretion plays a more important role in the outer disc since equilibrium temperatures are lower, meaning cooling is more efficient. Therefore, planets that form in the outer disc should accrete significant atmospheric mass fractions before migrating inwards. In the inner disc, planets still accrete atmospheres, albeit at a slower rate. It was noted that even planets forming at ~ 0.1AU would undergo accretion so efficiently that they should grow into hot-Jupiters. To explain the abundance of close-in super-Earths/sub-Neptunes, which clearly do not host atmospheres that dominate their mass, Lee et al. (2014); Lee and Chiang (2015a) argue that accretion may be delayed via dynamical friction. Note that I discuss this issue in-depth in Chapter 5. Nonetheless, the addition of an atmospheric-mass fraction to the problem of internal structure of exoplanets only compounds the degeneracy at hand.

Despite the aforementioned problems of internal planet structure, one can theoretically break the degeneracies if the evolution of the planet in known. This allows one to 'rewind the clock' on evolution and find the planet properties, and therefore histories, that are consistent with the observations. This concept is a strong theme in this thesis. Atmospheric evolution provides one such opportunity for this idea, particularly atmospheric escape. As discussed, there is strong observational evidence to support that highly irradiated planets tend to have higher bulk-densities and therefore are not likely to host large atmospheres. This suggests that the trends seen in exoplanet demographics are largely imprints of evolution, instead of formation. However, understanding evolution of observed planets is difficult since typical system ages for surveys such as *Kepler* are ~ 5 Gyr. Nonetheless, observational evidence does exist for escaping atmospheres.

1.4 Escaping Exoplanet Atmospheres

Direct observational evidence for escaping atmospheres has come from transit spectroscopy. In this technique, stellar light is absorbed by certain atomic lines which are used to probe the presence of extended/escaping atmospheres. The most popular of these is the Lyman-α line, which has a cross-section at line-centre of $\sigma_{\mathrm{Ly}\alpha} = 5.9 \times 10^{-14}$ cm^2 at 10^4 K gas (e.g. Hansen and Oh, 2006). Following Owen (2019), in order to provide unity optical depth, the number density of neutral hydrogen required in a transit is:

$$n_{\mathrm{H}} \approx \frac{1}{\sigma_{\mathrm{Ly}\alpha} R_{\mathrm{p}}} = 2.4 \times 10^2 \text{ cm}^{-3} \left(\frac{R_{\mathrm{p}}}{R_{\mathrm{J}}} \right)^{-1}, \tag{1.4.1}$$

where R_{J} is Jupiter's radius. Thus, a very small amount of neutral hydrogen is required to produce a large transit depth, particularly at large distances from the planet. One of the best examples of such an observation is for the warm sub-Neptune GJ436b from Ehrenreich et al. (2015), in which a 56% transit depth was detected in Ly-α when compared to the 0.69% in a standard broadband optical transit. The evolution of the Ly-α line for GJ436b is shown in Figure 1.4. The first feature to notice is the large absorption in the Ly-α line-core which is due to the interstellar medium. Despite this hindrance, there is still information that remains in the line-wings. In particular, one can see an additional absorption in the blue-wing of the Ly-α line when measured in-transit and post-transit (red and green lines respectively). This is indicative of a comet-like tail trailing the planet, with material moving towards the observer at speeds of up to 100km s^{-1}. This confirms that hydrogen gas is escaping beyond the planet's Roche lobe. This observation, amongst many others of hot Jupiters and sub-Neptunes (e.g. Vidal-Madjar et al., 2003; Lecavelier Des Etangs, 2007; Kulow et al., 2014; Lavie et al., 2017;

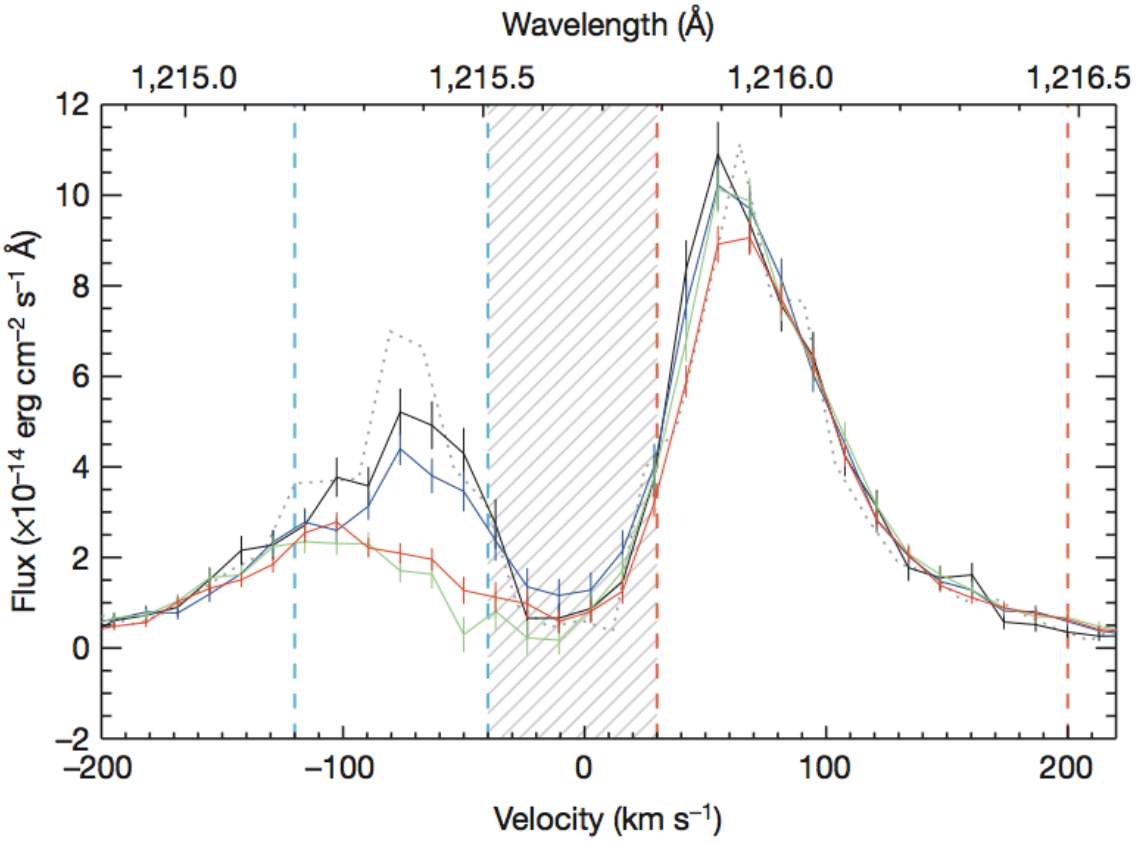


Figure 1.4: Flux for the Ly-α transit spectroscopy of GJ436b shown against wavelength and velocity. Averaged transits are shown in black, blue, red and green for out-of-transit, pre-transit, during-transit and post-transit respectively with 1σ uncertainties shown. The line core, shown in hatched region, demonstrates severe absorption due to the interstellar medium. The blue-wing however demonstrates an excess absorption during and post-transit which is suggestive of material escaping the planet (towards the observer) in a comet-like tail, absorbing Ly-α in the process. Figure taken from Ehrenreich et al. (2015).

Bourrier et al., 2018; dos Santos et al., 2020; Zhang et al., 2022) demonstrate a large absorption post optical-transit i.e. after the planet has left the stellar disc. Other lines can also be used to probe atmospheric escape such as Hα (e.g. Jensen et al., 2012; Yan and Henning, 2018), He 1083 nm (e.g. Spake et al., 2018; Ninan et al., 2020) or other UV metal lines (e.g. Vidal-Madjar et al., 2004; Loyd and France, 2014; dos Santos et al., 2019). It is therefore in no doubt, that many close-in planets are undergoing atmospheric escape.

What is causing atmospheres of such planets to escape? To understand this, one can compare the ratio, Λ, of gravitational binding energy of a gas particle to its thermal energy:

$$\Lambda = \frac{GM_{\rm p}\mu m_{\rm H}}{k_{\rm B}TR_{\rm p}}. \tag{1.4.2}$$

When this ratio is large i.e. for high mass, small and cold planets, material is tightly bound to the planet and can only escape via a process referred to as Jeans escape (Jeans, 1925). In this process, gas particles escape in the upper regions of the atmosphere, which are so underdense that particles no-longer collide with each other. If a particle has sufficient thermal energy such that it exceeds the escape velocity, it can leave the gravitational influence of the planet unhindered. On the other hand, if an atmosphere is less bound and, more specifically, the gas particles undergo many collisions over one scale height, then the atmosphere can be treated as a fluid and material is lost via a hydrodynamic outflow. When compared to Jeans escape, this

form of mass-loss yields much higher mass-loss rates due to the bulk flow of material. For the case of close-in exoplanets, the intense UV radiation from the host star is capable of dissociating the molecular hydrogen. As a result, the upper atmosphere is dominated by atomic hydrogen which is heated to typical temperatures of 10^4 K and therefore weakly bound to the planet. Since the atmosphere can be treated as collisional, the resulting mass-loss is hydrodynamic, yielding high mass-loss rates when compared to Jeans escape. This is explored in-detail in Chapter 2.

Hydrodynamic mass-loss takes the form of a transonic wind. Originally, this was applied to the stellar wind in Parker (1958), in which the isothermal fluid equations of motions were solved in non-hydrostatic equilibrium to show that the only physical solution is one in which the fluid velocity increases with distance, often referred to as a 'Parker wind'. The point at which the flow becomes supersonic is referred to as the sonic (or Bondi) radius and is controlled by the mass of the planet and the temperature of the outflow. Although originally utilised to describe the stellar wind, this solution is in-fact more appropriate to planetary outflow since strong magnetic fields have an important effect on the stellar wind. The theory of transonic winds is covered in Chapters 2 and 5. In order to launch a transonic wind from a planet, heating from high energy flux is required. Furthermore, in order to understand the dynamics of such a wind, hydrodynamic simulations are required.

In an attempt to understand UV-driven atmospheric loss, Watson et al. (1981) presented the notion of 'energy-limited' mass-loss, which at the time was applied to the hydrodynamic atmospheric escape of Earth and Venus. In this prescription, high-energy UV radiation is absorbed in a thin layer corresponding to the point where the gas becomes optically thick i.e. opaque to such photons. Watson et al. (1981) argued that, instead of adiabatically expanding and cooling, the material rising towards this point was heated by conduction. The mass-loss rate is thus 'energy-limited' by the balance between conductive heating and adiabatic cooling. In Lammer et al. (2003), this prescription was used to estimate the hydrodynamic escape rates for hot Jupiters. Later however, works from Waite et al. (1983); Yelle (2004) showed that the conduction invoked in Watson et al. (1981) provided a very small contribution to the heat flux of the problem, particularly for highly irradiated planets. As a result, the mass-loss rate estimates from Watson et al. (1981); Lammer et al. (2003) were overestimates. Importantly, one must consider a more accurate picture of radiative transfer to understand the problem. Indeed the works that followed, including Yelle (2004); Tian et al. (2005) and García Muñoz (2007), considered this to arrive at more sensible mass-loss estimates. In Murray-Clay et al. (2009), a coupled radiative-transfer and hydrodynamic scheme was used to show that radiative recombination of hydrogen i.e. protons capturing electrons, has an important role in mass-loss. During electron capture, UV photons are produced which act to cool the hydrogen gas, however these UV photons can also go on to ionise other hydrogen atoms. As such, for high-fluxes, the ionisation structure of the atmosphere can be altered. Murray-Clay et al. (2009) and subsequent works by (Bear and Soker, 2011) and (Owen and Alvarez, 2016) have shown that the outflow can be recombination-limited if the recombination timescale is shorter than the flow timescale, or energy-limited if this is not the case. In the former case, mass-loss rate scales approximately with square root of UV flux, whilst in the latter it scales linearly with UV flux.

Nowadays, the aforementioned energy-limited mass-loss prescription is removed from the original definition from Watson et al. (1981) that described the balance between conductive heating and adiabatic cooling. It now refers to high-energy flux doing work to remove atmospheric mass. In time Δt, one can assume that high-energy stellar flux $F_{\rm HE}$ removes an amount of mass, ΔM, to infinity i.e does work of $GM_{\rm p}\Delta M/R_{\rm p}$ at some efficiency, η. Since the flux is received over an

area of $\pi R_{\rm p}^2$, one can rearrange to show that mass-loss rate, $\dot{M} \equiv \Delta M/\Delta t$, is therefore:

$$\dot{M} = \eta \frac{\pi R_{\rm p}^3 F_{\rm HE}}{G M_{\rm p} K_{\rm eff}}, \tag{1.4.3}$$

where $K_{\rm eff}$ is a factor introduced by Erkaev et al. (2007) to account for the fact the material need not be removed to infinity, but only to the Roche radius in order to completely escape the planet's gravitational influence. In order to use this formalism, one must understand the full extent of $F_{\rm HE}$ i.e. which wavelengths are appropriate for atmospheric escape, as well as understand the nature of the escape efficiency η.

In the literature discussed thus far, the high-energy flux was only considered to be the extreme ultra-violet (EUV), which has typical energies of $h\nu \gtrsim 13.6$eV. X-rays however, which are generated in the stellar corona, are also capable of ionising hydrogen atoms, with typical energies of $h\nu \approx 0.1 - 10$keV (e.g. Erkaev et al., 2007; Lopez and Fortney, 2013). For a star on the main-sequence, X-rays contribute a negligible fraction to a star's bolometric output, however for young stars $\lesssim 100$ Myrs, this is not the case. As shown in Figure 1.5, observations from multiple works (e.g. Ribas et al., 2005; Jackson et al., 2012; Shkolnik and Barman, 2014; Tu et al., 2015; McDonald et al., 2019; Johnstone et al., 2021) have shown that high-energy photon production is at its highest whilst the star is young with the its highest rotation period. During this phase, referred to as the saturation period, stars have a typical ratio of X-ray and EUV (XUV) to bolometric luminosity of $L_{\rm XUV}/L_{\rm bol} \sim 10^{-3}$ (e.g. Güdel, 2004). Beyond ~ 100 Myr however, the star's rotation begins to slow as material is removed in the form of a stellar wind, causing angular momentum loss (e.g. Hempelmann et al., 1995; Wright et al., 2011). During this spin-down period, the mechanism that generates high-energy photons becomes less efficient and $L_{\rm XUV}/L_{\rm bol}$ rapidly falls on a timescale that is quicker than 1/age. This stellar evolution suggests that the mass-loss history, particularly for low-mass planets, may be dictated by the XUV flux in the first ~ 100 Myr (García Muñoz, 2007; Lammer et al., 2013, 2014, 2016; Erkaev et al., 2013, 2016). Note that recent work from King and Wheatley (2021) has suggested that EUV flux may decay slower than X-ray flux, meaning it may be important for planet evolution on Gyr timescales. I address this in Chapter 6.

In light of the potential high-impact of X-ray flux for planets around young stars, Owen and Jackson (2012) performed hydrodynamic simulations to differentiate between the X-ray-dominated and EUV-dominated regime of mass-loss. As discussed, EUV heating occurs via ionisation of atomic hydrogen, and cooling is controlled by hydrogen recombination which also produces further EUV photons. As such, EUV heated hydrogen gas thermostats to $\sim 10^4$ K. In the X-ray regime however, heating is controlled by photo-electron generation by the K-shells of metals such as carbon and oxygen. The heating timescale for this is much quicker than EUV since the photo-electrons are typically at $\sim 10^6$ K. Cooling in X-ray heated gas is also controlled by the metals, specifically through line-cooling. Since X-rays are more energetic, they will tend to penetrate, and thus heat, further into the atmosphere. Owen and Jackson (2012) showed that X-ray dominated escape is relevant in the case where the sonic point is within the inner X-ray heated region, whilst EUV dominated escape occurs when the sonic point lies in the outer EUV heated region. They also showed that as a host star spins down and generates less X-rays, the outflow will transition from X-ray to EUV dominated. Regardless, it is the X-ray driven outflow that dominates the atmospheric evolution due to its higher mass-loss rates.

As discussed, another important aspect of the energy-limited mass-loss formalism is the escape efficiency η. Based on their hydrodynamic models, Owen and Jackson (2012) provided calculations of this by comparing the energy input by X-ray heating with the mechanical luminosity of the wind. They found that the efficiency is far from constant, and varies for planets of different masses and sizes with an upper limit of $\sim 20\%$. As such, results that assume a constant mass-loss efficiency are to be taken with caution.

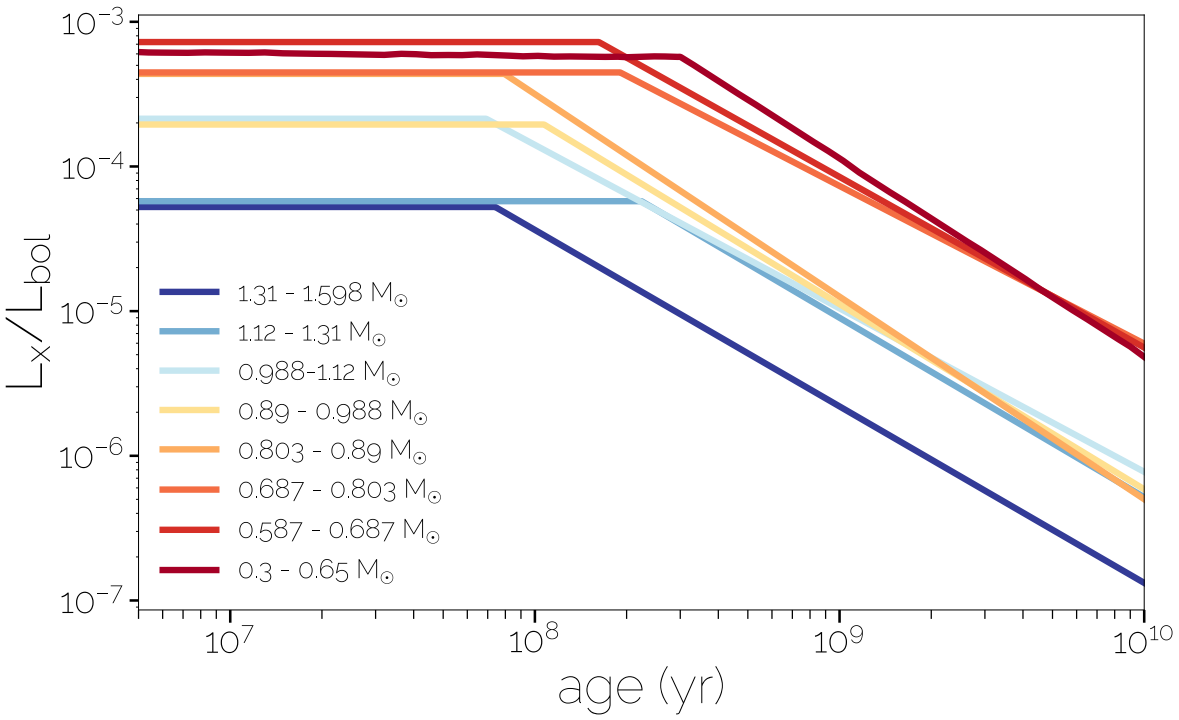


Figure 1.5: The evolution of X-ray to bolometric luminosity ratio is shown for different stellar types. For the first $\sim$ 100 Myr, stars begin with a high rotation period and a 'saturated' X-ray output, which is generated in the stellar corona. However, beyond this point, strong magnetic stellar winds cause the star to spin-down, which dramatically reduces the X-ray output. Figure taken from McDonald et al. (2019) with data from Jackson et al. (2012) and Shkolnik and Barman (2014).

With regards to the hydrodynamic simulations, only studies involving one-dimensional calculations have been discussed thus far. In reality, most of the small, close-in exoplanets will be tidally locked i.e. with a permanent day and night-side. Dramatic temperature and pressure gradients may cause rapid day-to-night flow that affect the outflow structure. Indeed this was confirmed in two and three-dimensional calculations from Stone and Proga (2009); Owen and Adams (2014); Tripathi et al. (2015), although it was noted that the average mass-loss rates were similar to that of one-dimensional simulations. This is not the case for planets with significant magnetic fields however, as shown in works from Adams (2011); Trammell et al. (2011); Owen and Adams (2014); Khodachenko et al. (2015); Arakcheev et al. (2017). In this case, the ionised outflow is strongly coupled to the planetary dipole field, resulting in reduced mass-loss, particularly near the closed field lines near the planet equator. Mass-loss is possible however towards the poles since thermal pressure opens the field lines. Nonetheless, in the case of a field strength of $\sim$ 5 Gauss, these works have shown that mass-loss can be reduced by up to an order of magnitude. Currently, there are no observational constraints on the magnetic fields of small, close-in exoplanets, hence work on this particular topic is still in its infancy. Nonetheless, provided that magnetic fields are not strong, one can safely assume that 1D models provide a good description for mass-loss of exoplanets.

1.5 The Radius Gap

One of the most significant theoretical predictions from atmospheric escape studies came from the works of Owen and Wu (2013) and Lopez and Fortney (2013), in which the long-term evolution for populations of small, close-in exoplanets were considered in the X-ray dominated mass-loss regime. Previous studies into atmospheric escape had only considered short-term evo-

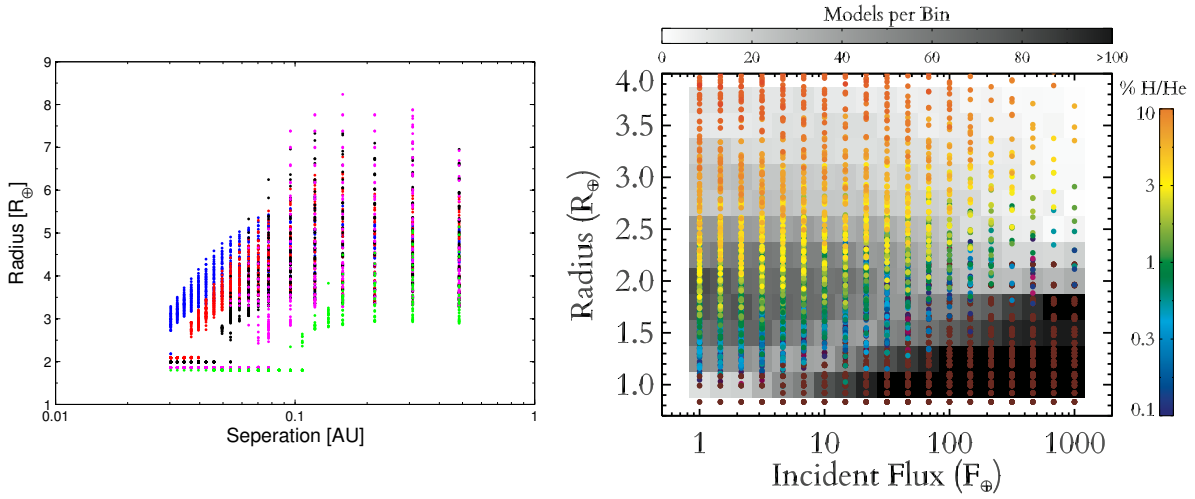


Figure 1.6: Predictions of the radius gap from Owen and Wu (2013) (left) and Lopez and Fortney (2013) (right). Both studied the long-term thermal and mass-loss atmospheric evolution of small, close-in exoplanets to reveal a dearth of planets in the orbital separation-radius plane or, equivalently, the flux-radius plane. This arises since planets with low core mass and high irradiation will be completely lose their atmosphere through photoevaporation, whilst others will maintain a atmospheric mass fraction of $\sim 1\%$.

lution, in-part because numerical hydrodynamic simulations are computationally expensive to run over long timescales. However, the evolution of such close-in exoplanets over Gyr timescales undergoing dramatic atmospheric escape had not been fully modelled. To overcome this, Owen and Wu (2013) used a grid of mass-loss rates and hence accurate photoevaporation efficiencies from the hydrodynamic models of Owen and Jackson (2012), which were implemented in the stellar evolution code `MESA` (Paxton et al., 2011, 2013, 2015, 2018) to perform long models which couple thermal and mass-loss evolution for multiple close-in planets. In Lopez and Fortney (2013), a simpler energy-limited approach was adopted, building upon previous work in Lopez et al. (2012) in which long-term evolution calculations were performed for the Kepler-11 system. Despite different methodologies, both of these works considered the evolution starting in the saturation phase of young stars, ranging to typical ages of *Kepler* targets ~ 5 Gyr. At the beginning of the evolution, the planets are young and high in entropy, thus beginning life with inflated atmospheres. This increased cross-sectional area, coupled with the high levels of X-rays generated by the host stars, suggested that the mass-loss history for such planets is dominated during the first ~ 100 Myr. The first conclusive result from these studies was that photoevaporation had little effect on the evolution of hot-Jupiters, owing to their massive and self-gravitating atmospheres. However, for super-Earths and sub-Neptunes, there were certain cases in which atmospheres could be completely stripped from the planet. Both works predicted the existence of a critical area of parameter space, specifically planets with low core mass and high irradiation, in which planets could not maintain an atmosphere during the first 100 Myr. The subsequent prediction from this were two populations of planets: those which could maintain atmospheres and those which could not. In between these two populations was an area that was devoid of planets at $\sim 2R_\oplus$, which is shown in Figure 1.6 for both studies. This predicted sparsity of planets was referred to as the 'photoevaporation valley' or 'radius gap'. This prediction was corroborated by subsequent works of Jin et al. (2014); Howe and Burrows (2015); Chen and Rogers (2016).

At the time, *Kepler* had only revealed a peak in planet occurrence at radii $\sim 2R_\oplus$ (Fressin et al., 2013; Petigura et al., 2013). Note however that the typical radii measurement uncertainties were $\sim 40\%$, meaning that the radius gap, present or not, could not be resolved. As shown in Eq.

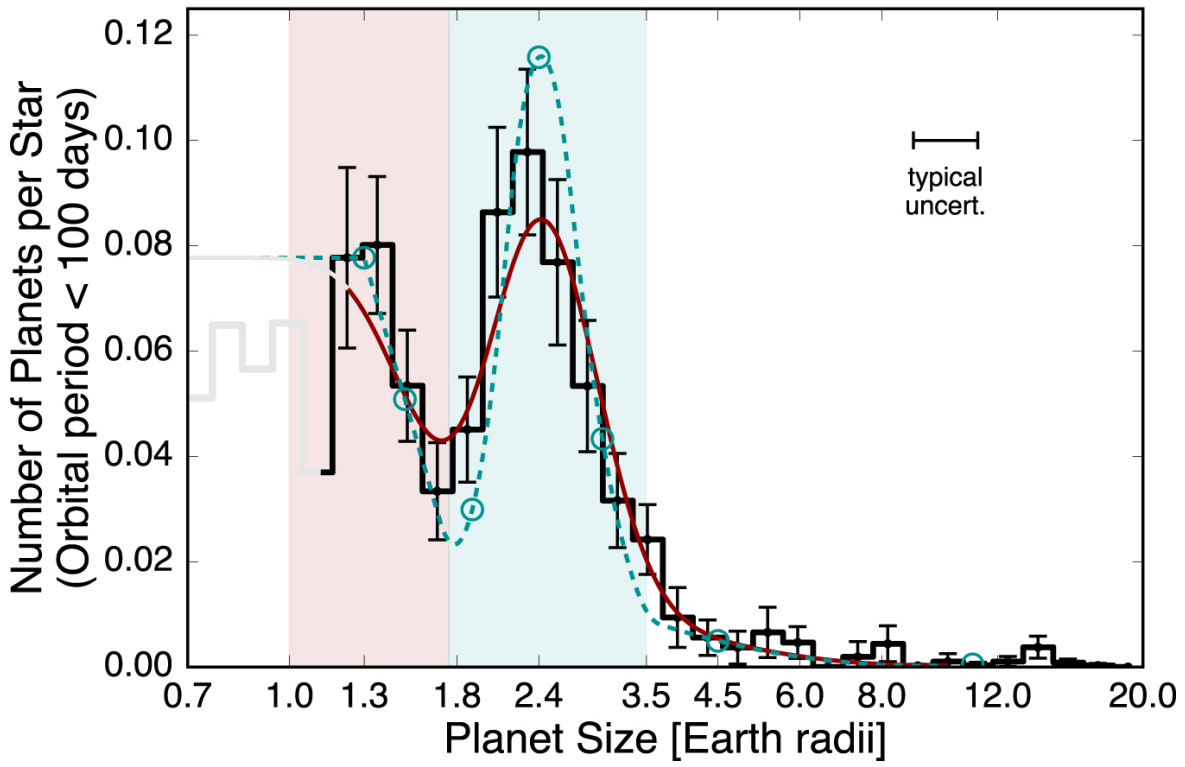


Figure 1.7: The radius distribution of small, close-in exoplanets taken from Fulton et al. (2017). A bimodal distribution around $\sim 1.8R_\oplus$ provides a clear distinction between super-Earths which lie on the left (shaded in red) and sub-Neptunes on the right of the radius gap (shaded in blue).

1.1.9, recall that the signal-to-noise ratio of a transit detection scales with the ratio of planet-to-stellar radius squared. As such, the uncertainty in stellar radius propagates forward to the uncertainty in planet size. For the first *Kepler* data releases, stellar radii were calculated from the *Kepler Input Catalogue* (KIC) which measured the stellar effective temperature through photometric surveys of host stars to an accuracy of $\sim 30\%$, which was then used in stellar models to calculate their size (Brown et al., 2011). To remedy this statistical constraint however, the *California Kepler Survey* (CKS) was established, which performed high-resolution spectroscopy of 1305 planet-hosting *Kepler* stars (Petigura et al., 2017; Johnson et al., 2017a). This added spectroscopic constraints to stellar parameters, meaning that typical stellar radii uncertainties dropped to $\sim 10\%$. In Fulton et al. (2017), these updated stellar parameters were used to perform a new demographic analysis of 900 small, close-in exoplanets, which now had typical radii measurement uncertainty of $\sim 12\%$. In this study, a bimodal distribution in the radii of planets was indeed observed with peaks at $\sim 1.3R_\oplus$ and $\sim 2.4R_\oplus$. A dearth of planets was observed in-between these populations at $\sim 1.8R_\oplus$, which was consistent with the predictions from photoevaporation. This result is shown in Figure 1.7. Note that the two populations at $1.3R_\oplus$ (shown in red) and $2.4R_\oplus$ (shown in blue) additionally provided a convenient way to distinguish super-Earths from sub-Neptunes, which is used as nomenclature for the remainder of this thesis. The result from Fulton et al. (2017) also demonstrated a negative slope in the period-radius plane (or equivalently a positive slope in the incident flux-radius plane) which additionally supported the atmospheric escape scenario shown in Figure 1.6.

In subsequent work from Van Eylen et al. (2018), the radius gap was again resolved, however with a smaller sample of *Kepler* planets for which the host stars had undergone asteroseismic follow-up to further constrain the stellar radius. Asteroseismology is the study of stellar oscillations, which for high-cadence and sufficient signal-to-noise ratio observations can lead to

exceptionally good inferences on properties such as stellar size and mass (Chaplin and Miglio, 2013). As a result, the typical planet size uncertainty for this sample was $\sim 3\%$, which yielded a radius gap that was completely devoid of planets when shown in the period-radius plane. When considering the gap as a power law, the slope in this study was found to have a gradient of $d\log R_{\rm p}/d\log P = -0.09^{+0.02}_{-0.04}$. In addition, the CKS collaboration further refined their data set in Fulton and Petigura (2018) by incorporating *Gaia* DR2 parallax measurements (Gaia Collaboration et al., 2018), allowing radii measurement uncertainty to reduce from $\sim 10\%$ to $\sim 5\%$. This further refined the radius gap, increasing its statistical significance. The *Gaia* stellar parameter constraints were also applied to the entire *Kepler* sample in Berger et al. (2020a), which showed tentative evidence for the fact that the ratio of super-Earths to sub-Neptunes increased with time, supporting the notion that planets evolve from atmosphere-hosting to stripped. This is also reinforced by the work of Van Eylen et al. (2021), in which masses and radii were considered for planets around M-dwarfs: they found that planets below the gap were consistent with a rocky composition, whilst planets above were consistent with hosting a $\sim 1-2\%$ H/He atmosphere.

1.6 CKS-Cool

The observational confirmation of the radius gap in Fulton et al. (2017) then allowed the introduction of comparisons between atmospheric escape models and the exoplanet population. In Owen and Wu (2017), a simple analytic model for XUV photoevaporation was shown to accurately reproduce the results from evolution models from Owen and Wu (2013). Using this analytic model, they provided predictions for the slope of the radius gap to be $d\log R_{\rm p}/d\log P \sim -0.16$, however noting that this slope depended on many model parameters. In addition, Owen and Wu (2017) also provided analytic arguments to suggest that the radius gap should move to larger sizes around smaller mass stars when keeping incident flux constant. Note that the physics behind the radius gap slope is explored in-detail in Chapter 3. At the time however, the CKS result was only presented for a very narrow range in stellar mass $0.8-1.4M_{\odot}$ (Fulton et al., 2017).

To explore how the radius gap changed as a function of stellar mass, the CKS collaboration endeavoured to extend their sample to cooler host stars and thus extended their stellar range down to $0.5M_{\odot}$. As part of this collaboration, hereafter referred to as CKS-Cool, I performed demographic analyses on the new data-set to understand how the radius gap changed under various conditions (Petigura et al., 2022). As part of the CKS-Cool collaboration, I developed a tool to characterise the position and slope of the radius gap in an robust manner. Note that in the study by Van Eylen et al. (2018), the radius gap was quantified using a Support Vector Machine (SVM) algorithm which classified the population of super-Earths from sub-Neptunes. This method was appropriate since the radius gap in this sample was devoid of planets and as such the algorithm could successfully classify the two populations in the period-radius plane. In the CKS sample however, the gap is not completely devoid of planets and instead presents as a significant reduction in planet detection/occurrence, due to statistical scatter caused by larger radii measurement uncertainties. As SVM algorithms were clearly not appropriate for the new CKS-Cool data release, I developed an algorithm which takes sequential slices at constant orbital period and finds the radius which minimises the planet detection/occurrence probability. These values therefore approximate the location of the gap as a function of orbital period (or other variables such as incident flux, stellar mass etc.), meaning a power law of the following form can be fit:

$$P(R_{\rm p}) = R_{\rm p,0}\left(\frac{P}{10\text{ days}}\right)^{m}, \tag{1.6.1}$$

where $R_{\rm p,0}$ and $m = d\log R_{\rm p}/d\log P$ are free parameters. As shown in Figure 1.8, I find that

the slope has a value of $d\log R_{\rm p}/d\log P = -0.10 \pm 0.03$ and does not change as a function of stellar mass. This is consistent with the prediction of XUV photoevaporation from Owen and Wu (2017), confirming that planets at further orbital separations experience less irradiation, meaning that the maximum size of stripped cores decreases. As with orbital period, the slope of the radius gap can be determined with respect to other variables, which are also shown in Figure 1.8. I find that the slope with incident flux, S, is $d\log R_{\rm p}/d\log S = -0.06 \pm 0.01$, with tentative evidence that the location of the radius gap decreases in size for larger stellar masses. As discussed previously, this is also consistent with photoevaporation, although a more rigorous analysis of this trend is presented in Chapter 3. With stellar mass, the slope is found to be $d\log R_{\rm p}/d\log M_* = 0.18 \pm 0.08$. This positive slope is indicative of the fact that larger stars will have larger luminosities, meaning that more massive planets can be stripped. However, it is interesting to note that the average sizes of sub-Neptunes increases with stellar mass, yet stays constant for super-Earths. This can be explained with a correlation between core mass and stellar mass, and is discussed in-detail in Chapter 3. Finally, the slope with respect to stellar metallicity, [Fe/H], is found to be $dR_{\rm p}/d[{\rm Fe/H}] = 0.01 \pm 0.06$. Theoretically, if one assumes that stellar metallicity is correlated with planet metallicity (since they both formed from the same circumstellar disc), then planetary atmospheres should therefore have higher opacity. Such atmospheres will cool less efficiently, meaning that one might expect sub-Neptunes to be more inflated around high metallicity stars. This appears not to the be the case, as shown in Figure 1.8. Note that Owen and Murray-Clay (2018) also highlight that metals such as carbon and oxygen allow efficient radiative line-cooling for hydrogen-dominated gas heated to $\sim 10^4$ K, which may inhibit mass-loss around higher metallicity stars. This is all, however, complicated by the fact that *Kepler* was a magnitude-limited survey, meaning that stellar mass, metallicity and age are heavily correlated. Thus, interpretation of results is non-trivial. I comment on these correlations in Chapter 3.

Despite the huge advances that have come since the first demographic studies from *Kepler*, there is a limit as to how much one can learn in simply comparing predictions of models, such as radius gap slope, with the population of observed exoplanets. The more exciting prospect is the ability to do *physics-based inference*, in which one exploits a physical process, in this case atmospheric escape, to break the degeneracy in internal structure of exoplanets. This is akin to studies in the field of cosmology, in which models such as Λ-CDM are used in physics-based inferences with the cosmic microwave background data to place constraints on parameters that are known to control our Universe; such as the Hubble constant, H_0, and cosmological parameters for matter and dark energy, $\Omega_{\rm m}$, Ω_Λ. The question is, instead of these variables, what do we learn from physics-based inference on exoplanet populations?

1.7 Physics-Based Inference

As discussed, the distribution of small, close-in exoplanets is successfully explained by XUV photoevaporation. The presence of a demographic feature and a robust evolutionary model opens up the exciting possibility of breaking the degeneracy in the mass-radius diagram using physics-based inference. One of the first attempts to perform this was in-fact applied to the Kepler-36 system in Owen and Morton (2016). This system contains two planets at very similar orbital periods ($\sim$ 14 and $\sim$ 16 days) yet very different bulk densities ($\sim$ 7.5 g cm^{-3} and $\sim$ 0.9 g cm^{-3}). As shown previously in Lopez and Fortney (2013), this system architecture can be explained under the photoevaporation model if the inner planet, that receives more irradiation, has a smaller core mass which causes the atmosphere to be removed from the planet. In Owen and Morton (2016), a grid of `MESA` models, similar to that of Owen and Wu (2013) was calculated for the Kepler-36 system to find which initial conditions were required to fit the planets that are observed today. In doing so, they used the accurate TTV mass

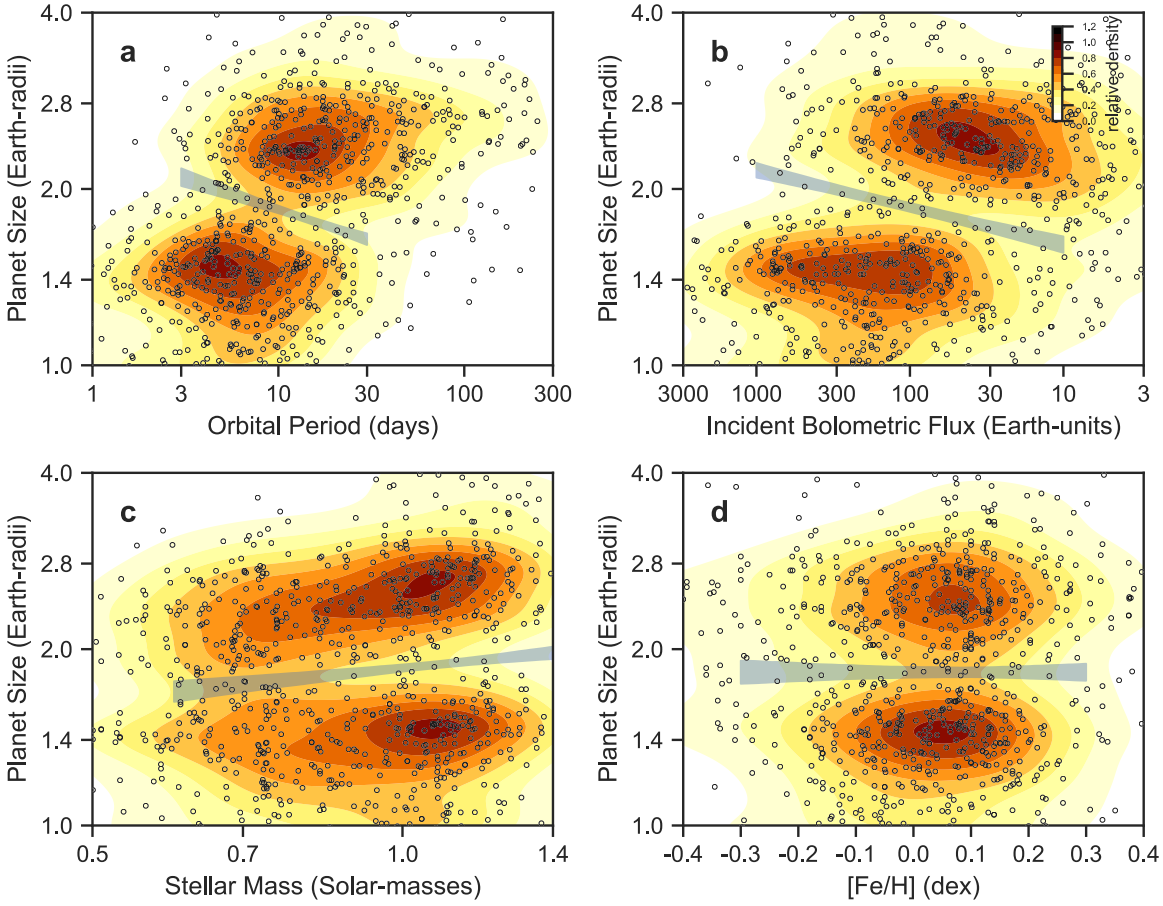


Figure 1.8: Planet sizes are shown as a function of orbital period, incident flux, stellar mass and metallicity in panels a-d from the CKS-Cool data release in Petigura et al. (2022). Orange contours show a Gaussian Kernel Density Estimation for number density of exoplanet detections. As part of the CKS-Cool collaboration, I calculated the position and slope of the radius gap, with results shown in blue. Uncertainties are calculated by bootstrapping over the planet detections and shown to 1σ.

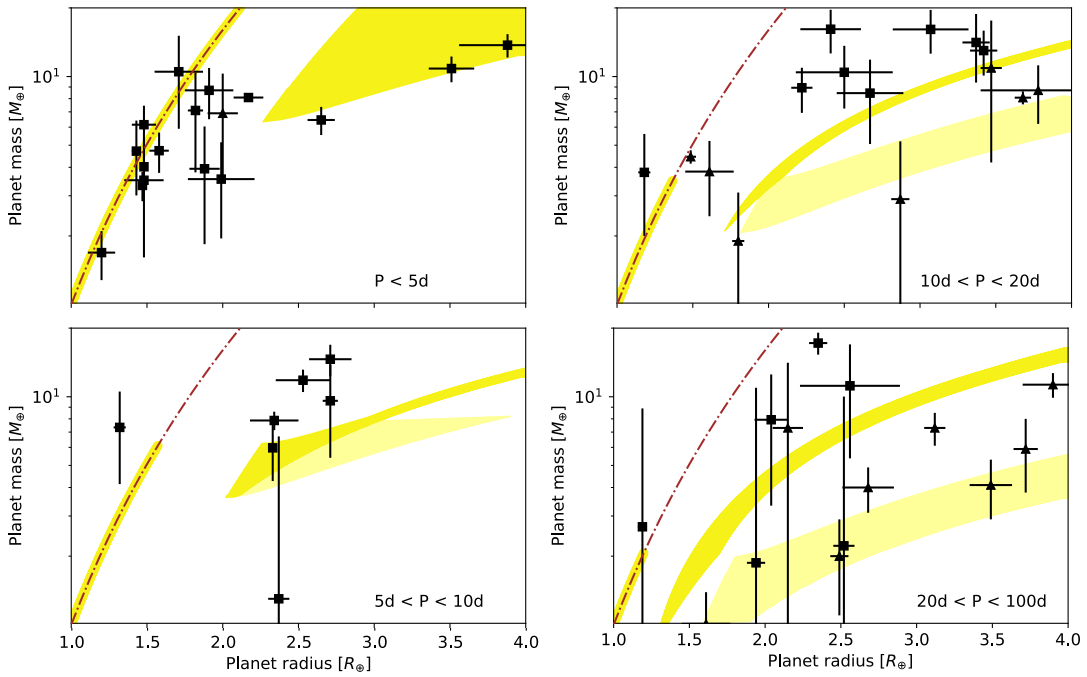


Figure 1.9: Planets from Wolfgang et al. (2016) are shown in the radius-mass plane for varying orbital period ranges. Yellow and light-yellow regions represent the predicted minimum and maximum atmospheric sizes of planets from Jankovic et al. (2019), assuming they go through in-situ formation, gas accretion and XUV photoevaporation. For many planets, their size is smaller than the minimum value, suggesting that their predicted atmospheric mass fraction is over-predicted by gas accretion. Figure taken from Jankovic et al. (2019).

measurements from Carter et al. (2012) to place constraints on the core mass, core composition and initial atmospheric mass of the two planets, since these parameters are required to set the atmospheric evolution of the planet. Firstly, the core mass sets the gravitational potential and so controls atmospheric contraction and also mass-loss rates; the core composition sets the size of stripped cores; and finally the initial atmospheric mass fraction sets the initial size of the atmosphere and hence cross-section to incoming XUV radiation. Note that these parameters are not observable using standard techniques and so constraints on them are invaluable in breaking the internal structure degeneracy in the mass-radius diagram. Owen and Morton (2016) performed a Bayesian inference model which suggested that both planet cores were consistent with an Earth-like composition and also both started with a significant atmospheric mass fraction, however the closer of the two planets with a smaller core mass, was completely stripped to become a super-Earth.

Another example of physics-based inference comes from Thorngren and Fortney (2018), in which interiors models were applied to a population of hot-Jupiters to determine the cause of their inflation (e.g. Miller and Fortney, 2011; Laughlin et al., 2011). They showed that the results were consistent with Ohmic dissipation (e.g. Batygin and Stevenson, 2010), which transfers magnetic energy from the atmosphere of the planet into the interior. Although not applicable to the super-Earths and sub-Neptunes which are the focus of this thesis, this study was the first example of physics-based inference being applied to a population of exoplanets to understand the physics that controls them.

As discussed, an analytic model for XUV photoevaporation and thermal evolution was developed in Owen and Wu (2017). The benefit of this simple model, which is presented in Chapter 2, is that it is computationally cheap to evaluate and can therefore be used to perform demographic comparisons. Similar to Kepler-36, Owen and Wu (2017) showed that the core composition of

all super-Earths and sub-Neptunes must be approximately the same as Earth. This is since ice/iron rich cores are less/more compressed and therefore have larger/smaller radii for a given mass. In order for the population of super-Earths to be consistent with the data from Fulton et al. (2017), they found (by eye) that this particular Earth-like composition was required. In Wu (2019), the same analytic model was used in a χ^2-minimisation study, to fit the planet sizes of the XUV photoevaporation model to the CKS data (Fulton et al., 2017). In the process, they placed constrains on the core mass and initial atmospheric mass fraction distributions, further confirming that physics-based inference was key to understanding the internal structure of such exoplanets. However, these works were heuristic in nature, and not statistically rigorous. Firstly, in order to compare models with data, one must properly incorporate the incompleteness of survey into the models. Statistically speaking, the correct way to do this is with a Bayesian hierarchical model. Additionally, Wu (2019) only used the radii of planets in their comparison. In reality, there is a wealth of information stored in the orbital periods of such planets, since the proximity to the host star plays a key role in determining atmospheric evolution. Finally, the distributions for core mass and initial atmospheric mass fraction were constrained to a specific functional form (in this case a log-Gaussian). Since we have no other constraints on these distributions, this may have not been an applicable model for the data.

In light of this, one of the main goals for this thesis is to perform the inference problem in-full: a Bayesian hierarchical model to infer the internal structure of small, close-in exoplanet by exploiting XUV photoevaporation. This is presented in Chapter 4, where statistically rigorous constraints are placed on the core mass, core composition and initial atmospheric mass fraction distributions. I discuss the implications these have on planet formation and evolution.

It is worth noting that, although the XUV photoevaporation model does an excellent job in explaining the origin and structure of the exoplanetary distribution we see today, it is in tension with core-nucleated accretion theory (e.g. Perri and Cameron, 1974; Mizuno, 1980; Pollack et al., 1996). In Jankovic et al. (2019), this was demonstrated by simulating the formation of planetary cores, followed by core-nucleated accretion and finally photoevaporation. The final mass and radii of these planets were then compared to current observational data to show that these values do not agree, with the model producing a higher atmospheric mass fraction than is observed. This is shown in Figure 1.9, in which minimum and maximum planet sizes are shown for planets in the radius-mass diagram. One can see that many detected planets (taken from the Wolfgang et al. (2016) sample) have radii smaller than the predicted minimum values from Jankovic et al. (2019). In summary, core-nucleated gas accretion theory predicts an excess of mass that photoevaporation cannot remove in order to fit the data we see today. There are three possibilities for a solution: classical core accretion over-predicts atmospheric mass; photoevaporation underestimates atmospheric mass-loss; or there is an intermediary mass-loss mechanism that brings the two models into alignment. With regards to the first option, Lee et al. (2014) suggested that gas accretion may be delayed due to dynamical friction, meaning that significant accretion can only occur as the disc is dispersing. As a result, planets will host a smaller atmospheric mass than in the standard formalism. In Chapter 5 however, I explore the last of these three options: that there is an intermediary mass-loss mechanism that removes mass prior to the onset of XUV photoevaporation. In Owen and Wu (2016), a mass-loss mechanism was proposed that occurs during protoplanetary disc dispersal, referred to as "boil-off". Although discs survive for $\sim 3-10$ Myr, the inner disc disperses in $\sim 10^5$ yrs (e.g. Kenyon and Hartmann, 1995; Ercolano et al., 2011; Koepferl et al., 2013). As a result, the confining pressure that once held the planet's atmosphere in equilibrium is quickly removed, inducing rapid mass-loss since the planet cannot remain in hydrostatic equilibrium. In Owen and Wu (2016), evolutionary models were performed with `MESA` that showed that up to an order of magnitude in atmospheric mass can be removed. However, two important assumptions were made that I address in Chapter 5: firstly the disc was assumed to disperse instantaneously; and

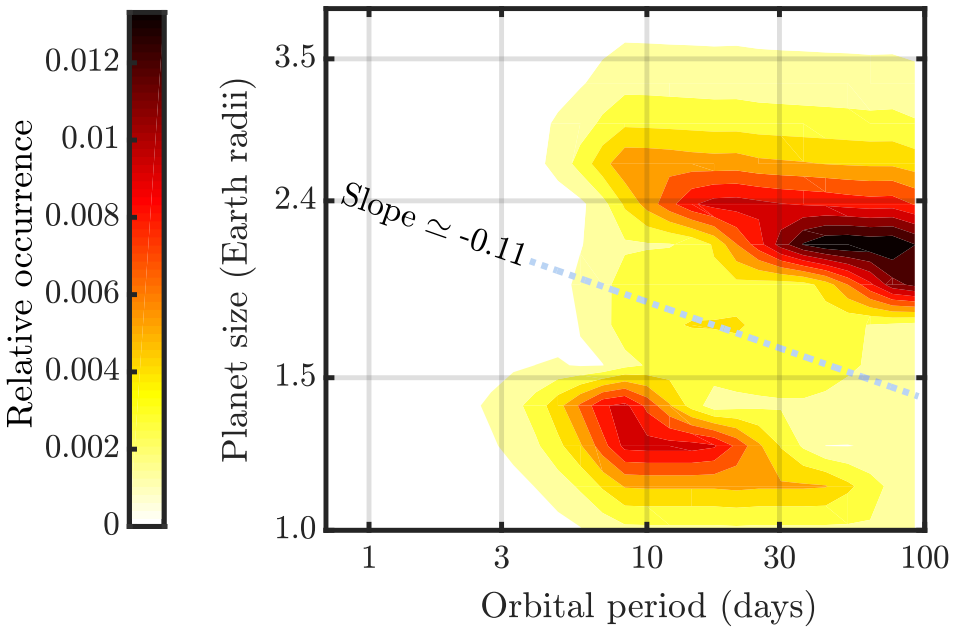


Figure 1.10: Relative occurrence of planets shown with the orbital period and planet size for the core-powered mass-loss model of Ginzburg et al. (2018); Gupta and Schlichting (2019). Similar to XUV photoevaporation, this model can accurately reproduce the radius gap as observed in multiple works (e.g. Fulton et al., 2017; Van Eylen et al., 2018). Figure taken from Gupta and Schlichting (2019).

secondly the planet began in a maximal entropy state. These assumptions are discussed in-detail in Chapter 5, however imply that the mass-loss rates determined in Owen and Wu (2016) were upper limits. I remedy these assumptions by self-consistently modelling the gas accretion and mass-loss of the planet through the lifetime of the disc in order to determine whether the discrepancy between photoevaporation and core-nucleated gas accretion can be resolved.

However, prior to physics-based inference in Chapter 3 and boil-off models in Chapter 4, I first focus on a complication on the topic of atmospheric escape. In the work from Ginzburg et al. (2018); Gupta and Schlichting (2019), an alternative model to XUV driven mass-loss was presented. They argue that high-energy flux is in-fact not required to form the radius gap, but instead the remnant energy from formation stored in the solid core inflates the atmosphere. This heating, catalysed by the stellar bolometric irradiation, was shown to be capable of inducing mass-loss and reproducing the radius gap, as shown in Figure 1.10. This model is referred to as 'core-powered mass-loss'. Since one of the main goals of this thesis is to perform physics-based inference with atmospheric escape, the first step is to determine which of these mechanisms is dominant in atmospheric evolution. To do this, observational tests must be performed that can easily distinguish between these two models. In Chapter 2, I present a theoretical discussion of atmospheric structure and mass-loss, including a comparison between XUV photoevaporation and core-powered mass-loss. Then, in Chapter 3, I present an observational test which can be used to distinguish between the two models, which utilises how the XUV flux from host stars varies with stellar mass.

Chapter 2

The Theory of Atmospheric Mass-Loss and the Origin of the Radius Gap

As discussed in Chapter 1, atmospheric escape is likely sculpting the distributions we see in current exoplanet surveys. In this chapter I shall provide a theoretical overview of atmospheric structure relevant to super-Earth and sub-Neptunes. In addition, I shall discuss how mass-loss is modelled in more detail, before discussing how the aforementioned 'radius-gap' can be sculpted by XUV photoevaporation or core-powered mass-loss.

2.1 Collisional Fluids

I begin by discussing the atmospheric structure of a small planet with a solid core and hydrogen dominated atmosphere, which is thought to be typical of the majority of Kepler planets (e.g. Rogers, 2015; Wolfgang and Lopez, 2015; Jontof-Hutter et al., 2016). It is first convenient to define some important length scales for an atmosphere. The pressure scale-height H is the change in altitude required to decrease the pressure by a factor of e. For the simple case of an isothermal atmosphere, as discussed in Section 2.2, this can be shown to be:

$$H = \frac{k_{\rm B}T}{\mu m_{\rm H} g}, \tag{2.1.1}$$

where $k_{\rm B}$ is Boltzmann's constant, T is the temperature, μ is the mean molecular weight, equal to 2.35 for molecular hydrogen of solar abundance, $m_{\rm H}$ is the mass of a hydrogen atom and $g = GM_{\rm p}/r^2$ is the local gravitational acceleration at a radius of r for planet of mass $M_{\rm p}$.

The planet's *Hill radius* $R_{\rm H}$ is defined as the point away from a planet at which the gravitational force becomes dominated by another astronomical object. For a planet of mass $M_{\rm p}$ that orbits a star of mass M_* at a semi-major axis of a, the Hill radius is defined as:

$$R_{\rm H} \approx a \sqrt[3]{\frac{M_{\rm p}}{3M_*}}. \tag{2.1.2}$$

Finally, the *Bondi radius*, $R_{\rm B}$, is the point at which atmospheric gas has sufficient thermal kinetic energy to escape the planet's gravitational influence:

$$R_{\rm B} = \frac{GM}{2c_{\rm s}^2}, \tag{2.1.3}$$

where c_s is the gas sound-speed. The first step in understanding the physical properties of an atmosphere is to determine whether the gas is *collisional*. This is quantified via the collisional mean-free-path l_p, defined as the average distance a particle travels between two consecutive collisions. The *exobase* of an atmosphere is then defined as the radius at which the mean-free-path is equal to one pressure scale-height. Above this radius, the collisions are so infrequent that gas with sufficient kinetic energy to exceed the escape velocity i.e. at the high-end tail of the Maxwell velocity distribution, can leave the gravitational influence of an atmosphere without collisions hindering its outward trajectory. This form of mass-loss is referred to as *Jeans escape* (e.g. Jeans, 1925; Tian et al., 2005; Volkov et al., 2011; Owen, 2019). If however, we wish to model an atmosphere as a fluid, and hence governed by fluid equations of motion, we must ensure it is collisional up to the Bondi radius. As we shall see in Section 2.3, this is because the Bondi radius also holds a strong importance for hydrodynamic outflow; it is the point at which atmospheric gas will become supersonic and thus causally disconnected to the base of the atmosphere.

Before continuing, let us confirm that the atmosphere is indeed collisional up to the Bondi radius. The mean-free path is given by:

$$l_p = \frac{1}{n\sigma}, \tag{2.1.4}$$

where n is the number density and σ is the scattering cross section. For a typical super-Earth at an equilibrium temperature of 1000K, the atmosphere has the following properties at the Bondi radius: $n \approx 10^{13}$ cm^{-3} and $\sigma \approx 10^{-15}$ cm^2, hence $l_p \approx 100$ cm. From Eq. 2.1.1, a typical scale height for such an atmosphere $H \approx 10^8$ cm, which is far greater than the mean-free path and hence a collisional fluid is a good approximation for an atmosphere up to the Bondi radius.

2.2 Fluid Equations and Hydrostatic Equilibrium

Since we now know that our atmosphere can be treated as a collisional fluid, let us now consider the fluid equations. First, we have the mass continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}) = 0, \tag{2.2.1}$$

where ρ is the mass density and $\boldsymbol{u}$ is the velocity. In addition, we have momentum conservation:

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = -\frac{1}{\rho}\nabla P + \boldsymbol{g}, \tag{2.2.2}$$

where P is the gas pressure and $\boldsymbol{g}$ is the gravitational force on a parcel of gas. We can instantly begin to simplify these equations with the statement of hydrostatic equilibrium. In this scenario the atmosphere is in force balance, with gas pressure resisting gravitational collapse, requiring $\boldsymbol{u} = \boldsymbol{0}$ and time derivatives equal to zero. This is a good approximation for planets which are not subject to strong forces that may introduce inflow or outflow ($\boldsymbol{u} \neq \boldsymbol{0}$) such as photoevaporative winds (as discussed in 2.3), Roche overflow (e.g. Ford and Rasio, 2006; Valsecchi et al., 2014; Jackson et al., 2016) or giant impacts (e.g. Genda and Abe, 2003, 2005; Schlichting et al., 2015). Following works such as Piso and Youdin (2014); Lee et al. (2014), in the case of hydrostatic equilibrium and spherical symmetry, it is convenient to rewrite the mass continuity equation in terms of mass:

$$\frac{\partial m}{\partial r} = 4\pi r^2 \rho, \tag{2.2.3}$$

and the momentum equation becomes:

$$\frac{\partial P}{\partial r} = -\frac{Gm}{r^2}\rho. \tag{2.2.4}$$

where m refers to the mass enclosed in a sphere of radius r. In addition, we require an equation to describe how temperature varies with altitude:

$$\frac{\partial T}{\partial r} = \frac{T}{P}\frac{\partial P}{\partial r}\nabla, \tag{2.2.5}$$

where ∇ is the temperature gradient, defined as:

$$\nabla \equiv \frac{\partial \ln T}{\partial \ln P}. \tag{2.2.6}$$

Finally, an equation is needed to determine the luminosity structure of the planet:

$$\frac{\partial L}{\partial r} = 4\pi r^2 \rho(\epsilon + \epsilon_{\rm g}), \tag{2.2.7}$$

where the gas is heated due to gravitational contraction at a rate of $\epsilon_{\rm g} = -TdS/dt$ per unit mass, where S is the specific entropy. The internal heat per unit mass on the other hand, such as that generated by radionuclides, is generated at a rate of ϵ. These four equations i.e. Eqs. 2.2.3, 2.2.4, 2.2.5 and 2.2.7, in combination with an equation of state (EOS), can be solved to find the full structure of a planet in hydrostatic equilibrium and are often referred to as the *stellar structure equations* since they apply to stars as well.

The form of ∇ is crucially important with regards to atmospheric structure and depends on how energy is transported i.e. radiation or convection. We will revisit this in detail in Section 2.4 and 2.5, however for now we employ an isothermal ideal gas equation of state, which greatly simplifies the problem:

$$P = c_{\rm s}^2 \rho. \tag{2.2.8}$$

As a result, the temperature is trivially solved for all r and we can separate and solve Eq. 2.2.4:

$$\int_{\rho_0}^{\rho} \frac{d\rho'}{\rho'} = -\frac{GM}{c_{\rm s}^2}\int_{r_0}^{r} \frac{d\rho'}{\rho'^2}, \tag{2.2.9}$$

where our boundary condition is set by $\rho = \rho_0$ at r_0. Finally, our density profile solution for isothermal hydrostatic equilibrium is:

$$\rho = \rho_0 \exp\left\{\frac{GM}{c_{\rm s}^2}\left(\frac{1}{r} - \frac{1}{r_0}\right)\right\}. \tag{2.2.10}$$

Assuming an isothermal ideal gas equation of state, the sound speed $c_{\rm s}$ is given by:

$$c_{\rm s} = \sqrt{\frac{k_{\rm B}T}{\mu m_{\rm H}}}, \tag{2.2.11}$$

and from inspection of the hydrostatic equilibrium density profile, we see that the value of r required to reduce the density by a factor of e is indeed the scale-height H of Eq. 2.1.1.

This case of isothermal hydrostatic equilibrium is a good place to start. However, as we shall see in Section 2.5, the assumption of an isothermal temperature profile is only valid in certain atmospheric regimes. In addition, the effect of stellar heating can induce mass-loss, often referred to as a wind, which means that our initial assumption of zero velocity in the momentum equation must be reassessed.

2.3 Isothermal Parker Wind

Before diving into a more accurate picture of atmospheric structure, let us go off on a short tangent and solve the isothermal fluid equations without assuming hydrostatic equilibrium i.e. solving for a non-zero velocity $\boldsymbol{u}$. This analysis was first proposed by the late, great Eugene Parker (Parker, 1958), when considering the dynamics of the solar corona. This gave rise to the concept of the stellar wind, which we may use as an analogy to a planetary wind. In order to solve analytically, we will again set time derivatives to zero, meaning we will arrive at a steady-state solution. Following Cranmer (2004) and assuming 1D spherically-symmetric flow, the mass conservation equation of Eq. 2.2.1 now takes the form:

$$\frac{1}{r^2}\frac{\partial}{\partial r}(\rho u r^2) = 0, \tag{2.3.1}$$

which boils down to the statement that mass-loss rate is constant with radius:

$$\dot{M} = 4\pi\rho u r^2 = \text{const.} \tag{2.3.2}$$

Momentum conservation on the other hand now takes the form:

$$u\frac{\partial u}{\partial r} + \frac{1}{\rho}\frac{\partial P}{\partial r} + \frac{GM}{r^2} = 0. \tag{2.3.3}$$

As before, we use an isothermal ideal gas equation of state to eliminate P and also note that from mass conservation:

$$\frac{1}{r^2}\left(ur^2\frac{\partial \rho}{dr} + \rho r^2\frac{\partial u}{\partial r} + 2\rho u r\right) = 0. \tag{2.3.4}$$

Hence we have that:

$$\frac{1}{\rho}\frac{\partial \rho}{\partial r} = -\frac{2}{r} - \frac{1}{u}\frac{\partial u}{\partial r}, \tag{2.3.5}$$

which is substituted into Eq. 2.3.3 to find the equation of motion for a steady state flow:

$$\left(u - \frac{c_\mathrm{s}^2}{u}\right)\frac{\partial u}{\partial r} = \frac{2c_\mathrm{s}^2}{r} - \frac{GM}{r^2}. \tag{2.3.6}$$

Note that this equation has a trivial solution where $u = c_\mathrm{s}$. In this case, the left-hand side becomes zero, allowing us to solve for the location of this solution. We find that the flow velocity is equal to the sound speed when:

$$r = R_\mathrm{B} \equiv \frac{GM}{2c_\mathrm{s}^2}, \tag{2.3.7}$$

which is the same definition of the Bondi radius as in Eq. 2.1.3. Rewriting the equation of motion in terms of the Bondi radius yields:

$$\left(u - \frac{c_\mathrm{s}^2}{u}\right)\frac{\partial u}{\partial r} = 2c_\mathrm{s}^2\left(\frac{1}{r} - \frac{R_\mathrm{B}}{r^2}\right). \tag{2.3.8}$$

Whilst this can be solved numerically, it can also be separated and integrated by-hand to find a transcendental equation in the following way:

$$\int\left(u - \frac{c_\mathrm{s}^2}{u}\right)du = 2c_\mathrm{s}^2\int\left(\frac{1}{r} - \frac{R_\mathrm{B}}{r^2}\right)dr \tag{2.3.9}$$

which evaluates and rearranges to:

$$\ln\left(\frac{u}{c_\mathrm{s}}\right)^2 - \left(\frac{u}{c_\mathrm{s}}\right)^2 = -4\ln r - 4\frac{R_\mathrm{B}}{r} - \ln c_\mathrm{s}^2 + C \tag{2.3.10}$$

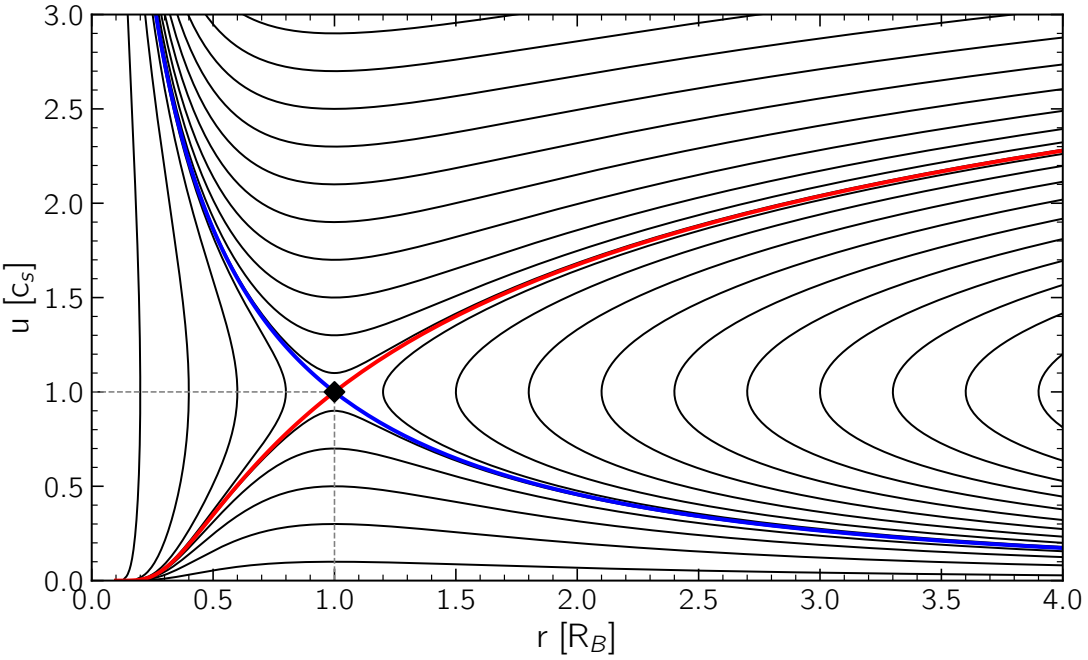


Figure 2.1: Solutions to the isothermal wind problem, presented in Section 2.3. Transonic solutions in red and blue represent an out-flowing wind and an in-flowing accretion solution respectively. In addition, breeze solutions are those that remain subsonic for all radii. Non-physical solutions are also shown, such as those that are multi-valued for a given radii, or those that are supersonic for all radii.

where we set the integration constant C with the boundary condition $u = u_{\rm s}$ at $r = R_{\rm B}$. Solutions to this equation can be written down in closed-form with the use of Lambert W functions (Cranmer, 2004), which are the multi-valued inversions to xe^x. There are an infinite many branches to the Lambert W function, satisfying:

$$W_k(z)e^{W_k(z)} = z, \tag{2.3.11}$$

where z is some arbitrary complex number and k is the branch label. If we require z to be real however, then the only branches that exist are $k = -1, 0$. The power of Lambert W functions become apparent when noting that the solutions to:

$$\ln(A + Bx) + Cx = \ln D, \tag{2.3.12}$$

can be written as:

$$x = \frac{1}{C} W\left[\frac{CD}{B}\exp\left(\frac{AC}{B}\right)\right] - \frac{A}{B}. \tag{2.3.13}$$

Since the transcendental equation of motion for the flow is of this form, we may write the solutions as follows:

$$u(r) = c_{\rm s}\sqrt{-W_k(-D(r))}, \tag{2.3.14}$$

where $D(r)$ is the exponential of the right-hand side of Eq. 2.3.10:

$$D(r) = \exp\left\{-4\ln r - 4\frac{R_{\rm B}}{r} - \ln c_{\rm s}^2 + C\right\}. \tag{2.3.15}$$

Different choices of branch k and boundary condition at $r = R_{\rm B}$ determine the nature of the solutions, which are shown in Figure 2.1. For example, the transonic solution is shown in red

and is subsonic at small radii and supersonic above R_B uses the $k = 0$ branch for $r < R_\mathrm{B}$ and $k = -1$ for $r > R_\mathrm{B}$. This solution represents a hydrodynamic outflow and is crucial for understanding exoplanet mass-loss. The counter-part to this solution is the accretion solution, as shown in blue, which accelerates as it approaches the planet, also becoming supersonic at the Bondi radius. This solution uses the $k = -1$ branch for $r < R_\mathrm{B}$ and $k = 0$ for $r > R_\mathrm{B}$. The other set of physical solutions are those that remain subsonic for all radii, known as breezes. These solutions are valid for systems with a sufficient pressure at infinity, which the transonic solutions do not require. For a planet in a vacuum this is appropriate, however this is not the case for planets immersed in a gaseous protoplanetary disc. These breeze solutions will be explored in detail in Chapter 5. All other solutions are unphysical, since they are either multi-valued for a given radius or supersonic for all radii.

The transonic outflow solution however, often referred to as a *Parker wind* (Parker, 1958), sheds light on outflow from planets and stars alike. We can see now, why we must ensure that the atmosphere is collisional up to Bondi radius, since if we are to model the outflow using fluid equations as we have done here, this condition must be met. In addition, beyond this altitude the flow becomes supersonic and is causally disconnected from the base of the flow. Since mass-loss rate is constant with radius, we therefore do not need to model the flow beyond this point and can assume the material is removed from the planet's gravitational influence. The case of Parker winds and non-hydrostatic equilibrium outflow in general will be revisited in detail in Chapter 5.

2.4 Radiation Transport

Let us now return to the case of hydrostatic equilibrium and therefore in the absence of mass-loss. Thus far, we have only considered an isothermal atmosphere. In reality however, we know that atmospheres are not exclusively isothermal e.g. Earth's atmospheric temperature drops with altitude near it's surface[1]. Since the temperature profile will have a strong influence on the atmospheric structure, we must take a closer look at the energy transport mechanisms that determine temperature.

The two mechanisms that are relevant to this problem are convection and radiation[2]. In general, the dominant mechanism is determined by the opacity of the material of interest. As we shall see, convection is a far more efficient method for energy transport but is only viable under certain conditions. To understand these conditions, I first present the formalism of opacity and radiation transport in general.

Radiative energy transport is the consideration of the interactions of photons with matter. The first place to start is with the *photon mean-free path* l_ph, not to be confused with the mean-free path l_p of gas particles in Section 2.1. In general, the photon mean-free path is a function of radiation frequency and position, since the material of interest may change properties in different locations. The *optical depth* τ can be thought of as the number of mean-free paths a photon travels whilst traversing a medium. If this number is very large i.e. $\tau \gg 1$, then the photon has undergone many interactions with the medium, which is thus described as *optically thick*. On the other hand, if $\tau \ll 1$ then the material is said to be *optically thin*.

We may define the optical depth between two locations s_0 and s_1 more rigorously with the extinction coefficient $\alpha_\nu \equiv 1/l_\mathrm{ph}$ for a given radiation frequency ν. Consider photons passing

[1]Note however that Earth's atmosphere is optically thin, whereas this is not the case for the majority of exoplanets that will be discussed in this thesis.

[2]Conduction is more appropriate for solids (such as Earth's core) and quantum-degenerate material (such as the interiors of white dwarfs).

through a gas cloud of density $\rho(s)$, the optical depth is given as:

$$\tau_\nu = \int_{s_0}^{s_1} \alpha_\nu(s)\, ds = \int_{s_0}^{s_1} \rho(s)\kappa_\nu(s)\, ds, \tag{2.4.1}$$

where κ_ν is the gas *opacity*, defined as $\kappa_\nu = \sigma_\nu/m$ where σ_ν is the cross-section for a photon interacting with a gas atom/molecule and m is the atom/molecule mass. Intuitively, the opacity is a measure of how easily photons can pass through a material. The higher the opacity, the harder it is for photons to pass through. Determining opacity is notoriously difficult since it requires full quantum mechanical calculations and/or laboratory experiments to determine the interaction of photons with atomic, molecular or ionic electron configurations. Although practically challenging, determining opacities is crucial as it encodes how much radiative energy is deposited in a layer of atmosphere. If more energy is absorbed in a given layer, it will be heated and increase in pressure. The resulting expansion is what determines the structure of an atmosphere.

In practise, opacities are integrated across all frequencies and presented in tables as a function of temperature and pressure for planetary material (e.g. Ferguson et al., 2005; Freedman et al., 2008). These are known as *Planck* or *Rosseland* opacities. For analytic models, it is convenient to fit power laws to these tables in the form of:

$$\kappa = \kappa_0 \left(\frac{P}{P_0}\right)^\alpha \left(\frac{T}{T_0}\right)^\beta, \tag{2.4.2}$$

where κ_0, P_0 and T_0 are the opacity, pressure and temperature of the gas at an outer boundary at $R_{\rm out}$.

The value of opacity of an atmosphere depends on the contributions from different photon absorption processes. In general, there are three types of interaction: bound-bound, bound-free and free-free. For bound-bound interactions, a photon excites a bound electron to a higher energy state. The absorption thus occurs at very specific frequencies due to the quantum nature of the transitions. Bound-free transitions on the other hand occur when a photon ionises a bound electron. In this case, the absorption can occur at any frequency above the critical ionisation level and the electron leaves with the excess kinetic energy. As a result, bound-free absorptions can increase the opacity dramatically below a certain wavelength (this is the cause of Lyman or Balmer breaks). Finally, free-free interactions occur when an electron absorbs a photon when interacting with an ion. This only becomes important at very high temperatures as it requires a high ionisation fraction.

In the case of super-Earth atmospheres, the upper atmosphere has low pressures and temperatures, meaning that free-free absorptions have a small contribution to the opacity. As a result, bound-free and bound-bound interactions play an important role in the overall opacity, as well as atomic/molecular rotation or vibrations modes (e.g. Ferguson et al., 2005; Freedman et al., 2008). Note however that dust species, if present, can play a important contribution to the opacity since they have broadband absorption features, whilst bound-bound interactions only absorb photons as specific wavelengths. For a highly irradiated, low-mass and solar metallicity H/He atmosphere, Rogers and Seager (2010) fitted the opacity tables of Freedman et al. (2008) to determine that $\alpha = 0.68$ and $\beta = 0.45$. Note that the opacity tables of Freedman et al. (2008) consider a dust-free atmosphere.

In the deep atmosphere, dust species are in the gas-phase so that the opacity is dominated by bound-free and free-free absorptions of H^- ions which are produced due to thermal ionisation at the high temperatures and pressures. In this case, Lee et al. (2014) extrapolated the opacity tables of Ferguson et al. (2005) to find that $\alpha \approx 0.50$ and $\beta \approx 3.80$.[3]

[3]Note that, in order to determine values for α and β from Lee et al. (2014), I have used an ideal gas equation

Now that we have an analytic form to the opacity laws, let us use them to determine whether an atmosphere is convective or radiative. To do so we must consider the convection instability and radiative-convective equilibrium.

2.5 Radiative-Convective Equilibrium

To determine which parts of the atmosphere are convective or radiative, let us follow Kippenhahn and Weigert (1994) and consider a parcel of gas inside an atmosphere. We assume that this parcel can move and hence have different thermodynamic properties to its surroundings. The difference in some property X between the parcel and its surroundings is $DX = X_p - X_s$ where subscripts refer to the parcel and surroundings respectively. We can assume that the parcel remains in pressure balance with the surroundings, hence $DP = 0$. If the temperature of the parcel is greater than surroundings, we expect $DT > 0$. For an ideal gas, $T \propto 1/\rho$, hence this implies that $D\rho < 0$. As a result, the parcel is less dense that its surroundings and a buoyancy force will push the parcel upwards. In other words, it is convectively unstable.

For a parcel at a given radius r, we can test for convective stability by perturbing its location upwards by Δr. As such, the change in density is now:

$$D\rho = \left[\left(\frac{\partial \rho}{\partial r}\right)_p - \left(\frac{\partial \rho}{\partial r}\right)_s\right]\Delta r. \tag{2.5.1}$$

This expression gives the difference in density per volume of a parcel and its surroundings. The buoyancy force is thus:

$$F_{\text{buoy}} = -|g|D\rho. \tag{2.5.2}$$

We see that the buoyancy force will be upward if $D\rho < 0$, meaning that the parcel will continue upwards, resulting in an growing perturbation i.e. the convective instability. If $D\rho > 0$ on the other hand, the the parcel will sink in the atmosphere, resulting in a stable equilibrium. As such, we can write the criterion for convective stability as:

$$\left(\frac{\partial \rho}{\partial r}\right)_p - \left(\frac{\partial \rho}{\partial r}\right)_s > 0 \tag{2.5.3}$$

This criterion is illuminating, however not practical, so we aim to rewrite in terms of the temperature gradient ∇ introduced in Eq. 2.2.6.

The first law of thermodynamics is written as:

$$dq = du + Pdv, \tag{2.5.4}$$

where q, u and v refer to heat, internal energy and volume per unit mass. From this and an equation of state $\rho = \rho(P, T, \mu)$ and $u = u(\rho, T)$, we may define derivatives:

$$\begin{aligned}
\psi &\equiv \left(\frac{\partial \ln \rho}{\partial \ln P}\right)_{T,\mu} = -\frac{P}{v}\left(\frac{\partial v}{\partial P}\right)_{T,\mu}, \\
\delta &\equiv -\left(\frac{\partial \ln \rho}{\partial \ln T}\right)_{P,\mu} = \frac{T}{v}\left(\frac{\partial v}{\partial T}\right)_{P,\mu}, \\
\phi &\equiv -\left(\frac{\partial \ln \rho}{\partial \ln \mu}\right)_{P,T},
\end{aligned} \tag{2.5.5}$$

of state. Despite not being a perfect description of the atmosphere at these depths, it is still illuminating when considering the convective stability in these regions (see Section 2.5).

where μ is mean molecular weight and its derivatives refer to composition gradients. Note that for an ideal gas $\rho \propto P\mu/T$, $\psi = \delta = \phi = 1$. Using these definitions, we can rewrite the equation of state in differential form:

$$\frac{d\rho}{\rho} = \psi\frac{dP}{P} - \delta\frac{dT}{T} + \phi\frac{d\mu}{\mu}, \tag{2.5.6}$$

and hence rewrite the criterion for convective stability as:

$$\left(\frac{\psi}{P}\frac{dP}{dr}\right)_p - \left(\frac{\delta}{T}\frac{dT}{dr}\right)_p + \left(\frac{\phi}{\mu}\frac{d\mu}{dr}\right)_p - \left(\frac{\psi}{P}\frac{dP}{dr}\right)_s + \left(\frac{\delta}{T}\frac{dT}{dr}\right)_s - \left(\frac{\phi}{\mu}\frac{d\mu}{dr}\right)_s > 0. \tag{2.5.7}$$

Since we assume that the mean molecular weight will not change radially within the parcel of gas, the second term is ignored. Additionally, as we assume $DP = 0$, the first and third terms cancel. We now utilise a more general definition of the pressure scale height:

$$H = -\frac{dr}{d\ln P} = -P\frac{dr}{dP}. \tag{2.5.8}$$

Note that applying this to the isothermal solution to hydrostatic equilibrium yields the same result as in Eq. 2.1.1, showing that $H > 0$. As a result, it may be used to rewrite the condition for stability as:

$$\left(\frac{d\ln T}{d\ln P}\right)_s < \left(\frac{d\ln T}{d\ln P}\right)_p + \frac{\phi}{\delta}\left(\frac{d\ln \mu}{d\ln P}\right)_s. \tag{2.5.9}$$

Defining derivatives as:

$$\nabla \equiv \left(\frac{d\ln T}{d\ln P}\right)_s,\ \nabla_p \equiv \left(\frac{d\ln T}{d\ln P}\right)_p,\ \nabla_\mu \equiv \left(\frac{d\ln \mu}{d\ln P}\right)_s, \tag{2.5.10}$$

we can write the stability condition for convection as:

$$\nabla < \nabla_p + \frac{\phi}{\delta}\nabla_\mu. \tag{2.5.11}$$

If we assume that the parcel exchanges no thermal energy with its surroundings, we may state that $\nabla_p = \nabla_{ad}$ where ∇_{ad} is the temperature gradient for an adiabatic gas. We finally arrive at the *Ledoux criterion* for convective stability:

$$\nabla < \nabla_{\rm ad} + \frac{\phi}{\delta}\nabla_\mu. \tag{2.5.12}$$

For an ideal gas we have that $P^{1-\gamma}T^\gamma = \text{constant}$, where γ is the adiabatic index and hence $\nabla_{ad} = (\gamma - 1)/\gamma$. We see that if heavier elements exist at lower radii, then $\nabla_\mu > 0$ which tends to suppress convection since the right-hand side of Eq. 2.5.12 increases. However, under the assumption of a homogeneous composition profile, this condition becomes:

$$\nabla < \nabla_{\rm ad} \equiv \frac{\gamma - 1}{\gamma}, \tag{2.5.13}$$

which is referred to as the *Schwarzschild criterion* for convection stability. Our goal now is to determine ∇ for the atmosphere and test whether it is stable to the convection instability.

As shown in Figure 2.2, consider an inert planetary core of radius $R_{\rm c}$ surrounded by a H/He atmosphere of photospheric radius $R_{\rm ph}$. By definition, below $R_{\rm ph}$ the photon mean free path is much smaller than the atmospheric scale height, and hence behaves as an optically thick gas. The appropriate timescales to consider for energy transport are the cooling timescale, often referred to as the *Kelvin-Helmholtz timescale*:

$$t_{\rm KH} \sim \frac{GM_{\rm p}M_{\rm atm}/R_{\rm p}}{L}, \tag{2.5.14}$$

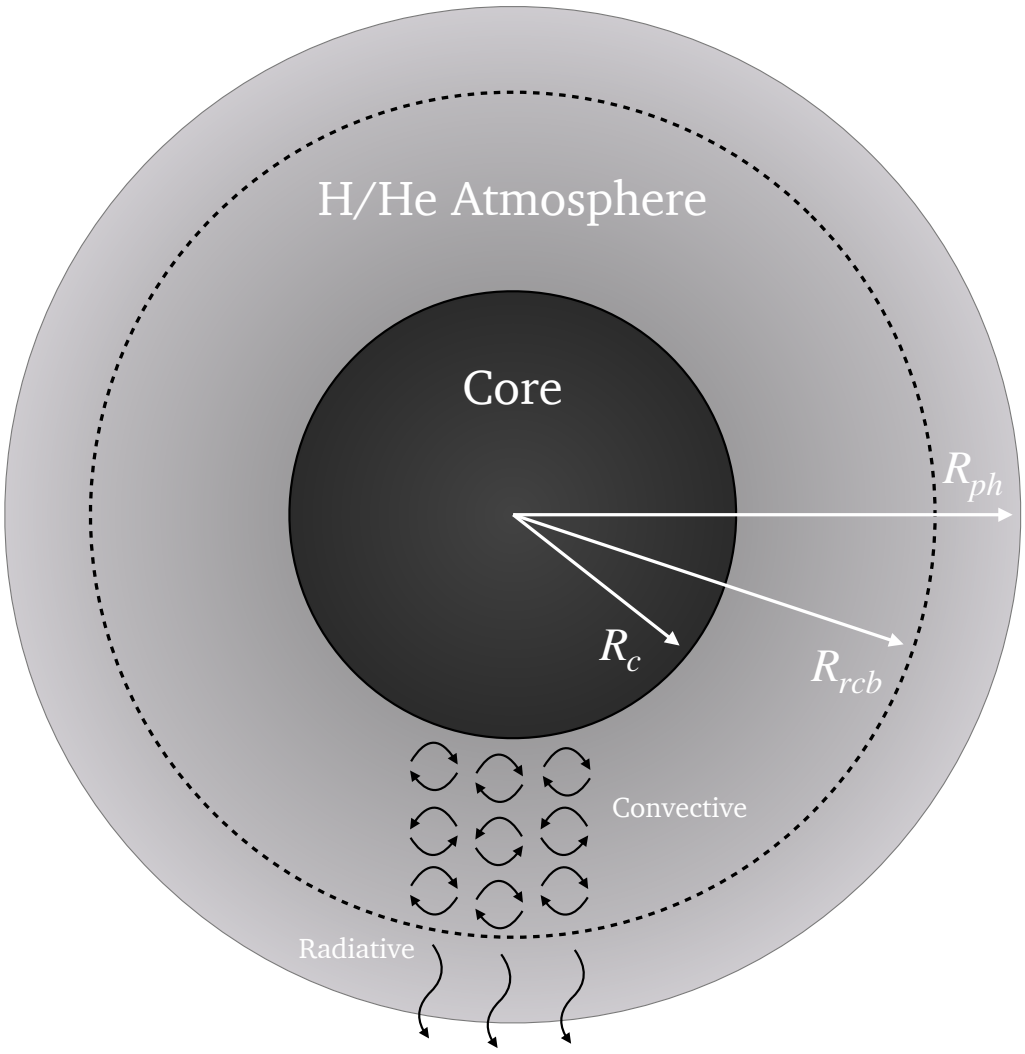


Figure 2.2: Schematic of a planet with inert core of radius R_c, a H/He dominated atmosphere with inner convective layer and outer radiative layer. The boundary between the these two layers, referred to as the radiative-convective boundary, has radius of R_{rcb}. The photospheric radius R_{ph} is the point at which the atmosphere becomes optically thick to outgoing radiation.

where L is the internal cooling luminosity of the planet which arises due to contraction. This timescale increases with with planet age, simply because planets cool and contract, resulting in a reduced radius and luminosity. Typically, for the planets of interest, it has a value $t_{\rm KH} \gtrsim 10^6$ yrs (see discussion in Section 2.6). Radiation transport on the other hand will act on timescales according to photon mean-free path $l_{\rm ph}$ from Section 2.4. Near the photosphere of a planet, typical densities are $\sim 10^{-6}$ g cm^{-3}. Typical opacities are of order $\kappa \sim 1$ cm^2 g^{-1}. The mean free path is hence $l_{\rm ph} = (\rho\kappa)^{-1} \sim 10^6$cm. Finally, the photon collision timescale is $t_{\rm ph} \sim 10^{-4}$ s. Hence we can safely state that:

$$t_{\rm KH} \gg t_{\rm ph}. \tag{2.5.15}$$

This means that the planet does not have time to thermally react to the incoming radiation and is instead in radiative equilibrium with the incoming stellar irradiation. As a result, the temperature of the photosphere is set by the equilibrium temperature caused by the host star (see Section 2.6 for derivation of equilibrium temperature). We can also see that the photon mean-free path is much shorter than the characteristic length scale of the atmosphere, which is $R_{\rm ph} - R_{\rm core} \sim 10^9$ cm. This implies that we can also assume that the energy transport occurs as a diffusion process. Following again from Kippenhahn and Weigert (1994) and by analogy of particle diffusion, the diffusive flux $\boldsymbol{J}$ is given by:

$$\boldsymbol{J} = -D\nabla n, \tag{2.5.16}$$

where n is the particle number density and D is the diffusion coefficient:

$$D = \frac{1}{3} v l_{\rm p}, \tag{2.5.17}$$

which is the standard result for particle diffusion and where v is mean velocity and $l_{\rm p}$ is the particle mean-free path. For photons, we replace v with c and $l_{\rm p}$ with $l_{\rm ph}$. The diffusive flux of radiative energy $\boldsymbol{F}$ is obtained by replacing the number density n with energy density U, defined as:

$$U = \frac{4\sigma}{c} T^4. \tag{2.5.18}$$

Furthermore, we invoke the spherical symmetry of the problem which reduces the equation to:

$$\frac{\partial U}{\partial r} = \frac{16\sigma}{c} T^3 \frac{\partial T}{\partial r}. \tag{2.5.19}$$

Using definitions for diffusion coefficient and photon mean-free path, we thus find the radiative flux is:

$$F = -\frac{16\sigma}{3} \frac{T^3}{\kappa\rho} \frac{\partial T}{\partial r}. \tag{2.5.20}$$

Rearranging and substituting the flux $F = L/(4\pi r^2)$:

$$\frac{\partial T}{\partial r} = -\frac{3\kappa\rho}{16\sigma T^3} \frac{L}{4\pi r^2}, \tag{2.5.21}$$

where the luminosity L arises due to the thermal contraction of the atmosphere. This equation is typically referred to as the *diffusion approximation.* Note that we can assume hydrostatic equilibrium from Eq. 2.2.4 and take the derivative of Eq. 2.5.21 to find the radiative temperature gradient:

$$\nabla_{\rm rad} = \frac{3}{4\pi\sigma G} \frac{\kappa L P}{m T^4}. \tag{2.5.22}$$

This approximation is only valid in the optically thick regime however, meaning that beyond the photosphere, a different approach must be taken. As we shall see in Chapter 5, this is particularly important when considering outflows that extend beyond the photosphere. For now

however, we restrict ourselves to the conditions of the optically thick atmosphere. Now, following from Rafikov (2006) and combining the diffusion approximation with that of hydrostatic equilibrium in Eq. 2.2.4, an opacity scaling relation from Eq. 2.4.2 and integrating from (P, T) to (P_0, T_0), we find that the P-T profile for a radiative atmosphere is as follows:

$$\left(\frac{P}{P_0}\right)^{\alpha+1} - 1 = \frac{8\nabla_0}{3}\frac{R_\mathrm{B}\lambda_0}{R_\mathrm{L}^2}\left[\left(\frac{T}{T_0}\right)^{4-\beta} - 1\right], \tag{2.5.23}$$

where $\nabla_0 \equiv (1+\alpha)/(4-\beta)$, $\lambda_0 = (\kappa_0\rho_0)^{-1}$, $R_\mathrm{B} = GM_\mathrm{c}/2c_\mathrm{s}^2$ is the Bondi radius and $R_\mathrm{L} = (L/16\pi\sigma T_0^4)^{1/2}$ is the luminosity radius, which is the radius of a radiating body with luminosity L and effective temperature T_0. In this optically thick case, we have that $\lambda_0 \ll R_\mathrm{B}$ and $R_\mathrm{L}^2 \ll \lambda_0 R_\mathrm{B}$.

We may now calculate the temperature gradient ∇ from Eq. 2.5.23 as follows:

$$\nabla(P,T) = \frac{3}{8}\frac{R_\mathrm{L}^2}{R_\mathrm{B}\lambda_0}\left(\frac{T}{T_0}\right)^{\beta-4}\left(\frac{P}{P_0}\right)^{\alpha}. \tag{2.5.24}$$

Then, substituting Eq. 2.5.23 back in, we obtain ∇ as a function of T:

$$\begin{aligned}\nabla(T) &= \nabla_0\left\{1 - \left(\frac{T_0}{T}\right)^{4-\beta}\left[1 - \frac{3}{8\nabla_0}\frac{R_\mathrm{L}^2}{R_\mathrm{B}\lambda_0}\right]\right\}. \\ &\approx \nabla_0\left[1 - \left(\frac{T_0}{T}\right)^{4-\beta}\right],\end{aligned} \tag{2.5.25}$$

where we have utilised the assumption that $R_\mathrm{L}^2 \ll \lambda_0 R_\mathrm{B}$ in an optically thick atmosphere. Hence, as we travel deeper into the atmosphere, T will increase above T_0 meaning that ∇ will become closer and closer to ∇_0. The hydrogen dominated atmosphere of concern will mostly consist of dissociated hydrogen atoms meaning that $\gamma = 5/3$ and hence $\nabla_\mathrm{ad} = 0.40$. As discussed in Section 2.4, the deep atmosphere's opacity is dominated by H^- ions, resulting in $\alpha \approx 0.5$ and $\beta \approx 3.8$ (Lee et al., 2014). As a result, $\nabla \approx 7.50$, implying that $\nabla > \nabla_\mathrm{ad}$ and is hence convectively unstable. In the upper atmosphere however, where $T \approx T_0$, we see from Eq. 2.5.25 that the temperature gradient will be very small and hence $\nabla < \nabla_\mathrm{ad}$, implying the upper atmosphere is convectively stable. In summary, the deep atmosphere is convective whilst the upper layer is not. The boundary between the two regions is referred to as the *radiative-convective boundary* (RCB). Note that here we have assumed an ideal equation of state and simple opacity scaling relations, which has revealed that that the inner atmosphere is convective, whilst the outer layer is radiative. Full treatment of this problem, involving numerically solving the stellar structure equations with tabulated equations of state and opacity tables, also confirm this result. An example of this from Piso et al. (2015) is shown in Figure 2.3, however this has been shown in multiple other works (e.g. Chabrier and Baraffe, 1997; Baraffe et al., 2003; Fortney et al., 2007; Rogers et al., 2011; Lee et al., 2014; Piso and Youdin, 2014; Lee and Chiang, 2015a).

The next question is how does temperature vary in these two layers? Inspection of Eq. 2.5.23 reveals the nature of the upper, convectively stable atmosphere i.e. above the RCB. One can see that the coefficient of the right-hand side is very large in the optically thick case, and hence a large change in pressure results in a small change in temperature. For the upper atmosphere, where $T - T_0 \lesssim T_0$, this suggests that despite a large change in pressure, the outer atmosphere is approximately isothermal (Mizuno et al., 1978), with the temperature set by the incoming stellar irradiation.

The temperature structure beneath the RCB requires a little more thought. Since the atmosphere is convective, we have that parcels of fluid are rising due to perturbations in density,

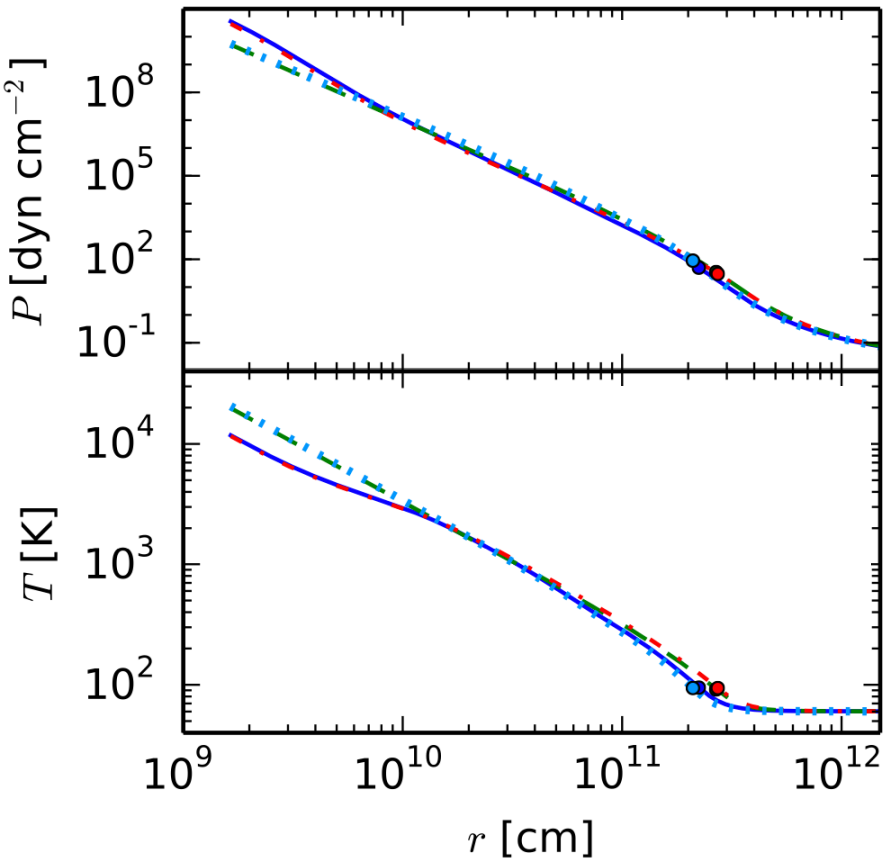


Figure 2.3: Pressure and temperature profiles for a giant planet with core of mass $10M_{\oplus}$ at 5AU for different equations of state, taken from Piso et al. (2015). Green dashed is an ideal gas equation of state, whilst solid blue is a realistic equation of state taken from Saumon et al. (1995). Dashed light-blue and dot-dashed red represent hybrid equations of state where an ideal gas was used below and above 500K respectively. Coloured circles represent the location of the radiative-convective boundary in these four cases. All models demonstrate an internal convective zone which approximately follows an adiabat. An outer radiative zone exhibits an approximately isothermal temperature profile with a rapid drop in pressure.

exchanging energy with surroundings, before finally becoming indistinguishable from the surroundings. Typically in stellar and planetary evolution numerical schemes, *mixing length theory* is implemented to model convection, which is a local theory that treats small parcels of gas as they rise over a characteristic 'mixing length' due to fluctuations in density. The mixing length encodes our ignorance of the details of convection (e.g. Böhm-Vitense, 1958; Henyey et al., 1965; Cox and Giuli, 1968).

In general, the total energy flux is the sum of radiative flux and convective flux. Following again from Kippenhahn and Weigert (1994), their sum is the gradient required to transport the whole flux via radiation:

$$F_{\rm rad} + F_{\rm con} = \frac{16\sigma G}{3}\frac{T^4 m}{\kappa P r^2}\nabla_{\rm rad}, \tag{2.5.26}$$

where m refers to the mass enclosed at radius r. However, in this case, some of the flux is transported via convection. As a result, the temperature gradient ∇ is altered and at this point, remains unknown. In this case, the radiative flux is:

$$F_{\rm rad} = \frac{16\sigma G}{3}\frac{T^4 m}{\kappa P r^2}\nabla. \tag{2.5.27}$$

To evaluate the convective flux, we must consider again a parcel of fluid with $DP = 0$ and $DT > 0$ with respect to its surroundings. It will ascend at a velocity of v, providing a flux of:

$$F_{\rm con} = \rho v c_{\rm P} DT, \tag{2.5.28}$$

where $c_{\rm P}$ is the heat capacity of the gas at fixed pressure. In mixing length theory, we state that the parcel travels a distance of $l_{\rm m}$ before mixing completely with its surroundings. However, different parcels passing through a radius r will have different values of v and DT, hence we average the mixing length to be $l_{\rm m}/2$. As a result:

$$\begin{aligned}\frac{DT}{T} &= \frac{1}{T}\frac{\partial(DT)}{\partial r}\frac{l_{\rm m}}{2} \\ &= (\nabla - \nabla_{\rm p})\frac{l_{\rm m}}{2}\frac{1}{H},\end{aligned} \tag{2.5.29}$$

where H is the pressure scale height and $\nabla_{\rm p}$ is the temperature gradient for the parcel. Similarly, from Eqs. 2.5.6 and 2.5.1, the density difference is:

$$\frac{D\rho}{\rho} = -\delta\frac{DT}{T}. \tag{2.5.30}$$

Then, for a radial buoyancy force per unit mass of $k_{\rm buoy} = -gD\rho/\rho$, the work done to move the parcel upwards by $l_{\rm m}/2$ is:

$$\frac{1}{2}k_{\rm r}\frac{l_{\rm m}}{2} = g\delta(\nabla - \nabla_{\rm p})\frac{l_{\rm m}^2}{8H}. \tag{2.5.31}$$

Assuming that half of this work goes into kinetic energy i.e. $v^2/2$ and the other is is used to displace the surroundings, then the average velocity of parcels passing through a layer of radius r is:

$$v^2 = g\delta(\nabla - \nabla_{\rm p})\frac{l_{\rm m}^2}{8H}. \tag{2.5.32}$$

Now inserting this back into the equation for convective flux yields:

$$F_{\rm con} = \rho c_{\rm p} T\sqrt{g\delta}\frac{l_{\rm m}^2}{4\sqrt{2}}H^{-3/2}(\nabla - \nabla_{\rm p})^{3/2}. \tag{2.5.33}$$

Since $DT > 0$ for the parcel, it will radiate away energy with a local flux f. According to the radiative flux of Eq. 2.5.20, we have that:

$$f = \frac{16\sigma T^3}{3\kappa\rho}\left|\frac{\partial T}{\partial r}\right|, \tag{2.5.34}$$

where $\partial/\partial r$ refers to change in temperature from parcel to surroundings. If we assume the parcel is spherical with diameter d, then the temperature gradient is approximately $\partial/\partial r \approx 2DT/d$. As such, the radiative loss λ per unit time from the surface area of the parcel S is:

$$\lambda = Sf = \frac{32\sigma T^3}{3\kappa\rho} DT \frac{S}{d}. \tag{2.5.35}$$

We may finally use this relation to consider the change in temperature of the parcel as it rises. Contributions for this will come from adiabatic expansion and radiative exchange with surroundings. From the above, the change in temperature per unit length as the parcel rises is $\lambda/\rho V c_{\rm P} v$, where V is the parcel volume. Hence:

$$\left(\frac{dT}{dr}\right)_{\rm p} = \left(\frac{dT}{dr}\right)_{\rm ad} - \frac{\lambda}{\rho V c_{\rm p} v}, \tag{2.5.36}$$

Then, multiplying by H/T, we get:

$$\nabla_{\rm p} - \nabla_{\rm ad} = \frac{\lambda H}{\rho V c_{\rm p} v T}. \tag{2.5.37}$$

Combining the above with the definition of λ from Eq. 2.5.35 and the average DT from Eq. 2.5.29, we find that:

$$\frac{\nabla_{\rm p} - \nabla_{\rm ad}}{\nabla - \nabla_{\rm p}} = \frac{24\sigma T^3}{\kappa\rho^2 c_{\rm P} l_{\rm m} v}, \tag{2.5.38}$$

where we have used the literature value for the factor $l_{\rm m} S/Vd \approx 9/2l_{\rm m}$ (e.g. see Böhm-Vitense, 1958; Cox and Giuli, 1968). We now have have five equations: 2.5.26, 2.5.27, 2.5.32, 2.5.33 and 2.5.38 that can be solved to determine the quantities $F_{\rm rad}$, $F_{\rm con}$, v, $\nabla_{\rm p}$ and ∇. Once ∇ is determined, it can be used in the stellar structure equations, particularly in:

$$\frac{\partial T}{\partial r} = \frac{T}{P}\frac{\partial P}{\partial r}\nabla, \tag{2.5.39}$$

to determine the atmospheric temperature structure in the convective regions. The term $(\nabla - \nabla_{\rm ad})$ is referred to as the *super-adiabaticity* and refers to how much the temperature gradient increases above the that of an adiabat. As an order of magnitude estimate as to how large the adiabaticity is, one can assume that $\nabla_{\rm p} \approx \nabla_{\rm ad}$, replace $F_{\rm con} = L/4\pi r^2$ in Eq. 2.5.32. and rearrange for the adiabaticity. This is then the adiabaticity required for convection to transport the entire radiative flux of the planet. In the deep atmospheres of typical super-Earths we have that $L \sim 10^{25}$ erg s^{-1} (Lopez and Fortney, 2014) and inserting this and other appropriate values shows that $(\nabla - \nabla_{\rm ad})/\nabla_{\rm ad} \approx 10^{-6}$, meaning that convection is very efficient at transporting radiative energy i.e. only a very small adiabaticity is required. In the deep convective atmosphere, it is then safe to assume that $\nabla \approx \nabla_{\rm ad}$. In the top layers of the convective zone i.e. approaching the RCB, the play-off between convection and radiative transport becomes more important. Firstly, the local gravitational acceleration becomes weaker further out of the potential-well, which increases the local scale height. As a result, the diffusion process becomes more efficient (recall that diffusion works due to the fact that the scale-height is much larger than the photon mean-free path). In addition, opacity decreases as we move to larger radii since temperatures and pressures are smaller. Hence, as the opacity decreases as we approach the RCB, it becomes easier to transport luminosity via radiative diffusion. As a result, the atmosphere becomes more stable to convection until eventually we pass through the RCB and the atmosphere is completely radiative. Mixing length theory provides the corrections to the temperature gradient near the RCB that address these effects.

2.6 An analytic model for atmospheric structure

Thus far, we have taken a deep dive into the structure of an optically thick atmosphere in hydrostatic equilibrium. I have presented an analytic formalism to determine which regions are convectively unstable and additionally a way to determine the resulting temperature structure. However, I have intentionally over-simplified the analysis with approximations such as an ideal gas equation of state. As discussed, one should solve the stellar structure equations with suitable equations of state for for all species present in a hydrogen dominated atmosphere (e.g. Saumon et al., 1995; Ferguson et al., 2005) (see Figure 2.3). To go further than this however, one can solve the full hydrodynamic problem (i.e. not assume hydrostatic equilibrium) with correct equations of state and include additional physics such as mixing length theory (e.g. Chabrier and Baraffe, 1997; Owen and Wu, 2013, 2016). This is indeed performed in Chapter 5 using *Modules for Experiments in Stellar Astrophysics*, `MESA` (Paxton et al., 2011, 2013, 2015, 2018). However, for a large portion of this thesis, one of the key tools is a simple analytic model for atmospheric structure which can be coupled to mass-loss and is computationally cheap to evaluate. This comes from Owen and Wu (2017) and is presented in the following.

We begin by making assumptions of the atmospheric structure in light of the previous discussion. The atmosphere is split into two layers: the upper layer is in radiative equilibrium with the incoming stellar irradiation and assumed to be isothermal with the photospheric equilibrium temperature set by the stellar insolation. To find this temperature, we can consider a planet-hosting star of mass M_* with luminosity L_*. The flux received at an orbital separation of a is $F_* = L_*/4\pi a^2$. Then, assuming the planet re-emits all the radiation is absorbs i.e. it is a perfect black-body, we may use the Stefan-Boltzmann law to show that:

$$\frac{F_*}{4} = \sigma T_{\rm eq}^4, \tag{2.6.1}$$

where $T_{\rm eq}$ is the planet's equilibrium temperature and the factor of $1/4$ comes from the fact that the planet is exposed to $1/2$ of the stellar surface and is irradiated on $1/2$ of the its own surface i.e. the day-side. As such, we set the equilibrium temperature as:

$$T_{\rm eq} = \left(\frac{L_*}{16\sigma\pi a^2}\right)^{1/4}, \tag{2.6.2}$$

where we find the stellar luminosity from a mass-luminosity relation:

$$L_* = L_\odot \left(\frac{M_*}{M_\odot}\right)^{\zeta}, \tag{2.6.3}$$

where $\zeta > 0$ and $\zeta \approx 3.2$ for main-sequence stars (e.g. Cox, 2000). Note that the influence of the value of ζ is discussed in Chapter 3. Below this radiative layer, we assume a fully convective layer which is modelled as adiabatic. In this sense, we ignore the effects of super-adiabaticity near the RCB at $R_{\rm rcb}$ and do not invoke mixing length theory, which provides a considerable simplification. We are able to do this since the specific luminosity of the planet is very small with typical effective temperatures of ~ 50K (e.g. see Figure 5 of Lopez and Fortney, 2014). From the diffusion approximation, we see that the radiative flux is proportional to the temperature gradient, which will be very small as a result. The small corrections of mixing length theory will have a very small effect of the flux transported through the RCB. Hence, solving Eq. 2.5.23 with an adiabatic equation of state:

$$P = K\rho^\gamma, \tag{2.6.4}$$

where K is a positive constant, one arrives at the density profile for the convective layer of the atmosphere:

$$\rho = \rho_{\rm rcb}\left[1 + \nabla_{\rm ad}\left(\frac{GM_{\rm c}}{c_{\rm s}^2 R_{\rm rcb}}\right)\left(\frac{R_{\rm rcb}}{r} - 1\right)\right]^{\frac{1}{\gamma-1}}, \tag{2.6.5}$$

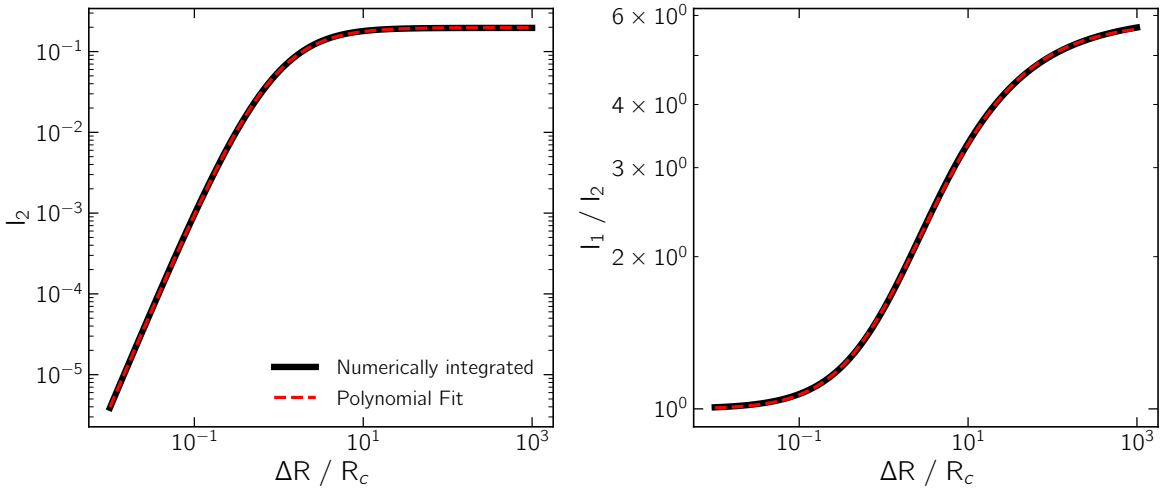


Figure 2.4: The integrals I_2 and I_1/I_2 required for the analytic model of Owen and Wu (2017) are shown in the left and right-hand panel respectively. These come from Eqs. 2.6.8 and 2.6.15. These integrals are evaluated numerically using scipy.integrate.quad in black and fit with an exponential analytic form for computational speed in red dashed.

where the isothermal sound speed for an ideal gas $c_s^2 \equiv \partial P/\partial\rho|_T = P/\rho$ is evaluated at the RCB. Here, the boundary condition is such that $\rho = \rho_{\rm rcb}$ at $r = R_{\rm rcb}$. For atmospheres that are gravitationally bound, we expect the sound speed to be much less than the escape velocity (i.e. $v_{\rm esc}^2 \sim GM_c/R_p \gg c_s^2$). As a result, the unity term in the previous equation can be ignored when compared to the second term. Hence, we may approximate the density profile as:

$$\rho \simeq \rho_{\rm rcb}\left[\nabla_{\rm ad}\left(\frac{GM_c}{c_s^2 R_p}\right)\left(\frac{R_{\rm rcb}}{r}-1\right)\right]^{\frac{1}{\gamma-1}}. \tag{2.6.6}$$

This expression may now be used to evaluate the atmospheric mass fraction $X \equiv M_{\rm atm}/M_c$, which is the ratio of atmospheric mass $M_{\rm atm}$ to core mass M_c. Since the density profile for an isothermal atmosphere falls exponentially, one can assume that the mass of the atmosphere is dominated by the convective region. As such, we can calculate the atmospheric mass and hence atmospheric mass fraction by integrating the above expression from the core to the RCB:

$$\begin{aligned} M_{\rm atm} &= \int_{R_c}^{R_{\rm rcb}} 4\pi r^2 \rho \, {\rm d}r. \\ &\simeq 4\pi R_{\rm rcb}^3 \rho_{\rm rcb}\left(\nabla_{\rm ad}\frac{GM_c}{c_s^2 R_{\rm rcb}}\right)^{1/(\gamma-1)} I_2(R_c/R_{\rm rcb},\gamma), \end{aligned} \tag{2.6.7}$$

where the dimensionless integral I_2 is as follows:

$$I_2(R_c/R_{\rm rcb},\gamma) = \int_{R_c/R_p}^{1} x^2\left(\frac{1}{x}-1\right)^{\frac{1}{\gamma-1}} dr, \tag{2.6.8}$$

which is shown in Figure 2.4. This integral was fit with a exponential analytic formula, since this is computationally cheaper to evaluate than a numerical integral. The final expression for atmospheric mass fraction will depend on $\rho_{\rm rcb}$ which still needs to be calculated. This is done by matching the temperature gradient from both sides of the RCB. In the adiabatic layer this is given by $\nabla_{\rm ad}$. From the stellar structure equations i.e. Eq. 2.2.4, we thus have that:

$$\frac{d\ln T}{dr} = \nabla_{\rm ad}\frac{d\ln P}{dr} = -\nabla_{\rm ad}\frac{GM_c}{R_{\rm rcb}^2 c_s^2}. \tag{2.6.9}$$

From the radiative layer however, the gradient is given by radiative diffusion. From Eq. 2.5.21, it follows that:

$$\frac{d\ln T}{dr} = -\frac{L}{4\pi R_{\rm rcb}^2}\frac{3\kappa\rho_{\rm rcb}}{16\sigma T_{\rm eq}^4}, \tag{2.6.10}$$

where it is assumed that the temperature at the RCB are the same as the surface i.e. $T_{\rm rcb} = T_{\rm eq}$. Matching these two gradients allows us to calculate $\rho_{\rm rcb}$ and hence $M_{\rm atm}$ from Eq. 2.6.7. The final variable required to close these equations is the cooling luminosity L, which controls how quickly the atmosphere contracts. In essence, this contraction is a change in the atmospheric binding energy,

$$L = \frac{dU}{dt} \approx \frac{U}{t_{\rm KH}} \approx \frac{1}{t_{\rm KH}}\left(\int_{R_{\rm c}}^{R_{\rm p}} \frac{GM_{\rm core}\rho}{r}\, 4\pi r^2\, dr\right), \tag{2.6.11}$$

where $t_{\rm KH}$ is the Kelvin-Helmholtz cooling timescale, defined as

$$t_{\rm KH} \approx \frac{|E|}{L} = \frac{GM_{\rm c}M_{\rm atm}}{R_{\rm p}L}. \tag{2.6.12}$$

The value of $t_{\rm KH}$ is typically of order the age of the planet. The heuristic argument for this is because if the cooling timescale of a planet was significantly shorter than its own age, it would rapidly cool, resulting in an increased cooling timescale. This argument naturally implies that the cooling timescale must be approximately the planet's age. However, this is not the case after rapid mass-loss processes such as during a 'boil-off' phase, which increases the cooling timescale of a young planet beyond its age (Owen and Wu, 2016; Ginzburg et al., 2016). This concept is explored in detail in Chapter 5. For now however, we follow that of Owen and Wu (2017) and assume that the cooling timescale behaves as follows:

$$t_{\rm KH} = \begin{cases} 10^8 \text{ yrs} & \text{for } t < 10^8 \\ t & \text{for } t \geq 10^8. \end{cases} \tag{2.6.13}$$

With the use of Eq. 2.6.6 and 2.6.7, the internal luminosity can hence be approximated as:

$$L \approx \frac{1}{t_{\rm KH}}\frac{GM_{\rm c}M_{\rm atm}}{R_{\rm p}}\frac{I_1(R_{\rm c}/R_{\rm p})}{I_2(R_{\rm c}/R_{\rm p})}, \tag{2.6.14}$$

where an additional dimensionless integral I_1, shown in Figure 2.4, is given as:

$$I_1(R_{\rm c}/R_{\rm p}, \gamma) = \int_{R_{\rm c}/R_{\rm p}}^{1} x\left(\frac{1}{x} - 1\right)^{\frac{1}{\gamma-1}} dr. \tag{2.6.15}$$

To summarise, this prescription can provide a planet's atmospheric mass fraction X as a function of $(M_{\rm core}, T_{\rm eq}, R_{\rm p}, t)$. In practise however, this model is used for the inverse problem i.e. finding the radius of a planet for a given $(M_{\rm core}, T_{\rm eq}, X, t)$. Since there is a small radiative (and hence isothermal) outer layer to the atmosphere, we add an isothermal density profile to the top of the convective zone, such as that in Eq. 2.2.10 with $R_{\rm out} = R_{\rm rcb}$ and $\rho_0 = \rho_{\rm rcb}$. An important question to answer is: what defines the planet radius? Thus far, I have referred to the photospheric radius $R_{\rm ph}$ as the point where the atmosphere becomes optically thick. But how do we actually locate the photosphere?

Following Carroll and Ostlie (2006) and assuming that the opacity and temperature of the atmosphere are constant in the upper layers, as well as taking the *plane-parallel* approximation i.e. considering radiation to travel perpendicular to the atmosphere (valid since the planet's radius of curvature is far larger than the thickness of the thin radiative layer), one can arrive at the following expression for temperature variation very near to the top of the atmosphere:

$$T^4 = \frac{3}{4}T_{\rm eq}^4(\tau + \frac{2}{3}). \tag{2.6.16}$$

Note that in this calculation, one must assume that opacity is independent of wavelength (referred to as a *gray atmosphere*) and also assume that radiation field is split into independent inward and outward travelling fields (referred to as the *Eddington approximation*). One can see that if $\tau = 2/3$, then $T = T_{\rm eq}$. In other words, the surface of the planet that emits black-body radiation at $T_{\rm eq}$ is at an optical depth of 2/3, which we define to be the photosphere[4]. Hence, we have that:

$$\tau_{\rm ph} = \int_{R_{\rm ph}}^{\infty} \kappa\rho \, dr = \bar{\kappa} \int_{R_{\rm ph}}^{\infty} \rho \, dr = \frac{2}{3}, \tag{2.6.17}$$

where $\bar{\kappa}$ is the averaged opacity for the atmosphere above photosphere. We may integrate the equation of hydrostatic equilibrium from Eq. 2.2.4 to show that the pressure at the photosphere is:

$$P_{\rm ph} = \int_{R_{\rm ph}}^{\infty} g\rho \, dr = g_0 \int_{R_{\rm ph}}^{\infty} \rho \, dr, \tag{2.6.18}$$

where g is the gravitational acceleration and is assumed to be constant i.e. $g \approx g_0 \equiv GM_{\rm c}/R_{\rm rcb}^2$ since the majority of the mass is below the RCB and hence g will not vary considerably above this altitude. We can replace the integral in Eq. 2.6.18 with that of Eq. 2.6.17 to show that:

$$P_{\rm ph} = \frac{2}{3}\frac{g_0}{\bar{\kappa}}. \tag{2.6.19}$$

This specific pressure is located using the isothermal radiative atmosphere structure to finally locate the photospheric radius.

2.7 An analytic model for atmospheric mass-loss

Thus far, this model addresses contraction of a planet purely due to thermal cooling. However, the effect of XUV photoevaporation is the removal of mass via a hydrodynamic outflow, which changes the radius evolution of the planet. As shown in Section 2.3, a transonic outflow describes the physics of this problem. Furthermore, as discussed in Chapter 5, a non-zero flow velocity at a radius r requires a reduced density profile for a constant mass-loss i.e. to satisfy conservation of mass. However, in this simple analytic model, we do not need to address this, since below the photosphere, the Mach number would be very small and hence the advection term in Eq. 2.2.2 is very small. It is only once you approach the Bondi radius, that these effects must be addressed. Hence, in the simple analytic model of Owen and Wu (2017), we may consider 'snapshots' of the atmospheric structure that are in hydrostatic equilibrium. To address the change in mass in between these snapshots, the mass-loss timescale is defined as

$$t_{\dot{X}} \equiv \frac{X}{\dot{X}} = \frac{M_{\rm atm}}{\dot{M}_{\rm atm}}, \tag{2.7.1}$$

where the mass-loss rate $\dot{M}_{\rm atm}$ is calculated by assuming that XUV flux goes into doing PdV work at some efficiency η:

$$\dot{M}_{\rm atm} = \eta \frac{L_{\rm XUV}}{4\pi a^2} \frac{\pi R_{\rm p}^2}{GM_{\rm p}/R_{\rm p}} = \eta \frac{L_{\rm XUV}}{4a^2} \frac{R_{\rm p}^3}{GM_{\rm p}}, \tag{2.7.2}$$

where a is the planet semi-major axis and $L_{\rm XUV}$ is the luminosity in X-ray ($h\nu \approx 0.1 - 10$keV) and EUV ($h\nu \gtrsim 13.6$eV) wavelengths (e.g. Erkaev et al., 2007; Lopez and Fortney, 2013). These high-energy photons are produced in the stellar corona and all have sufficient energy to

[4]Note that this derivation, as shown in Carroll and Ostlie (2006), assumes that the planet is not irradiated, which is not the case for the planets of interest. Nevertheless, the actual value of optical depth at the photosphere is $\sim 2/3$ and any correction would be negligible when compared to total planet size.

photo-ionise hydrogen atoms, resulting in gas temperature of a few hundred K to $\sim 10^4$ K (e.g. Glassgold et al., 2004; Owen and Jackson, 2012). This in turn induces an outflow as summarised in Section 2.3. This 'energy-limited' mass-loss formalism is discussed in Section 2.9.

As discussed in Chapter 1, the stellar XUV luminosity $L_{\rm XUV}$ is measured and quantified via its ratio with bolometric luminosity $L_{\rm bol}$. As shown in multiple observational works (e.g. Ribas et al., 2005; Jackson et al., 2012; Wright et al., 2011; Tu et al., 2015; McDonald et al., 2019), this ratio demonstrates an evolution as follows:

$$\frac{L_{\rm XUV}}{L_{\rm bol}} = \begin{cases} \left(\frac{L_{\rm XUV}}{L_{\rm bol}}\right)_{\rm sat} & \text{for } t < t_{\rm sat} \\ \left(\frac{L_{\rm XUV}}{L_{\rm bol}}\right)_{\rm sat} \left(\frac{t}{t_{\rm sat}}\right)^{-1-a_0} & \text{for } t \geq t_{\rm sat}, \end{cases} \tag{2.7.3}$$

where $a_0 = 0.5$ is a positive constant and $t_{\rm sat}$ represents the time at which the star begins to spin down and X-ray generating mechanisms in the corona begin to diminish. The time-integrated luminosity is thus dominated by contributions before $t_{\rm sat}$, which is typically $\sim$ 100Myr, implying that any mass loss experienced by a planet occurs in this same time frame. The dependence of $t_{\rm sat}$ and $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$ on stellar mass is in fact imprinted on the exoplanet population under this model and is explored in Chapter 3.

Finally, the ordinary differential equation to find $X(t)$ is thus

$$\frac{dX}{dt} = -\frac{X}{t_{\dot{X}}}. \tag{2.7.4}$$

By solving this ODE numerically (as described in Chapter 3), one can determine the mass and radii evolution of a planet undergoing mass-loss via XUV photoevaporation. Example evolution tracks are shown in Figure 2.5 for a $5M_\oplus$ core planet around a solar mass star at a 10 day orbital period. From this, we can start to see why the radius gap naturally emerges as result of mass-loss, as is explained in the following discussion.

2.8 Carving the Radius Gap

Figure 2.5 demonstrates the atmospheric evaporation of a $5M_\oplus$ core orbiting at 0.1AU around a solar-mass star for varying initial atmospheric mass fraction. From this Figure, it is evident that planets either have their atmosphere completely stripped and thus evolve into super-Earths, or they are shepherded to $X \sim 0.01$ and form sub-Neptunes. Owen and Wu (2017) showed that this mass fraction corresponds to an atmosphere that approximately doubles the radius of the planet with respect the core. The first takeaway here is that the relation between photospheric radius and atmospheric mass is highly non-linear; a small amount of mass can highly inflate the planet, as shown in Figure 2.6. Crucially, it is no coincidence that a mass fraction that doubles the core radius results in a maximum mass-loss timescale for the planet. To understand this, let us find the dependencies of mass-loss timescale from Eqs. 2.7.1 and 2.7.2:

$$t_{\dot{X}} = \frac{4GM_{\rm c}^2 aX}{\eta L_{\rm XUV} R_{\rm p}^3}, \tag{2.8.1}$$

where we have assumed that $M_{\rm p} \approx M_{\rm c}$. Now consider a planet with a large inflated atmosphere that dominates its radius; due to its increased cross-sectional area, it will receive more flux from the host star and thus undergo rapid mass-loss with a short mass-loss timescale. This can be seen with the strong dependency of mass-loss timescale with planet size $t_{\dot{X}} \propto R_{\rm p}^{-3}$. As the size of the planet consequently decreases, the mass-loss timescale will increase until the point where the atmosphere no longer dominates the size of the planet i.e. a mass fraction such that

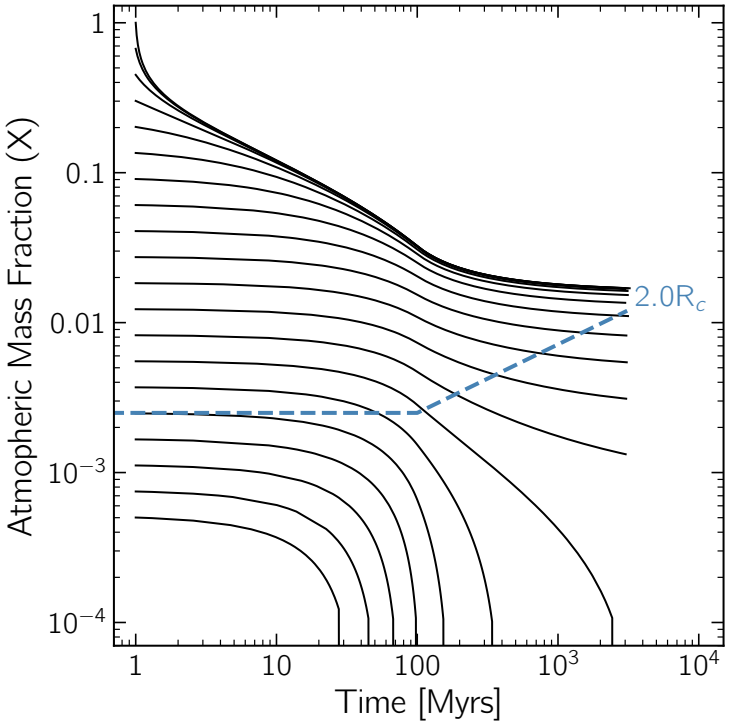


Figure 2.5: Solid lines show the evolution of varying initial atmospheric mass fractions for a $5M_\oplus$ core, orbiting a solar mass star at 10 day orbital period. Calculations are performed using the analytic model of Owen and Wu (2017). The dashed line tracks an atmospheric mass fraction required to double the core radius.

$R_\mathrm{p} \sim 2R_\mathrm{c}$. If the mass-loss timescale at this point is longer than t_sat (the time at which XUV luminosity rapidly shuts off) then the planet will retain its atmosphere. If this is not the case however, mass-loss timescale will continue to decrease, simply since there is less material to remove. Again, this is apparent since $t_{\dot{X}} \propto X$ and so a smaller atmospheric mass results in a shorter mass-loss timescale. As a result, planets are most stable to photoevaporation with an atmospheric mass that doubles its size. This timescale argument is demonstrated in Figure 2.6.

From this argument, we see that the radius gap is a natural outcome of atmospheric mass-loss. Sub-Neptunes form if a initially large planet, undergoes mass-loss to a point of stability in which its radius is doubled by its atmosphere, typically $X \sim 0.01$. If the mass-loss timescale at this point is greater than t_sat, then the planet will maintain this atmospheric mass. Super-Earths on the other hand will form if this timescale is shorter than t_sat, which is the case for shorter orbital periods or smaller core masses. This results in continuing mass-loss until the planet loses its primordial atmosphere. These two end results in planet radii correspond to the peaks on either side of the radius gap. Under this model we expect a clean gap i.e. an area of parameter space devoid of planets around evolved systems since after $\sim$ 100 Myr, the vast majority of mass-loss has already occurred. Thus the only way to observe a planet in the gap around an evolved star would be if its orbital and planetary parameters (e.g. semi-major axis and core mass) were finely tuned to mean mass-loss is still ongoing. This can be seen in Figure 2.5 with very few planets evolving to an atmospheric mass fraction $10^{-4} < X \lesssim 0.01$, corresponding to planets in the gap.

The radius gap is, of course, a demographic feature and depends on the distributions that underlay the planet population, such as that of the core masses or initial atmospheric masses. In Chapter 4, this analytic model is fit to the data in a rigorous statistical framework, resulting in constraints with regards to planet formation and evolution.

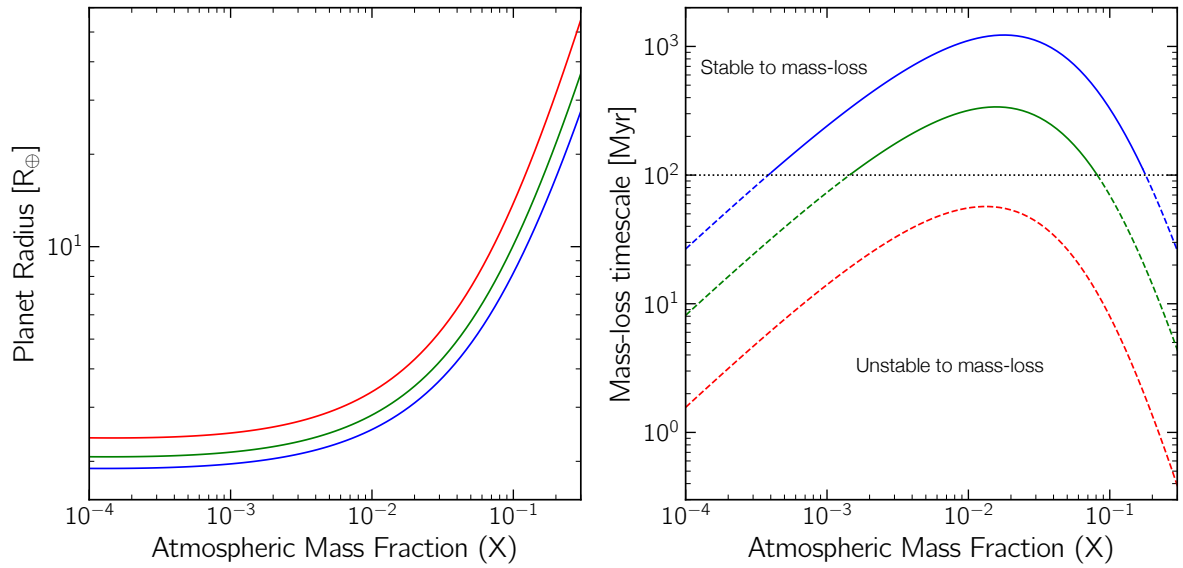


Figure 2.6: Left: Planet size is shown as a function of atmospheric mass fraction for a planet of $2M_\oplus$ core, around a solar-mass star. Red, green and blue curves represent planets at an orbital period of 2, 5 and 10 days respectively. The addition of a small amount of atmosphere shows the non-linear increase in planet size. Planets closer to their host star are more inflated due to a higher equilibrium temperature. Right: the mass-loss timescale is shown for the same planets. As atmospheric mass is removed (moving right to left), the mass-loss timescale increases since the size of the planet decreases (see Eq. 2.7.2). For planets at larger orbital separations, this timescale eventually grows larger than $100Myr$, which is the typical saturation time t_{sat} of stellar X-ray luminosity evolution. As a result, mass-loss stalls and the planet maintains $X \sim 0.01$. Planets closer to the host star do not enter this region of stability and continue to lose their entire atmosphere.

2.9 Model Limitations

The presented analytic model is exceptionally useful for population modelling as it is simple and therefore computationally cheap to evaluate. However, this benefit does come at a cost in terms of accuracy. In reality, one should perform numerical models with realistic equations of state and energy transport to understand this planetary evolution (e.g. Owen and Jackson, 2012; Owen and Wu, 2013; Lopez and Fortney, 2013). One major limitation in the model is the 'energy-limited' approach to mass-loss i.e. the mass-loss rate calculated in Eq. 2.7.2. The efficiency factor η has been shown to be non-constant in multiple works (Murray-Clay et al., 2009; Owen and Jackson, 2012; Salz et al., 2016; Caldiroli et al., 2021) meaning that, again, full hydrodynamic simulations are required to model this physics. Nonetheless, in this model, we upgrade to a non-constant η by adopting a scaling relation from Owen and Wu (2017) for η as follows:

$$\eta = \eta_0 \left(\frac{v_{\rm esc}}{25 {\rm km\ s^{-1}}} \right)^{-\alpha_\eta}, \tag{2.9.1}$$

where $v_{\rm esc} = \sqrt{2GM_{\rm c}/r}$ is the escape velocity of the planet and η_0 and α_η are positive constants. This form comes from Owen and Wu (2017), who fit mass-loss rates of Owen and Jackson (2012), however the numerical values of η_0 and α_η are discussed and constrained in Chapter 4.

2.10 Core-Powered Mass-Loss

As shown in previous discussion, the radius gap in a natural outcome of atmospheric mass-loss. However, this is not exclusive to XUV photoevaporation. An alternative model is also able to reproduce the observed demographics, namely core-powered mass-loss (e.g. Ginzburg et al., 2018; Gupta and Schlichting, 2019). The principle behind this mechanism is that the hydrodynamic outflow is not due to high energy flux, but instead a result of bolometric luminosity of the host star in conjunction with accretion luminosity from the core. This energy reserve builds during core formation - as solid material falls onto the forming protoplanet, binding energy is released in the form of heat. As this luminosity is slowly released into the atmosphere, it heats the gas from below and inflates the planet. If left in isolation, this heating would not be enough to drive an outflow i.e. the planet will remain in hydrostatic equilibrium with an increased altitude for the RCB. However, when this inflation is coupled with bolometric luminosity from the host star, a parcel of gas in the upper atmosphere has sufficient kinetic energy to overcome the gravitational potential of a planet. Similarly to photoevaporation, this then induces a hydrodynamic wind. A similar analytic model has been put forward in Gupta and Schlichting (2019) which is effective in demonstrating the validity of this mechanism. It relies on an identical treatment of atmospheric structure i.e. radiative convective equilibrium, however the mass-loss rates are adapted for the differing physics at play. I present this model here, since it is utilised in Chapter 3.

According to Gupta and Schlichting (2019), the mass-loss rate $\dot{M}_{\rm atm}$ in the core-powered mass-loss mechanism is given by:

$$\dot{M}_{\rm atm} = {\rm minimum} \left\{ \dot{M}^E_{\rm atm},\ \dot{M}^B_{\rm atm} \right\}. \tag{2.10.1}$$

where $\dot{M}^E_{\rm atm}$ is the energy-limited rate i.e. the maximum mass-loss rate achievable assuming all cooling luminosity from the planet goes into driving atmospheric mass-loss,

$$\dot{M}^E_{\rm atm} \simeq \frac{L_{\rm rcb}(t) R_{\rm c}}{G M_{\rm c}}, \tag{2.10.2}$$

where $L_{\rm rcb}$ is the luminosity of the planet at the radiative-convective boundary. Whereas the origin of this luminosity in photoevaporation is due to a change in binding energy e.g. Eq.

2.6.12, in this case a second contribution comes from the core, meaning the total thermal and gravitational energy available for cooling is as follows:

$$E_{\rm cool} = \frac{GM_{\rm c}}{R_{\rm c}^2}\Delta R\left(\frac{\gamma}{2\gamma-1}M_{\rm atm} + \frac{1}{\gamma}\frac{\gamma-1}{\gamma_{\rm c}-1}\frac{\mu}{\mu_{\rm c}}M_{\rm c}\right), \tag{2.10.3}$$

where $\Delta R = R_{\rm rcb} - R_{\rm c}$ is the radial depth of the atmosphere, $\gamma_{\rm c}$ is the adiabatic index for the convective core and μ and $\mu_{\rm c}$ are the mean molecular weights for the atmosphere and core respectively. In this case, the first term represents energy from the atmosphere and the second from the core. Alternatively, $\dot{M}^B_{\rm atm}$ is the Bondi-limited rate i.e. the maximum mass-loss rate that is physically achievable given the thermal energy of the escaping gas molecules,

$$\dot{M}^B_{\rm atm} = 4\pi R_{\rm B}^2 c_{\rm s}\rho_{\rm rcb}\exp\left(-\frac{GM_{\rm p}}{c_{\rm s}^2 R_{\rm rcb}}\right). \tag{2.10.4}$$

Following Gupta and Schlichting (2019), the energy required to lose the entire atmosphere is $E_{\rm loss} \approx gM_{\rm atm}R_{\rm c}$. Comparing this to available cooling energy from Eq. 2.10.3, we see that $E_{\rm cool} \lesssim E_{\rm loss}$ for $X > \mu/\mu_{\rm c} \sim 5\%$ i.e. heavy atmospheres, whereas $E_{\rm cool} \gtrsim E_{\rm loss}$ for $X < \mu/\mu_{\rm c} \sim 5\%$ i.e. light atmospheres. Note that we have ignored γ and $\gamma_{\rm c}$ for simplicity. The radius gap in this picture is then a natural outcome of energetics and directly related to the initial atmospheric mass fraction after disc dispersal. For planets with a large initial atmosphere, there is insufficient energy to remove the atmosphere and since $R_{\rm rcb}$ will be large, the exponential term in the Bondi-limited mass-loss rate of Eq. 2.10.4 will remain very small for these planets. In addition, the dependence in this equation of isothermal sound speed also introduces a dependency of incident stellar flux from the host star. On the other hand, planets with an initially small atmospheric mass-fraction will, from an energy point of view, continue to lose mass in a runaway process since, like photoevaporation, it is easier to remove a smaller amount of atmosphere.

Similar to photoevaporation, this model can reproduce the radius gap as seen in exoplanet surveys as shown in Figure 1.10. One major difference between the two models is the timescale on mass-loss occurs. Whereas the majority of photoevaporation is achieved in the first 100Myrs, core-powered mass-loss works over Gyr timescales (Gupta and Schlichting, 2019, 2020). This may prove to be key in determining which of these mechanisms is dominant. In Chapter 3, I present an additional way of distinguishing these models and lay out a path forward to answering this fundamental question.

Chapter 3

Photoevaporation Vs. Core-Powered Mass-Loss: Model Comparison with the 3D Radius Gap

3.1 Introduction

This chapter focuses on the comparison between XUV photoevaporation and core-powered mass-loss. Despite the difference of energy source for each mechanism, the two models predict very similar observable signatures in the exoplanet demographics. As shown in multiple works, both models predict that the core compositions of super-Earths and sub-Neptunes are Earth-like (e.g. Owen and Wu, 2017; Jin and Mordasini, 2018; Wu, 2019; Gupta and Schlichting, 2019, see Chapter 4). They also predict similar slopes to the radius valley as a function of orbital period and incident bolometric flux. Additionally, inference analysis suggests that the core mass distribution of small, close-in exoplanets should be peaked at a few Earth masses for both models (Gupta and Schlichting, 2019, see Chapter 4); however, it should be noted that the core-mass function derived from these models are different, as discussed in Section 3.2.2. More detailed information as to how these predictions are made can be found in Chapter 4.

If one is to determine which mechanism (if indeed there is one) drives the evolution of close-in exoplanets, one must focus on predictions that differ between the models. One possible avenue is the temporal evolution of the radius distribution of super-Earths and sub-Neptunes. Whilst photoevaporation predicts that the majority of mass-loss occurs in the first few 100 Myrs after disc dispersal when the X-ray flux is stronger, core-powered mass-loss predicts this to occur on $\sim$ Gyr timescales. It follows that one would expect to observe stripped cores (i.e. super-Earths) around stars $\sim$ 100 Myrs old if photoevaporation is taking place, but only after billions of years for core-powered mass-loss (Gupta and Schlichting, 2020, see Chapter 4). Recent attempts to constrain the evolution of the radius gap have found that results are currently consistent with both models, in part because there is a significant lack of planets observed around young stars (Berger et al., 2020a; Sandoval et al., 2020). With surveys such as THYME (Newton et al., 2019; Rizzuto et al., 2020; Mann et al., 2020; Newton et al., 2021) additional young planetary systems will be revealed, allowing these predictions to be investigated further. Note that recent work from King and Wheatley (2021) suggests that EUV stellar flux may subside slower than stellar X-ray flux, meaning that it may cause photoevaporative mass-loss on $\sim$ Gyr timescales. I address this point in Chapter 6.

In this chapter, I present a new method for model comparison, which utilises a key difference between the two models. As discussed in Chapter 2, the energy source for photoevaporation

stems from the host star, specifically from EUV/X-rays. Crucially, and as shown in Figure 1.5, the time-integrated X-ray exposure of a planet at a fixed incident bolometric flux decreases with increasing stellar mass (e.g. McDonald et al., 2019; Johnstone et al., 2021). This relationship arises due a combination of negative trends in stellar saturation time (i.e. the time at which stars' ratio of high-energy to bolometric output begins to decline) with stellar mass, as well as in the ratio of X-ray-to-bolometric luminosity also with stellar mass (Ribas et al., 2005; Wright et al., 2011; Jackson et al., 2012; Tu et al., 2015; Lopez and Rice, 2018; McDonald et al., 2019). As a result, photoevaporation predicts that lower mass stars will be able to strip larger planets at fixed incident flux. Core-powered mass-loss on the other hand does not predict a trend in stellar mass at constant incident flux, providing a useful difference in demographic predictions. Note that since transit surveys like *Kepler* and *TESS* typically search for planets over a fixed range in orbital period, care must be taken to work at constant incident flux, not fixed orbital period, or orbital period ranges for this analysis to work correctly.

This chapter begins with a formalism to describe this new method of model comparison in Section 3.2, followed by a proposed method of extracting such trends from data in Section 3.2.5. Results from synthetic and real surveys are presented in Section 3.3.

3.2 Method

3.2.1 The Radius Gap in 3D

As discussed in Chapter 2 and shown in Figure 3.1, the models for photoevaporation and core-powered mass-loss are both capable of reproducing the general demographics of close-in exoplanets. Determining a method for model comparison therefore depends on finding observable signatures that are predicted to differ substantially from one model to another. In previous works, the radius gap has been quantified as a line in 2D parameter space (e.g. as a function of orbital period P, stellar mass, M_*, or incident bolometric flux, S Fulton et al. 2017; Van Eylen et al. 2018; Berger et al. 2020a). However, both models predict the slopes in these planes to be very similar, and essentially indistinguishable with current and near-future exoplanet surveys (Loyd et al., 2020). Crucially this is not the case when looking at the radius gap in higher dimensions, i.e. a plane as a function of stellar mass and incident bolometric flux. As discussed, this arises because the models disagree on how the radius gap behaves as a function of stellar mass at fixed incident bolometric flux. In order to quantify this, I define the position of the radius gap $R_{\rm val}$ as the size of the largest super-Earths. This choice is made as this is the fundamental quantity that can be theoretically calculated from mass-loss models: namely, the largest core that can be entirely stripped at a given position in parameter space[1]. Thus, for an arbitrary evolutionary model, one can describe the 3D radius gap as a joint power law of the form:

$$R_{\rm val} \propto S^{\alpha}\, M_*^{\beta}, \tag{3.2.1}$$

where α and β are crucial mass-loss parameters and can be used to distinguish between photoevaporation and core-powered mass-loss. Specifically they are defined as:

$$\begin{aligned} \alpha &\equiv \left(\frac{\partial \log R_{\rm val}}{\partial \log S}\right)_{M_*} \\ \beta &\equiv \left(\frac{\partial \log R_{\rm val}}{\partial \log M_*}\right)_{S}. \end{aligned} \tag{3.2.2}$$

From theoretical predictions (discussed in Section 3.2.3) one would expect α to be similar for both models, however one would expect β to be negative for photoevaporation and zero for core-

[1] Although as described later, this quantity cannot always be extracted from a noisy exoplanet population in an unbiased way.

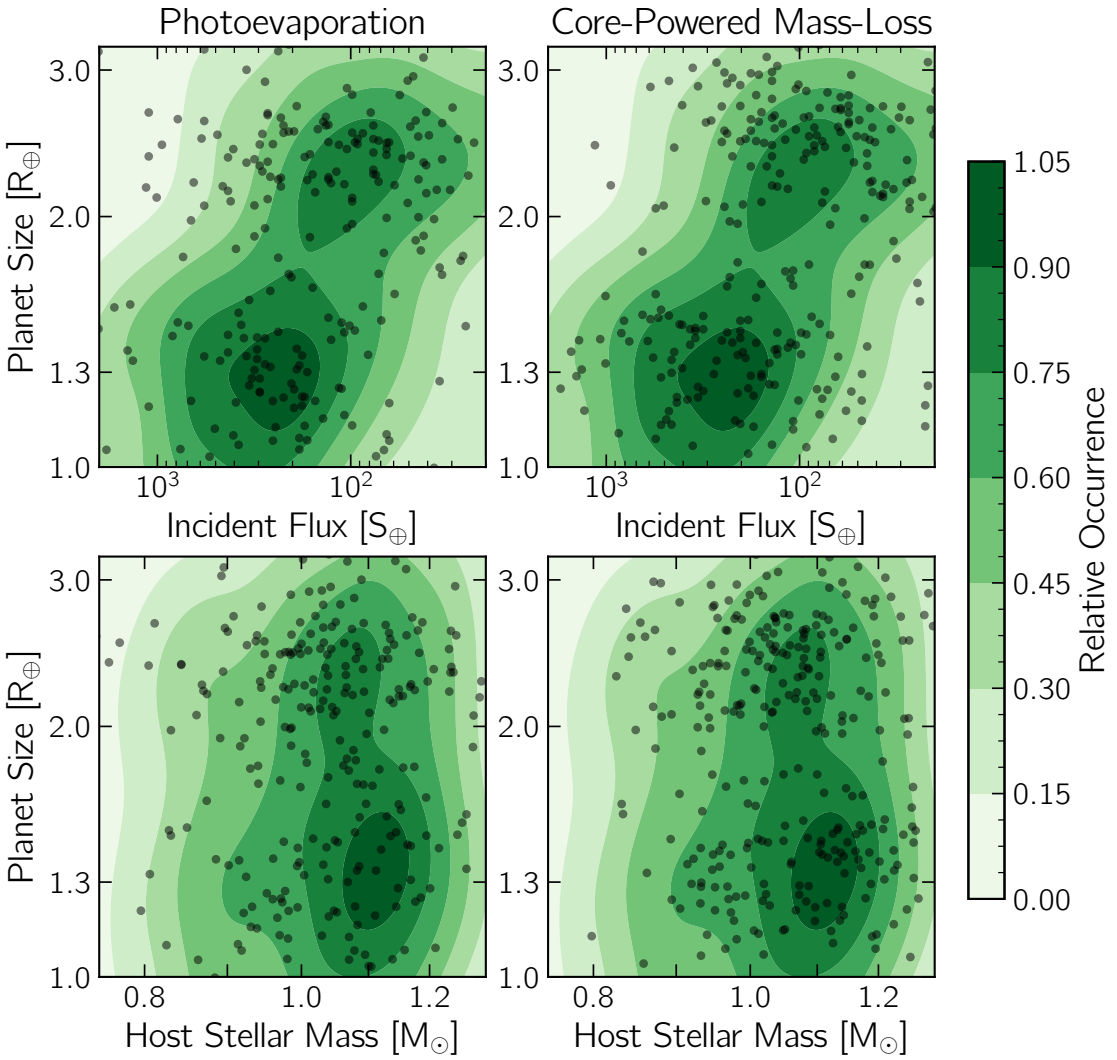


Figure 3.1: The radius gap is shown as function of incident bolometric flux (top row) and host stellar mass (bottom row). Here, green contours represent relative occurrence of the observed CKS planets (e.g. Fulton et al., 2017). Black points represent forward modelled planets from the photoevaporation model (e.g. Owen and Wu, 2013) (left) and core-powered mass-loss model (e.g. Ginzburg et al., 2016) (right), the latter of which were provided by A. Gupta (private communication). Note that modelled planets have appropriate bias and noise added such that they are representative of CKS data. Clearly, both models can reproduce the distribution of observed planets with the radius gap in the correct location and similar slopes. In order to determine which of the two models best describes the data, one must look to 3D representations of the radius gap, as laid out in Section 3.2.1. Details of prescriptions for each model are presented in Section 3.2.2.

powered mass-loss. This is because the high-energy exposure at fixed bolometric flux decreases as function of stellar mass, which means that smaller stars will be able to strip larger cores at fixed incident bolometric flux in the photoevaporation model, thus implying a negative value of β. This crucial fact allows one to differentiate between the evolutionary histories, and use it as a tool to find signatures in the data.

Before I use analytic models to predict theoretical values of α and β, I first address the radius gap in the M_*-R_p plane. Specifically, I wish to show why using this plane to perform model comparisons is extremely difficult due to strong degeneracies in the models. I justify this claim with a simplistic approach. I start with the joint power law of Eq. 3.2.1 in an inequality form, defining the maximum size of super-Earths for a given model, at that position in parameter space:

$$R_\mathrm{val} \leq A\ S^{\alpha}\ M_*^{\beta}, \tag{3.2.3}$$

where A is some arbitrary constant. Next I aim to eliminate S such that one can find the slope in the M_*-R_p plane. To do so, I use that $S \propto L_*/a^2$ where L_* is the bolometric luminosity, and utilise a stellar luminosity relation $(L_*/L_\odot) \propto (M_*/M_\odot)^{\zeta}$ as well as Kepler's third law $a^3 \propto P^2 M_*$ to find:

$$R_\mathrm{val} \leq A'\ M_*^{\alpha(\zeta-\frac{2}{3})+\beta}\ P^{-\frac{4}{3}\alpha}, \tag{3.2.4}$$

where A' is a re-scaled constant. Noting that the period distribution of close-in exoplanets is approximately independent as a function of stellar mass (Fressin et al., 2013; Dressing et al., 2015; Lee and Chiang, 2017; Petigura et al., 2018) and then differentiating one can derive the radius gap slope in the P-R_p plane as:

$$\frac{\mathrm{d}\log R_\mathrm{val}}{\mathrm{d}\log P} \approx -\frac{4\alpha}{3}. \tag{3.2.5}$$

Similarly, and more pertinent to this work, one can derive the radius gap slope in the M_*-R_p plane as:

$$\frac{\mathrm{d}\log R_\mathrm{val}}{\mathrm{d}\log M_*} \approx \alpha(\zeta - \frac{2}{3}) + \beta. \tag{3.2.6}$$

Inspection of this result shows that the slope is not only controlled by α and β but also ζ and hence the stellar luminosity relation. Since $\zeta \gg |\beta|$ and $\zeta \gg \alpha$ for both models (as shown in Sections 3.2.3 and 3.2.3) the value of this slope is heavily controlled by the stellar luminosity relation (Gupta and Schlichting, 2020; Loyd et al., 2020).

As an additional complication and as explored first by Wu (2019), one can invoke a core mass scaling with stellar mass which further contributes to the degeneracy in this result (see also Gupta and Schlichting 2020). Although difficult to incorporate into Eq. 3.2.6, the introduction of such scaling means that multiple combinations of α, β, ζ and the unknown core-mass–stellar-mass scaling can yield the same value of the slope in the M_*-R_p plane. To illuminate the effect of such a core-mass–stellar-mass scaling, Figure 3.2 shows planet distributions of the photoevaporation and core-powered mass-loss models, calculated using the models presented in Section 3.2.2. Core-powered mass-loss models were provided by A. Gupta (private communication). In the top row, I show a set of planets for each model with a core-mass function that is independent of stellar mass, which clearly shows a positive slope in the M_*-R_p plane. However, in the middle and bottom rows, I introduce core mass scalings $M_\mathrm{c} \propto M_*^{\pm 2}$ with the positive scaling in the middle row and negative in the bottom row. Clearly, the slope of the radius-gap in this plane can be altered dramatically with the inclusion of such scalings, further contributing to the degeneracy in the slope. One should note however that the magnitudes of the scalings chosen/implemented here are not consistent with the current observations and are used to simply demonstrate how correlations between stellar and planet masses can affect the slope.

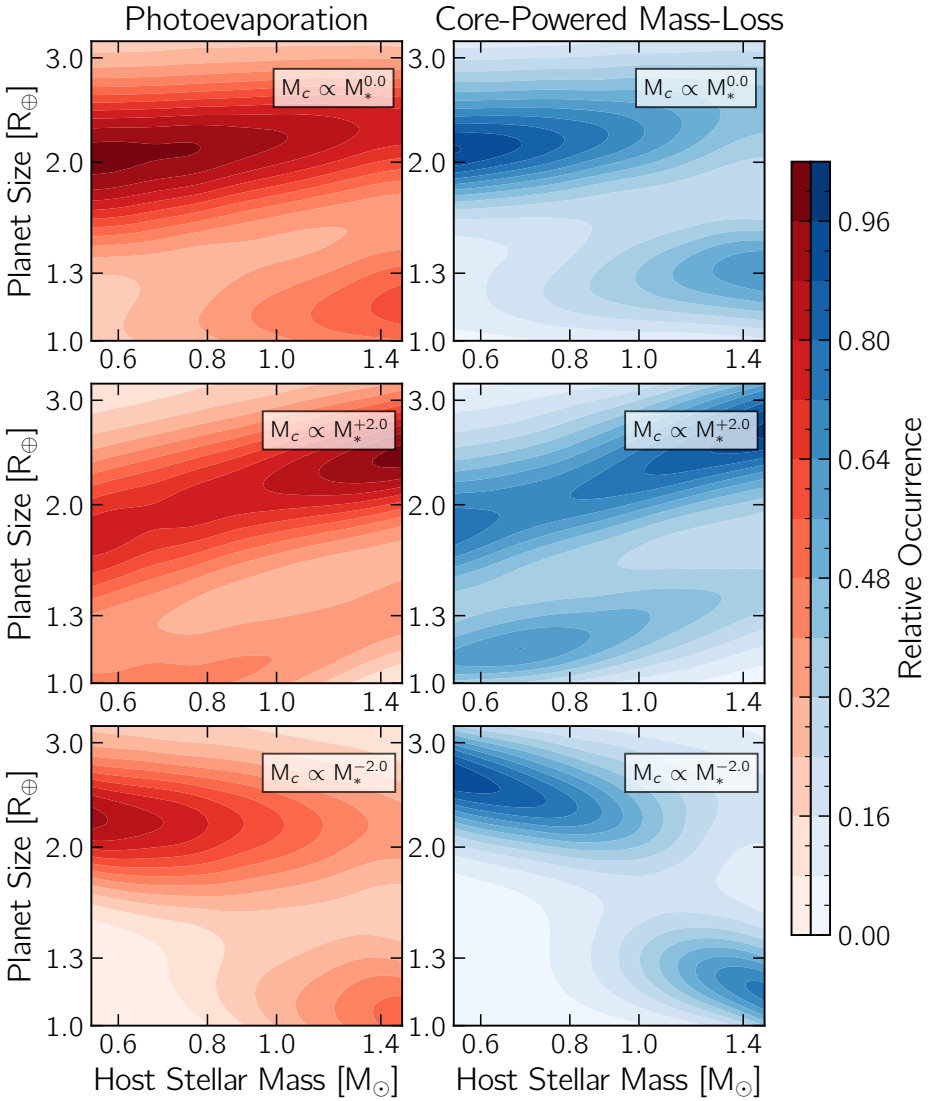


Figure 3.2: The radius gap is shown as a function of host stellar mass for the photoevaporation models (left) and core-powered mass-loss models (right). Here, contours represent relative occurrence and each slice in stellar mass is normalised to unity to better highlight the location of the radius gap. In the top row, I present the models for each of the two mass-loss schemes, where the core-mass function is independent of stellar mass. Both models predict a positive slope, however the exact value of slope is strongly controlled by the stellar mass-luminosity relation (see Section 3.2.1, specifically Eq. 3.2.6). In the middle and bottom rows, I incorporate scalings of core mass to stellar mass into the models $M_c \propto M_*^{\pm 2}$, with positive and negative scalings in the middle and bottom row respectively. This clearly demonstrates that the radius gap slope can be fundamentally altered by the inclusion of such scalings. Although I do not argue scalings of these magnitudes are physical, I stress that attempting model comparisons and indeed demographic inferences in this parameter space are fraught with degeneracies. All core-powered mass-loss models were provided by A. Gupta (private communication).

Indeed the analysis of Wu (2019), in which demographic inferences were performed in the M_*-R_p plane using the photoevaporation model, with the added introduction of a core mass scaling with stellar mass, argued for a linear relation between core-mass and stellar mass. However, note that the implemented bolometric luminosity relation (i.e. ζ in Eq. 3.2.6) was weaker than used in this work (and implemented in Figure 3.4). I choose to adopt a value of $\zeta = 4.5$ due to the fact that the CKS and GKS dataset are a magnitude-limited rather than a volume-limited sample, thereby including more high luminosity/mass stars (Petigura et al., 2017; Gupta and Schlichting, 2020). In addition, there exists a correlation between stellar mass and stellar metallicity in the *Kepler* data (e.g. Eker et al., 2018; Owen and Murray-Clay, 2018; Petigura et al., 2018; Gupta and Schlichting, 2020) which, in combination with the latter, increases the value of ζ in the stellar luminosity relation from the standard empirical relation of $L_* \propto M_*^{3.2}$ (e.g. Cox, 2000) that was adopted in Wu (2019). Since a lower value of ζ yields a shallower radius gap slope in the M_*-R_p plane (see Eq. 3.2.6), the analysis of Wu (2019) required a stronger core mass scaling in order to match the the observed value, compared to what one would infer from my implementation. Thus, if this analysis were repeated with a larger value of ζ, one can speculate that the inferred stellar-to-core-mass scaling from the photoevaporation model would be weaker. Nevertheless, this type of inference will be useful in the future once constraints are placed as to which mechanism dominates atmospheric mass-loss, which I discuss in Section 3.4.

Thus, the clear benefit of analysing the radius gap in 3D is that in determining α and β, one does not integrate through S or M_* axes (as in the derivation for Eq. 3.2.6) and hence avoids many of the degeneracies produced as a result. Additionally, the slopes α and β are, to the first-order, independent of any stellar mass scalings. This is because the plane is defined as the maximum size of stripped cores for a given incident flux S and stellar mass M_*. Hence, at constant S and M_*, the only way that core mass to stellar mass scalings can manifest in the data is if there is an upper limit on the sizes of cores for a given stellar mass (i.e. higher mass stars can host planets up to a higher mass cutoff). If this were the case, super-Earths simply would not exist at certain stellar masses. Nonetheless, for scaling relations such as those considered in Figure 3.2, this 3D method provides a way of circumventing the effects of such scalings. I discuss this claim in Section 3.4.4.

3.2.2 The Models

In order to explore these effects, I use modelled populations of planets from the analytic models for XUV photoevaporation from Owen and Wu (2017) and core-powered mass-loss[2] from Gupta and Schlichting (2019) as described in Chapter 2. In order to do so, I require underlying planet distributions for the following properties: orbital period P, stellar mass M_*, core density ρ, core mass M_c and initial atmospheric mass fraction X. An in-depth discussion of these distributions and how to determine them is provided in Chapter 4.

Following previous observational studies (e.g. Fressin et al., 2013; Petigura et al., 2018), both photoevaporation and core-powered mass-loss models assume that planets have the following underlying period distribution:

$$\frac{dN}{d\log P} \propto \begin{cases} P^2, & P < 8 \text{ days, and} \\ \text{constant}, & P > 8 \text{ days.} \end{cases} \tag{3.2.7}$$

In addition, both models assume that the cores are Earth-like in composition. On the other hand, the models use different core mass distributions and initial atmospheric mass fraction distributions, which were determined through inference work focused on planets around a narrow

[2]Note that all core-powered mass-loss models were provided by A. Gupta (private communication).

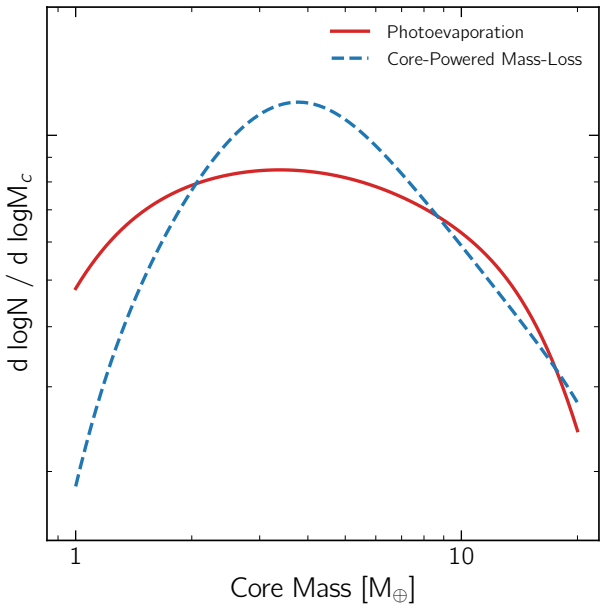


Figure 3.3: The core mass distributions are shown for the photoevaporation model (solid line) and core-powered mass-loss model (dashed line). These are taken from the inference analyses of (Gupta and Schlichting, 2019) and that of Chapter 4, in which the respective models were fit to the CKS data (Fulton et al., 2017).

range of Sun-like stars (Gupta and Schlichting, 2020; Rogers and Owen, 2021, see Chapter 4). As a result, both models represent current best-fits to exoplanet observations. I show the core mass distributions of both models in Figure 3.3 and note that both distributions are consistent with current data. I stress that that there would be little point performing this analysis with identical initial distributions as it would require at least one model to be inconsistent with the observations. Rather, this difference makes clear the importance of determining the dominant mass-loss model, as the inferred core-mass functions differ. In other words, if the observational tests proposed in this Chapter reveal that one mechanism dominates the planet evolution over the other mechanism, there will be little room to question such a conclusion as the models are already optimised to the observations. Further details of these distributions are discussed in the subsequent sections.

EUV/X-ray Photoevaporation Model

For the photoevaporation model, I make small improvements upon that which is presented in Chapter 2. Specifically, to address the aforementioned effect of stellar mass on stellar high energy evolution, I assume that the ratio of XUV to bolometric luminosity follows a broken power law:

$$\frac{L_{\rm XUV}}{L_{\rm bol}} = \begin{cases} \left(\frac{L_{\rm XUV}}{L_{\rm bol}}\right)_{\rm sat} \left(\frac{M_*}{M_\odot}\right)^{-0.5} & \text{for } t < t_{\rm sat}, \\ \left(\frac{L_{\rm XUV}}{L_{\rm bol}}\right)_{\rm sat} \left(\frac{M_*}{M_\odot}\right)^{-0.5} \left(\frac{t}{t_{\rm sat}}\right)^{-1-a_0} & \text{for } t \geq t_{\rm sat}, \end{cases} \tag{3.2.8}$$

where $a_0 = 0.5$ and $(L_{\rm XUV}/L_{\rm bol})_{\rm sat} = 10^{-3.5}$ (Wright et al., 2011; Jackson et al., 2012). The saturation time $t_{\rm sat}$ follows:

$$t_{\rm sat} = 10^2 \left(\frac{M_*}{M_\odot}\right)^{-1.0} \text{ Myr}. \tag{3.2.9}$$

Note the departure from Eq. 2.7.3 with the introduction of negative stellar mass scalings in Eq. 3.2.8 and Eq. 3.2.9, which is motivated by observational studies of (Jackson et al., 2012;

Shkolnik and Barman, 2014) and (McDonald et al., 2019; Johnstone et al., 2021). I do not at this stage, however, incorporate the observed spread in $L_{\rm XUV}/L_{\rm bol}$ (e.g. Tu et al., 2015; Johnstone et al., 2021). This is instead discussed in Chapter 6. Finally, in order to calculate $L_{\rm XUV}$, I use MIST stellar evolution tracks (Dotter, 2016; Choi et al., 2016) to accurately model the pre-main sequence and main-sequence evolution of the host star bolometric luminosity $L_{\rm bol}$. As discussed in Section 3.2.4, I eventually aim to simulate populations of planets that are akin to those in either the *California-Kepler Survey* (CKS) (Petigura et al., 2017; Fulton et al., 2017; Fulton and Petigura, 2018; Petigura et al., 2022) or *Gaia-Kepler Survey* (GKS) (Berger et al., 2020a). As the CKS sample is a subset of the GKS sample with spectroscopic follow-up, I take the metallicities and masses of 100 randomly sampled host stars from this survey and model them through MIST evolution tracks. I then interpolate $L_{\rm bol}$ evolution on a regular grid of time and stellar mass which is used in simulations of exoplanet photoevaporation. Not only does this improve the computational speed (instead of having to individually evolve each host star), but also means one naturally incorporates the intrinsic stellar luminosity relation into the simulations. For completeness, I also simulated a set of planets through photoevaporation in which the stellar luminosity evolution was taken from Baraffe et al. (1998). This yielded a difference in final planetary radii of $< 0.2\%$, which is much less than typical measurement uncertainties.

The chosen core mass distribution and initial atmospheric mass fraction for photoevaporated planets comes from the inference analysis presented in Chapter 4, in which $5^{\rm th}$-order Bernstein polynomials are used to constrain their functional forms. The best fit distribution for core mass is peaked at $\sim 4M_\oplus$ with sharp drops in occurrence towards higher and lower mass cores. Similarly, the initial atmospheric mass fraction distribution is sharply peaked at $\sim 2\%$. Note that these two distributions are independent of each other i.e. initial atmospheric mass fraction is independent of core mass, which is not the case for the core-powered mass-loss model.

For a given planet, I solve the ODE evolution equation of 2.7.4 with an RK45 adaptive-step numerical integrator with a error tolerance of 10^{-5} for 3 Gyrs, which is typical of the *Kepler* planet population. This provides a population of evolved planets which can be used to produce synthetic transit surveys (see Section 3.2.4).

Core-Powered Mass-Loss Model

In the core-powered mass-loss mechanism, models are based on the from the work of Gupta and Schlichting (2019). These were performed by A. Gupta (private communication, see Rogers et al. (2021) for details). The core mass distribution peaks at $\sim 4M_\oplus$ and can be expressed as

$$\frac{dN}{d\log M_{\rm c}} \propto \begin{cases} M_{\rm c}^2 \exp\left(-M_{\rm c}^2/(2\sigma_{M_{\rm c}}^2)\right), & M_{\rm c} < 5\ M_\oplus \\ M_{\rm c}^{-1}, & M_{\rm c} > 5\ M_\oplus, \end{cases} \tag{3.2.10}$$

where $\sigma_{M_{\rm c}} \sim 3M_\oplus$. Following Ginzburg et al. (2016), in which gas accretion and subsequent mass-loss are modelled during disc dispersal (see Chapter 5 for a discussion of this), I assume that a planet's initial atmospheric mass fraction follows:

$$X \simeq 0.05(M_{\rm c}/M_\oplus)^{1/2}. \tag{3.2.11}$$

3.2.3 Model Predictions

Photoevaporation Predictions

For photoevaporation, I derive α and β in a similar manner to the work of Owen and Wu (2017), in which a scaling of the radius gap was determined as a function of orbital period and mean core density. In this case however, I aim to find a joint power-law of the form of Eq. 3.2.1.

To begin, one can state that for a given core mass, the longest evaporation timescale $t_{\dot{X}}^{\max}$ is approximately equal to the stellar saturation time $t_{\rm sat}$ i.e. the largest cores to be stripped are those which take the longest time, approximately equal to the saturation time of the host star (see Figure 2.6 in Chapter 2). As argued by Owen and Wu (2017) and Mordasini (2020) planets that have not lost their atmosphere entirely by the end of the saturation period, retain them for the remainder of their lifetime. Combining definitions from Eq. 2.7.1 with Eq. 2.7.2 and discarding order-unity constants, one can write the longest evaporation timescale as:

$$t_{\dot{X}}^{\max} \sim \frac{a^2 M_{\rm p}^2 X_2}{\eta R^3 L_{\rm XUV}}, \tag{3.2.12}$$

where I set $X = X_2$, which is the atmospheric mass-fraction required to double a planet's core radius, and therefore yields the longest evaporation timescale (Owen and Wu, 2017). Equating this relation to $t_{\rm sat}$ and rearranging, one finds a condition for the location of the radius gap:

$$\frac{M_{\rm val}^2 X_2}{R_{\rm val}^3} \sim \eta t_{\rm sat} \frac{L_{\rm XUV}}{a^2}. \tag{3.2.13}$$

I now incorporate all the necessary scaling of these variables with $R_{\rm p}$, S and M_* such that one can find a joint power law of the form of Eq. 3.2.1. From its definition in Eq. 17 in Owen and Wu (2017), I write X_2 as:

$$X_2 \propto T_{\rm eq}^{-0.24}\, M_{\rm c}^{0.17} \propto S^{-0.06}\, R_{\rm p}^{0.68}, \tag{3.2.14}$$

where I have implemented a mass-radius relationship $M_{\rm c} \propto R_{\rm p}^4$ (Valencia et al., 2006) for bare cores. I take the photoevaporation efficiency η to scale with escape velocity, as in Eq. 4.2.7, and hence:

$$\eta \propto R_{\rm c}^{-3}. \tag{3.2.15}$$

Combining these relations with the scaling of $L_{\rm XUV}/L_{\rm bol}$ and $t_{\rm sat}$ with stellar mass (Eq. 3.2.8 and 3.2.9 respectively), one finds that:

$$R_{\rm val} \propto S^{0.12}\, M_*^{-0.17}. \tag{3.2.16}$$

Finally, one finds α and β:

$$\begin{aligned} \alpha &\equiv \left(\frac{\partial \log R_{\rm val}}{\partial \log S}\right)_{M_*} \simeq 0.12, \\ \beta &\equiv \left(\frac{\partial \log R_{\rm val}}{\partial \log M_*}\right)_{S} \simeq -0.17. \end{aligned} \tag{3.2.17}$$

Putting these values into the prediction for the slope of the radius gap in the M_*-$R_{\rm p}$ plane from Eq. 3.2.6, one finds a predicted value of:

$$\frac{{\rm d}\log R_{\rm val}}{{\rm d}\log M_*} \approx 0.29, \tag{3.2.18}$$

where I have assumed a stellar luminosity relation $(L_*/L_\odot) \propto (M_*/M_\odot)^\zeta$ with $\zeta = 4.5$ consistent with the CKS data set (Petigura et al. 2017; Gupta and Schlichting 2020, see Section 3.2.1). Additionally, and as shown in (Owen and Wu, 2017), the value of the slope in the P-$R_{\rm p}$ plane from Eq. 3.2.18 is predicted to be:

$$\frac{{\rm d}\log R_{\rm val}}{{\rm d}\log P} \approx -0.16, \tag{3.2.19}$$

Note that the assumptions made in this photoevaporation model, particularly the scaling of $t_{\rm sat}$ and $L_{\rm XUV}/L_{\rm bol}$ with stellar mass, as well as the efficiency scaling were chosen to yield the most

conservative estimate of β. This is because I wish to create a scenario in which differentiating between the two models is as challenging as reasonably possible. Changing these scalings to produce a more extreme value of β may falsely suggest that the models differ more substantially. As the evolution of high energy stellar luminosity still remains uncertain, I choose to adopt values that produce a prediction of β that is conservative, whilst still physically justifiable. I discuss these assumptions in Section 3.3.4.

Core-Powered Mass-Loss Predictions

To determine a joint power-law of the form Eq. 3.2.1 and hence derive α and β for core-powered mass-loss, I appeal to the argument previously presented by Gupta and Schlichting (2019) in which the slope of the radius valley in the planet distribution is set by the condition when the cooling timescale equals the mass-loss timescale, $t_{\rm cool} = t_{\dot{X}}$. (Gupta and Schlichting, 2019) showed that because $t_{\dot{X}}$ has an exponential dependence whereas $t_{\rm cool}$ doesn't, this equality can be approximated to $GM_{\rm c}/c_{\rm s}^2 R_{\rm rcb} =$ constant. If one then substitutes for $c_{\rm s}$ and uses the mass-radius relation $M_{\rm c} \propto R_{\rm p}^4$ (Valencia et al., 2006) and approximates $R_{\rm rcb} \simeq R_{\rm p} \sim 2R_{\rm c}$, this relation reduces to $R_{\rm p}^3 T_{\rm eq}^{-1} \simeq$ constant. Changing variables from equilibrium temperature to incident bolometric flux yields:

$$R_{\rm val}^3 T_{\rm eq}^{-1} \sim R_{\rm val}^3 S^{-1/4} \simeq \text{constant}. \tag{3.2.20}$$

Hence the predictions of α and β for core-powered mass-loss are as follows:

$$\begin{aligned} \alpha &\equiv \left(\frac{\partial \log R_{\rm val}}{\partial \log S}\right)_{M_*} \simeq 0.08, \\ \beta &\equiv \left(\frac{\partial \log R_{\rm val}}{\partial \log M_*}\right)_{S} \simeq 0.00, \end{aligned} \tag{3.2.21}$$

Note that due to lack of a stellar mass term in Eq. 3.2.20, the predicted value of β is zero. As with photoevaporation in Eq. 3.2.18, one can use these values to predict the slopes of the radius gap in the M_*-$R_{\rm p}$ and P-$R_{\rm p}$ planes from Eqs. 3.2.6 and 3.2.5, yielding:

$$\begin{aligned} \frac{{\rm d}\log R_{\rm val}}{{\rm d}\log M_*} &\approx 0.32, \\ \frac{{\rm d}\log R_{\rm val}}{{\rm d}\log P} &\approx -0.11, \end{aligned} \tag{3.2.22}$$

which are consistent with the analytical and numerical results of (Gupta and Schlichting, 2020), with the slope in M_*-$R_{\rm p}$ plane arising simply as a by-product of the stellar mass-luminosity correlation. Comparing these predictions to values of 0.29 and -0.16 for photoevaporation (Eqs. 3.2.18 and 3.2.19) further demonstrates that performing model comparisons in either of these 2D parameter spaces is extremely challenging.

3.2.4 Synthetic Transit Surveys

Before attempting to measure α and β from the real data, I first seek to determine whether they can be measured from synthetic surveys created using either the photoevaporation or core-powered mass-loss models. Evidently, the larger the survey i.e. the more planets that are observed, the better the constraints on α and β are likely to be. However, unlike with analysis of the radius gap in 2D (e.g. Fulton et al., 2017; Van Eylen et al., 2018; Berger et al., 2020a), more planets are needed to analyse the gap in 3D due to geometric scaling of the statistical problem. Additionally, effects such as the stellar mass distribution, measurement uncertainty and false positive contamination will also have an effect on the extraction of α and β.

In order to understand all these effects, I produce synthetic transit surveys which model the completeness and noise of such an observation. By controlling the number of planets that are observed, I therefore also control the Poisson noise. To do this, I model 100,000 planets and follow a similar prescription to Chapter 4, in which a high accuracy probability density function (PDF) for planet occurrence is calculated from each model in S-M_*-R_p space using a Kernel Density Estimation. I create the PDF using the kernel density estimator *fastKDE* of O'Brien et al. (2014, 2016) as this is computationally faster than standard techniques such as using a Gaussian kernel.

I adopt an underlying stellar mass distribution as a Gaussian of mean μ_{M_*} and standard deviation σ_{M_*}. I investigate the effect of changing this distribution in Section 3.3. The PDF for planet occurrence is then multiplied by the completeness map for the *Kepler* survey taken from (Fulton et al., 2017), but now extended to 3D space to match the domain of the planet occurrence PDF (see Chapter 4 for details on this calculation). Note that in doing so, I have assumed completeness is approximately constant as a function of stellar mass in the range $0.5M_\odot < M_* < 1.5M_\odot$. This multiplication has the effect of biasing the synthetic survey to preferentially observe planets with large radii and high incident bolometric flux. The new PDF is therefore a high accuracy representation of planet detection for a *Kepler*-like transit survey. The final task is to draw a desired number of planets N_p from this 3D PDF and incorporate measurement uncertainty. These are incorporated by adding a random Gaussian perturbation to planet parameters with zero mean and standard deviation equal to the associated error. Finally, I take these synthetically observed planets and measure α and β as laid out in Section 3.2.5. In Section 3.3, I aim to quantify how measurement uncertainty, false positive contamination and Poisson noise effect results of α and β. Hence to quantify this, I repeat the above process 10,000 times to produce joint-posterior distributions for α and β for a given survey setup.

3.2.5 Gradient Calculation

Determining the position and slopes of a radius gap in real data is an ill-defined statistical problem. Whereas one typically wishes to find where data trends exist, I now aim to quantify a trend in the absence of data. Here I present a methodology for isolating and extracting α and β from data in the 3D space of S-M_*-R_p. This builds upon work performed as part of the CKS-Cool collaboration, as discussed in Chapter 1. Instead of a line, the valley is now characterised as a plane in log-space in the following manner:

$$\log R_\mathrm{val} = \left(\frac{\partial \log R_\mathrm{val}}{\partial \log S}\right)_{M_*} \log S + \left(\frac{\partial \log R_\mathrm{val}}{\partial \log M_*}\right)_{S} \log M_* + \log R_0, \tag{3.2.23}$$

where $\log R_0$ is the normalisation of the plane and a nuisance parameter in this problem. One can then use definitions of α and β from Eq. 3.2.2 to simplify:

$$\log R_\mathrm{val} = \alpha \log S + \beta \log M_* + \log R_0. \tag{3.2.24}$$

As one would expect the radius valley to be flat in log-space, I expect α and β to be constants and can therefore extract them to quantify the slope for a given set of planets. The methodology for this fitting is as follows: I first use a population of observed planets (such as those produced from the synthetic surveys in Section 3.2.4) to calculate a PDF for planet detection in S-M_*-R_p space. As with Section 3.2.4, I create the PDF using *fastKDE* of O'Brien et al. (2014, 2016). I label this PDF as $\lambda(S, M_*, R_\mathrm{p})$. I then take slices at constant stellar mass $M_{*,i}$ through this PDF and locate the radius gap in the S-R_p plane for each 2D slice. To do this, I define the radius gap at slice i as:

$$\log R_\mathrm{val} = \alpha_i(\log S - 2) + \log R_{\mathrm{p},100,i} \tag{3.2.25}$$

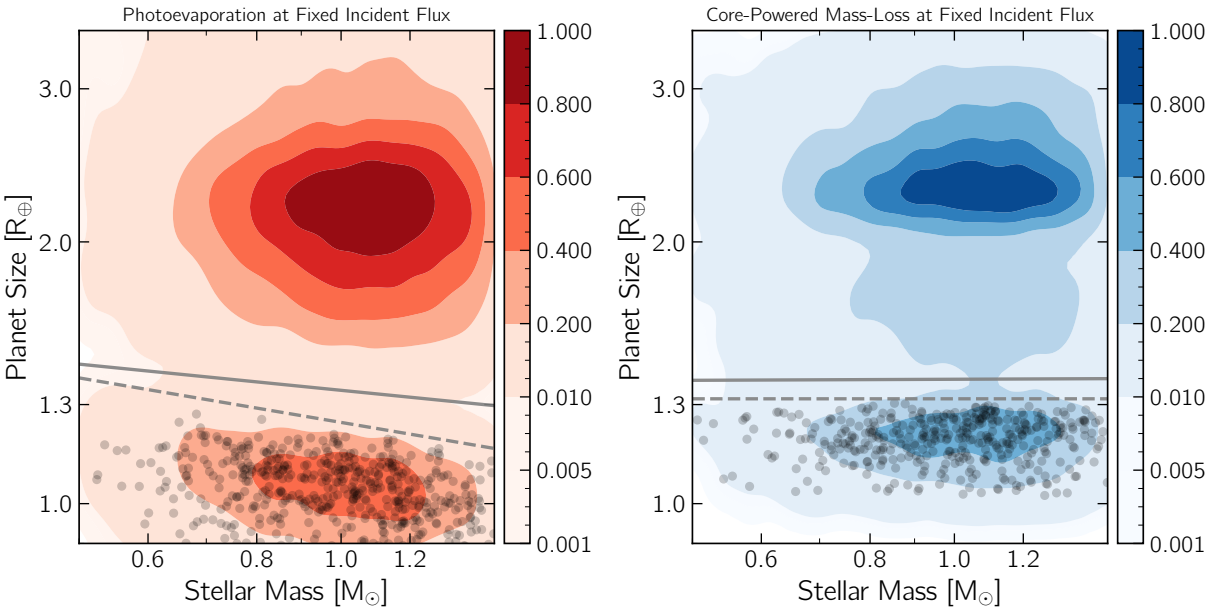


Figure 3.4: The underlying planet distributions for photoevaporation (left) and core-powered mass-loss (right) in the M_*- R_p plane at constant incident flux $S = 50S_\oplus$. By holding incident flux constant, the radius gap slope is representative of β. Here contours show relative occurrence. Dashed lines show theoretical predictions of β from Sections 3.2.3 and 3.2.3, whilst solid lines show values of β measured using method described in Section 3.2.5. These two lines are offset because the theoretical radius gap defines the size of the largest super-Earths and hence is at lower radii than gap measured by the proposed algorithm. Individual super-Earths are plotted for each model in black i.e. those with no atmosphere present after 3 Gyrs of evolution, in the incident flux range $45\ S_\oplus < S < 55\ S_\oplus$. Core-powered mass-loss models are provided by A. Gupta (private communication, see Rogers et al. (2021) for details).

where $\log R_{\mathrm{p},100,i}$ is the intercept at $S = 100S_\oplus$. I extract parameters α_i and $\log R_{\mathrm{p},100,i}$ for each slice by evaluating the line integral $\mathcal{I}$ through $\lambda(S, M_*, R_\mathrm{p})$:

$$\mathcal{I} = \int_{\mathcal{C}} \lambda(S, M_{*,i}, R_\mathrm{p})\ \mathrm{d}S, \tag{3.2.26}$$

where $\mathcal{C}$ is the path defined by the line in Eq. 3.2.25. The exact parameters are determined by minimising $\mathcal{I}$ using *scipy.optimize*, which provide α_i and $\log R_{\mathrm{p},100,i}$ for each slice. The final value of α is calculated as the mean α_i for all slices, whilst the value for β is determined by fitting a straight line through the values of $\log R_{\mathrm{p},100,i}$ as a function of stellar mass $M_{*,i}$. The slope of this line provides a value for β. In the event that this method misidentifies the radius gap for a given slice, the value of α_i or $\log R_{\mathrm{p},100,i}$ will be an outlier compared to those of other slices in the same PDF. I employ a simple outlier detection scheme to remove these erroneous values, by removing any α_i or $\log R_{\mathrm{p},100,i}$ that lies 2σ away from the mean value, where σ is the standard deviation of $\{\alpha_i\}$ or $\{\log R_{\mathrm{p},100,i}\}$. I choose $N = 15$, with slices logarithmically spaced between $1 - \sigma_{M_*}$ and $1 + \sigma_{M_*}$ where σ_{M_*} is the standard deviation of the stellar mass distribution. It was determined that extracted values of α and β converged with $N \gtrsim 10$, which is reflected in the choice of N.

Although in this work I apply this method to the 3D space of S-M_*-R_p, it can be used to locate the radius gap in any chosen parameter space. To demonstrate its effectiveness and also its limitations, Figure 3.4 shows the gradients calculated in the 2D M_*-R_p plane at a constant incident flux for photoevaporation and core-powered mass-loss models. Note that by holding incident flux constant the radius gap slope is representative of β, not that of Eq. 3.2.6 and those shown in Figure 3.2. Solid lines show the slopes extracted by minimising the PDF integral through the M_*-R_p plane, whilst dashed lines represent the theoretical predictions of β

for the respective models. This highlights a limitation of the methodology, which is particularly visible for the photoevaporation data. Note that whilst the predicted and measured slopes are equal for core-powered mass-loss data, they differ for photoevaporation, with the measured value being shallower than theoretically predicted. This issue arises because the numerical method for determining the slope locates the line that bisects the two populations of super-Earths and sub-Neptunes, whilst the theoretical predictions are determined by finding the size of the largest super-Earths. This subtle distinction implies that the one expects the extracted value of β to be shallower[3] than the theoretical prediction of $\beta = -0.17$. The implications of this effect are discussed in Section 3.4.3. I emphasise that in this work I present a *possible* method for extracting the values of α and β from an observed population of exoplanets that successfully works on real data, albeit in a biased fashion. Therefore, while the theoretical ideas underpinning the observational test are independent of the method used, the accuracy and precision to which these values can be extracted from the same data-set is method dependent. Since my best method is biased, I encourage further work to develop a precise and unbiased method.

3.3 Results

The goal of this work is to determine the qualities of a hypothetical transit survey suitable for determining which model best describes the exoplanet data. Therefore, I wish to ascertain approximate values for the required number of detected planets $< 4R_\oplus$, measurement uncertainty threshold, false positive contamination level and also the host stellar mass distribution for such a survey. To begin, I reproduce a CKS (Fulton and Petigura, 2018; Petigura et al., 2022) or GKS (Berger et al., 2020a) style survey which both benefit from stellar radii constraints from *Gaia* (Gaia Collaboration et al., 2018). As a result, I model the typical measurement uncertainties of 5% for planet radius, 5% for stellar mass and 8% for incident bolometric flux. I perform surveys with a host stellar mass distribution approximated as a Gaussian function with mean $\mu_{M_*} = 1.0 M_\odot$ and varying standard deviation $\sigma_{M_*} = \{0.1, 0.2, 0.3, 0.4\} M_\odot$. Initially I also assume that there are no false positives in the survey. Recall that the two parameters of interest are $\alpha = (\partial \log R_{\rm val} / \partial \log S)_{M_*}$ and $\beta = (\partial \log R_{\rm val} / \partial \log M_*)_S$. As β is the parameter expected to differ the most between the two models and quantifies how the radius gap behaves as a function of stellar mass, it follows that the wider the stellar mass distribution, the better the constraint. Recall that for a survey of size $N_{\rm p}$ planets, I quantify the uncertainty in these parameters by redrawing $N_{\rm p}$ planets and extracting α and β 10,000 times. This is performed twice: once for planets modelled using the photoevaporation scheme and again for planets modelled using the core-powered mass-loss scheme.

Figure 3.5 shows a summary of posteriors of α and β for varying size of survey and host stellar mass distribution width. Red contours represent the 1σ and 2σ bounds for synthetic surveys with photoevaporated planets, whilst blue contours show the same for planets that have undergone core-powered mass-loss. Black crosses represent the theoretical values expected from the each model, given by Eq. 3.2.17 for photoevaporation and Eq. 3.2.21 for core-powered mass-loss. As expected, the size of the posteriors reduces as the size of the survey increases (i.e. moving right in the figure). This can be understood because the increase in detected planets results in the Poisson noise having less of an effect on the planet distribution. Additionally, as the width of the stellar mass distribution is increased (i.e. moving down in the figure), the contours move closer to the theoretical predictions of α and β. Again, this is expected, as a wider distribution provides more information on how the radius gap changes as function of

[3]Note that if the line of smallest sub-Neptunes is parallel to the line of largest super-Earths then the measured value will be approximately equal to that which is theoretically predicted. This is the case for α and β in the core-powered mass-loss model but only α for photoevaporation.

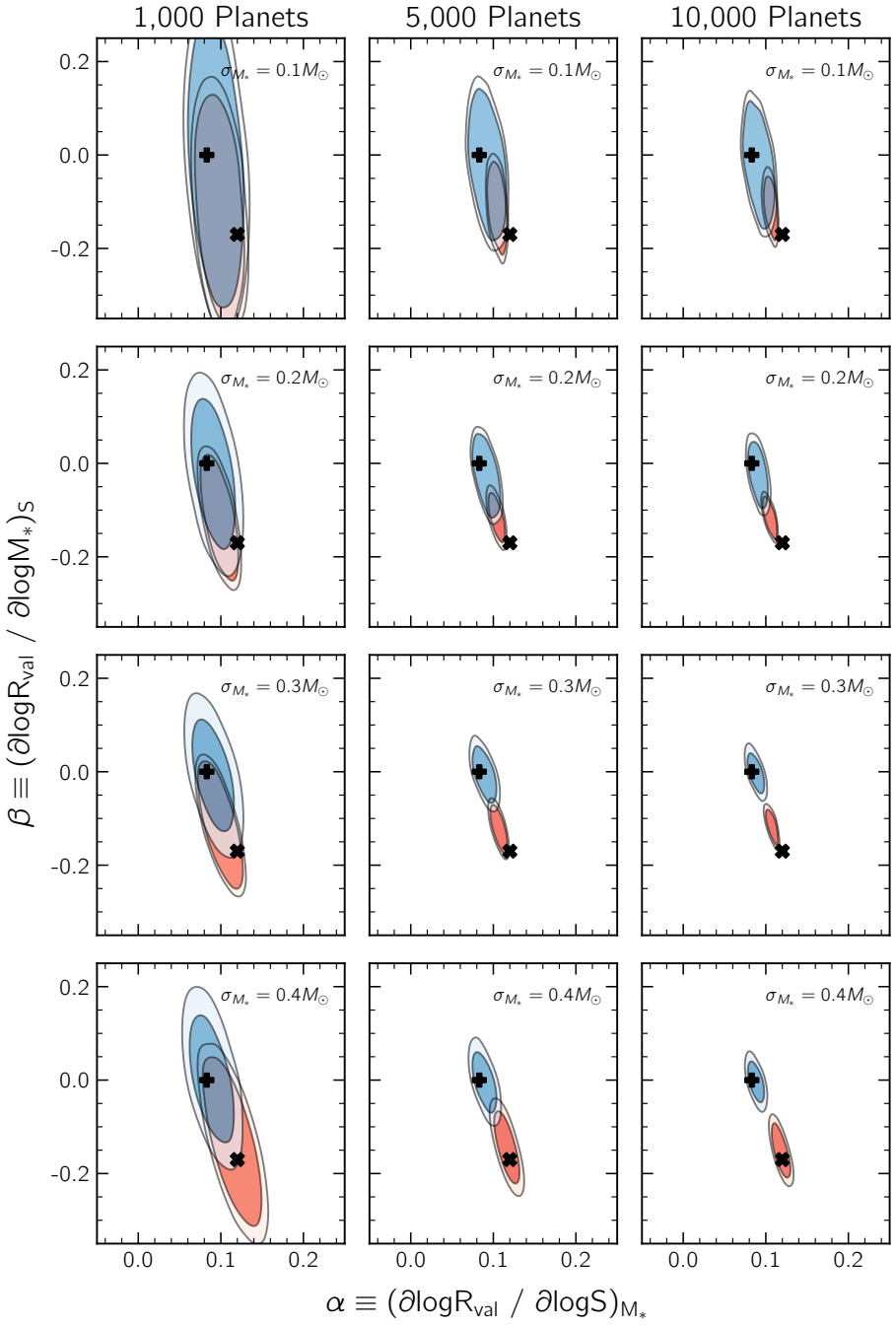


Figure 3.5: Posterior distributions are shown of key radius gap parameters α and β for synthetic transit surveys with planets modelled through photoevaporation evolution in red and core-powered mass-loss evolution in blue. Contours represent the 1σ and 2σ levels, with expected values from theoretical predictions shown as a '+' for core-powered mass-loss and a '×' for photoevaporation. One can see that α and β are better constrained for larger surveys (moving from left to right) and wider stellar mass distribution widths (moving from top to bottom). Note that the posteriors are not always consistent with theoretical predictions due to limitations in the gradient extraction method (see Sections 3.2.5 and 3.4.3).

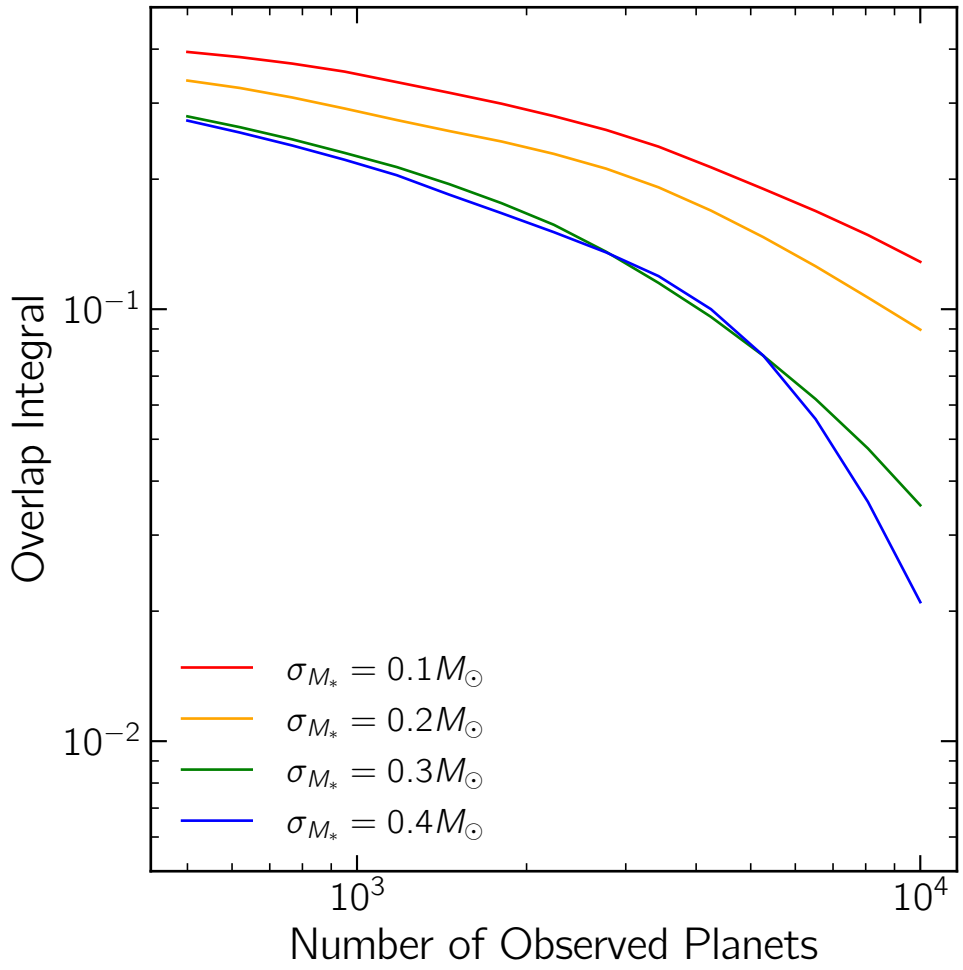


Figure 3.6: The overall limitations of extracting radius gap parameters α and β for synthetic transit surveys are demonstrated with the overlap integral between posteriors of Figure 3.5 as function of survey size. This captures the "precision" of the statistical analysis i.e. how much do the posteriors overlap and prevent separation of models. Essentially this is a measure of the probability that one would fail to distinguish between the models in a given experiment. Different coloured lines show the overlap integral for varying widths in stellar mass distribution.

stellar mass. Note how the posteriors for photoevaporation are consistently finding a shallower value for β than theoretically predicted, as discussed in Section 3.2.5, due to the inherent bias in extracting these slopes; see Section 3.4.3 for further discussion.

To quantify the effectiveness of the proposed statistical method of extracting α and β, I determine the overlap integral of the posteriors shown in Figure 3.5. The overlap integral assesses the "precision" of the method i.e. how well can photoevaporation be distinguished from core-powered mass-loss. This is shown in Figure 3.6 as a function of survey size and stellar mass distribution width. It can be seen that the overlap decreases with increasing size of survey as well as increasing stellar mass distribution width, confirming the general trends seen in Figure 3.5. Additionally, the overlap integral estimates the significance level at which the two posteriors can be distinguished. As an example, for the survey with stellar mass width $\sigma_{M_*} = 0.3M_\odot$, one can differentiate between the two models 90% of the time with a survey of $\gtrsim 5000$ planets.

3.3.1 Measurement Uncertainty

An important question is at what threshold of measurement uncertainty can α and β be sufficiently extracted from a given survey. In order to do this, I take a baseline survey with a stellar mass distribution of a Gaussian function, centred at $1M_\odot$ with a standard deviation of $0.3M_\odot$ and then vary the fractional measurement uncertainty to see how this effects the analysis. I consider three scenarios: the first in which percentage errors for each of the quantities (S, M_*, R_p) are $(4\%, 2\%, 2\%)$, a second scenario with $(8\%, 5\%, 5\%)$ and finally a third with $(12\%, 8\%, 8\%)$. Note that uncertainties in incident flux S from transit surveys are typically larger as this quantity is determined via stellar modelling which is used to calculate the bolometric luminosity and hence incident flux.

Figure 3.7 demonstrates the effect of changing the measurement uncertainties of a survey by considering the overlap integral. In green, it shows that reducing percentage uncertainties to $(4\%, 2\%, 2\%)$ can dramatically improve the distinction between photoevaporation and core-powered mass-loss models by up to $\sim 50\%$. Increasing the uncertainties on the other hand, to $(12\%, 8\%, 8\%)$, has the opposite effect with a larger overlap between the two posteriors shown in red. The result can be understood because measurement uncertainty can stochastically shift planets into the radius gap and thus mask the underlying values of α and β.

3.3.2 False Positives

Similar to measurement uncertainty, the presence of false positives in a sample can place data points into the radius gap, and thus reduce a survey's purity. To investigate this effect, I model false positive contamination as data points randomly drawn within $S \in [10^{-1}, 10^4]S_\oplus$, $M_* \in [0.5, 1.5]M_\odot$ and $R_\mathrm{p} \in [1.0, 4.0]R_\oplus$. For a given survey of size N_p, I remove a fraction of these planets and replace them with false positives. I show the results for a 5% and 10% contamination in dashed and dot-dashed lines respectively in Figure 3.7. Interestingly, the injection of false positives has little, to no effect on the the ability to extract α and β and thus on the distinction between photoevaporation and core-powered mass-loss. One can understand this as follows: whereas measurement uncertainty effects every data point in the survey and crucially those near the radius gap, false positives are injected randomly across the whole domain, meaning a large number of false positives will be placed far from the gap. Thus, up to at least a 10% contamination level, any false positives found in the gap will not prevent α and β from being extracted accurately.

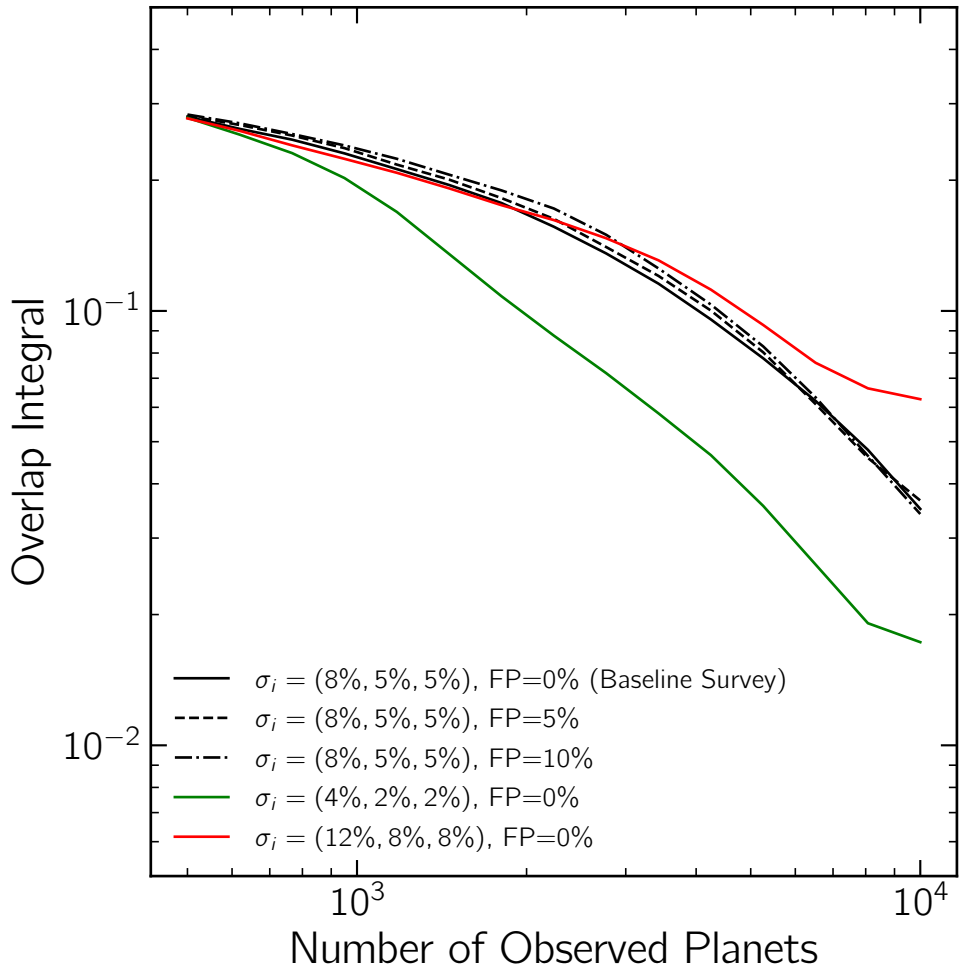


Figure 3.7: Same as Figure 3.6, whereas here the overlap integral is calculated for varying survey quality as a function of survey size. Here, σ_i refers to the percentage errors in each of three measurements required for this statistical analysis (S, M_*, R_p), whilst FP represents the percentage of false positives present in the survey. The baseline survey (black solid line) has a stellar mass distribution standard deviation of $0.3M_\odot$ (same as green line in Figure 3.6). Dashed and dot-dashed lines represent the baseline survey with a 5% and 10% false positive contamination. Red and green lines represent the baseline survey with measurement uncertainties of $(4\%, 2\%, 2\%)$ and $(12\%, 8\%, 8\%)$ respectively.

3.3.3 Changing the Stellar Mass Distribution

So far in this study, I have restricted the stellar mass distribution to be centred at $1M_\odot$, as this is typical of the CKS or GKS datasets (Fulton et al., 2017; Berger et al., 2020a; Petigura et al., 2022). Further improvements can be made in our ability to distinguish between the two models if the stellar mass distribution shifts to lower masses. To understand this, one must take a closer look at the radius gap. As seen in models of photoevaporation and core-powered mass-loss in Figure 3.1, the radius gap extends to high incident flux, whereupon it opens into a large sparsity of planets referred to as the "Neptune-desert". When α and β are being extracted from the 3D data (as in Section 3.2.5), this wedge-shaped feature frequently causes the extracted value of β to be diminished. However, as incident flux is strongly correlated with stellar mass, if one moves to lower mass stars, the Neptune-desert becomes less prevalent in the data. As a result, the algorithm can determine β more accurately.

To demonstrate this, Figure 3.8 shows the overlap integrals of a selection of the baseline surveys in Figure 3.6 compared with surveys in which the stellar mass distribution is shifted to a Gaussian function centred at $0.75M_\odot$. One clearly sees that shifting the distributions to lower stellar masses reduces the overlap integral and thus improves ability to distinguish photoevaporation from core-powered mass-loss. For a survey with stellar mass distribution centred at $0.75M_\odot$ and width $0.2M_\odot$, photoevaporation can be distinguished from core-powered mass-loss with a survey of $\gtrsim 3000$ planets 90% of the time. I also include a survey centred at $1.0M_\odot$ with a large width of $0.4M_\odot$ to show that one can achieve the same, if not a better, result if one surveys a narrow mass range of stars but biased to lower masses.

3.3.4 Changing Model Parameters

Photoevaporation

As discussed in Section 3.2.3, the predicted value of β can be altered if one assumes different model parameters for the high-energy stellar evolution in the photoevaporation model. As a result, I intentionally designed the model such that the predicted value of β was as small in magnitude (and therefore closest to the value predicted by core-powered mass-loss) as possible, whilst still being physically justified. This results in shifting the posterior of photoevaporation closer to that of core-powered mass-loss, and hence systematically increases the overlap integral. By presenting a conservative model for photoevaporation, I make the task of performing this statistical analysis more challenging, but can state more securely that if this were done on real data, any definitive result cannot be negated by large changes in the models of photoevaporation.

With that in mind however, if photoevaporation is the driving mechanism of atmospheric mass-loss, it may in reality result in a larger magnitude of β. To demonstrate the effect this would have, Figure 3.9 shows the overlap integral calculated for two surveys: one in which the photoevaporation model is the conservative one presented in Section 3.2.2 with $\beta = -0.17$, whilst the other is a more extreme model, where $\beta = -0.23$ is the new predicted value. To do this, I change the stellar mass scalings of X-ray to bolometric luminosity $L_{\rm XUV}/L_{\rm bol}$ to be:

$$\frac{L_{\rm XUV}}{L_{\rm bol}} = \begin{cases} \left(\frac{L_{\rm XUV}}{L_{\rm bol}}\right)_{\rm sat} \left(\frac{M_*}{M_\odot}\right)^{-1.0} & \text{for } t < t_{\rm sat}, \\ \left(\frac{L_{\rm XUV}}{L_{\rm bol}}\right)_{\rm sat} \left(\frac{M_*}{M_\odot}\right)^{-1.0} \left(\frac{t}{t_{\rm sat}}\right)^{-1-a_0} & \text{for } t \geq t_{\rm sat}. \end{cases} \tag{3.3.1}$$

Note that before (as in Eq. 3.2.8), the index on stellar mass was -0.5 as opposed to -1.0 in this case. As is clear in Figure 3.9, the overlap decreases assuming this form of photoevaporation is taking place, meaning that distinguishing between it and core-powered mass-loss could potentially occur for a smaller survey.

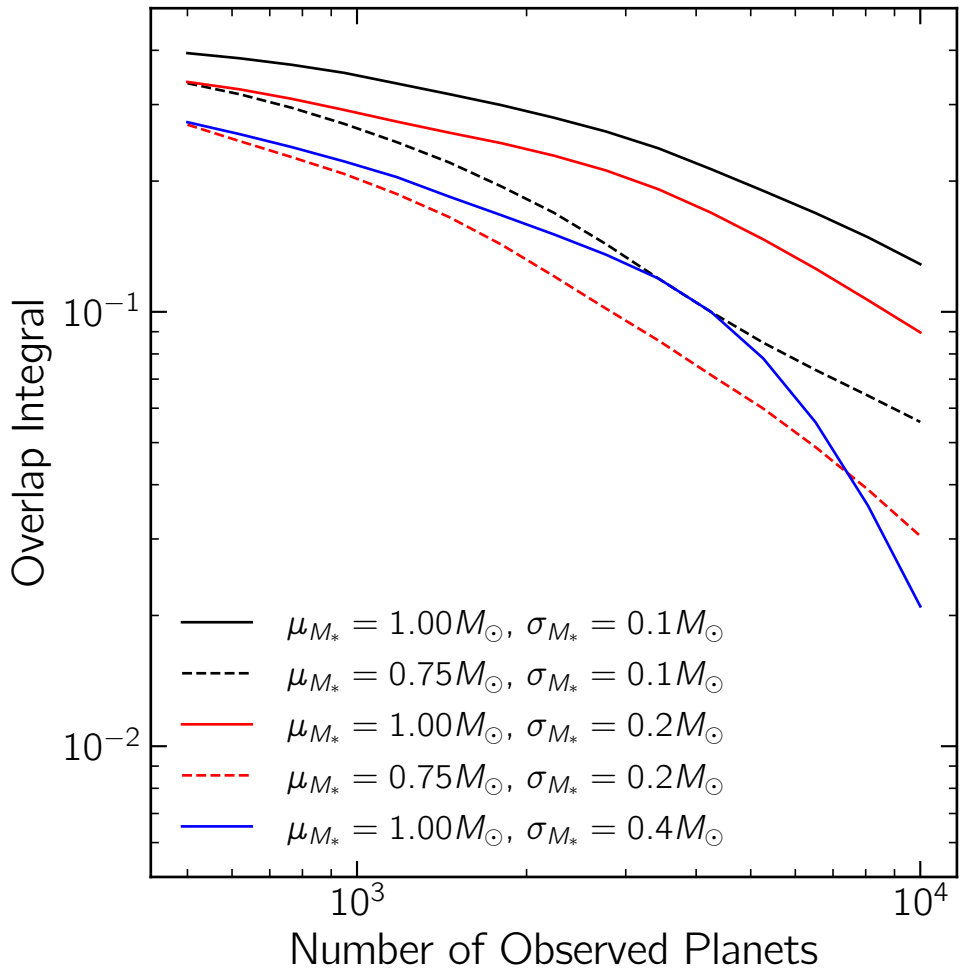


Figure 3.8: Same as Figures 3.6 and 3.7, whereas here the overlap integral is calculated for different stellar mass distribution means μ_{M_*} and standard deviations σ_{M_*}. Dashed lines represent surveys where the stellar mass distribution is centred at lower masses, which has the effect of reducing the overlap integral for the same distribution width. A survey with a distribution centred at $1.0M_\odot$ and large width of $0.4M_\odot$ is shown in blue for comparison.

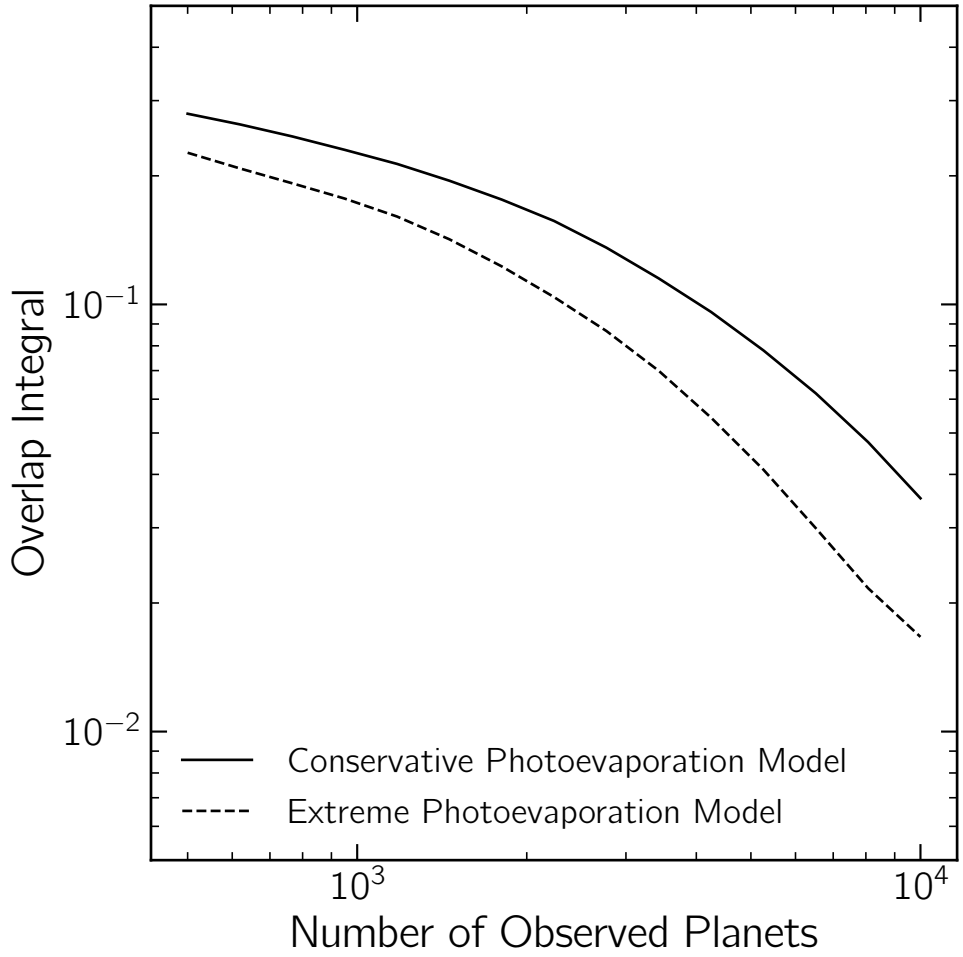


Figure 3.9: Same as Figures 3.6, 3.7 and 3.8, whereas here the overlap integral is calculated under different assumptions for the photoevaporation model. Both surveys have a stellar mass distribution centred at $1.0M_{\odot}$ *and width of* $0.3M_{\odot}$*. The conservative model (i.e. as presented in Section 3.2.2) is shown in the solid black line, whereas a more extreme model with a larger magnitude of predicted* β *is shown in the dashed line.*

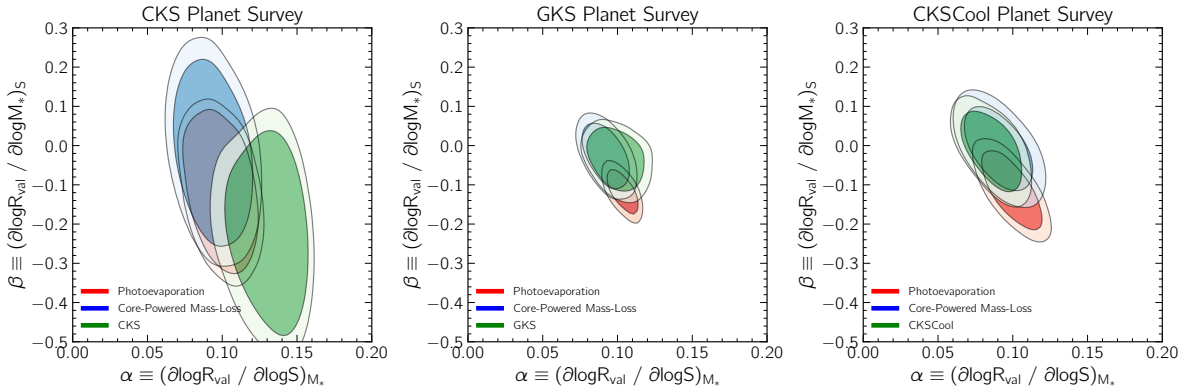


Figure 3.10: Posteriors for α and β are shown in green for CKS (data-release 1), which has 921 planets (Fulton et al., 2017; Fulton and Petigura, 2018) (left), GKS survey with 3760 planets (Berger et al., 2020a) (centre) and CKS-Cool (CKS data-release 2) (Petigura et al., 2022) (right), with contours representing 1σ and 2σ confidence levels. Red and blue contours represent synthetic surveys using the photoevaporation model and core-powered mass-loss model respectively. The synthetic surveys match the number of detected planets and assume a stellar mass distribution centred at $1M_{\odot}$ with widths of $0.15M_{\odot}$, $0.2M_{\odot}$ and $0.3M_{\odot}$ for CKS, GKS and CKS-Cool respectively. Note that I assume measurement uncertainties in incident flux, stellar mass and planet size as (8%,5%,5%) and adopt the completeness map from the CKS survey.

Core-Powered Mass-Loss

Unlike photoevaporation, the predicted values of α and β for core-powered mass-loss are more certain. One possible avenue to consider however, is the metallicity trend of host stars (Gupta and Schlichting, 2020). Let us suppose that host star metallicities and masses are positively correlated: $Z_* \propto M_*^a$ where $a > 0$ is some constant. In addition, following (Gupta and Schlichting, 2020), one can assume that the metallicities of primordial planet atmospheres follow that of their host stars such that $[\mathrm{Fe/H}] \simeq [Z_*/Z_{\odot}]$, which is not unreasonable because metal-rich stars are expected to have metal-rich discs. It then follows that more massive stars will host planets with metal enhanced atmospheres. Furthermore, assuming that metal-rich atmospheres have higher opacities (κ) such that $\kappa \propto Z_*$, implies that more massive or metal-rich stars are likely to host planets with higher opacity atmospheres. Consequently, more massive stars are likely to host planets that are inflated in sizes for their age because of their longer cooling timescales.

For planets hosting significant H/He atmospheres i.e. sub-Neptunes, this will result in a positive slope in their size as a function of stellar mass. Recall however, that α and β are defined as slopes in the maximum size of stripped cores i.e. super-Earths, as a function of S and M_*. Therefore, a slope in the sizes of sub-Neptunes will not effect either parameter. This is true in practise as well, I find that the extracted values of α and β are consistent and independent of metallicity trends. Figure 3.4 demonstrates that the gradient extraction method, when applied to core-powered mass-loss, follows the trends in the sizes of super-Earths. If one were to incorporate metallicity trends, the sub-Neptune slope would change but this would only negligibly effect the slope of super-Earths and hence the extracted values of α and β. Note that since both models incorporate atmospheric cooling, this conclusion also likely holds for photoevaporation but is less certain due to the critical role of metals in the absorbing of high-energy radiation and in regulating cooling when in the escaping layers of the atmosphere (Owen and Murray-Clay, 2018).

3.3.5 Results for Existing Surveys

To test the methodology, I perform the analysis on three surveys which both derive from *Kepler* data. The first data-release of the CKS dataset (Petigura et al., 2017; Fulton et al., 2017; Fulton and Petigura, 2018) is a sample of 921 planets with high purity (i.e. low levels of false positive contamination with further cuts on effective temperature, impact parameter and removal of giant stars, see Fulton et al. (2017) for details). I model it's stellar mass distribution as a Gaussian function centred at $1M_{\odot}$ and width of $0.15M_{\odot}$ (see Chapter 4). The GKS survey on the other hand (Berger et al., 2020a) has a sample of 3760 planets but includes planet candidates which are (as of yet) not confirmed, even after the appropriate cuts. As a result, the level of false positive contamination may be significantly larger than that of CKS. This survey has a slightly broader stellar mass distribution, which I model as being Gaussian with centre of $1M_{\odot}$ and width of $0.2M_{\odot}$. Finally, I perform tha analyis on the second data-release of the CKS dataset, hereafter CKS-Cool, which extends the original sample to planets around smaller mass stars (Petigura et al., 2022). This sample has a very high purity level due to stringent cuts and as such contains 970 planets. The stellar mass distribution is modelled with centre of $1M_{\odot}$ and width of $0.3M_{\odot}$. I construct synthetic transit survey with photoevaporation and core-powered mass-loss models assuming the same number of planets as the individual surveys and the appropriate stellar mass distributions. Note that, as shown in Section 3.3.2, false positive contamination has little effect on determining α and β, hence I do not include them in the synthetic surveys. In addition, all surveys benefit from *Gaia* stellar mass constraints and thus I model the measurement uncertainty in incident flux, stellar mass and planet size as {8%, 5%, 5%} for both synthetic surveys. One caveat is that the completeness map from the CKS data is used for all synthetic surveys. However, as the data-sets derive from the same underlying planet population and were observed using the same telescope, this is a fair approximation to make.

Figure 3.10 shows the posteriors for the CKS (left), GKS (centre) and CKS-Cool (right) data-sets alongside the appropriate synthetic surveys. As expected, the posterior for GKS is significantly smaller than CKS, owing to its larger survey size and wider stellar mass distribution. Despite CKS-Cool's small survey size however, its posterior is similar in size to GKS, in part due to it's very wide stellar mass range and higher purity. Extending this latter style of survey to larger numbers is the ideal goal for the future, as discussed in Section 3.4.1. I measure $\alpha = 0.13^{+0.03}_{-0.05}$ and $\beta = -0.21^{+0.33}_{-0.39}$ for the CKS data, $\alpha = 0.10^{+0.03}_{-0.02}$ and $\beta = -0.03^{+0.10}_{-0.12}$ for the GKS data and $\alpha = 0.09^{+0.03}_{-0.03}$ and $\beta = -0.01^{+0.14}_{-0.16}$ with all uncertainties quoted as 2σ. All surveys are consistent with each other, and also consistent with synthetic surveys for photoevaporation and core-powered mass-loss. As a result, one can not conclude the data favours one model over the other at this stage.

3.4 Discussion

In this work I have analysed the radius gap in 3D parameter space of incident flux, stellar mass and planet size. I have shown that it is suitable for model comparisons, particularly because $\beta = (\partial \log R_{\rm val}/\partial \log M_*)_S$ is predicted to be significantly different between the two models. Comparing this to model comparisons made in the M_*-$R_{\rm p}$ plane, in which the slope of the radius valley is fraught with degeneracies of α, β as well as any scaling between core mass and stellar mass, it follows that the adopted 3D method far outperforms the 2D equivalent.

3.4.1 The Perfect Survey

From the results presented, the hypothetical perfect survey to understand which of photoevaporation or core-powered mass-loss is the driving mechanism in exoplanet evolution would have

the following traits:

- Biased towards host stellar masses $\leq 1.0 M_{\odot}$, but with a typical spread of $\geq 0.2 M_{\odot}$.
- Emphasis on reducing measurement uncertainty, with typical percentage errors $\leq 8\%$ for incident flux, $\leq 5\%$ for stellar mass and $\leq 5\%$ for planet radius.
- A false positive contamination of $\leq 10\%$.

In this scenario, core-powered mass-loss may be distinguished from a conservative model of photoevaporation 90% of the time with a survey of $\gtrsim 3000$ planets. However, given that the majority of currently observed planets are centred around $\sim 1.0 M_{\odot}$, a more realistic survey which includes these may be able to achieve this goal with $\gtrsim 5000$ planets.

For future surveys, I propose the following process. Determine α and β in a similar manner to that of Section 3.2.5 and repeat by re-sampling planet measurements within the associated error range. The priority is that the adopted method must be robust i.e. it consistently extracts α and β for any resample of the data and the values appear to be correct by-eye. Then, a hypothesis test should be performed, with the null hypothesis being that $\beta = 0.0$ (i.e. consistent with core-powered mass-loss). This can then be rejected in favour of the alternative hypothesis $\beta < 0.0$ (i.e. favouring photoevaporation) if the data suggests so, to some significance level. Note that the posteriors presented in this work from synthetic models will differ from real surveys, since in this work I am determining the uncertainty in α and β by performing $10,000$ individual surveys to understand the Poisson error. For real data however, uncertainties in α and β are determined by resampling over the individual planet measurements.

3.4.2 The Effect of Survey Completeness

All surveys, including transit surveys, have inherent bias folded into the data. In this case, the bias arises due to the geometric probability of transit and also pipeline efficiency for a given exoplanet. As a result, detections are biased to be at short orbital periods (large incident flux) and large planet radii. Although this is unavoidable, it is worth noting that this does have an effect on the values of α and β that can be extracted from data. Figure 3.11 shows posteriors in colour for photoevaporation and core-powered mass-loss for a survey of 5000 planets, measurement uncertainty in $(S, M_*, R_{\rm p})$ of $(8\%, 5\%, 5\%)$ and a stellar mass distribution of a Gaussian function centred at $1.0 M_{\odot}$ with a width of $0.3 M_{\odot}$. Recall that in order for a synthetic transit survey to be produced, the underlying planet distribution is biased with the CKS completeness map i.e. all radius gap analysis thus far has been performed in detection space (see Section 3.2.4). To demonstrate the effect of bias, I also show the posteriors in solid black and dashed contours for α and β extracted from the underlying distributions i.e. no bias and no measurement uncertainty, labelled as the occurrence. The effect is clearly seen for core-powered mass-loss: the posterior for the occurrence is centred correctly on the theoretically predicted value, whereas the posterior for the biased and noisy data is larger and shifted to larger values of α and smaller values of β. The increase in posterior size is due to measurement uncertainty (also demonstrated in Figure 3.7), whereas the shift in position arises due to survey bias. As survey incompleteness is not uniform i.e. it preferentially biases towards larger S and larger $R_{\rm p}$, it introduces biases in the slope of the radius gap. This is similar to photoevaporation (left panel of Figure 3.11) in that the posterior of the biased survey is shifted to larger values of α and smaller values of β. The difference however, is that the posterior distribution in fact gets closer to the theoretically predicted values. This is due to an unrelated but inherent limitation of the gradient extraction method that is further discussed in Section 3.4.3.

Despite the introduction of systematic errors and as further argued in Chapter 4, it is important

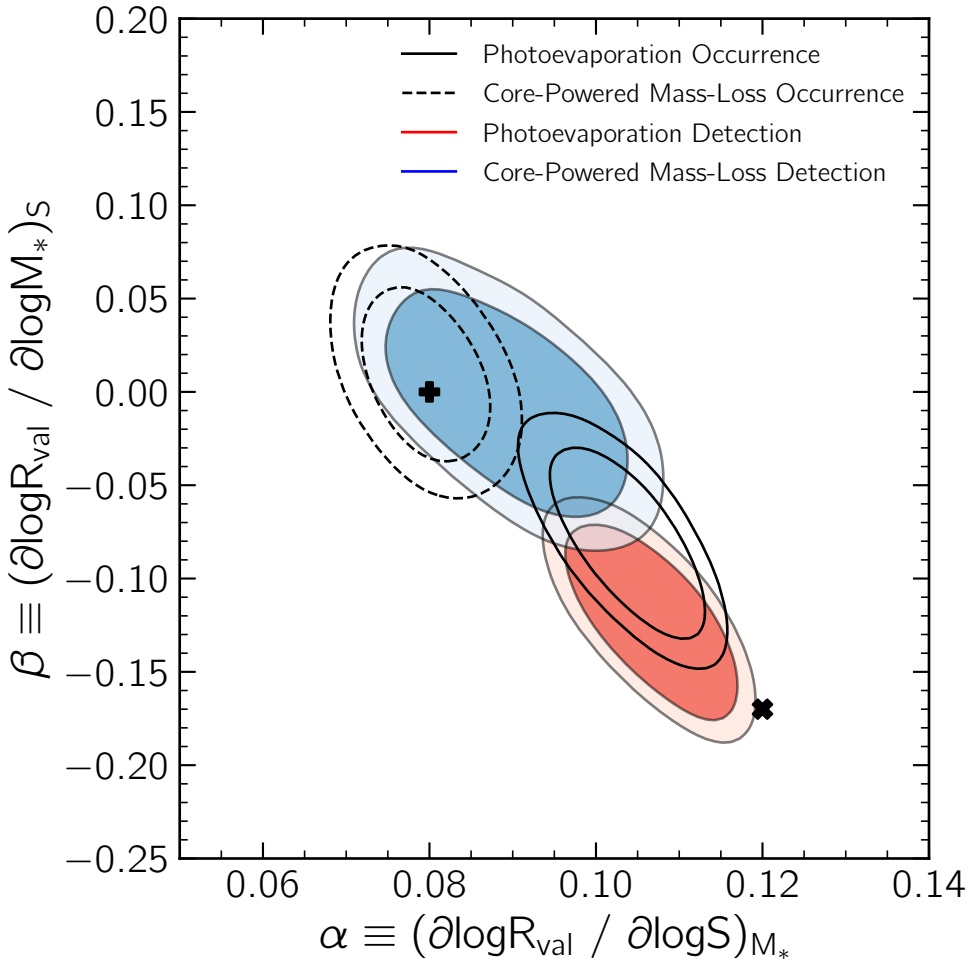


Figure 3.11: Posteriors for α and β are shown for photoevaporation and core-powered mass-loss for a survey with host stellar mass distribution centred at $1.0M_{\odot}$ with a width of $0.3M_{\odot}$. Coloured contours represent the synthetic survey measurements for photoevaporation (red) and core-powered mass-loss (blue) with observational bias, selection effects and measurement uncertainty incorporated. Black contours represent the posteriors when α and β are extracted from the underlying planet distributions provided by the models, referred to as the occurrence, with solid lines representing photoevaporation and dashed lines for core-powered mass-loss. Black markers represent the theoretical predictions for each model.

to do this analysis in detection space as opposed to performing completeness corrections and performing analysis in occurrence space. This is because planets detected with low completeness (i.e. low probability of being detected such as those with large orbital periods), are given large weights and significantly increase the inferred occurrence in that region. Instead, statistical analysis should be done in detection space, whilst predictions from the models should be calculated by forward modelling planets through a synthetic transit survey so that they can be compared. Note that this then requires completeness maps to be calculated for a given survey so that they can be used with the theoretical models.

3.4.3 Limitations of the Statistical Analysis

One of the main drawbacks of this method is its partial ineffectiveness in extracting the correct value of β from the photoevaporation models. This is discussed in Section 3.2.5 and can be seen in Figures 3.4, 3.5 and in the left side of Figure 3.11, in which the posterior distributions for photoevaporation are inconsistent with the theoretical predictions. The cause of this limitation originates from a subtle difference in the definition of the radius gap between the models and the data. In the analytic models, the valley slope is defined by locating the size of the largest super-Earths as a function of either orbital period, incident flux or stellar mass (see Sections 3.2.3 and 3.2.3). The adopted computational method of extracting these slopes however is to find a plane in the three-dimensional space of incident flux, stellar mass and planet size which intersects with the fewest number of planets and tends to bisect the two populations of super-Earths and sub-Neptunes (see Section 3.2.5). Whilst these two definitions are similar, the theoretical value is consistently steeper for photoevaporation. As a result, the value extracted is shallower in β than the predicted values. Overall, this issue stems from the fact that characterising the radius gap is an ill-posed statistical problem. Note that this also holds true for other methods of radius gap characterisation, such as *gapfit* from (Loyd et al., 2020). When performing this analysis on future surveys, other, potentially more sophisticated statistical frameworks may be set up in order to extract α and β. This new framework should, as done in this work, be implemented on modelled data as well as surveys such that observational data is not directly compared to theoretical predictions, but synthetic surveys instead. For now though, this is beyond the scope of this work.

3.4.4 Independence of Core Mass Scaling

As discussed in Section 3.2.1, analysing the gap in 3D is independent of scalings of core mass with stellar mass. This is contrary to the M_*-R_p plane, in which scalings can alter the slope of the radius gap (e.g. Wu, 2019) and further increase the degeneracy in its physical origin (i.e. Figure 3.2). Whilst in theory this claim is true when applied to the analytic models, in practise the applied method of gradient extraction suffers from small systematic shifts in measured β when scalings of core mass to stellar mass are introduced. This originates from the same issues discussed in Section 3.4.3, in which the theoretically predicted radius gap is not what is extracted from the survey data. This is therefore further evidence that refining the statistical method of radius gap detection to be closer to the theoretical definitions will improve future analysis.

3.4.5 A Route Forward

Eventually, determining which of photoevaporation or core-powered mass-loss determines exoplanet evolution will become a multi-faceted endeavour. In this Chapter, I have presented a new avenue to explore, in which the radius gap is extended to 3D space and mass-loss parameters α and β are extracted. This method should be combined with other techniques capable of differentiating between the models, such as radius gap evolution (e.g. Berger et al., 2020a; Sandoval

et al., 2020) and direct observation of atmospheric mass-loss through transit spectroscopy (e.g. Gupta and Schlichting, 2021) of Lyman-α (e.g. Vidal-Madjar et al., 2003; Ehrenreich et al., 2015), Hα (e.g. Yan and Henning, 2018), He 1083 nm (e.g. Spake et al., 2018; Ninan et al., 2020) or other metal lines (e.g. Vidal-Madjar et al., 2004).

In addition, further observational work is needed in order to constrain the X-ray evolution of stars, particularly how X-ray exposure changes with stellar mass (e.g. McDonald et al., 2019), in addition to the level of intrinsic scatter of X-ray exposure for a given stellar mass (e.g. Tu et al., 2015), as discussed in Chapter 6. This has a direct effect on the predicted value of β for photoevaporation, hence further constraints will aid in making model predictions.

Finally, further theoretical work is needed to construct a combined model of photoevaporation and core-powered mass-loss. A major assumption in this work is that the physics of the two models are mutually exclusive. However, in reality, it is likely that both mechanisms have a part to play in exoplanet evolution. Constructing a joint model which encapsulates the physics of both mechanisms would therefore be of great value and is also discussed in Chapter 6.

Once knowledge is gained as to which mechanism dominates the evolution of close-in exoplanets, one can exploit the evolutionary process to 'rewind the clock' and infer underlying demographic distributions that are hidden in standard surveys, which is the focus of Chapter 4.

3.5 Conclusions

In this work, I have presented a new method of differentiating between the atmospheric mass-loss models of EUV/X-ray photoevaporation (e.g. Owen and Wu, 2013, 2017) and core-powered mass-loss (e.g. Ginzburg et al., 2016; Gupta and Schlichting, 2019) using exoplanet demographics. This involves analysing the radius gap in demographic data in the 3D parameter space of incident flux S, host stellar mass M_* and planet size R_p. I define two new parameters, $\alpha = (\partial \log R_\mathrm{val}/\partial \log S)_{M_*}$ and $\beta = (\partial \log R_\mathrm{val}/\partial \log M_*)_S$ and a statistical method of extracting them from transit data. I have shown that performing model comparisons in this 3D space is superior to considering the gap in 2D (such as the M_*-R_p plane) as it avoids degeneracies of α, β as well as scalings of core mass with stellar mass (e.g. Wu, 2019; Gupta and Schlichting, 2020). I have performed synthetic transit surveys for both models of varying size and quality in order to determine the best route forward to determine which of the two models best describes the observed exoplanet population. The main conclusions are as follows:

- The predicted values of α are ~ 0.08 and ~ 0.12 for core-powered mass-loss and photoevaporation respectively. The predicted values of β are ~ 0.00 and ~ -0.17 respectively. This difference arises because the energy source for photoevaporation originates from the high energy flux exposure from the star. At fixed incident bolometric flux, this has been shown to decrease with host stellar mass (McDonald et al., 2019; Johnstone et al., 2021), resulting in a negative value of β. On the other hand, the energy source for the core-powered mass-loss is a combination of core luminosity and stellar bolometric luminosity, implying β is predicted to be zero, from its definition.

- I predict that transit surveys with $\gtrsim 5000$ planets will have the statistical power necessary to extract α and β such that one can begin to determine which model best describes the data. This applies to a transit survey with a stellar mass distribution centred at $1.0 M_\odot$ and standard deviation of $0.3 M_\odot$, as well as typical measurement uncertainties of $\leq 8\%$ for incident flux, $\leq 5\%$ for stellar mass and $\leq 5\%$ for planet radius.

- I find that false positive contamination of up to 10% does not effect the ability to perform

this analysis on a given transit survey, whereas reducing the measurement uncertainty has a major positive impact. Additionally, emphasis should be put on increasing the width of the host stellar mass distribution, as this provides greater statistical power to determine β. Biasing towards host stellar masses $< 1M_{\odot}$ also reduces the negative impact of the Neptune-desert on the analysis.

- This analysis was performed on three existing surveys: the CKS data-set (Fulton and Petigura, 2018), the GKS data-set (Berger et al., 2020a) and CKS-Cool data-set (Petigura et al., 2022). I find that the extracted values for α and β from these surveys are consistent with each other and also consistent with both photoevaporation and core-powered mass-loss synthetic surveys.

- Additionally, I show that performing model comparisons in the M_*-$R_{\rm p}$ plane is fraught with degeneracies between α, β and the presence of any stellar mass-core mass scalings. This motivates the analysis to be performed in 3D S-M_*-$R_{\rm p}$ space, as performed in this work, in order to break these degeneracies.

- Finally, and as argued in Chapter 4, I further stress that comparisons between data and models should be performed in detection space as opposed to completeness-corrected occurrence space, which can bias results due to planets in low-completeness regions being given large weights.

Looking forward to new surveys, the introduction of data from missions such as *K2* (e.g. Howell et al., 2014; Zink et al., 2020) and *TESS* (e.g. Ricker et al., 2015) are encouraging for this avenue of model comparison. In addition, future survey missions such as *PLATO* (Rauer et al., 2014) will provide observations of thousands of additional exoplanets. In light of these results, I would encourage future demographic missions to put emphasis on sampling a broad range of host stellar masses and utilising high accuracy planetary radii and age determination (e.g. asteroseismology) for planets near or inside the radius gap. For now however, the fundamental question still remains: what caused close-in exoplanets to lose their atmospheres? To answer this, we will require further work to be done on both observational and theoretical sides of the field. Besides the demand for larger surveys, further constraints are needed on the X-ray evolution of young stars, which will aid in the understanding of the photoevaporation models. Additionally, pursuing other model comparison avenues such as transit spectroscopy and radius gap time evolution will aid in answering this question. Finally, theoretical work is needed to further develop both evolutionary models. Both of the adopted formalisms include a series of theoretical assumptions that require further investigation. In addition it is unclear how photoevaporation and core-powered mass-loss compete and which, if any, dominates under varying conditions.

Chapter 4

Unveiling the Planet Population at Birth

4.1 Introduction

In this chapter I present a novel physics-based inference formalism that uses atmospheric evolution to infer underlying planet properties of *Kepler* planets. I show that one can rewind the clock on planet evolution to place constraints on their initial conditions and how these link to planet formation models. This method is thus a way of breaking the degeneracies in the mass-radius plane, as discussed in Chaper 1. Unlike the previous Chapter in which models were used to make predictions, here I adopt the XUV photoevaporation model and fit it to the observed data. Note that this inherently assumes that photoevaporation is dominant over its competing mechanism, namely core-powered mass-loss.

As discussed in Chapter 1, previous works of Owen and Wu (2017) and Wu (2019) have heuristically made inferences of planet properties, such as core composition, core mass and initial atmospheric mass fractions required to be consistent with the data. In this work, I construct a rigorous statistical approach that is able to make meaningful constraints on planet formation.

4.2 Method

This study invokes the photoevaporation process to infer properties of *Kepler* planets that are undetectable from standard survey techniques. Specifically, I wish to infer the following population demographics: core-mass distribution, $f(M_{\rm core})$, initial atmospheric mass fraction distribution[1], $f(X_{\rm atm}^{\rm init} \equiv M_{\rm atm}/M_{\rm core})$, and core density distribution, $f(\rho_{\rm core})$. Whilst directly computing exact synthetic populations of planets from these distributions is impossible, one can sample from the distributions and thus synthesise planet populations and compare with real data. To this end, I forward model an ensemble of planets, each with a separate core mass, $(M_{\rm core})$, initial atmospheric mass fraction $(X_{\rm atm}^{\rm init})$, and core density $(\rho_{\rm core})$ drawn from the above distributions. By then performing a synthetic transit survey on the evolved planets and comparing to real *Kepler* data, I determine how well the population distributions can reproduce the data. As with Chapter 3, one crucial philosophy I adopt in this work is that any inference of demographic properties should be done by comparing synthetic measurements with real observations, as opposed to completeness corrected data compared directly to the models. As discussed, completeness corrections are performed by assigning a weight to individual planet

[1]I.e. the atmospheric mass fraction distribution at the start of the photoevaporation process, typically after protoplanetary disc dispersal.

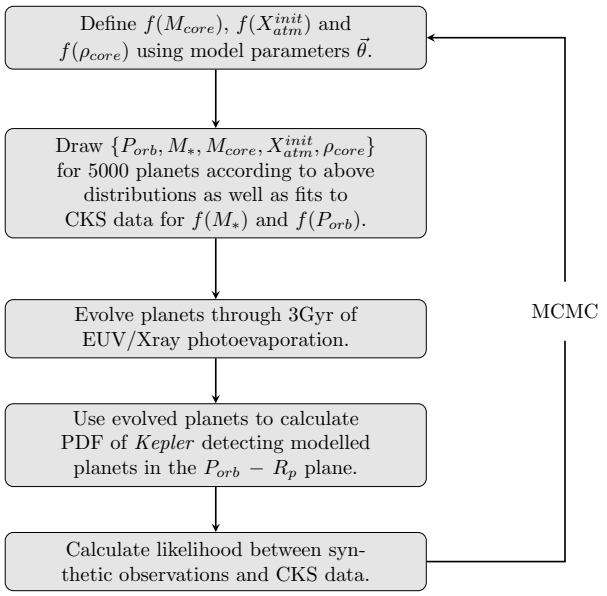


Figure 4.1: Flowchart to demonstrate the hierarchical inference process. Each MCMC iteration involves the construction of planetary distributions from model parameters, from which 5000 planets are evolved through EUV/X-ray photoevaporation and used to calculate a PDF for planet detection in the orbital period - radius plane. Finally, a likelihood is calculated between the synthetic data and the CKS data set. Each of these five-steps are described separately in Sections 4.2 - 4.2.3.

observations based on their probability of occurring. However, for transit surveys, planets at large orbital periods and small sizes can be given very large weights, which can skew the occurrence calculations. In my adopted methodology however, I accurately incorporate the noise and biases inherent in the data directly into a model and hence have a better handle on the underlying property of question. The adopted method is thus akin to Bayesian Hierarchical Model[2].

The natural choice for the data set comes from the California Kepler Survey (CKS, Petigura et al. 2017; Johnson et al. 2017b) - the sample is large, with well defined cuts and is of high purity. Additionally, the derived planetary radii are sufficiently accurate to resolve a bimodality in planet sizes between $\sim 1 - 4R_\oplus$ (Fulton et al., 2017). I discuss how additional data sets, particularly RV/TTV mass measurements could be incorporated into the model in Section 4.5. One advance over the study of Wu (2019), which only used the planetary radius distribution in their inference method, is that I choose to compare my models with data in the orbital period - radius plane, which provides far greater leverage on the data. Photoevaporation is after all a period dependent process. Figure 4.1 shows a schematic outline of the inference problem, which is split into five sections; (1) Constructing the distributions, (2) Drawing the planets, (3) Evolving the planets, (4) Observing the planets and (5) Calculating a likelihood. I will now follow the order of this outline and describe each of these individual processes in the following sections.

Constructing the Distributions

In order to forward model exoplanets in this framework, I require distributions for core mass, initial atmospheric mass fraction and bulk core density. For the latter, I choose to quantify this by adopting the approach from Owen and Morton (2016), whereby core density is interpreted

[2]Although it is technically not a Bayesian Hierarchical Model: this is because the model of the underlying exoplanet population cannot be computed explicitly, and is thus generated by random sampling (as discussed in the following sections).

by a single parameter $\tilde{\rho} \in [-1, 1]$, which tracks the linear fraction of ice, rock and iron as used in mass-radius relationships from Fortney et al. (2007). A composition of $\tilde{\rho} \leq 0$ signifies a ice-rock mixture, with $\tilde{\rho} = -1$ implying a 100% ice core, ranging to $\tilde{\rho} = 0$ implying a 100% rocky core. Similarly, $\tilde{\rho} \geq 0$ relates to a rock-iron mixture, with $\tilde{\rho} = 1$ resulting in a 100% iron core. In this parameterisation the Earth has a value of 1/3 implying a 1/3 iron, 2/3 rock mass fraction in the (Fortney et al., 2007) mass-radius relations. This parameterisation allows a single parameter to specify the core density. I emphasise that the actual quantity I am constraining is the bulk core density, yet I choose to interpret this density in terms of a composition according to a specific mass-radius relationship. In Section 4.5, I expand this parameterisation and constrain the full water/rock/iron distribution of the cores.

Choosing a functional form for these planet properties, namely core mass, initial atmospheric mass fraction and core density is a challenge as there is insufficient observational or theoretical information to make an informed choice for any distribution. In Wu (2019), the functional forms for core mass and initial atmospheric mass fraction are log-Gaussians, and all planets share the same mean core composition. In the first part of this work, I aim to perform a similar analysis to Wu (2019), but comparing with the data in both radius and period, rather than radius alone. Thus the distributions for $M_{\rm core}$ and $X_{\rm atm}^{\rm init}$ are set to be log-Gaussians:

$$\begin{aligned} \frac{{\rm d}N}{{\rm d}\log M_{\rm core}} &\propto \exp\left[-\frac{(\log M_{\rm core} - \log \mu_M)^2}{2\sigma_M^2}\right], \\ \frac{{\rm d}N}{{\rm d}\log X_{\rm atm}^{\rm init}} &\propto \exp\left[-\frac{(\log X_{\rm atm}^{\rm init} - \log \mu_X)^2}{2\sigma_X^2}\right]. \end{aligned} \tag{4.2.1}$$

The model parameters for these distributions are thus the means and standard deviations for each respective function (e.g $\{\mu_M, \sigma_M, \mu_X, \sigma_X\}$). Thus, with regards to the planet distributions, the model parameters I wish to infer are $\boldsymbol{\theta} = \{\tilde{\rho}, \mu_M, \sigma_M, \mu_X, \sigma_X\}$. I label this simulation as MODEL I.

I then go on to relax the constraint of a specific functional form for core mass and initial atmospheric mass fraction to allow these distributions to be completely arbitrary. Despite being a more challenging inference problem, this choice is necessary as one can now infer new and unexplored features in the distributions of interest, e.g. are the distributions truly peaked, are they skewed? To do so, I employ Bernstein polynomials to define non-parametric distributions for $M_{\rm core}$ and $X_{\rm atm}^{\rm init}$. Similar to other expansions, Bernstein polynomials are capable of approximating any well-behaved function and are an attractive choice as they have convenient degree reduction properties that simplify inference problems. They are formed as an n^{th} order expansion of Bernstein basis polynomials:

$$B_n(x) = \sum_{\nu=0}^{n} \beta_{\nu,n}\, b_{\nu,n}(x), \tag{4.2.2}$$

where β_ν are the Bernstein coefficients and the basis functions $b_{\nu,n}(x)$ are given by:

$$b_{\nu,n}(x) = \binom{n}{\nu} x^\nu \left(1 - x\right)^{n-\nu}. \tag{4.2.3}$$

For a given function $f(x)$ to be approximated, the Bernstein coefficients are calculated by:

$$\beta_{\nu,n} = f\left(\frac{\nu}{n}\right), \tag{4.2.4}$$

hence, combing Eqs. 4.2.2, 4.2.3, 4.2.4 leads to an approximation of $f(x)$ to n^{th} order:

$$B_n(f)(x) = \sum_{\nu=0}^{n} f\left(\frac{\nu}{n}\right) \binom{n}{\nu} x^\nu \left(1 - x\right)^{n-\nu}. \tag{4.2.5}$$

More information about Bernstein polynomials can be found in Farouki (2012) or Ning et al. (2018).

In this relaxed inference problem problem, I use Bernstein polynomials to model the initial atmospheric mass fraction distribution $f(X_{\rm init})$ and core mass distribution $f(M_{\rm core})$. For the latter, the Bernstein coefficients simply control the PDF of $\log M_{\rm core}$ in the chosen domain between $0.6M_\oplus$ and $100M_\oplus$. The lower limit was chosen such that a pure ice composition would be at the current detection limit. The upper limit on the other hand was required to be sufficiently large to avoid boundary issues with the Bernstein polynomials. Whilst the coefficients for core mass control a PDF, the coefficients for $\log X_{\rm atm}^{\rm init}$ control an inverse CDF. In other words, I construct a Bernstein polynomial function in the domain $[0, 1]$ and range $[\log X_{\rm min}, \log X_{\rm max}]$. By drawing a random number $x \in [0, 1]$, I can therefore read-off an initial atmospheric mass fraction according to the desired distribution. By adapting 4.2.5, I can approximate the desired inverse CDF $C(x)$ as:

$$C(x) \approx \left\{ \frac{y_{\rm max} - y_{\rm min}}{B(1) - B(0)} \times \sum_{\nu=0}^{n} C\left(\frac{\nu}{n}\right) \binom{n}{\nu} x^\nu \left(1 - x\right)^{n-\nu} \right\} + y_{\rm min} \tag{4.2.6}$$

where $B(0)$ and $B(1)$ are Eq. 4.2.5 evaluated at $x = 0, 1$ for the Bernstein coefficients $\beta_{\nu,n} = C(\nu/n)$ to be determined in the inference problem. The choice of PDF vs. CDF between $f(X_{\rm init})$ and $f(M_{\rm core})$ came about due to convergence issues of MCMC chains (see Section 4.2.3) with both distributions controlled by a CDF or both with a PDF. I found that the core mass and initial atmospheric mass fraction distributions were effectively constrained if they were fit with a PDF and CDF respectively.

Thus, in addition to a mean core composition $\tilde{\rho}$, I now include model parameters for the Bernstein polynomial coefficients that define the probability density functions for $\log M_{\rm core}$ and $\log X_{\rm atm}^{\rm init}$. I assume uniform priors on these coefficients in the domain $[0, 1]$, which thus leads to a log-uniform prior on the core mass and atmosphere mass fraction distributions. The relaxed inference problem thus has the following parameters to fit: N_M coefficients for the core mass PDF, N_X coefficients for initial atmospheric mass fraction PDF and one value for the mean core composition. While their is a strong theoretical prejudice to suspect the initial atmospheric mass fraction correlates with core mass (e.g. Ginzburg et al., 2016), I choose not to implement this coupling here. This correlation will be explored in future work, see Chapter 6. This simulation is labelled as MODEL II.

The third and final model is similar to MODEL II in that I use Bernstein polynomials to constrain the core mass and initial atmospheric mass fraction distribution. However, unlike MODEL II, I also attempt to fit the core composition distribution with a Gaussian function. As shown in Owen and Wu (2017), the core density of a planet controls the maximum size at which it can be stripped (at a given core mass and orbital period) and therefore strongly controls the location of the radius gap. The effect of allowing a range of core densities in a given population is to 'smear' out the radius gap and thus reduce its depth. In order to maintain the observed sparsity of planets $\sim 1.8R_\oplus$ in my model, I expect the width of this Gaussian function to be very narrow (Owen and Wu, 2017; Van Eylen et al., 2018); however it is yet to be quantified. The final simulation, including the Bernstein polynomials for core mass and initial atmospheric mass fraction, as well as a Gaussian distribution for core composition is labelled as MODEL III.

Note that, for all three models, I additionally require distributions for orbital period and host stellar mass. These are determined in separate inference problems, as described in Section 4.2.3.

4.2.1 Evolving the Planets

The evolution of the planets is determined with the analytic model for XUV photoevaporation from Owen and Wu (2017); Owen and Campos Estrada (2020) as described in Section 2.6 of Chapter 2. As with Chapter 3, I solve the ODE evolution equation 2.7.4 with an RK45 adaptive-step numerical integrator with a error tolerance of 10^{-5}, which can handle the sharp changes in X as seen in Figure 2.5. Note that I assume a planet has been completely stripped when it's atmospheric mass fraction falls below 10^{-4}, a value at which the planet's radius is indistinguishable from the core's radius in transit observations. In this model, each planet is evolved for 3 Gyr to match the typical ages of Kepler systems. Recall from Section 2.6 that the majority of mass-loss occurs during the first few 100 Myr, so changing the total evolution time by even Gyrs has little to no effect on the final size distribution of the simulated planets as seen in Figure 2.5.

Like with the Chapter 3, I introduce some changes to the model which are appropriate for this problem. Firstly, I adopt the mass-loss efficiency from Section 2.9 i.e. an approximate fit to mass-loss simulations from Owen and Jackson (2012), as used in Owen and Wu (2017); Wu (2019):

$$\eta = \eta_0 \left(\frac{v_{\rm esc}}{25 {\rm km\ s^{-1}}} \right)^{-\alpha_\eta} \tag{4.2.7}$$

where $v_{\rm esc}$ is the escape velocity of the planet, η_0 is the normalisation and α_η is the power-law index. Whilst the value for η_0 is taken to be 0.17; motivated by hydrodynamic simulations (Owen and Jackson, 2012), I choose to add α_η to the inference parameters in MODEL III. The reason for not incorporating η_0 is due to a strong degeneracy shared with core composition (Owen and Adams, 2019), which would require either perfect numerical photoevaporation models, or planet mass data (either RV or TTV) in order to break. Therefore, as described in (Owen and Adams, 2019; Mordasini, 2020), my constraints on core-composition are completely dependent on the accuracy of photoevaporation simulations; as discussed further in Section 4.5.

With regards to XUV luminosity evolution, I take a simpler approach to Chapter 3, particularly in the dependence with stellar mass which was crucial for that work. In this model, I take the $L_{\rm XUV}$ magnitude and evolution from Owen and Wu (2017):

$$L_{\rm XUV} = \begin{cases} L_{\rm sat} & \text{for } t < t_{\rm sat}, \\ L_{\rm sat} \left(\frac{t}{t_{\rm sat}} \right)^{-1-a_0} & \text{for } t \geq t_{\rm sat}, \end{cases} \tag{4.2.8}$$

where $a_0 = 0.5$, $t_{\rm sat} = 100$ Myr and the saturation luminosity follows

$$L_{\rm sat} \approx 10^{-3.5} L_\odot \left(\frac{M_*}{M_\odot} \right), \tag{4.2.9}$$

which is motivated from both observational and theoretical studies (e.g. Güdel et al., 1997; Ribas et al., 2005; Wright et al., 2011; Jackson et al., 2012; Tu et al., 2015; McDonald et al., 2019). This simpler approach is suitable for this inference problem, since the CKS data from Fulton et al. (2017); Fulton and Petigura (2018) has a narrow range in stellar mass $\sim 0.15 M_\odot$, meaning that the effects of stellar mass will be negligible. In future work however, where surveys with large ranges in stellar mass (e.g. Petigura et al., 2022) are implemented, more accurate XUV evolution models will be utilised.

4.2.2 Detecting the Planets

Once the planets have been evolved through 3 Gyr of photoevaporation, I perform a synthetic *Kepler* transit survey with the intention of then comparing it with real CKS data. As with

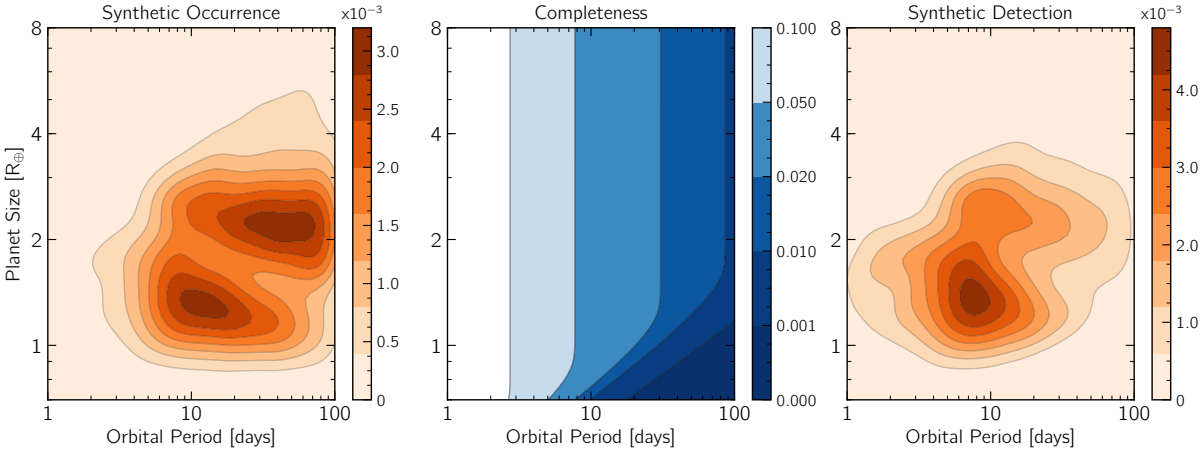


Figure 4.2: The left-hand panel shows an example PDF for synthetic occurrence of exoplanets in the orbital period - radius plane. This is produced by forward-modelling an ensemble of planets through photoevaporation, producing the bimodal distribution of super-Earths and sub-Neptunes. The central panel shows the completeness of the Kepler survey in the same domain. Contours represents the product of $p_{transit} \cdot p_{det}$, which is increased for close-in, large planets and reduced for far-out, small planets. In the right-hand panel I show the synthetic detection PDF $\lambda(P, R_p)$, which is the product of the the occurrence and the completeness and represents the planet occurrence weighted by the Kepler survey biases. One can see that planets with a large orbital period and hence a small $p_{transit}$ are heavily disfavoured to be detected.

Chapter 3, I add errors to the simulated photospheric radii of the evolved planets as well as introduce observational biases that match the real data. To quantify the radii errors of the real data, I take the typical fractional error of the CKS planets ($\sim 6\%$) and add a similar fractional error to the evolved planet radii by adding a random Gaussian perturbation with zero-mean and standard deviation of $\sigma = 0.06$. This has the effect of adding noise to the data - which shifts some planets into the radius valley, as opposed to a clean valley as seen in asteroseismic surveys (Van Eylen et al., 2018), see Petigura (2020). I choose not to add fractional errors to the orbital period measurements as these were measured with extreme precision by *Kepler*, typically $\sim 10^{-6}$. As a result, the main source of noise in the CKS data set is from the planet size measurements, as well as sampling uncertainty (Poisson shot noise).

I now introduce the biases associated with the *Kepler* survey. The most rigorous method to achieve this would be to directly inject the modelled planets into the *Kepler* pipeline and determine whether they would have been detected around the CKS stars. However, this task is computationally unfeasible for a large hierarchical model such as this. I therefore decide to take an average approach, following the prescription from Fulton et al. (2017), in which forward models of planets are biased by their associated probability of being detected. As discussed in Chapter 1, the first contribution to this derives from the geometric probability of transit:

$$p_{\rm transit} = b_{\rm cut} \frac{R_*}{a}, \tag{4.2.10}$$

where $b_{\rm cut} = 0.7$ represents the cutoff in transit impact parameter chosen in Fulton et al. (2017) to avoid grazing transits. The second contribution to the completeness is the probability of detection, resulting from the Kepler pipeline efficiency:

$$p_{\rm det} = \frac{1}{N_*} \sum_{i}^{N_*} C(m_i) \tag{4.2.11}$$

where $N_* = 36,075$ is the number of stars in the Stellar17 catalogue (Mathur et al., 2017) that pass the *Kepler* pipeline (Christiansen et al., 2012, 2015, 2016) as well as CKS filters. The

fraction of injections C recovered in the *Kepler* pipeline is a function of injected signal-to-noise m_i. As in Fulton et al. (2017), I choose the pipeline efficiency $C(m_i)$ to be a Γ cumulative distribution function of form,

$$C(m_i; k, \theta, l) = \Gamma(k) \int_0^{\frac{m_i - l}{\theta}} t^{k-1}\, e^{-t}\, \mathrm{d}t, \tag{4.2.12}$$

with $k = 17.56$, $l = 1.00$ and $\theta = -0.49$. The product of $p_{\mathrm{transit}} \cdot p_{\mathrm{det}}$ is calculated to produce a mean completeness map for the entire CKS sample (similar to Figure 6 in Fulton et al., 2017) that captures the completeness probability of the survey for a given P_{orb} and R_{p}. This is shown in the centre panel of Figure 4.2.

In order to introduce the completeness map to the evolved and noisy (i.e. including R_{p} errors) synthetic planet data, I adopt the following process: first I perform a 2D Gaussian Kernel Density Estimation (KDE) on the synthetic planet population in P_{orb}-R_{p} space. Crucially, the bandwidth of the KDE generator $\sigma_{\mathrm{KDE}} = 0.05$ is chosen to be less than typical planet radii uncertainty, as this ensures the behaviour of the function is set by the planet distribution and not numerical effects from the construction of the KDE (discussed in Section 4.2.3, see Figure 4.5). This KDE can be thought to be an approximation to the PDF for planet occurrence. I then bias this PDF with the *Kepler* completeness map and normalise to unity, which results in a PDF for planet detections, which I label $\lambda(P_{\mathrm{orb}}, R_{\mathrm{p}})$. In other words, this PDF represents the probability of detecting a planet at a given orbital period and planet radius, given a set of model parameters. An example of synthetic occurrence and synthetic detection PDFs are shown in Figure 4.2. This PDF is then used as the basis for the likelihood calculation.

4.2.3 Likelihood Calculation

The final step in the inference problem is to compare the real CKS data to the noisy, incomplete synthetic data. As used in previous works (e.g. Foreman-Mackey et al., 2014; Hsu et al., 2018; Bryson et al., 2020), I assume that the detection of an exoplanet is an independent process. The appropriate statistical framework of this inference problem is based on that of an *inhomogeneous point Poisson process*. I therefore provide an overview, before discussing how they are implemented. In a typical *homogeneous* point Poisson process i.e. random and independent events, the Poisson rate λ is assumed to be constant. As such, the probability of n events occurring is then:

$$P(n) = e^{-\lambda} \frac{\lambda^n}{n!}. \tag{4.2.13}$$

An inhomogeneous process is one in which λ is not constant. In this case, I implement an inhomogeneous process to describe the detection of a planet at a given orbital period and planet radius, given a set of model parameters. As such, I calculate the total occurrence factor Λ:

$$\Lambda = N_* \cdot f_* \cdot \iint \lambda(P, R_{\mathrm{p}})\, \mathrm{d}P\, \mathrm{d}R_{\mathrm{p}}, \tag{4.2.14}$$

where $N_* = 36,075$ is again the number of stars in the filtered Stellar17 catalogue (Mathur et al., 2017) and f_* is the mean number of planets per star in the *Kepler* field. I allow f_* to vary as a model parameter in all simulations in the aim of inferring its value. Whilst in my model, f_* is a nuisance parameter, other works have derived this value in far more detail (e.g. Howard et al., 2012; Fulton et al., 2017; Zhu et al., 2018; Zink et al., 2019; Hsu et al., 2019). The probability of detecting n planets is:

$$P(n) = e^{-\Lambda}\, \frac{\Lambda^n}{n!}, \tag{4.2.15}$$

whereas the probability density of observing a planet at $(P_i, R_{\mathrm{p},i})$ is:

$$\rho\left(P_i, R_{\mathrm{p},i}\right) = \frac{\lambda(P_i, R_{\mathrm{p},i})}{\Lambda}. \tag{4.2.16}$$

Assuming that all observations are independent[3], the probability of detecting an ensemble of planets at $\pi = \{(P_1, R_{\mathrm{p},1}), ..., (P_n, R_{\mathrm{p},n})\}$ is:

$$P(\pi) = \prod_i^n \frac{\lambda(P_i, R_{\mathrm{p},i})}{\Lambda}. \tag{4.2.17}$$

The likelihood of observing this sample is thus:

$$L(\pi) = P(n) \cdot P(\pi) = e^{-\Lambda} \frac{\Lambda^n}{n!} \prod_i^n \frac{\lambda(P_i, R_{\mathrm{p},i})}{\Lambda}. \tag{4.2.18}$$

In this hierarchical inference model, I wish to sample a posterior with the assumption that the ensemble of planets π is the CKS sample, whilst the detection rate $\lambda(P, R_{\mathrm{p}})$ is calculated using a forward modelled sample of planets drawn from distributions defined by the model parameters $\boldsymbol{\theta}$. As a result, the final likelihood is:

$$\begin{aligned} L(\pi_{\mathrm{CKS}}) &= P(N_{\mathrm{CKS}}) \cdot P(\pi_{\mathrm{CKS}}) \\ &= e^{-\Lambda} \frac{\Lambda^{N_{\mathrm{CKS}}}}{N_{\mathrm{CKS}}!} \prod_i^{N_{\mathrm{CKS}}} \frac{\lambda(P_i, R_{\mathrm{p},i})}{\Lambda}. \end{aligned} \tag{4.2.19}$$

where N_{CKS} is the number of planets in the CKS sample and π_{CKS} is the set of values of $(P_{\mathrm{orb}}, R_{\mathrm{p}})$ for all CKS planets within the chosen domain. Similarly to Fulton et al. (2017), I restrict the orbital period to lie between 1 and 100 days, whilst the planet size at 3 Gyr can vary between $0.95R_\oplus$ and $6R_\oplus$. The lower bound here is chosen to avoid areas of low-completeness which are almost devoid of planets[4], whilst the upper bound is chosen to avoid modelling large planets hosting self-gravitating atmospheres which are not accounted for in the evolution model and may form in a different way (e.g. Owen and Lai, 2018). The likelihood function (Equation 4.2.19) can be interpreted in two parts: the Poisson pre-factor is maximised when the number of synthetically observed planets from the model matches the number observed in the CKS data set. Therefore this factor allows us to put realistic uncertainties on the distribution of interest due to the finite sampling of the exoplanet data. The product factor on the other hand is maximised when the shape of the modelled planet distribution in P_{orb}-R_{p} space matches that of the CKS planets. One therefore finds a high likelihood when the number of observed planets and shape of the planet distribution match the CKS data.

In order to sample from the posterior, the log-likelihood (i.e. $\ln L(\pi_{\mathrm{CKS}})$) is fed into a Monte Carlo Markov chain (MCMC) algorithm, specifically the EMCEE Python implementation (Foreman-Mackey et al., 2013) of the affine-invariant ensemble sampler from Goodman and Weare (2010). I run with 150 walkers and a sufficient number of iterations required for chain convergence, which varies for each of the models.

The end result of this MCMC inference problem are the constrained core mass, initial atmospheric mass fraction and composition distributions. However, as discussed in Section 4.2, I also require distributions for the host stellar mass M_* and orbital period P_{orb} in order to model the

[3]In this calculation, I am thus assuming single and multiple transiting systems are implicitly drawn from the same formation model. As a result this ignores the detection bias that occurs in multi systems (Zink et al., 2019).

[4]More specifically, this lower limit of $0.95R_\oplus$ is chosen as a compromise between inferring the core mass distribution to smaller masses whilst still allowing smaller cores to have an icy composition and still be detected.

atmospheric evolution for populations of planets. Whilst, strictly speaking, these distributions should be inferred as part of the full hierarchical model, I choose not to include them for a variety of reasons. Firstly, stellar mass measurements are essentially decoupled as they are directly measured as part of the CKS program. One can therefore fit this distribution independently without affecting the other distributions. The orbital period distribution on the other hand is coupled, albeit weakly, with the other model parameters, meaning it is more important to include in the full inference problem. However, when this was attempted, the MCMC chain became stuck in local minima, resulting in uninformative posteriors and computationally infeasible run-times. As this removed the possibility of fitting the period distribution in the full model, I choose to fit it independently, drawing upon multiple previous works that achieve the same task (e.g. Fressin et al., 2013; Howard et al., 2012; Petigura et al., 2018). The added benefit of not including the fit of either stellar mass of orbital period distributions is that is it reduces the number of model parameters and hence simplifies the inference problem. As previously stated, I choose to infer exoplanet demographic properties by comparing synthetic measurements with real observations. To this end, I fit the stellar mass and orbital period distribution by forward modelling the detections of such quantities, given an underlying population. For the orbital period, motivated by previous works (e.g. Fressin et al., 2013; Howard et al., 2012; Dressing and Charbonneau, 2015; Petigura et al., 2018), I choose the underlying distribution to be a smooth broken power law:

$$\frac{\mathrm{d}N}{\mathrm{d}P} \propto \frac{1}{\left(\frac{P}{P_0}\right)^{-k_1} + \left(\frac{P}{P_0}\right)^{-k_2}}, \tag{4.2.20}$$

Once normalised as a PDF, I bias this function with a 1D integral of the completeness map of the CKS data set (i.e. integrate middle panel of Figure 4.2 along the R-direction). In this way, I calculate the probability of detecting a planet for a given orbital period, which I label $\lambda(P)$. As with the likelihood for the main inference problem (Equation 4.2.19) I construct a likelihood from an inhomogeneous point Poisson process, now in 1D,

$$\begin{aligned} L(\pi_{\mathrm{CKS}}) &= P(N_{\mathrm{CKS}}) \cdot P(\pi_{\mathrm{CKS}}) \\ &= e^{-\Lambda} \frac{\Lambda^{N_{\mathrm{CKS}}}}{N_{\mathrm{CKS}}!} \prod_i^{N_{\mathrm{CKS}}} \frac{\lambda(P_i)}{\Lambda}, \end{aligned} \tag{4.2.21}$$

where,

$$\Lambda = N_* \cdot f_* \cdot \int \lambda(P)\, \mathrm{d}P, \tag{4.2.22}$$

is the total number of planets expected to be detected and π_{CKS} is the population of CKS planets. This likelihood is given to the affine invariant MCMC of Goodman and Weare (2010) from the EMCEE package (Foreman-Mackey et al., 2013) with 500 walkers in order to infer P_0, k_1 and k_2 from Equation 4.2.20 where I find the power law break at $P_0 = 5.75$ days, with $k_1 = 2.31$ and $k_2 = -0.08$, as shown in Figure 4.3. Whilst completeness effects are included in this fitting, I choose not to add noise to the modelled orbital periods as the fractional error from *Kepler* is typically $\sim 10^{-6}$ and hence negligible.

Similar to the orbital period distribution, I fit the stellar mass distribution of *Kepler* planets using an inhomogeneous point Poisson process. I do not, however, quantify the selection biases and instead choose to simply infer the underlying distribution as Gaussian function:

$$\frac{\mathrm{d}N}{\mathrm{d}M_*} \propto \exp \frac{-(M_* - \mu_{M_*})^2}{2\sigma_{M_*}^2}. \tag{4.2.23}$$

where μ_{M_*} and σ_{M_*} are the mean and standard deviation. In order to forward model the stellar mass errors, I take the typical fractional uncertainties from the CKS catalogue (typically $\sim$ 4%) and add them to a population of stellar masses drawn from a Gaussian function. I then

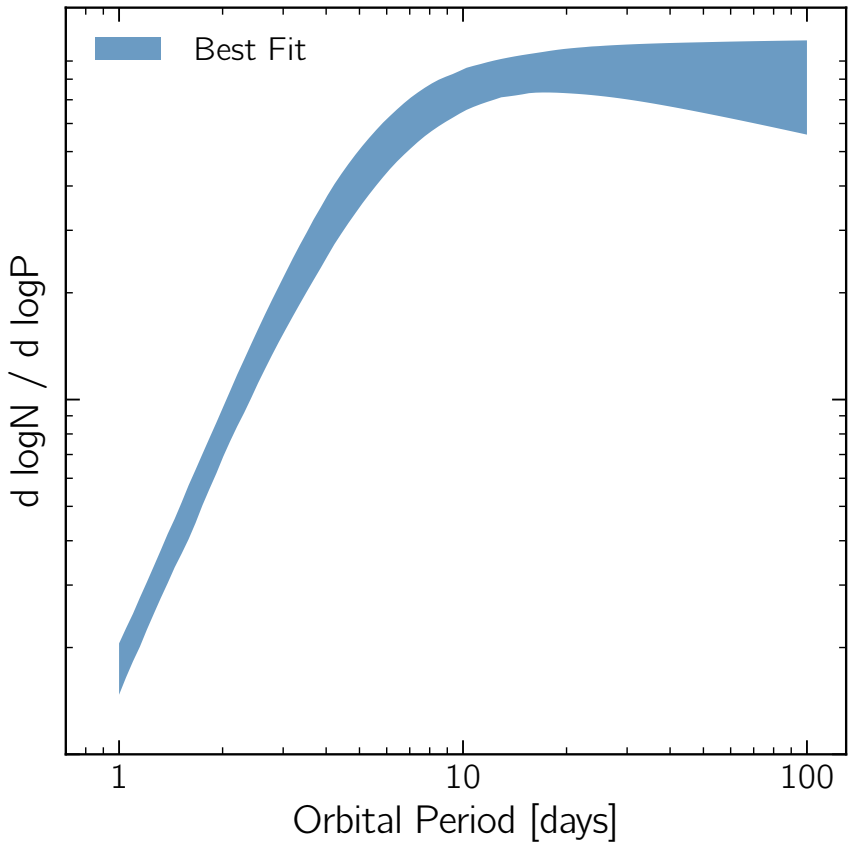


Figure 4.3: Best fit orbital period distribution when fit to the CKS planets, parameterised by a smooth broken power law (Equation 4.2.20). Shaded region shows 1σ *uncertainties.*

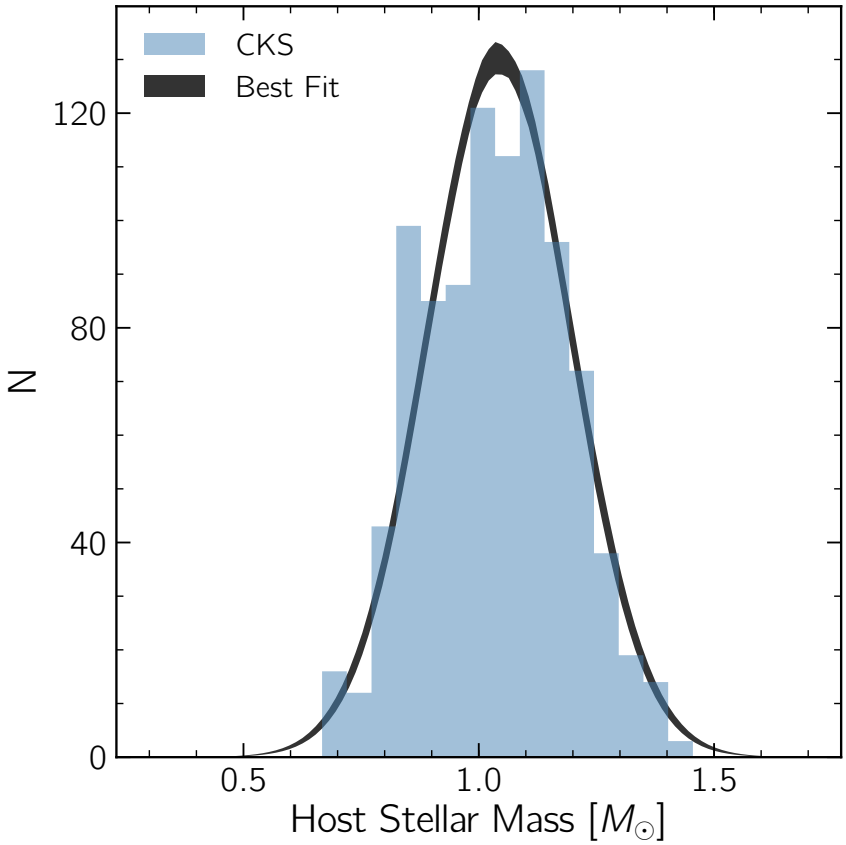


Figure 4.4: A histogram of CKS host stellar masses is plotted with best fit Gaussian distribution on top. This has mean $\mu_{M_*} = 1.04M_\odot$ and standard deviation $\sigma_{M_*} = 0.15M_\odot$. The width of the best-fit line represents 1σ uncertainties.

calculate a PDF from this new distribution of noisy stellar masses $\lambda(M_*)$ and, as with Equation 4.2.21, measure a likelihood by taking the value of this PDF for each of the CKS stellar mass measurements. I provide this likelihood function to EMCEE (Foreman-Mackey et al., 2013) in order to infer the mean and standard deviation of the underlying stellar mass distribution from the CKS catalogue. I find that $\mu_{M_*} = 1.04M_\odot$ and $\sigma_{M_*} = 0.15M_\odot$, which is shown in Figure 4.4. Now, including these fits into the planet evolution models allows one to evolve each planet with a core mass, initial atmospheric mass fraction, core density, host stellar mass and orbital period drawn from their respective distributions.

Likelihood Convergence

In order to ensure the planet occurrence PDF $\lambda(P, R_\mathrm{p})$, and hence likelihood is accurate and converged, I require a suitably large number of planets to be simulated for each iteration of the MCMC. This is so that the planet radii errors that I add in the synthetic detection process are integrated over and the PDF for planet occurrence/detections are controlled by the planet distribution shape, and not subject to sampling noise. On the other hand however, simulating a large number of planets (albeit semi-analytically) can become a computationally expensive process. I therefore wish to find a compromise. Figure 4.5 shows convergence tests for MODEL II. This is done by calculating the percentage difference of the likelihood from the 'true' value for a range of simulated planets. I approximate the true likelihood by evaluating the likelihood (Equation 4.2.19) with $100,000$ planets and thus achieving an extremely high accuracy. The error in the log-likelihood changes as a power-law with increasing number of planets indicating convergence. I choose to use 5000 planets as this provides an error of $\lesssim 0.3\%$ and provides a

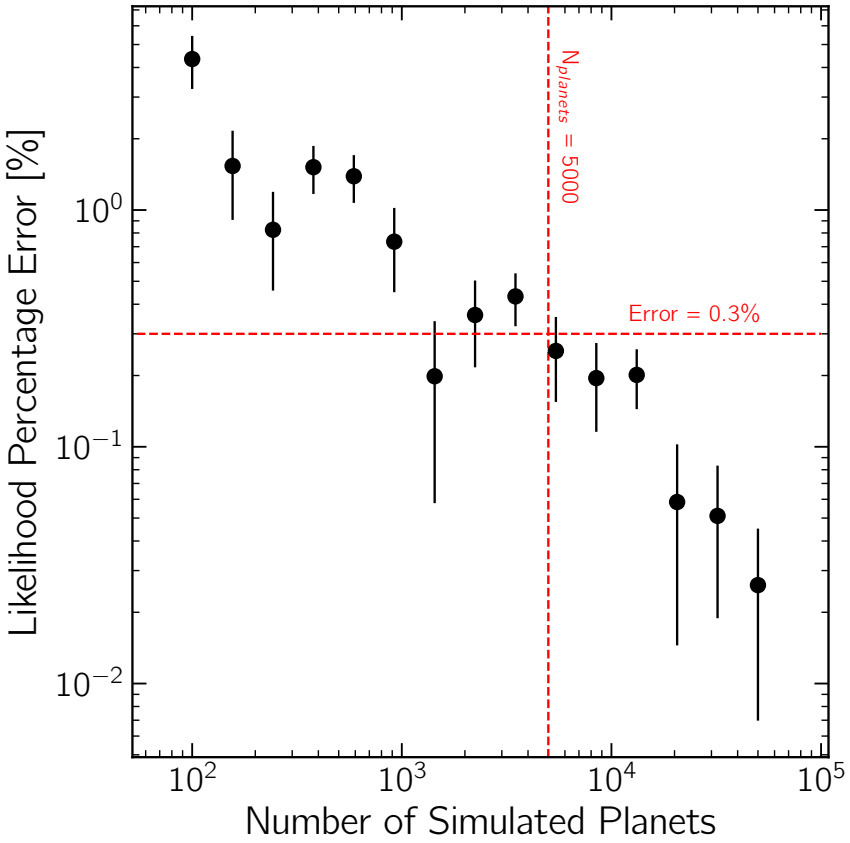


Figure 4.5: Percentage error in log-likelihood is plotted as a function of number of planets simulated for MODEL II. *The stochastic nature of the model introduces noise into error calculations, which is quantified by repeating each simulation 100 times and determining the standard deviation for each value of simulated planets. Dotted lines represents the adopted compromise between likelihood accuracy and computational speed, with a percentage error of* 0.3%*, achieved by simulating 5000 planets per MCMC iteration.*

balance between accuracy and computational efficiency.

4.3 Results

4.3.1 Model I

MODEL I is designed to test the effect of moving from the work Wu (2019) to the inference models presented in this work. I choose to adopt the same parametric forms for core mass, initial atmospheric mass fraction and core composition as in Wu (2019), with my new and improved inference methodology. I place flat priors on all parameters, with the exception of the log-Gaussian mean-values, which have log-flat priors. Values compared with that of Wu (2019) are shown in Table 4.1. Note that whilst the adopted methodologies are similar, there are distinct differences, hence one would expect to see difference in the results. I have therefore included this comparison to demonstrate how different results are found with different methodologies. One of the main differences which likely has the largest effect is that the model infers a 2D planet distribution, as opposed to a 1D radius histogram, which likely widens the constrained distributions. In addition, I adopt a more accurate completeness model and hence allow a larger fraction of smaller planets to be detected - which results in a lower mean core mass. Finally, in Wu (2019), core masses were scaled with stellar mass (see Chapter 3) which will also act to narrow constrained distributions (further discussed in Section 3.4). This comparison indicates that performing this inference in the 2D planet distribution is important.

Table 4.1: Comparing constrained parameters from MODEL I *and Wu (2019), in which core mass and initial atmospheric mass fraction are restricted to be log-Gaussian. Here, μ_M and σ_M are mean and standard deviation for core mass distribution, whilst μ_X and σ_X are mean and standard deviation for initial atmospheric mass fraction distribution (Equation 4.2.1). Quoted errors for* MODEL I *are the 1σ percentiles calculated from marginalised posteriors.*

Parameter	MODEL I	Wu (2019)
$\mu_{\rm M}$ $[M_\oplus]$	$3.72^{+0.45}_{-0.33}$	$7.70^{+1.50}_{-1.50}$
$\sigma_{\rm M}$	$0.44^{+0.06}_{-0.03}$	$0.29^{+0.06}_{-0.06}$
μ_X	$0.040^{+0.015}_{-0.016}$	$0.026^{+0.006}_{-0.006}$
σ_X	$0.51^{+0.20}_{-0.12}$	$0.29^{+0.06}_{-0.06}$

4.3.2 Model II

MODEL II relaxes the condition of a specific functional form for the core masses and initial atmospheric mass fractions. For both distributions, I use a $5^{\rm th}$-order Bernstein polynomial expansion with model parameters being the 5 polynomial coefficients. As with MODEL I, I also constrain a single value for core composition, and the mean number of planets per star. The left hand panel of Figure 4.7 demonstrates a good agreement of model and data by comparing radius histograms. This is produced by resampling 1000 sets of model parameters from the MCMC chain and evolving 5000 planets for each sample. By then calculating a PDF for planet detection (as laid out in Section 4.2.2), one can determine the observed radius distribution and thus 1σ errors across the resampled models. MODEL II can be seen to reproduce the bimodal distribution with peaks and gap in the correct positions.

The best fit core mass and initial atmospheric mass fraction distributions, produced from their associated Bernstein polynomial are shown in the middle and right-hand panel of Figure 4.7. Similar to Wu (2019), I find a peaked core mass distribution with most planets hosting a core of a few earth masses. Crucially, one can see this distribution is not symmetric, unlike a log-Gaussian, with a positive skewness of 0.04 ± 0.01 and kurtosis of $0.15^{+0.02}_{-0.03}$. This result thus confirms the requirement of relaxing assumptions from MODEL I. Similarly, and starkly different to the latter, the initial atmospheric mass fraction distribution is asymmetric and strongly peaked at $\sim 10^{-2}$, implying that the majority of Kepler planets were born with a significant H/He atmosphere of $\sim 1\%$ mass, instead of being born with a negligible atmosphere. The skewness and kurtosis of this distribution are $-0.16^{+0.12}_{-0.09}$ and $1.45^{+0.21}_{-0.13}$ respectively.

In addition to core mass and initial atmospheric mass fraction, I also constrain the mean number of planets per star to be $f_* = 0.72 \pm 0.03$ and the core composition to be $\tilde{\rho} = 0.18 \pm 0.09$. This composition equates to a bulk density for a $1M_\oplus$ core of $\rho_{M_\oplus} = 4.74^{+0.40}_{-0.36}$ g cm^{-3}, pertaining to a slightly lower density to that of Earth.

4.3.3 Model III

MODEL III is the most ambitious of the inference problems. I retain the fitting of core mass and initial atmospheric mass fraction distributions with $5^{\rm th}$ order Bernstein polynomials, but additionally choose to infer the index of photoevaporation efficiency scaling α_η (Equation 4.2.7) assuming a uniform prior, as well as the core composition distribution with a Gaussian function (also assuming uniform prior on mean and standard deviation). I show the detection PDF $\lambda(P, R)$ for best fit parameters alongside the real CKS planets in Figure 4.6, which demonstrates that the adopted likelihood is effective in capturing the shape of the planet distribution. As before, the best fit radius distribution for MODEL III is shown in the left hand panel of Figure 4.7.

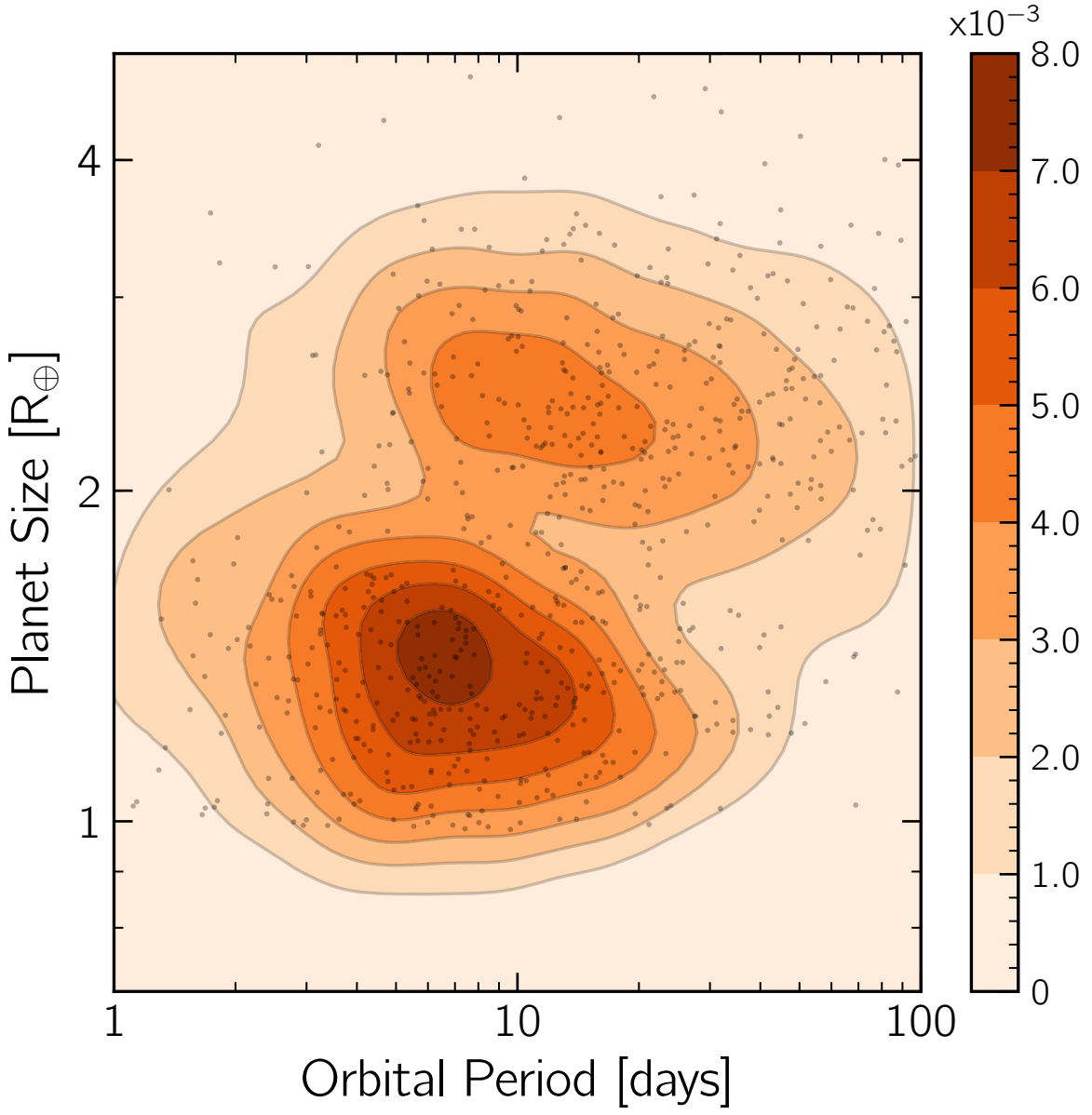


Figure 4.6: Synthetic detection PDF is shown in orange contours for best fit parameters from MODEL III. *CKS planets from Fulton et al. (2017) are also shown to demonstrate the goodness of fit of model and data.*

Parameter	MODEL I	MODEL II	MODEL III
Mean Number of Planets per Star (f_*)	$0.70^{+0.04}_{-0.03}$	$0.72^{+0.03}_{-0.03}$	$0.70^{+0.03}_{-0.03}$
Composition Mean	$0.27^{+0.08}_{-0.08}$	$0.18^{+0.09}_{-0.09}$	$0.26^{+0.08}_{-0.09}$
Composition Standard Deviation	-	-	< 0.16

Table 4.2: Comparison of constrained parameters from MODELS I, II *and* III *for core composition and mean number of planets per star in the domain of* $1 \leq P \leq 100$ *days and* $0.95 \leq R_p \leq 6$ $R_\oplus$. *Note that* MODEL III *is the only model which attempts to fit composition with a Gaussian function and hence constrains a value for composition distribution standard deviation.*

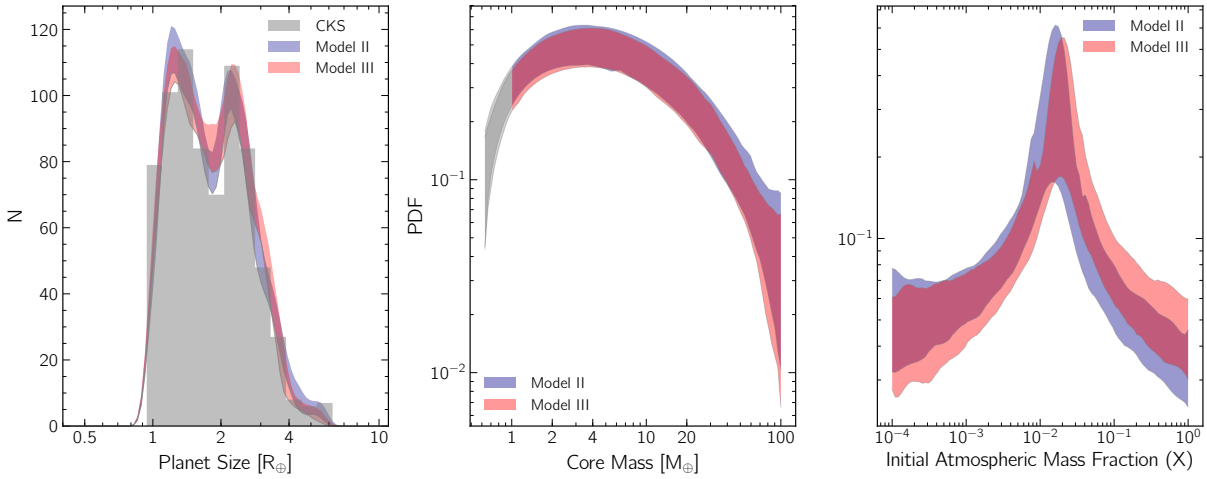


Figure 4.7: Left hand panel shows radius histogram comparison of CKS data with best fit from MODEL II *and* MODEL III. *The shaded line represents* 1σ *errors of* 1000 *forward-models resampled from the MCMC chains, whilst grey histogram bins are the CKS data from Fulton et al. (2017). Middle and right panels show constrained core mass and initial atmospheric mass fraction distribution for* MODEL II *and* MODEL III, *both of which are described by* 5^{th}*-order Bernstein polynomial expansions. Shaded regions are representative of a* 1σ *confidence interval, calculated from 1000 resamples from the MCMC chains. Grey region in core mass distribution represents masses* $< 1M_{\oplus}$ *and hence under the current detection limit.*

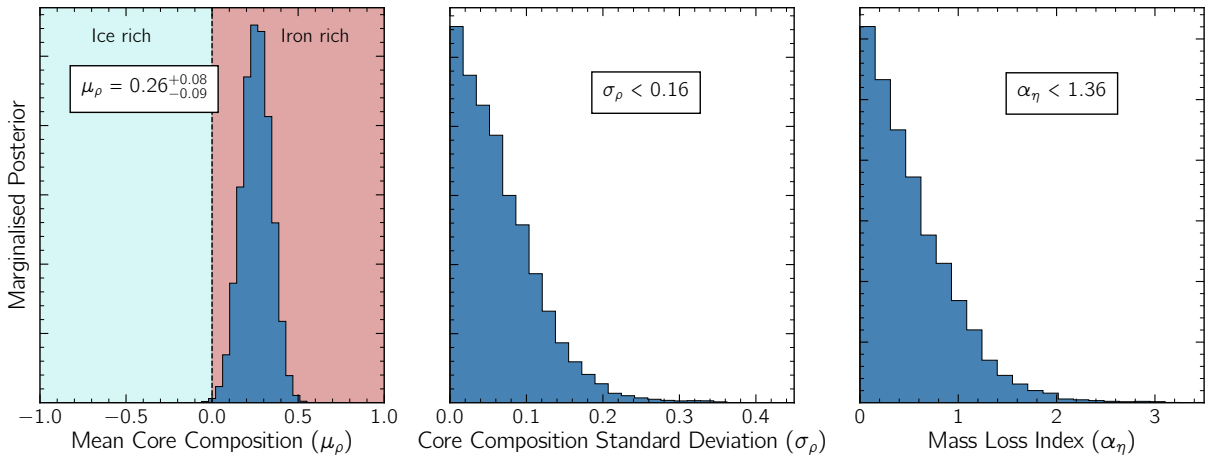


Figure 4.8: Marginalised posteriors from MODEL III. *Left-hand panel shows the mean* μ_ρ *for the Gaussian function used to fit the core composition distribution, with* 1σ *uncertainties quoted. The area* $\mu_\rho < 0$ *represents the region of parameter space corresponding to ice rich cores, whilst* $\mu_\rho > 0$ *corresponds to iron rich cores. Middle panel shows posterior for Gaussian composition distribution standard deviation, whilst right-hand panel shows posterior for photoevaporation mass-loss index from Equation 4.2.7. Note that due to uniform priors on core composition and mass-loss index, I am only able to place* 2σ *upper-bounds on quantities of interest.*

Similar to MODEL II, I get excellent agreement of model and data. As well as a mean number of planets per star of $f_* = 0.70 \pm 0.03$, the Bernstein coefficients (and hence distribution shape) for core mass and initial atmospheric mass fraction are consistent to 1σ between MODEL II and MODEL III. In particular, I find the skewness and kurtosis of the core mass distribution to be $0.04^{+0.02}_{-0.01}$ and 0.14 ± 0.02 respectively, and $-0.23^{+0.13}_{-0.08}$ and $1.52^{+0.30}_{-0.18}$ for initial atmospheric mass fraction, consistent within 1σ of those derived in MODEL II.

Figure 4.8 shows marginalised posteriors for three model parameters of interest. The left and middle panel show constraints I place on the mean and standard deviation of the composition distribution, under the assumption that it follows a Gaussian function. Consistent with MODEL I and II, I find a mean composition of $\tilde{\rho} = 0.26^{+0.08}_{-0.09}$, which is iron-rich/water-poor and consistent with that of Earth (which has a value ~ 0.33). As predicted in Section 4.2 and shown in Table 4.2, I also find an extremely narrow standard deviation of the core composition distribution with 2σ upper-limit of $\sigma_{\tilde{\rho}} < 0.16$, which is evidently consistent with zero. This points towards a singular formation pathway for small, close-in exoplanets, which is discussed in Section 4.5. Finally, as shown in the right-hand panel of Figure 4.8, I place a 2σ upper limit on the index for the photoevaporative mass-loss efficiency scaling $\alpha_\eta < 1.36$, which is consistent with the findings of Wu (2019).

4.4 Convergence and full posteriors

Convergence of MCMC chains is notoriously difficult to prove (e.g. Christian and George, 2010). At best one can demonstrate that there is no evidence that MCMC chains are not converged. I find justifying convergence based on the integrated auto-correlation time (e.g. Goodman and Weare, 2010) to be computational infeasible given the size of the inference problem.

Therefore, first I justify convergence by ensuring the auto-correlation tends to zero as a function of lag. I show the full posteriors and auto-correlation against lag plots in Figures 4.9 & 4.10 for MODEL I, 4.11 & 4.12 for MODEL II, and 4.13 & 4.14 for MODEL III. These Figures indicate that the posteriors appear well sampled and the auto-correlation against lag tends to zero.

In addition I perform the Gelman-Rubin diagnostic (Gelman and Rubin, 1992) on MODEL III and find a value of 1.01, which is sufficiently close to unity. This is computed by running two independent MCMCs with different initial positions for the *emcee* walkers (Foreman-Mackey et al., 2013). Furthermore, these two independent chains give the same results for the model parameters. Thus the tests with the Gelman-Rubin diagnostics and auto-correlation trends provide confidence that the MCMC chains are converged.

Finally I note that a copy of MODEL III was performed with 8^{th}-order Bernstein polynomials to investigate whether higher order approximations to the distributions of interest could capture more information from the data. I performed similar tests to justify convergence and, upon inspection, found the there was no benefit of increasing the order of the polynomials. I can therefore conclude that the results of MODEL III are independent of polynomial order and that the Bayesian hierarchical model is currently limited by the data.

4.5 Discussion

The analysis provides an illuminating view of the exoplanet population at birth. I have used the photoevaporation model to undo the billions of years of evolution experienced by the observed exoplanet population. While photoevaporation has evidently sculpted the exoplanet population and hidden its initial conditions, they have not been destroyed. In fact I have been able to leverage the photoevaporation model to uncover the "birth" properties of the close-in, low-mass exoplanet population. According to my results as shown in Figures 4.7 and 4.8, the population

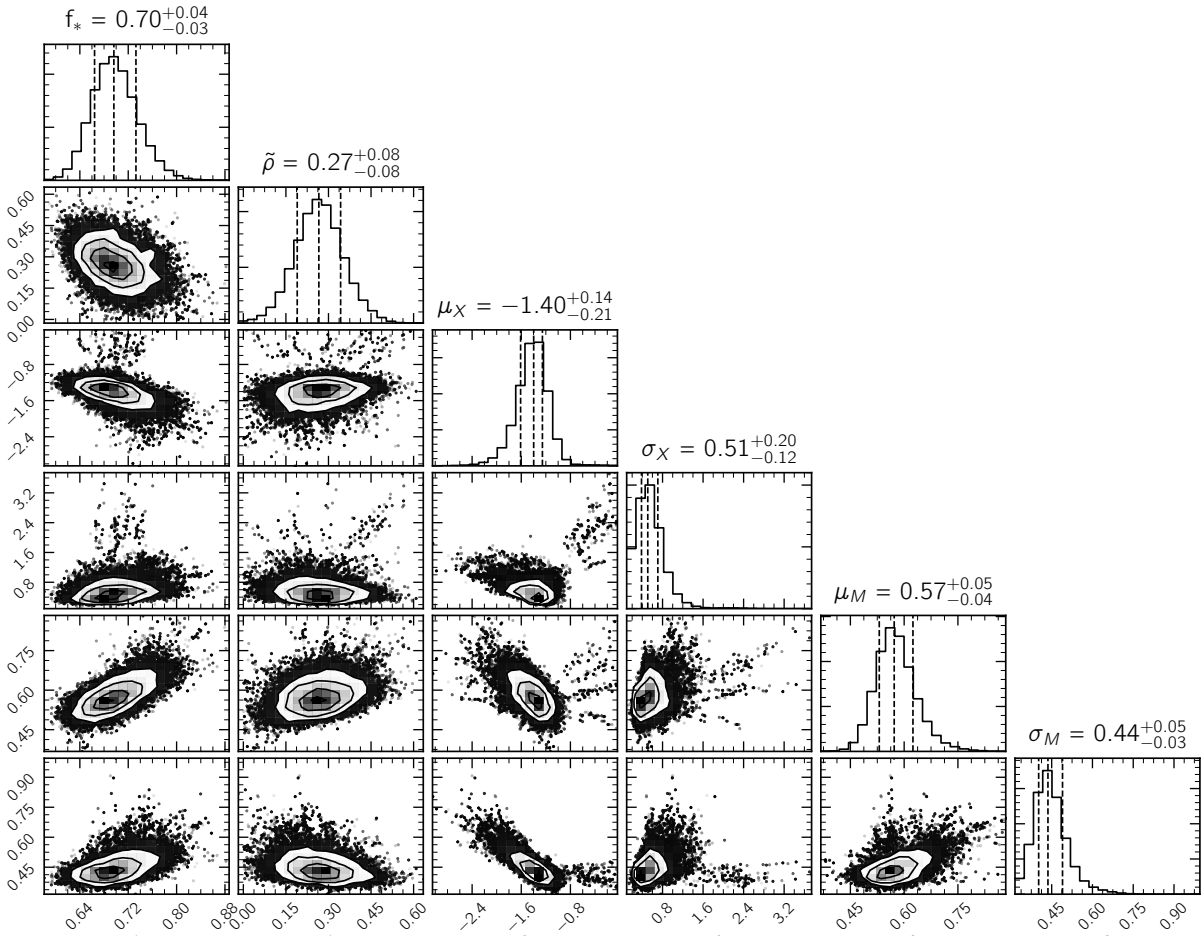


Figure 4.9: Corner plot for MODEL I. *Constrained parameters are the mean number of planets per star f_*, core composition $\tilde{\rho}$, mean and standard deviation for the initial atmospheric mass fraction fraction distribution $\{\log \mu_X, \sigma_X\}$ and the mean and standard deviation for the core mass distribution $\{\log \mu_M, \sigma_M\}$. Quoted uncertainties are representative of 1σ from the MCMC chain.*

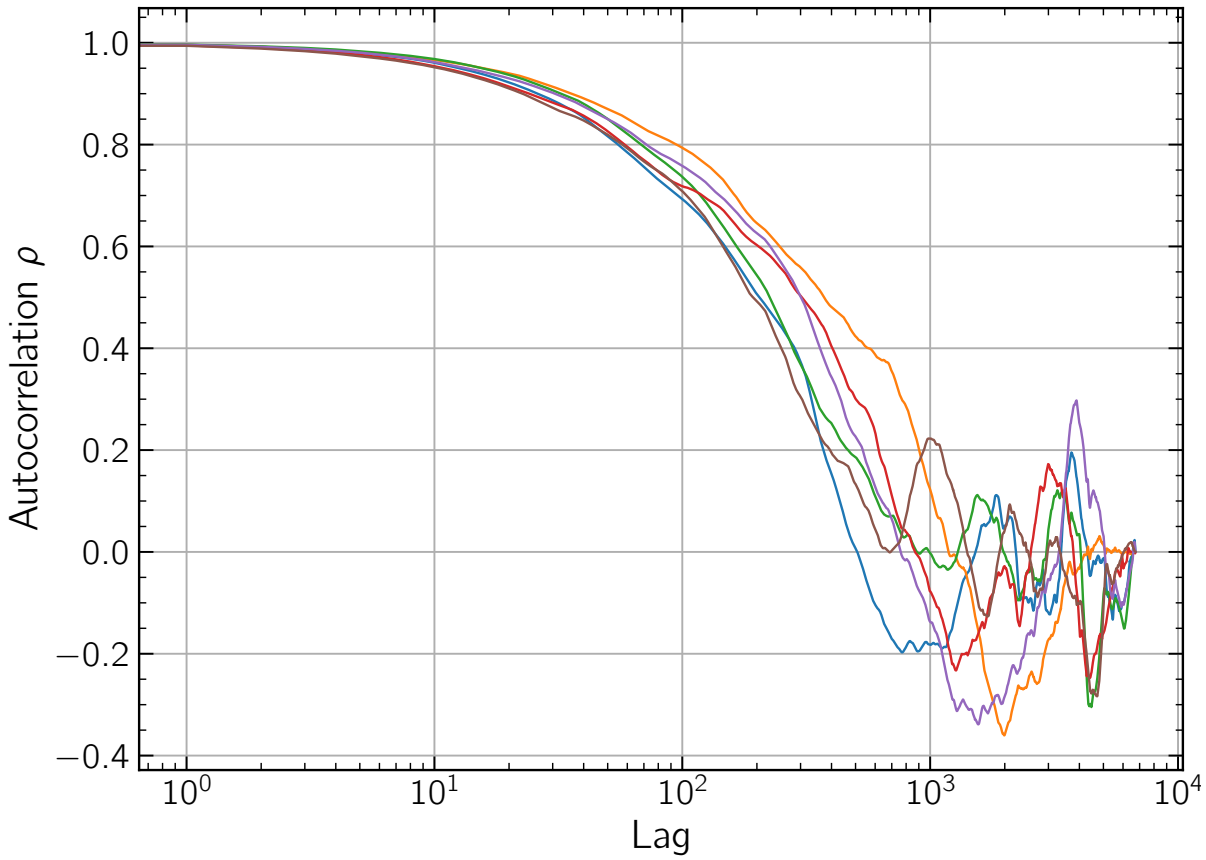


Figure 4.10: Autocorrelation ρ as a function of lag for MODEL I. *Coloured lines represent the different model parameters being inferred. One can see convergence (i.e. $\rho \sim 0$) at $\sim 10^3$ iterations.*

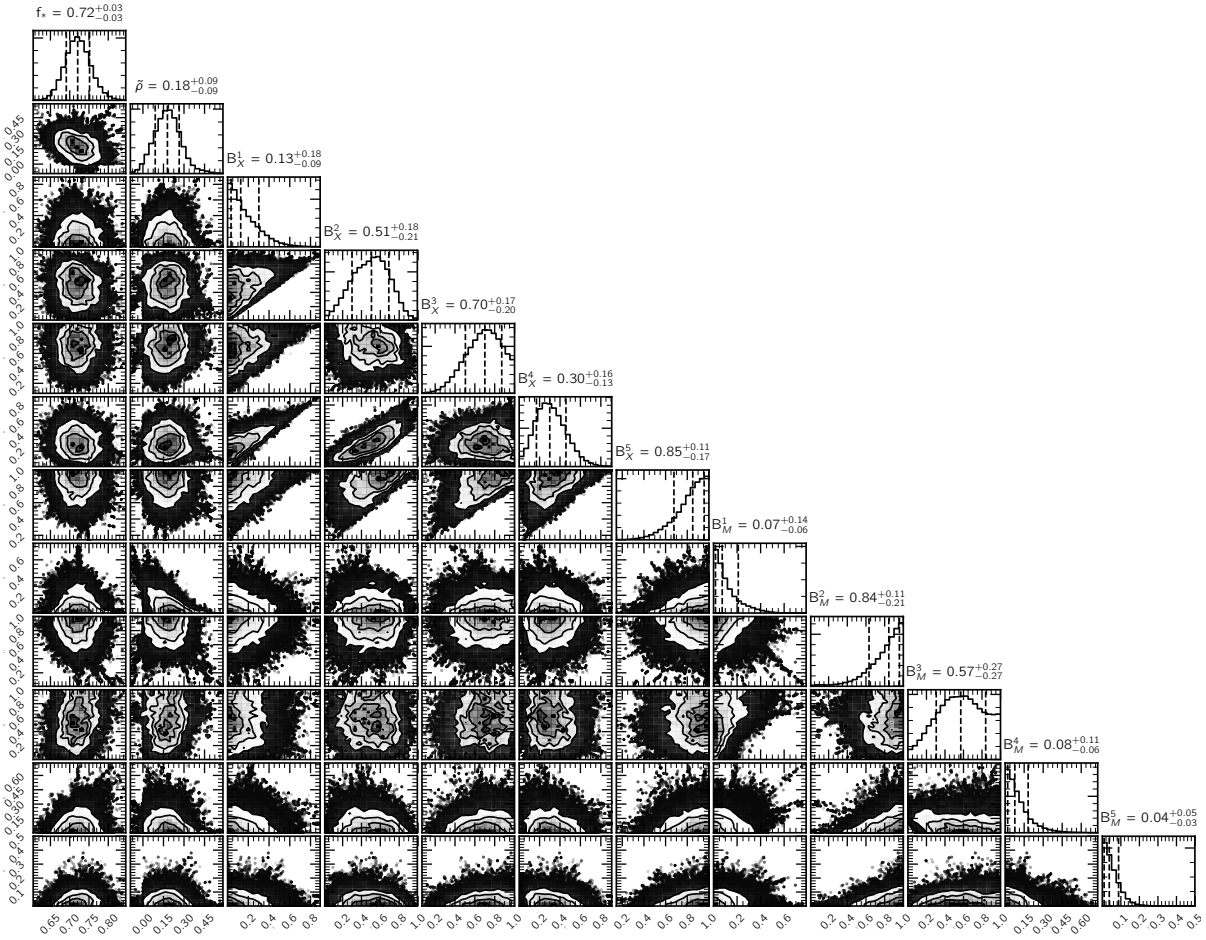


Figure 4.11: Corner plot for MODEL II. *Constrained parameters are the mean number of planets per star f_*, core composition $\tilde{\rho}$, coefficients B^i_X and B^i_M for the 5^{th}-order Bernstein polynomials describing the initial atmospheric mass fraction distribution and core mass distribution respectively. Quoted uncertainties are representative of 1σ from the MCMC chain.*

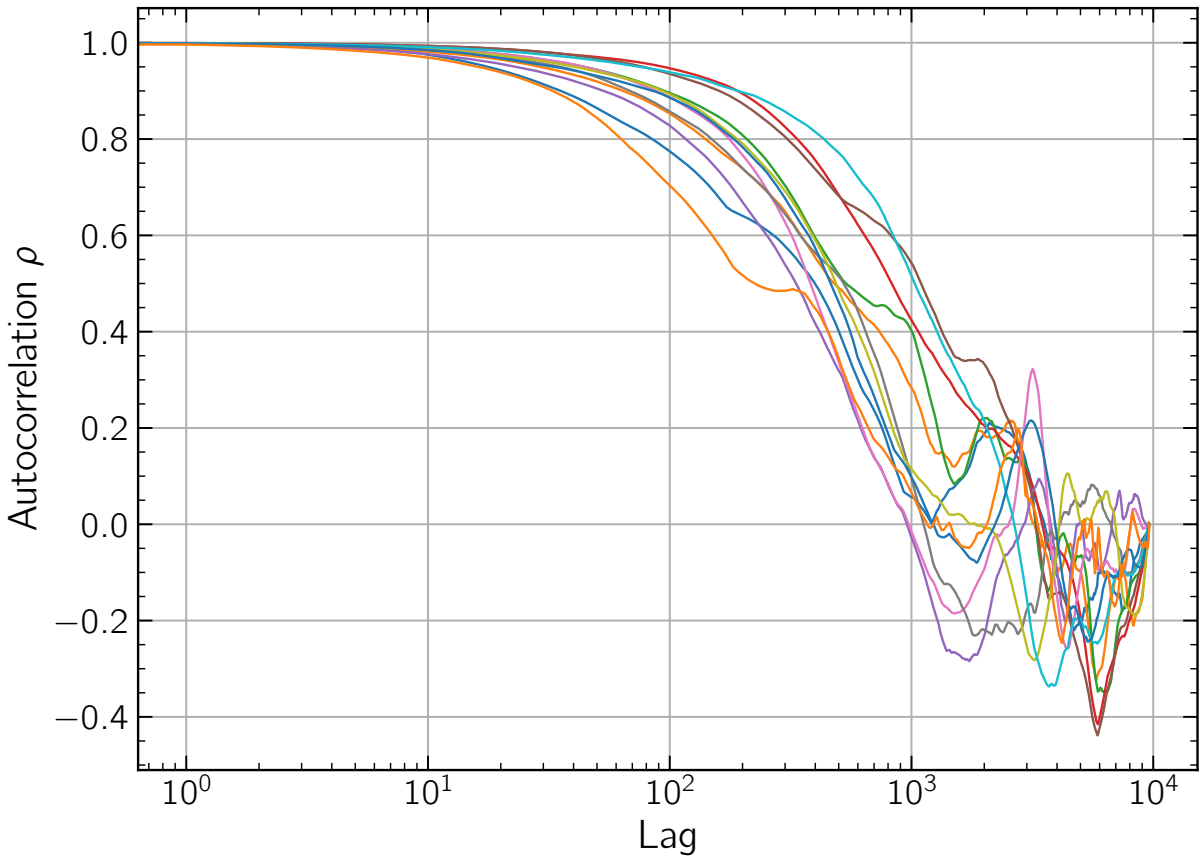


Figure 4.12: Autocorrelation ρ as a function of lag for MODEL II. *Coloured lines represent the different model parameters being inferred. One can see convergence (i.e. $\rho \sim 0$) at $\sim 3 \times 10^3$ iterations.*

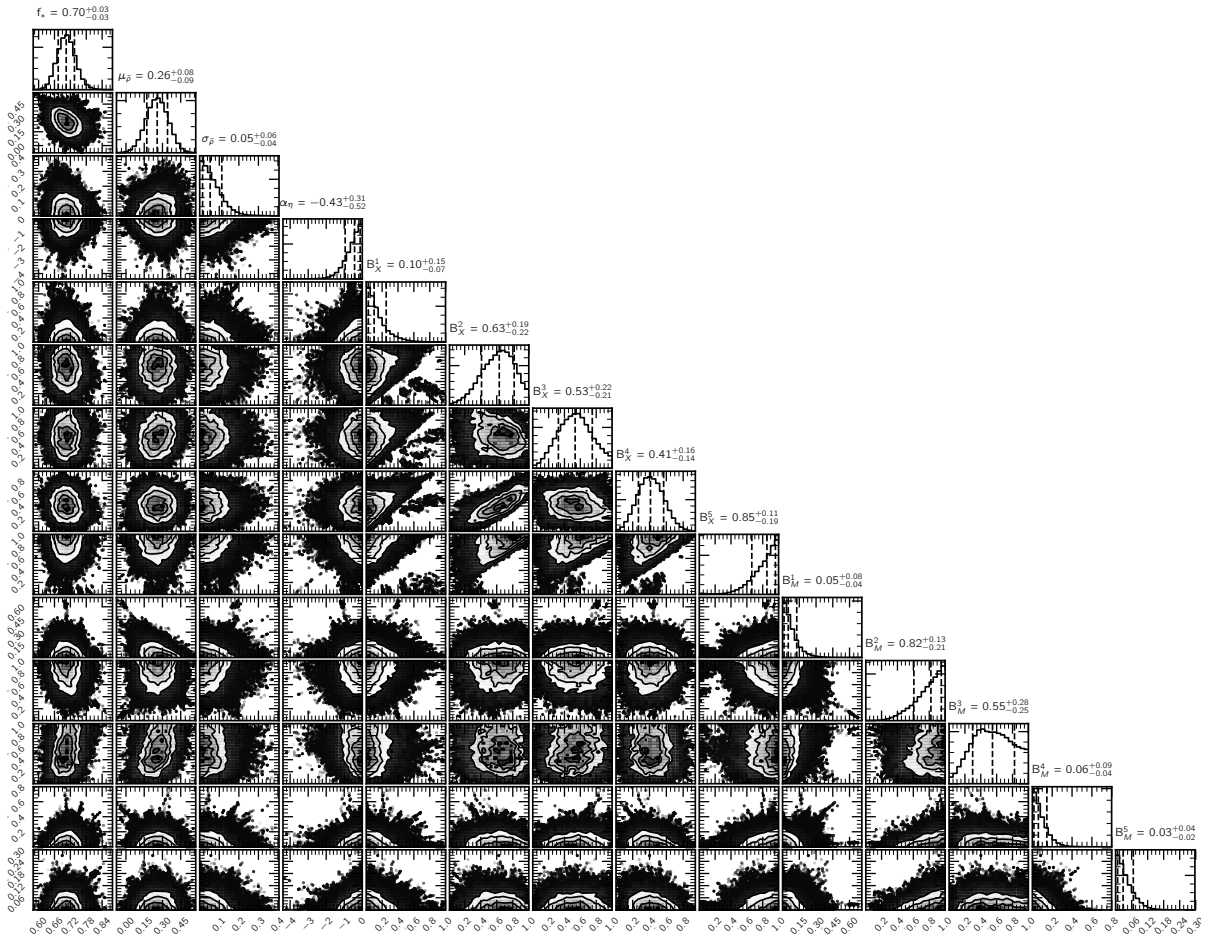


Figure 4.13: Corner plot for MODEL III. *Constrained parameters are the mean number of planets per star f_*, core composition mean $\mu_{\tilde{\rho}}$ and standard deviation $\sigma_{\tilde{\rho}}$, the photoevaporative mass-loss efficiency scaling α_η and the coefficients B^i_X and B^i_M for the 5^{th}-order Bernstein polynomials describing the initial atmospheric mass fraction distribution and core mass distribution respectively. Quoted uncertainties are representative of 1σ from the MCMC chain.*

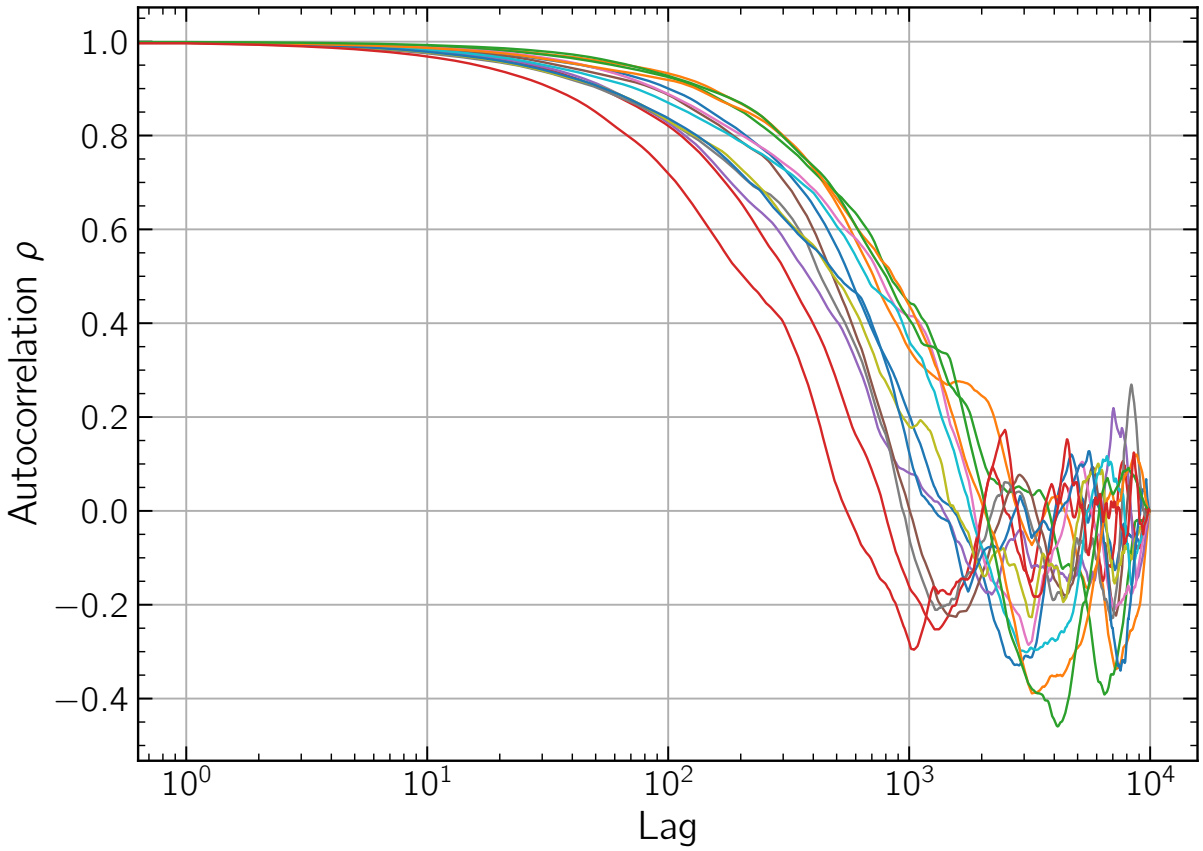


Figure 4.14: Autocorrelation ρ as a function of lag for MODEL III. *Coloured lines represent the different model parameters being inferred. One can see convergence (i.e. $\rho \sim 0$) at $\sim 4 \times 10^3$ iterations.*

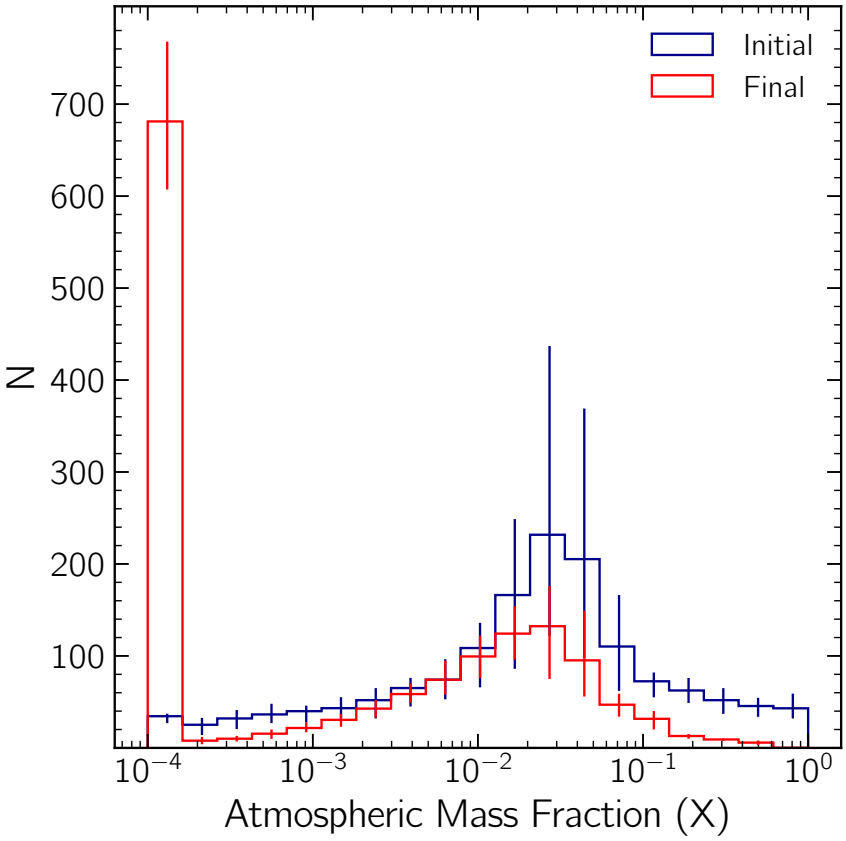


Figure 4.15: Evolution of atmospheric mass fractions from MODEL III. *Initial mass fractions shown in blue come from the constrained distribution shown in Figure 4.7, whilst final mass fractions have been evolved through 3 Gyrs of photoevaporation. Error bars come from resampling from MCMC chain. The final distribution shows a large spike at $X = 10^{-4}$, which corresponds to planets which have been completely stripped of their H/He atmospheres.*

of planets typically have cores $\sim 4M_{\oplus}$ and began with an atmospheric mass fraction of $\sim 1\%$. The core compositions are decidedly terrestrial, with a ratio of silicates to iron consistent with Earth's, and inconsistent with a significant amount of water. I now discuss the implication of such findings.

4.5.1 Atmospheric Mass Fraction Distribution

The inferred atmospheric mass fraction distribution is strongly peaked at $\sim 4\%$. Unlike the previous results in Owen and Wu (2017) who only put weak constraints on the initial mass fraction, I am able to constrain a precise distribution. This is due to the fact that photoevaporation becomes ineffective at longer orbital periods. Hence, the larger observed sub-Neptunes at longer periods in the CKS sample provide upper limits on the atmospheric mass fraction distribution. This result is further evidence that fitting the planet population in the period-radius plane (as opposed to the 1D radius distribution) is crucial in order to reveal more information of the underlying demographics.

Furthermore, the atmospheric mass fraction distribution points heavily towards the fact that the majority of Kepler planets were not "born terrestrial" i.e. with negligible atmosphere. Figure 4.15 demonstrates the evolution of atmospheric mass fractions according to MODEL III, with initial and final populations represented in blue and red respectively. As found in Wolfgang and Lopez (2015), the final distribution shows a peak at a few percent, as this corresponds to the population of sub-Neptunes, i.e. those with large extended H/He atmospheres that double

their core radius. On the other hand, the large spike of evolved planets with an atmospheric mass fraction of $X = 10^{-4}$ indicates that these planets have been stripped of their natal H/He atmosphere. As a result, these bare rocky cores form the large population of super-Earths that we observe today. In order to quantify this, I estimate the ratio between number of planets that were born rocky, and number of planets that evolved to become so. By integrating the initial and final distribution below $X = 10^{-3}$, I find that ~ 4 times as many planets evolved to become super-Earths as a result of photoevaporation, rather than be born with negligible atmosphere. Note that these results do indicate that at least some of the super-Earth's formed without a significant H/He atmosphere; however it is clearly a sub-dominant mode. One therefore gains insight into the origin of super-Earths and sub-Neptunes and can state that they both predominantly evolved from the same initial population. These results are consistent with Neil and Rogers (2020); Neil et al. (2022), in which three sub-populations of small, close-in exoplanets are identified: sub-Neptunes, super-Earths that were photoevaporated and super-Earths that were born terrestrial. Although each of the sub-populations was identified with a separate mass function, as opposed to my method which used a common function for all cores, both inference problems find a best fit to the data arises with when all three populations are present. Owen and Murray-Clay (2018) found evidence that the population of born terrestrial planets was common around lower-mass or lower-metallicity stars and speculated that the transition between accreting or not accreting a large H/He atmosphere was related to the supply of solids to the inner disc. It is worth emphasising that core accretion models (e.g. Lee and Chiang, 2015b) indicated a 1 $M_\oplus$ core will accrete a 1% H/He atmosphere by mass under typical nebula conditions. Therefore, while we see these born terrestrial planets with radii and masses above Earth's today, they must have been significantly less massive at the time of gas disc dispersal and continued accretion after disc dispersal, akin to the standard model for terrestrial planet formation in the Solar System.

4.5.2 Core Mass Distribution

The core mass distribution, shown in Figure 4.7, indicates a strong peak at $\sim 4M_\oplus$, additionally confirming previous works which required a peaked distribution to fit data (e.g. Owen and Wu, 2017; Wu, 2019). This feature implies that a common core mass and thus formation pathway is favoured in planetary systems. In addition, the steeper slope beyond $\sim 20M_\oplus$ indicates that very few cores are produced beyond this range. This result confirms the predictions that cores growing to this size will begin to host self-gravitating atmospheres and undergo runaway gas accretion (Pollack et al., 1996). As these planets grow towards Jovian masses, they reach the "thermal mass" and may carve gaps in their protoplanetary discs, perhaps slowing accretion. This idea was put forward by Wu (2019) as to the origin of the peak, which was strengthened by the fact they found this peak mass became smaller at lower-stellar mass. This trend is something that can be investigated within this inference framework, potentially with new surveys with wider stellar mass distributions such as Petigura et al. (2022). Nevertheless, I do require a non-zero fraction of cores $\gtrsim 20$ $M_\oplus$ to explain the data. This appears consistent with recent *NGTS* and *TESS* detections that have discovered a number of 30-40 $M_\oplus$ planets residing in the desert (West et al., 2019; Jenkins, 2019; Armstrong et al., 2020). While these planets are certainly consistent with photoevaporation, how they formed and why they did not accrete a massive H/He envelope remains a mystery. One recent suggestion from Ali-Dib et al. (2022) is that late-time collisions of giant planets could be sufficient to remove the atmospheres of such planets.

One benefit of the adopted EUV/X-ray photoevaporation evolution model is that it is capable of providing core mass, final atmospheric mass and photospheric radius for an ensemble of planets. Thus, one can produce a synthetic mass-radius diagram as shown in the left-hand panel of Figure 4.17. Here I have added 6% fractional errors to planet radii and 20% errors to planet masses.

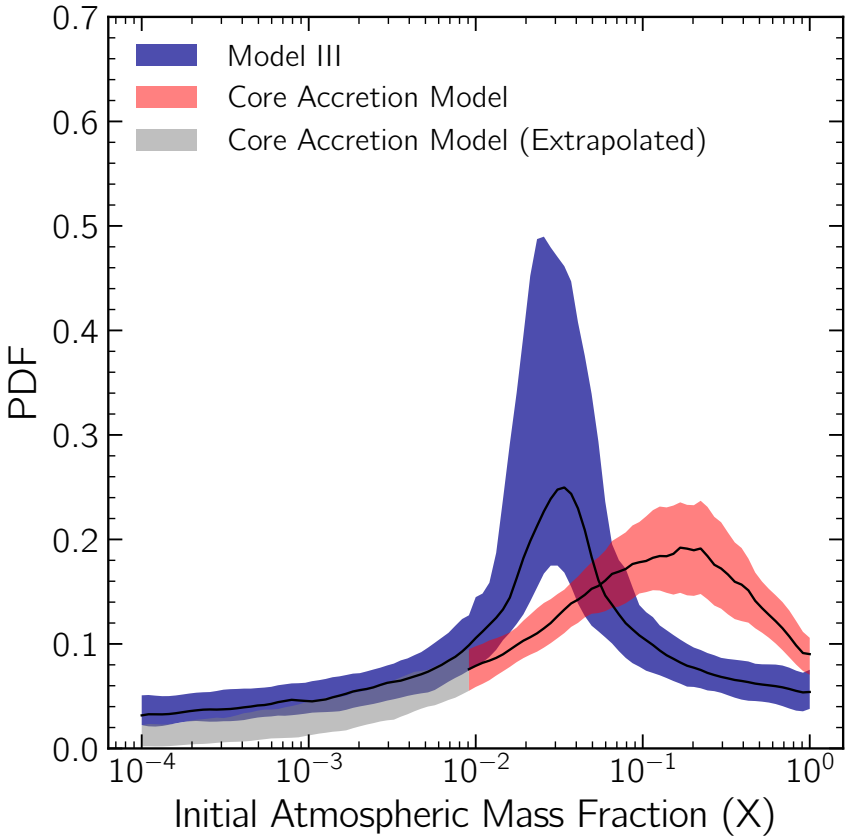


Figure 4.16: Comparison of inferred initial atmospheric mass fraction from MODEL III *(shown in blue) with core accretion model from Lee and Chiang (2015b), adapted in Jankovic et al. (2019) (Equation 4.5.1). Red region shows the predicted atmospheric mass accreted, assuming the core mass distribution found in* MODEL III. *Grey region represents accreted atmospheric mass fractions calculated by extrapolating the constrained core mass distribution to smaller cores masses ($< 0.6 M_{\oplus}$).*

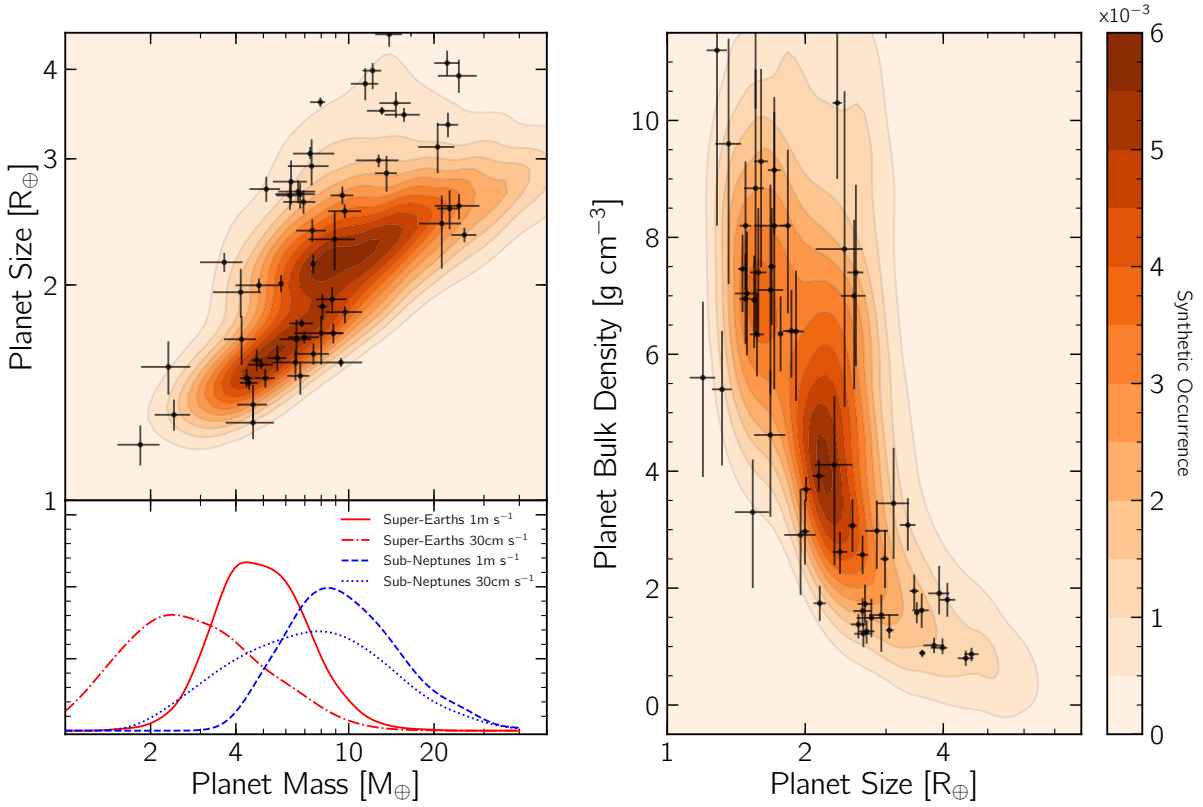


Figure 4.17: Left hand panel shows mass-radius plot for forward modelled planets from MODEL III *for which RV semi-amplitude is* $> 1ms^{-1}$*. Whilst I have added* 6% *fractional errors to planet radii and* 20% *errors to planet masses, I have not included survey biases similar to those in the main inference problem. In addition, planet mass PDFs are also provided for super-Earths and Sub-Neptunes, defined by* $R_p < 1.8R_\oplus$ *and* $R_p > 1.8R_\oplus$ *respectively with minimum RV semi-amplitude of* $1ms^{-1}$ *(solid and dashed respectively) as well as* $30cms^{-1}$ *(dash-dotted and dotted respectively). Black points represent confirmed planets taken from the NASA Exoplanet Archive, for which mass and radius measurement uncertainty is* $< 20\%$*. In addition I do not include biases of surveys, hence this should be interpreted as an underlying mass-radius distribution.*

In addition I have limited planets to have an RV semi-amplitude of $> 1\text{ms}^{-1}$, which represents current detection limits. Note however, that I have not included survey biases (as performed in the inference model) as these are currently not quantifiable for mass measurements. As a result, the mass-radius plot I provide should be interpreted as an underlying distribution with added noise. As observed in multiple works (e.g. Weiss and Marcy, 2014; Hadden and Lithwick, 2014; Dressing et al., 2015; Jontof-Hutter et al., 2016; Dorn et al., 2019), one can see a large population on a single composition line (i.e. super-Earths with terrestrial cores) and another population at a larger radius for a similar mass. This latter group corresponds to the sub-Neptune population, harbouring extended H/He atmospheres that approximately double their observed radius. As a result, the mass distribution, as observed with RV or TTV surveys, for super-Earths and sub-Neptunes will be different. This is demonstrated in the left-hand panel of Figure 4.17, in which the relative occurrence for super-Earths ($R_p < 1.8R_\oplus$) and sub-Neptunes ($R_p < 1.8R_\oplus$) is separated and plotted as PDFs, with sub-Neptunes typically occurring at larger masses. I also plot the same PDFs for a minimum RV semi-amplitude of 30cms^{-1} which represents the precision of next-generation RV instruments. One sees that by improving this minimum value allows smaller planets to be detected and thus both distributions extend further to lower mass. In addition I also take planets from the *NASA Exoplanet Archive*[5] with mass and radius measurement uncertainty $< 20\%$ and plot them on top of the expected occurrence. In the right-hand panel of Figure 4.17, I show the bulk density for the same population of modelled planets as a function of planet size. A clear bimodality is observed, with super-Earths centred at $\sim 7\text{ g cm}^{-3}$, owing to their terrestrial composition, whilst sub-Neptunes lie at a lower bulk-density $\sim 2\text{ g cm}^{-3}$ as a small increase in H/He atmosphere increases the planet's photospheric radius non-linearly.

An interesting question one can pose is, given the inferred core mass distribution, how much H/He should the cores accrete whilst immersed in a protoplanetary disc? Taking this further, I can determine if the constrained core mass distribution can predict the constrained initial atmospheric mass fraction distribution. To do so, I employ a core accretion scaling relation from Lee and Chiang (2015b) for a dust-free H/He atmosphere in order to calculate the accreted mass fraction for a given core mass[6]. As used in Jankovic et al. (2019), the scaling relation is adapted for varying gas surface density (Lee et al., 2018; Fung and Lee, 2018):

$$X(t) = 0.18 \left(\frac{t}{1\text{ Myr}}\right)^{0.4} \left(\frac{0.02}{Z}\right)^{0.4} \left(\frac{\mu}{2.37}\right)^{3.3} \times \left(\frac{M_{\text{core}}}{5M_\oplus}\right)^{1.6} \left(\frac{1600\text{ K}}{T_{\text{rcb}}}\right)^{1.9} \left(\frac{f_\Sigma}{0.1}\right)^{0.12} \tag{4.5.1}$$

where Z is atmospheric metallicity, μ is mean molecular weight, T_{rcb} is the temperature at the radiative-convective boundary inside the atmosphere and f_Σ is the ratio of gas surface density to that of the minimum mass solar nebula (Hayashi, 1981). Figure 4.16 shows the inferred initial atmospheric mass fraction distribution (from MODEL III) in blue, whilst the accreted mass fraction (according to Equation 4.5.1) in red. In order to produce a full range of atmospheric masses, the constrained core mass distribution is extrapolated below the range used in the inference model - this untrustworthy region of parameter space is shown in grey. One can see clearly that the atmospheric mass predicted by core accretion is ~ 5 times larger than the constrained distribution.

This discrepancy has been found in other previous works. As discussed in Chapter 1, in Jankovic et al. (2019), magneto-rotational instability (MRI) simulations were performed to reveal that the masses of cores forming in the inner disc would result in atmospheres that were too large to

[5] Accessed on 04/04/2020.

[6] Although, as shown by (Jankovic et al., 2019) choosing this specific model over others makes little difference

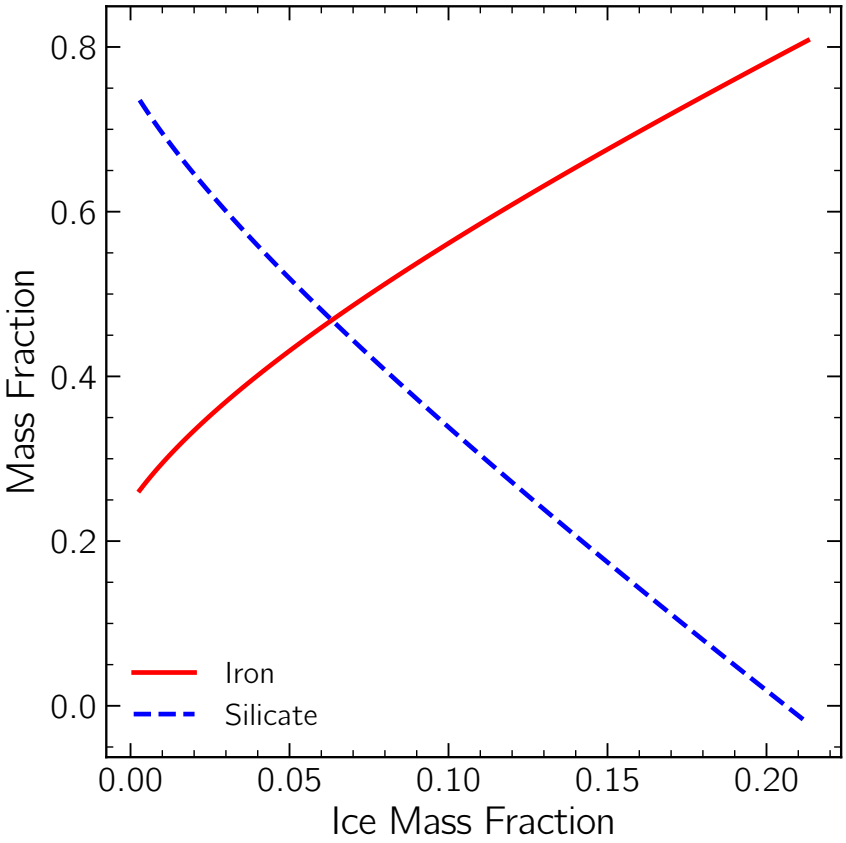


Figure 4.18: Mass fractions of iron and silicate as a function of ice mass fraction for a $4M_\oplus$ planet according to core interior models from Zeng and Seager (2008). I fix the core radius to be $\sim 1.63R_\oplus$ in this calculation as this requires that an ice-free core has a silicate-to-iron composition ratio of $\sim$ 3:1 and hence aligns with the constrained distributions.

be photoevaporated and thus evolve into the typical size of planet we observe today. Possible mechanisms to resolve this tension include forming the planets at the very end of the disc lifetime (Ikoma and Hori, 2012; Lee and Chiang, 2016), or alternatively improving the accuracy of gas accretion models to include the effects of giant mergers (Liu et al., 2015; Inamdar and Schlichting, 2016) which would result in potentially significant atmospheric mass-loss and the production of heat which would take typically kyrs to disperse. The inclusion of 3D simulations has additionally shown that recycling of high-entropy gas during the accretion phase can act to reduce the final atmospheric mass of the planet (Ormel et al., 2015; Fung et al., 2015; Cimerman et al., 2017; Ali-Dib et al., 2020; Chen et al., 2020).

Finally, a different approach is the hypothesised 'boil-off' mechanism (Owen and Wu, 2016) in which, during protoplanetary disc dispersal, low-mass exoplanets may launch a Parker wind (Parker, 1958) due to the quick reduction in confining gas pressure. This rapid mass-loss mechanism may account for the disagreement between the constrained atmospheric mass fraction distribution and that provided by core accretion theory. This is the focus of the work presented in the Chapter 5.

4.5.3 Core Composition

As with all models presented in this work, the inferred mean core composition is found in a silicate/iron dominated regime. Taking the value from MODEL III, I find a mean value of $\tilde{\rho} = 0.26^{+0.08}_{-0.09}$, which is equal to a bulk core density for a $1M_\oplus$ core of $\rho = 5.10^{+0.39}_{-0.40}$ g cm^{-3}. Interpreting this composition using the Fortney et al. (2007) mass-radius relations, suggests

that cores of this nature would have a silicate-to-iron ratio of ∼3:1, which is consistent with that of Earth and confirms findings of previous works (e.g. Owen and Wu, 2017; Wu, 2019). Additionally, the typical bulk core density of observed planets beneath the gap (and hence host negligible atmosphere) are consistent with this value (Dressing et al., 2015; Dorn et al., 2019). Taking this interpretation of the core composition further, it follows that in order to form cores dominated by silicate material and absent of ices, it is necessary to build them interior to the water-ice line (Chatterjee and Tan, 2014; Jankovic et al., 2019). A possible restriction in the adopted methodology is that the core composition analysis has implemented the Fortney et al. (2007) mass-radius relationships, which only allow two species to be present in a core at one time: either an ice-silicate mixture or a silicate-iron mixture. In reality, all three species are likely to coexist in cores and hence I use the exoplanet interior model from Zeng and Seager (2008) to calculate the allowed composition fractions for a planet with the inferred bulk core density $\sim$ 5g cm^{-3}. Figure 4.18 shows the iron and silicate mass fractions as a function of ice mass fraction for a $4M_{\oplus}$ planet with a $\sim 1.63R_{\oplus}$ core radius. This combination ensures matching bulk core densities with the inferred distributions and also produces a typical mass and radius for a super-Earth. One can see that in order to increase the ice mass fraction by a small amount, the iron content must increase dramatically. Furthermore, if one wishes to match ice mass fractions predicted by formation models beyond the water-ice line, the iron mass fraction must rise to unphysical levels and remove the vast majority of the silicate mass. I thus conclude that, although the adopted composition interpretation is limited to two species, only a negligible ice content would be introduced with more sophisticated interior models. Although this points towards the *in-situ* formation pathways, additional physics such as the presence of planetary magnetic fields (Owen and Adams, 2019) can act to suppress atmospheric mass-loss, leading to a lower inferred core density and hence ice-rich cores. This would, however, result in an inconsistency between inferred planet masses and those observed with RV or TTV surveys. Therefore, this result robustly points towards in-situ formation pathways for small, close-in planets.

4.5.4 Core Composition Spread

In fitting a Gaussian function to core compositions, MODEL III not only provides a mean-value, but also a width to the core composition distribution. As shown in Figure 4.8, one can only place a 2σ upper limit on this width to be $\sigma_{\tilde{\rho}} < 0.16$, which is evidently consistent with zero. This composition spread corresponds to 2σ upper limit in the variation of the density of an Earth-mass core of < 1.5 g cm^{-3}. It was shown in Owen and Wu (2017) that for a decrease in bulk core density for constant mass, a planetary core expands to a larger radius. As a result, a larger atmosphere can be stripped and thus the location of the radius gap moves to a larger value. To produce a clean gap i.e. one with a greater sparsity of planets, such as that found in Van Eylen et al. (2018), the spread in bulk core densities must be very narrow. As this has now been quantified in MODEL III, it suggests that, despite different host stars and stellar neighbourhoods, small, close-in exoplanet cores all attain a very similar core density. Whilst there exist strong degeneracies in core composition, which allow different fractions to equate to the same bulk core density, it still raises the question of why such fine-tuning of bulk density would arise in a stochastic core assembly process (Michel et al., 2020).

It is worth noting that whilst the cleaner gap found in the asteroseismic sample of Van Eylen et al. (2018) is more accurate and therefore desirable (Petigura, 2020), the survey is currently not large enough to yield meaningful results if included in this forward modelling methodology. In addition, in order for the model to be computationally efficient, a completeness map is required in order to calculate synthetic planet detections, which is not available for the Van Eylen et al. (2018) sample.

To check whether the model can produce a clean gap in the absence of radius errors, Figure

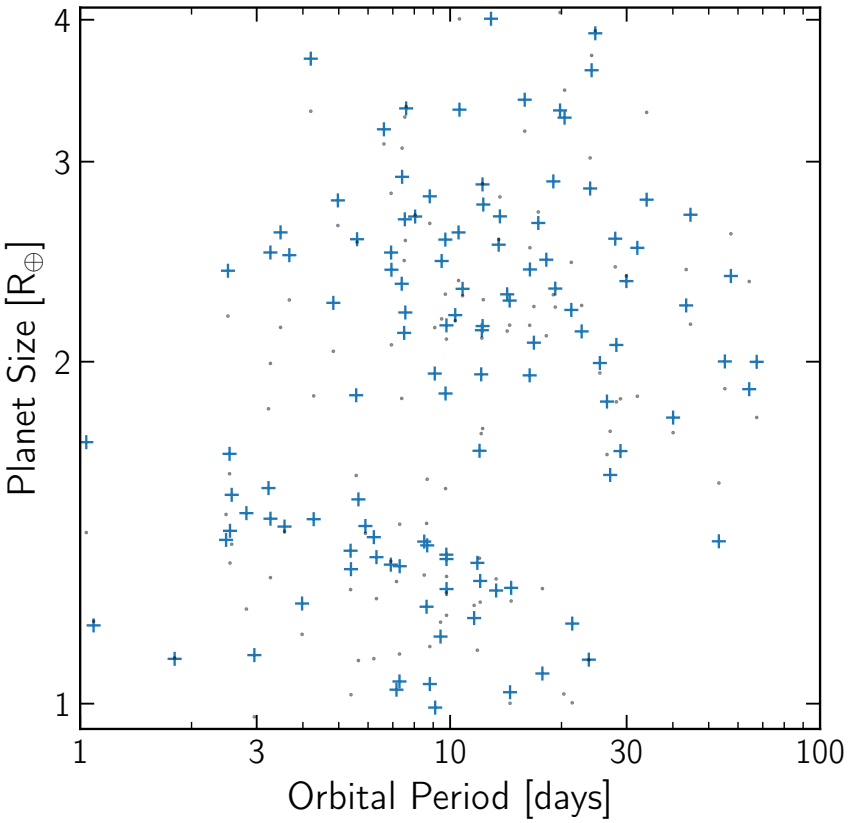


Figure 4.19: As a comparison to the asteroseismic survey of Van Eylen et al. (2018), here I show 117 forward modelled planets randomly selected from MODEL III. *Blue crosses are planets where no radii errors are added, hence I produce a clean gap. Black dots represents the same planets with typical CKS errors added instead, which can shift some planets into the gap. Hence, despite using a data set with larger radii uncertainties, I still predict a clean gap in the underling planet distribution.*

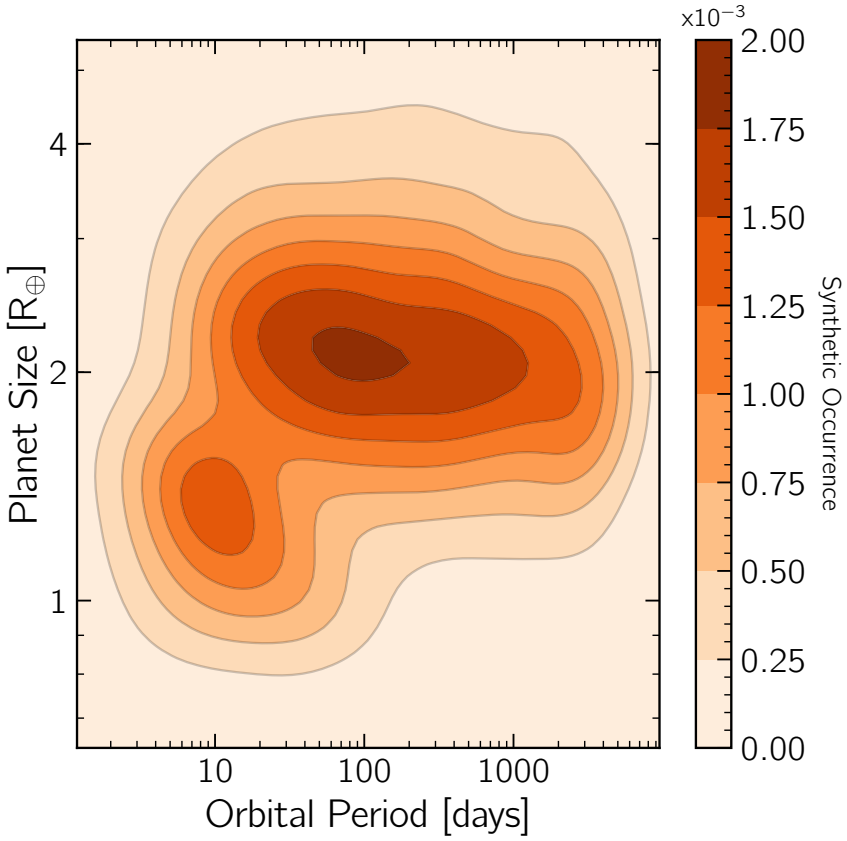


Figure 4.20: Planet size distribution as a function of orbital period for best fit MODEL III *parameters under the assumption that the inferred period distribution (Equation 4.2.20) extends beyond 100 days.*

4.19 shows a random draw of 117 planets (the same as the Van Eylen et al. 2018 sample) forward modelled using MODEL III with and without radius errors included. This shows that the underlying distribution of planets produces a clean radius gap, and it is the uncertainty in planetary radii that pushes planets into the valley.

4.5.5 Planet Distribution Beyond 100 day Orbital Period

Whilst current detection limits prevent surveys probing exoplanet distributions at larger orbital periods, one can postulate as to whether the distribution we observe today extends beyond 100 days. As shown in Figure 4.20, extrapolating the period distribution used in this work (Equation 4.2.20) and modelling planets according to MODEL III demonstrates that the sub-Neptune population would in-fact be far greater in number than that of the super-Earths. This is because planets at greater orbital separations are less prone to atmospheric mass-loss and therefore retain a few percent mass atmosphere. If, on the other hand, one were to observe a distribution in which the relative occurrence of super-Earths and sub-Neptunes remains approximately equal beyond 100 days, it would suggest that the number of planets that were born terrestrial might increase with orbital period. Future transit surveys such as *PLATO* (Rauer et al., 2014) may be able to probe to such orbital distances, which may shed light on this.

4.5.6 Changing the Evolution Model

As discussed in Section 4.2, all results presented in this Chapter hinge on the EUV/X-ray photoevaporation model and accurate photoevaporation rates, which forms a strong prior on on all constrained quantities. As discussed in Chapters 2 and 3, core-powered mass-loss is similarly capable of reproducing the observed data (Gupta and Schlichting, 2019, 2020). Whilst both

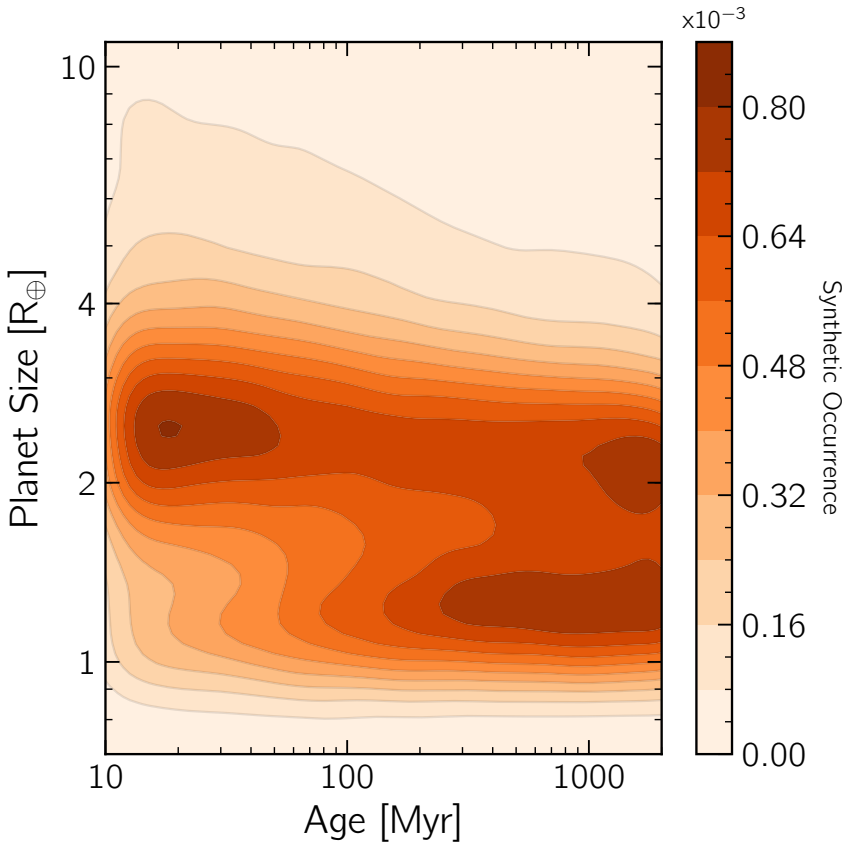


Figure 4.21: Planet size distribution as a function of time (or similarly host star age). The bimodality can be seen to emerge after 100 Myrs, typically when the star begins to spin down and the XUV luminosity drops off precipitously. When compared to core-powered mass-loss (i.e. Figure 10 in Gupta and Schlichting (2020)) one can see the timescale for radius gap emergence is of order Gyrs.

models have similar dependencies and hence inferred quantities (e.g. core composition), they differ in some regards, particularly in planet radius evolution. As the majority of the mass-loss in photoevaporation occurs during the first few 100 Myrs before the host star begins to spin down, the radius gap is predominantly formed at the end of this stage (see Figure 4.21). Core-powered mass-loss on the other hand invokes the core's accretion luminosity to provide the energy source required to induce atmospheric mass-loss. As the time scales for this energy-transfer are far greater than that of EUV/X-ray photoevaporation, the radius gap takes far longer to form, typically of order Gyrs. This difference in evolution may lead to different inferred demographics, especially initial atmospheric mass fraction. I therefore suggest that comparisons of inferred quantities from both models can act as further tests between the two mechanisms. As a very basic statistic, I measure the occurrence ratio of super-Earths to sub-Neptunes defined by number of planets with $R_p \leq 1.8R_\oplus$ and $R_p > 1.8R_\oplus$ respectively. I find ratio values of 0.77 ± 0.08 for stellar ages $t_* < 1$ Gyr and 0.95 ± 0.08 for $t_* > 1$ Gyr. This indicates that, while the majority of planets cross the gap on 100 Myr time-scales, there are a minority that do so at later times (as can be seen in Figure 4.21). Works are beginning to constrain the time-dependence of the exoplanet population (such as Berger et al., 2020a; Fernandes et al., 2022), but a single cut at a specific time is particularly sensitive to the underlying age distribution either side of this cut. However, trying to constrain the evolution of the ratio of super-Earths to sub-Neptunes as a function of time will provide a constraint on the driving mass-loss model.

4.5.7 Improving the Model

In general, this work has been a proof of concept that observationally unobtainable demographic quantities can be inferred from pre-existing data by utilising an evolutionary model. One key limitation is the amount of data included. As new transit surveys such as *TESS* (Ricker et al., 2015) begin to reveal a myriad of small, close-in exoplanets, the inclusion of the such data will only increase the constraining power of the model. Despite what I have shown is possible, the limited CKS sample (~ 900 planets) meant that the non-parametric Bernstein polynomial approximations of core mass and initial atmospheric mass fraction distributions had to be limited to 5^{th}-order. As discussed in Section 4.4, an 8^{th}-order model was also performed which yielded results that were consistent with the three models presented here. This confirms that the analysis is currently limited by the data. In addition to more transit data, mass measurements may be included such as from RV or TTVs, again improving the constraining ability of the model. I do note however, that the observational biases of any new data must be quantifiable in order to include it in the inference model.

In the presence of more data, more complex trends can be explored. One such trend that was probed in Wu (2019) was the dependency of core mass with stellar mass i.e. do larger stars host larger planets? In future work, I therefore aim to determine whether the peak of the inferred core mass distribution increases for larger host stars. Furthermore, one might expect the core mass and initial atmospheric mass fraction distributions to be correlated due to the fact that larger cores are able to accrete larger atmospheres. Furthermore, more data could be used to determine dependencies of core mass or initial atmospheric mass fraction distributions with stellar mass or orbital period, the latter of which might aid in placing constraints on core formation pathways, atmospheric accretion rates or 'boil-off' style mass-loss mechanisms.

A final note is that the success of this inference model is pivotal on the fact that the analytic model from Owen and Wu (2017), as described in Chapter 2, is capable of rapidly calculating planetary evolution due to photoevaporation. The down-side of this however, is that this is an approximation of full numerical simulations, such as those from Owen and Jackson (2012); Owen and Wu (2013). As a result, accurately modelling aspects such as mass-loss efficiency, atmospheric opacity and envelope equation of state is limited. More accurate simulations of atmospheric mass-loss are considered in Chapter 5. I also demonstrate how machine-learnt emulators may be used to incorporate accurate hydrodynamic models into physics-based inference in Chapter 6.

4.6 Conclusion

In this Chapter I have used an evolutionary model for EUV/X-ray photoevaporation and the California Kepler Survey (CKS) data to infer the core mass distribution, the initial atmospheric mass fraction distribution and the core composition distribution for small close-in exoplanets. This is achieved by invoking the photoevaporation model to evolve populations of exoplanets and then synthetically observe them with the observational biases and noise of the *Kepler Space Telescope*. By then comparing the modelled detections with the real CKS data, I infer which underlying planet demographics are required to match the model with the data. My main conclusions are as follows:

- The core mass distribution is peaked at $\sim 4M_{\oplus}$, with a steep decline of occurrence towards higher masses. This points towards a singular formation pathway of planetary cores, preferentially produced at a few earth masses.

- The core composition distribution is centred at a bulk core density for an Earth mass core of $\rho_{1\ \mathrm{M}\oplus} \approx 5.1$ g cm^{-3}. This corresponds to 4:1 to 3:1 silicate iron composition ratio,

which is consistent with Earth and heavily favours forming cores interior to the water ice-line.

- The core composition distribution has an extremely narrow width, which is consistent with zero. This suggests a fine tuning of bulk core density, despite difference in host stars and stellar neighbourhoods.

- The initial atmospheric mass fraction distribution is strongly peaked at $\sim 4\%$. Evolving this forward in time shows that approximately four times as many planets evolved to become stripped rocky cores, than those that were born rocky with no extended H/He atmosphere.

- Core accretion models over-predict the initial atmospheric masses of small exoplanets when compared to that which photoevaporation can strip. To resolve this tension, more sophisticated simulations of core accretion are needed. Alternatively, additional mass-loss mechanisms such as late-stage giant mergers or 'boil-off' phases may further reduce the atmospheric mass and hence bring these models into agreement. This latter scenario is discussed in Chapter 5.

These conclusions are all dependant on the fact that one assumes the main driver of small, close-in exoplanet evolution is EUV/X-ray photoevaporation. Changing the adopted evolution model to alternative theories, such as core-powered mass-loss may result in differing constrained distributions, which may provide tests between the models.

Chapter 5

Under the Light of a New Star: Evolution of Planets through Disc Dispersal and Boil-Off

5.1 Introduction

In this Chapter I discuss the atmospheric evolution of small, close-in exoplanets that are immersed in a protoplanetary disc. As the disc disperses, the young and inflated planet is irradiated by its star for the first time and rapid mass-loss can ensue, which is referred to as the 'boil-off' phase (Owen and Wu, 2016). This event has large implications for a planet's future evolution and as such it is an important process to understand. As discussed in the previous Chapter, it may also provide a solution to the discrepancy between core-nucleated gas accretion and XUV photoevaporation. This tension arises since core accretion models over-predict the amount of atmospheric mass that XUV photoevaporation strip in order to be consistent with the exoplanet population observed today. The boil-off phase is a mass-loss process that occurs in between gas accretion and XUV photoevaporation and hence may be able to bring these two mechanisms into alignment. To understand this, I first provide a brief overview of protoplanetary discs and planet formation, before discussing boil-off and its implications.

5.1.1 Protoplanetary Discs

Planets form as a natural consequence of stellar evolution. A gravitationally collapsing molecular cloud will have a finite angular momentum about some axis. As it continues to collapse and forms a central protostar, this angular momentum is conserved such that the material in the cloud spins up into a circumstellar disc perpendicular to the rotation axis (also known as a protoplanetary disc). To prevent rotational break-up of the star, magnetic braking removes angular momentum via a stellar wind (e.g. Bouvier et al., 2007; Li et al., 2014). Although the collapse of the system is halted due the balance of centrifugal and gravitational forces, the combination of gas (primarily molecular hydrogen/helium) and dust (silicates, ices etc.) slowly accretes onto the central star via angular momentum transport. In forming a disc and not spherically accreting onto the star, the dust is provided extra time to locally collapse inside the disc to form planets and eventually accrete atmospheres via core accretion. Observations suggest that after a few million years the gas in the disc disperses or is completely accreted onto the star (e.g. Haisch et al., 2001; Hernández et al., 2007; Mamajek, 2009; Richert et al., 2018), leaving behind one or more orbiting planets. As a result of this time line, understanding the nature of protoplanetary discs is of great importance to understanding planet formation.

Accretion discs are a family of astrophysical structures, of which the protoplanetary disc is a member, that are ubiquitous in our Universe; from planetary rings to quasars and even galactic discs. As discussed, they naturally form from material collapsing with a finite angular momentum around a deep gravitational well. Following from Pringle (1981); Ogilvie (2005); Armitage (2015), for a non-self-gravitating disc around a star $M_{\rm disc} \ll M_*$, the potential felt by a test particle in cylindrical coordinates (r, ϕ, z) is:

$$\Phi = -\frac{GM_*}{(r^2 + z^2)^{1/2}}. \tag{5.1.1}$$

The orbits for this potential are hence Keplerian, with azimuthal velocity $v_{\rm K}$ and angular frequency $\Omega_{\rm K}$:

$$v_{\rm K} = \sqrt{\frac{GM_*}{r}}, \tag{5.1.2}$$

$$\Omega_{\rm K} = \sqrt{\frac{GM_*}{r^3}}. \tag{5.1.3}$$

The main principle of an accretion disc is that energy is released by moving angular momentum outwards and mass inwards. Although the majority of the mass moves radially inwards, the angular momentum is in-fact transported to large radii via a small amount of mass. Put another way, there must be a dissipative mechanism which allows mass to be transported inwards, allowing it to eventually accrete onto the star. To investigate this, consider the simple case of vertical hydrostatic equilibrium for a steady-state, inviscid disc of density ρ:

$$\frac{1}{\rho}\frac{{\rm d}P}{{\rm d}z} = -\frac{{\rm d}\Phi}{{\rm d}z}, \tag{5.1.4}$$

where P is the gas pressure. For a thin disc i.e. $z \ll r$, one can show that ${\rm d}\Phi/{\rm d}z \approx \Omega^2 z$. Examining Eq. 5.1.4 in this new regime, one finds that

$$\frac{1}{\rho}\frac{P}{H} \sim -\Omega^2 H \rightarrow H\Omega \sim \sqrt{\frac{P}{\rho}} \equiv c_{\rm s}, \tag{5.1.5}$$

where H is the scale-height of the disc and $c_{\rm s}$ is the isothermal sound speed. The dimensions of dynamical viscosity μ are identical to that of P/Ω, hence we can state that

$$\mu = \alpha\frac{P}{\Omega}, \tag{5.1.6}$$

where α is the *dimensionless viscosity parameter* as introduced by Shakura and Sunyaev (1973) in a way to bypass the need for any physical explanation for the source for this viscosity, which is still an ongoing mystery today (e.g. Turner et al., 2014). The kinematic viscosity ν for the disc is thus:

$$\nu \sim \frac{\mu}{\rho} \sim \alpha\frac{c_{\rm s}^2}{\Omega} \sim \alpha c_{\rm s} H. \tag{5.1.7}$$

We may now define two important timescales for disc evolution. The dynamic timescale is the time in which bodies orbit the central star

$$t_{\rm dyn} \sim \frac{1}{\Omega}, \tag{5.1.8}$$

whilst the viscous timescale is the time in which material accretes radially onto the star and therefore represents an estimate of the disc lifetime

$$t_{\rm visc} \sim \frac{r^2}{\nu} \sim \alpha^{-1}\left(\frac{H}{r}\right)^{-2} t_{\rm dyn}. \tag{5.1.9}$$

For a thin disc $H/r \ll 1$ with $\alpha < 1$, we find that

$$t_{\rm dyn} \ll t_{\rm visc}, \tag{5.1.10}$$

which is indicative of the fact that the lifetime of the disc is far greater than the length of a single orbit. Finally, one can examine the azimuthal velocity of gas in the disc to show that it travels *sub-Keplerian*, meaning $v_\phi < v_{\rm K}$. Consider the radial force balance in a smooth disc:

$$\frac{v_\phi^2}{r} = \frac{1}{\rho}\frac{{\rm d}P}{{\rm d}r} + \Omega_{\rm K}^2 r, \tag{5.1.11}$$

where terms on the right-hand side represent gas pressure and gravity. On dimensional grounds, ${\rm d}P/{\rm d}r \sim -P/r$ (pressure is a decreasing function with radius) and since $P \sim c_{\rm s}^2\rho$ we may rewrite this as:

$$\frac{v_\phi^2}{r} = \Omega_{\rm K}^2 r - \frac{c_{\rm s}^2}{r}. \tag{5.1.12}$$

Using equation 5.1.5, we finally arrive at

$$v_\phi^2 = v_{\rm K}^2\left(1 - \mathcal{O}\left(\frac{H}{r}\right)^2\right), \tag{5.1.13}$$

which confirms that the gas in the disc orbits the central star at a sub-Keplerian velocity. The origin of this behaviour is due the gas pressure which acts to oppose the gravitational force.

Now consider the dust in the disc - it has sufficient inertia (when compared to the gas) to not feel a noticeable acceleration due to gas pressure and thus orbits at a higher azimuthal velocity than the gas. As a result, the dust experiences a gas headwind and thus loses angular momentum and slowly spirals in towards the star (Weidenschilling, 1977). The behaviour is referred to as *radial drift* and has been observationally confirmed on multiple occasions (Panić et al., 2009; Andrews et al., 2012; Isella et al., 2012; Powell et al., 2017). The net movement of dust is of crucial importance when considering planet formation, as it is the dust that gravitationally collapses to form planetesimals and eventually planetary cores.

Finally, it is convenient to define the gas surface density $\Sigma_{\rm g}$ for the disc:

$$\Sigma_{\rm g}(r) \equiv \int_{-\infty}^{+\infty} \rho(r)\, dz, \tag{5.1.14}$$

where ρ is the gas density. The surface density arises when solving for vertical hydrostatic equilibrium for a vertically isothermal disc from Eq. 5.1.4 (see Ogilvie, 2005) and is related to the mid-plane density $\rho_{\rm mid}$ as follows:

$$\rho_{\rm mid}(r) = \frac{\Sigma_{\rm g}(r)}{\sqrt{2\pi}H} = \frac{\Sigma_{\rm g}(r)}{c_{\rm s}}\sqrt{\frac{GM_*}{2\pi a^3}}, \tag{5.1.15}$$

where in the last step I have used relations for the vertical scale height H from Eq. 5.1.5 and Keplerian angular frequency $\Omega_{\rm K}$ from 5.1.3. Note that the surface density can also refer to the dust density $\Sigma_{\rm d}$, as opposed to gas. If one assumes that a planet orbits in the mid-plane of a disc, this relation allows one to determine the local density/pressure a planet is subject to in the presence of a protoplanetary disc.

5.1.2 Atmospheric Accretion

If a planetary core becomes sufficiently large before the disc disperses, surrounding gas is drawn in to form an atmosphere. Specifically, one requires the Bondi radius of the planet to extend

beyond the core radius (e.g. Ginzburg et al., 2016; Massol et al., 2016), since accreted material is only thermally bound to the planet within the Bondi sphere (Pollack et al., 1996). Early works also suggested that once a planet's mass is dominated by its atmosphere, it becomes self-gravitating and triggers a runaway accretion process (Perri and Cameron, 1974; Mizuno, 1980). One can understand the rate of accretion by examining the thermodynamics of the system. Consider a parcel of fluid in-falling onto the planet - as it approaches the planet it releases binding energy, which in turn heats the system and causes a finite expansion of the surrounding gas. To accrete more gas, the atmosphere must cool and contract before any gas can fit inside the Bondi radius. As a result, *a planet can only accrete as quick as its atmosphere can cool.* The relevant timescale for cooling and therefore accretion is thus the Kelvin-Helmholtz timescale of Eq. 2.6.12.

Since gas accretion can only occur if the planet can cool, this raises the question of what controls cooling? The internal luminosity of a small mass planet arises due to change in gravitational binding energy, as discussed in Chapter 2. The associated temperatures are thus very small, typically ~ 50K (Lopez and Fortney, 2014) meaning that this source does not prevent cooling. However, during solid core accretion, an additional source of luminosity $L_{\rm acc}$ is present that arises due to in-falling solid material releasing binding energy. This luminosity keeps the atmosphere heated and slows/prevents gas accretion. If this is the case, how then do planets build atmospheres?

A key fact for planets in the inner disc is that core formation via either pebble or planetesimal accretion occurs on very short timescales (Hansen and Murray, 2012; Chiang and Laughlin, 2013; Jankovic et al., 2019). Following a simple argument from Lee et al. (2014), if one assumes the disc produces a planetary core within an annulus at orbital separation of a, then the local dust surface density must be $\Sigma_{\rm d} \sim M_{\rm c}/a^2$. As such, the coagulation timescale $t_{\rm coag}$, i.e. the time taken to build the planet is:

$$t_{\rm coag} \sim \left(\frac{a}{R_{\rm c}}\right)^2 \Omega_{\rm K}^{-1} \sim 10^4 \text{ yr} \left(\frac{a}{0.1 \text{ AU}}\right)^{3.5} \left(\frac{1.6 R_\oplus}{R_{\rm c}}\right)^2. \tag{5.1.16}$$

This timescale is much shorter than the disc lifetime $t_{\rm disc}$, meaning planet formation in the inner disc happens rapidly. The main implication of this statement is that any luminosity produced as a result of solid core accretion i.e. binding energy released, $L_{\rm acc}$, will quickly shut off once the planet forms. As such, once a planet has been built, one may assume $L_{\rm acc} = 0$ and gas accretion may ensue until the disc disperses, or the local gas reservoir is depleted. Indeed, if accretion luminosity was thermodynamically important for the entire disc lifetime, then planetary cores would subsequently grow to be much larger and inconsistent with the observed masses of super-Earths (Lee and Chiang, 2015a). Note however that $t_{\rm coag}$ increases strongly with orbital separation, implying that $L_{\rm acc} \neq 0$ for a considerable portion of the disc lifetime in the outer disc. As such, when modelling planet formation at large orbital separations, one should assume that both solid and gas accretion may occur at the same time.

5.1.3 Disc Dispersal

Disc dispersal is the process by which gas in a protoplanetary disc is removed. It can either accrete onto the central star or be blown off via high-energy irradiation of the gas. This latter process is similar to that of atmospheric photoevaporation, such as that covered in previous chapters, but now relates to removal of material from the disc. As discussed, a disc evolves on the viscous timescale $t_{\rm visc}$ (see Equation 5.1.1) which is typically of order 10^6 yrs. Through population studies, it was observed that roughly two populations existed in colour-colour diagrams, such as that shown for Taurus in Figure 5.1 from Luhman et al. (2010). In this Figure, the ratio in infrared emission for protostars is measured for $K_{\rm s} - [24]$ and $K_{\rm s} - [8]$. Stars that host a protoplanetary disc are expected to emit an excess in infrared due to the dust. Once this disperses

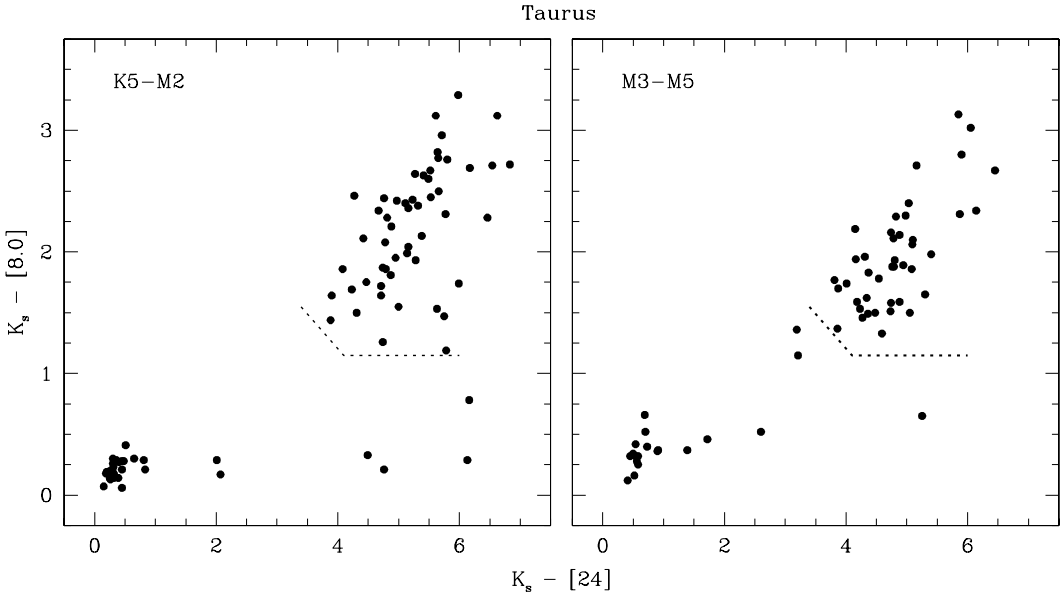


Figure 5.1: Colour-colour diagram for protostars in Taurus, with typical ages of ~ 1*Myr, for different spectral types. Dashed lines represent the boundary between primordial and dispersing discs. Figure taken from Luhman et al. (2010).*

however, one would expect this excess to subside and the spectral energy distribution to be consistent with simply the photosphere of the star. Indeed, the values of $K_{\rm s} - [24]$ and $K_{\rm s} - [8]$ in Figure 5.1 are normalised to the expected values for a Rayleigh-Jeans stellar photosphere of a given spectral type. One can see that there are two populations of young stars: those that host discs (upper right) and those that don't (bottom left). Since we observe very few stars in the intermediate regime, one can make a timescale argument to suggest that disc dispersal must occur very rapidly. Multiple works have shown that the dispersal timescale, $\tau_{\rm disp}$, is typically an order of magnitude quicker than the disc lifetime $\tau_{\rm disp} \sim 10^4 - 10^5$yrs (Kenyon and Hartmann, 1995; Ercolano et al., 2011; Luhman et al., 2010; Koepferl et al., 2013). This spawned the family of "two-timescale" dispersion models, of which the "photoevaporative switch" model successfully describes how the disc disperses on a much shorter timescale than its lifetime (e.g. Clarke et al., 2001).

This mechanism relies on the trade-off between gas accretion onto the central star, and gas being driven away vertically via photoevaporative winds generated by the high-energy flux from the star. As disc mass and therefore accretion rate drop towards the end of the disc lifetime, there will come a point where gas has a larger probability of being blown away than accreted. As a result, a gap is opened at this radius, typically $\sim$ 1AU, which disconnects an inner and outer disc - causing the two to evolve separately. Noting that as the viscous timescale (Equation 5.1.1) scales linearly with radius[1], it is apparent that the inner disc evolves on a much shorter time period. This is understood by recalling that, in order for accretion to occur, angular momentum must be transported to the edge of the disc with a small amount of mass. Before the gap opened, angular momentum had to be transported hundreds of AU away from the star in order for accretion to occur. Now, however, the inner and outer discs are disconnected, and angular momentum need only be transported to the inner disc edge in order for gas to be accreted onto the star. As a result, the inner disc quickly drains onto the star - the disc is now referred to as a *transition disc* due to the hole that has formed where the inner disc used to be. Finally, the hole provides a clear path for X-rays to reach the outer disc, which is then quickly

[1]For a constant α, the kinematic viscosity $\nu \sim r$ and hence the viscous timescale $t_{\rm visc} \sim \frac{r^2}{\nu}$ scales linearly with radius.

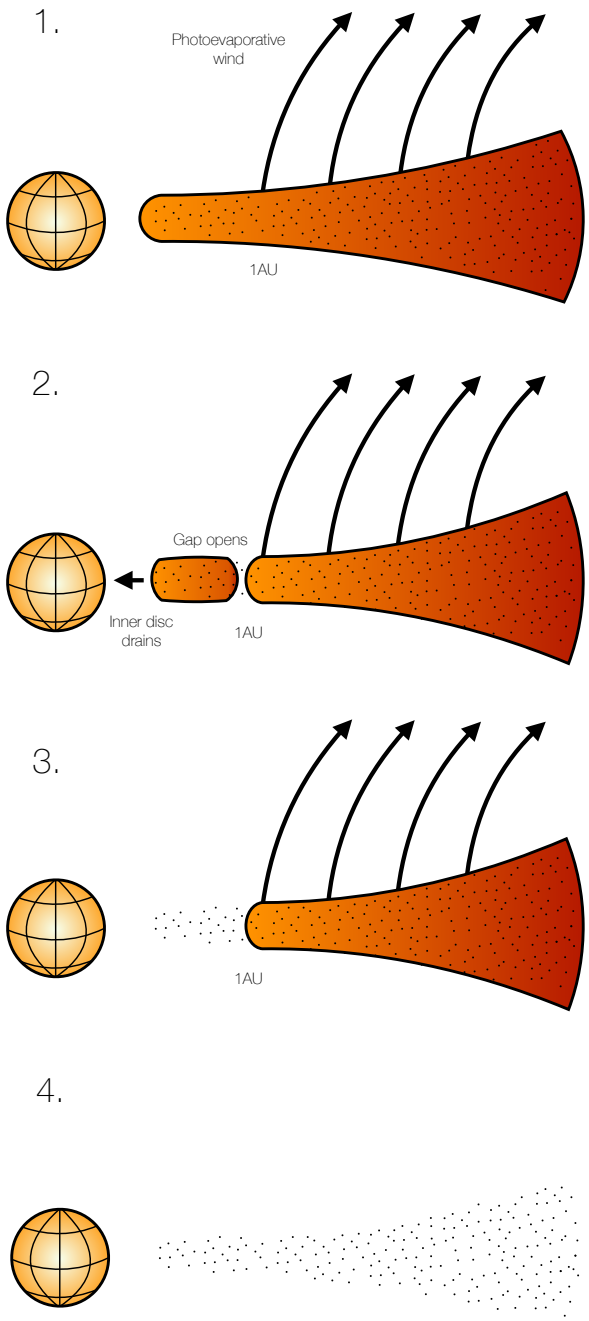


Figure 5.2: Schematic demonstrating disc dispersal (not to scale). In step 1, high energy stellar flux induces a wind that removes gas from the disc. As a result a gap is opened at $\sim 1AU$, separating an inner and outer disc. The inner disc quickly drains onto the central star (step 2) which allows stellar flux to reach the outer disc (step 3) causing it to evaporate (step 4).

evaporated. This process is outlined in Figure 5.2.

5.1.4 Boil-Off

Disc dispersal has strong implications for the atmospheric evolution of close-in exoplanets. At the point of disc dispersal $t_{\rm disp} \sim 3-10$ Myrs, the cooling timescales for a planet will have an upper limit of $t_{\rm KH} \sim t_{\rm disp} \sim 3-10$ Myrs since cooling timescales approximately track a planet's age (see Chapter 2) and planets can form at any point during the disc's lifetime. From observational studies, the disc disperses on timescales $\tau_{\rm disp} \sim 10^4 - 10^5$ yrs (Kenyon and Hartmann, 1995; Ercolano et al., 2011; Koepferl et al., 2013), which suggests that $t_{\rm KH}$ may be greater than $\tau_{\rm disp}$. If this is the case, a planet cannot thermally respond to any changes in external pressure, such as the disc dispersing, and will undergo adiabatic expansion on timescales of $\tau_{\rm disp}$. Since the stellar irradiation will heat the expanding material, atmospheric escape may occur via a rapid hydrodynamic outflow. The ensuing mass-loss is referred to as 'boil-off' (Owen and Wu, 2016).

In Ikoma and Hori (2012), gas accretion and mass-loss were self-consistently modelled for rocky cores surrounded by H/He atmospheres. They modelled disc dispersal as a continuous process, however assumed that the planet's atmosphere and surrounding gaseous disc remained in hydrostatic equilibrium for the entire process. As a result, they found that planets can lose their entire atmospheric mass, which is no surprise because the only way to remain in hydrostatic equilibrium with a vanishing disc is to host a vanishing atmosphere. As discussed however, this assumption may not be valid since the cooling timescale of the planet can be longer than the disc dispersal timescale. In Owen and Wu (2016), boil-off evolution models were performed using *Modules for Experiments in Stellar Astrophysics*, `MESA` (Paxton et al., 2011, 2013, 2015, 2018), which is described in-detail in Section 5.2.1. Instead of assuming hydrostatic equilibrium as is the case in Ikoma and Hori (2012), these models assumed the planet undergoes mass-loss via a hydrodynamic isothermal Parker wind. These models did not, however, self-consistently model the gas accretion phase prior to disc dispersal or a continuous disc dispersal process. Instead, planets were modelled from an initial inflated state in which $R_{\rm ph} \sim R_{\rm B}$. This can be thought of as a young planet that is inflated due to an accretion luminosity $L_{\rm acc}$. At the start of the evolution, this luminosity is switched off $L_{\rm acc} = 0$ (as argued in Section 5.1.2) and the outer boundary was exposed to stellar radiation at an equilibrium temperature $T_{\rm eq}$ associated with the planet's orbital separation. The results of Owen and Wu (2016) demonstrated that the planet's atmosphere will contract, lose mass and cool on a timescale $\sim 10^5$ yrs. The rapid contraction of the atmosphere provides sufficient luminosity to replenish the mass that has been lost in the upper atmosphere, resulting in continued mass-loss. This shuts off however once the photospheric radius shrinks to $R_{\rm ph}/R_{\rm B} \sim 0.1$. One can see this from Figure 2.1, in which the Mach number, and hence mass-loss rate, of a Parker wind falls to very small values for $R_{\rm ph}/R_{\rm B} \lesssim 0.1$. As a result, even small mass planets in the inner disc did not have their atmospheres completely removed, as was the case in hydrostatic equilibrium from Ikoma and Hori (2012).

It is worth noting that in Ginzburg et al. (2018), analytic arguments were made to suggest that in the presence of a sufficient core luminosity, planets can continue to lose mass after the boil-off phase since this energy keeps the atmosphere inflated. The bolometric luminosity then removes the outer layers of the atmosphere in a hydrodynamic outflow. This was in-fact the precursor to the theoretical prediction of core-powered mass-loss from Ginzburg et al. (2018); Gupta and Schlichting (2019), as discussed in Chapter 2. I discuss this in Section 5.4.5.

Another key finding from Owen and Wu (2016) was that advection of atmospheric material played a key role in cooling the planet. In the standard picture of an optically thick atmosphere, as presented in Chapter 2, it is the upper radiative zone that limits the cooling of the planet.

Although the models of Owen and Wu (2016) began with a fully convective atmosphere (an assumption that is addressed in this chapter), the planets cool and develop upper radiative layers, which would normally stifle further thermal contraction. In this case however, the planet is not in hydrostatic equilibrium and advection of material from the upper atmosphere removes energy, bypassing the standard radiative bottleneck. The internal luminosity, which had previously been converted into PdV work, drops substantially once mass-loss shuts off. Combining this with a extreme radial contraction results in a Kelvin-Helmholtz timescale that is much longer than the age of the planet, typically $t_{\rm KH} \sim 10^8$ yrs. Thermodynamically speaking, planets therefore appear older than their actual age after the boil-off phase due to premature cooling. Note that this explains the relation of Eq. 2.6.13 adopted in Chapters 2, 3 and 4. Owen (2020) argued that this result could be used to constrain the initial entropy of young planets: for a given measured mass, radius and age, an inferred smaller entropy (or equivalently larger cooling timescale) would suggest that planets did or did not go through a boil-off phase.

In this Chapter, I build upon the models of Owen and Wu (2016). Specifically, I wish to address two assumptions that were made to simplify their approach. The first was the adoption of an instantaneous disc dispersal, which was inherently assumed when exposing the highly inflated atmosphere to irradiation and vacuum boundary condition at the start of the simulation. Whilst the disc dispersal timescale is much shorter than the disc lifetime, the disc will require a finite time to reduce it's gas surface density such that the planet is exposed to a vacuum. A gradual reduction in confining pressure will make the boil-off process less extreme, since the planet has more time to thermodynamically respond. This in-turn means that the planet may be able to remain in a quasi-hydrostatic equilibrium with the disc i.e. an instantaneous pressure balance, as was assumed in Ikoma and Hori (2012), for at least part of the planet's evolution.

The second assumption in Owen and Wu (2016) that I will address is that the atmospheres began in a maximally entropic state i.e. heated with an artificial accretion luminosity such that $R_{\rm ph} \sim R_{\rm B}$. It is only once this is turned off (at the start of the simulation) that the combination of thermal cooling and irradiation from the host star forms an upper radiative zone with a characteristic isothermal temperature profile. In reality however, a planet's core will form and then, since any accretion luminosity quickly tends to zero in the inner disc, cooling will take place leading to atmospheric accretion. This cooling also leads to a radiative upper zone. As discussed, radiative zones tend to limit cooling (especially in the absence of advection), and so will also act to reduce the mass-loss. Crucially I intend to self-consistently model the gas accretion phase whilst immersed in the disc as well as the subsequent boil-off.

The combination of these two assumptions imply that the mass-loss rates in Owen and Wu (2016) were likely upper limits. The presence of an upper radiative zone and a non-instantaneous disc dispersal will act to reduce the mass-loss, hence they must be addressed to gain a better understanding of this evolution period. In this chapter I present new, more sophisticated simulations of the boil-off phase and discuss their implications for planet evolution.

5.2 Method

Similar to Owen and Wu (2016), the simulations in this Chapter are performed using `MESA` (Paxton et al., 2011, 2013, 2015, 2018). To begin, I provide an overview of this code and then discuss how it is modified for boil-off simulations.

5.2.1 MESA

Modules for Experiments in Stellar Astrophysics, `MESA` (Paxton et al., 2011, 2013, 2015, 2018) is an open source one-dimensional fluid code that originates in stellar evolution modelling. More recently however, it has been used to understand planet evolution (e.g. Owen and Wu,

2013, 2016; Chen and Rogers, 2016; Kubyshkina et al., 2020). The code solves the stellar structure equations as presented in Chapter 2 by implementing a Henyey style implicit numerical scheme (Henyey et al., 1965). A solution to the equations at a given time is found with a Newton-Raphson iterative method, which yields the radius, density, luminosity, temperature and composition structure of a stellar/planetary envelope as a function of mass. In addition, one has the option of including hydrodynamic terms into the stellar structure equations, which allows the study of advection and non-hydrostatic systems. In this sense, `MESA` is a fully implicit Lagrangian fluid solver i.e. it follows fluid elements with time, as opposed to an Eulerian solver which discretises space and tracks fluxes in and out of cells as a function of time. The benefit of an implicit numerical scheme is that large timesteps can be taken in order to model evolution on long timescales, which is not the case for an explicit scheme which will become numerically unstable for a large timestep. For each timestep in `MESA`, the total mass in the simulation domain is adjusted in accordance with any mass-loss/accretion. The system then relaxes by re-solving the stellar structure equations as described above. Timesteps are chosen based on a number of different criteria, but reflect the ability to make large jumps in time whilst still being able to converge to a solution is relatively few iterations.

`MESA` can accommodate studies of material spanning a very wide range in temperatures and densities. For a small mass exoplanet with H/He dominated atmosphere, it uses the SCVH equation of state from Saumon et al. (1995) and opacity tables from Freedman et al. (2008). If, however, one considers a system which is not covered by this combination of EOS and opacity tables, blending options are available to transition to other tables. As such, the thermodynamic state of the atmosphere is accurately modelled, particularly in finding the temperature gradient and its implication on the transition between convective and radiative zones. In addition, mixing-length theory is used to adjust the temperature profile in accordance with convection as discussed in Chapter 2.

An important limitation of `MESA` is that it performs radiative transfer using the diffusion approximation (see Chapter 2), which requires that the atmosphere be locally optically thick in the entire domain. Specifically, this requires that the photon mean-free path is much shorter than the pressure scale height for the entire domain. Additionally, the code is only able to solve the stellar structure equations in sub-sonic flow. This is because the implicit scheme can not resolve the necessary sound waves with respect to time, in order to accurately model shocks and super-sonic flow. A Parker wind, being the adopted model for mass-loss, is a transonic solution which also spans both optically thin and thick regimes i.e. the flow is launched in an optically thick region but extends beyond the photosphere. This poses a severe challenge to modelling mass-loss. A solution exists however, if one carefully considers where to place the outer boundary condition of the `MESA` model.

5.2.2 Outer Boundary Condition

When modelling boil-off and hydrodynamic escape in general, one must consider supersonic flow. As discussed, `MESA` sets two constraints that determine where the simulation outer boundary can be. Namely, the atmosphere below this point should be locally optically thick and subsonic. As described in the next Section, I choose to model the mass-loss in this problem via an isothermal Parker wind beyond this boundary, which adds a third constraint, which is that the atmosphere above the boundary should be isothermal. To satisfy these constraints, I position the outer boundary condition at a fixed pressure P_0 and temperature T_0. The choice in pressure is arbitrary, although is chosen to ensure the gas is locally optically thick and in the upper radiative (and therefore isothermal) layer. To ensure that this pressure satisfies these conditions, I show that the corresponding boundary condition radius $R_{\rm BC}$ at (P_0, T_0) is beneath the Bondi radius (subsonic), in the radiative layer (isothermal) and that the photon mean-free path is much shorter than the pressure scale height for the entire domain (locally optically thick) in Figure

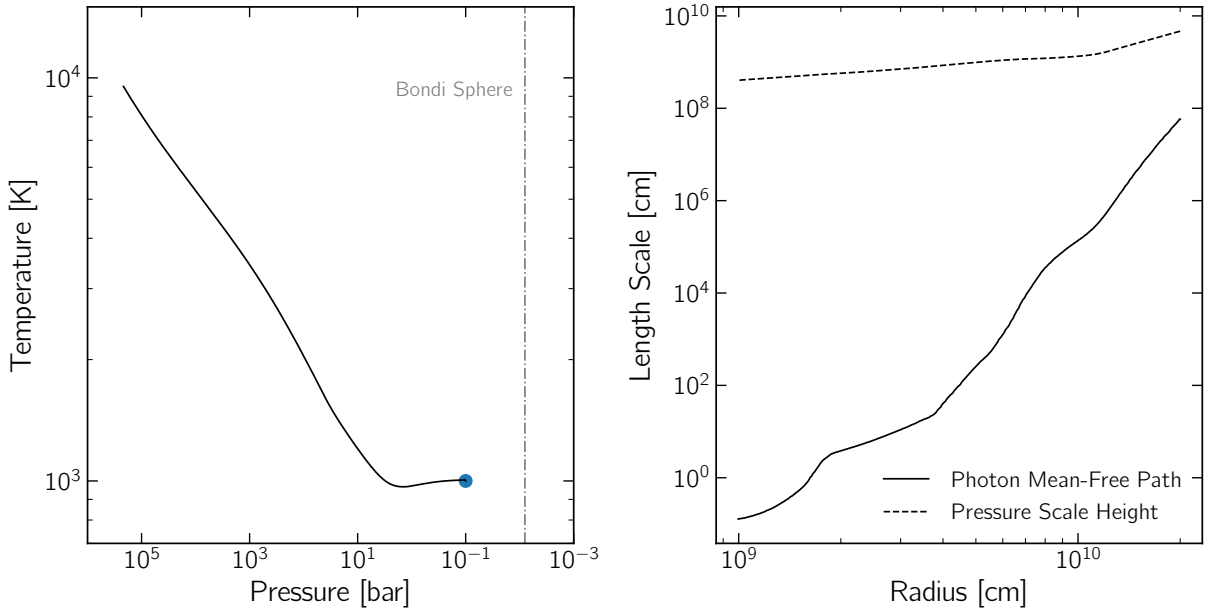


Figure 5.3: In the left-hand panel, a temperature-pressure profile is shown for an atmospheric model with $5M_\oplus$ core and atmospheric mass fraction of $\sim 10\%$. The boundary condition is set with $P_0 = 0.1bar$ and $T_0 = 1000K$, as shown in blue. One can see that the temperature in the inner atmosphere is approximately adiabatic, but flattens out in the outer radiative zone. Beyond this, the atmosphere can be assumed to isothermal. In addition the location of the Bondi sphere is shown in grey dot-dashed. This implies that MESA boundary condition is in an isothermal (radiative) and subsonic (interior to Bondi sphere) region. The right-hand panel shows the photon mean-free path, $\lambda_{ph} = 1/(\kappa\rho)$ (shown in black-solid) is much shorter than the pressure scale height (black-dashed) for the entire domain, meaning that the entire MESA domain is locally optically thick.

5.3.

The outer boundary temperature, which is in the outer radiative layer of the atmosphere, is set to a constant equilibrium temperature $T_0 = T_{\rm eq}$, since it is in radiative equilibrium with the host star/disc (see Eq. 2.6.2). In addition to a fixed temperature and pressure, `MESA` also requires the optical depth to out-going radiation at this boundary. This relies on knowing the density structure above $R_{\rm BC}$, which is determined by the mass-loss/accretion. This is discussed in the next Section.

5.2.3 Accretion and Mass-Loss

In order to model gas accretion and boil-off for a planet immersed in an protoplanetary disc, I require the ability to model both accretion and mass-loss self-consistently. As discussed, mass-loss in `MESA` is treated by adding/removing mass to the domain[2] and then allowing the system to relax. If one knows the mass-flux due to accretion/mass-loss then this can be prescribed at the outer boundary condition at $R_{\rm BC}$ discussed in the previous Section. The question is, how to determine the mass-flux?

The distinction between mass-loss and accretion will depend on the physics of the disc. If the planet is pressure-supported by the disc and the planet can cool, then mass should flow onto the planet. On the other hand, if the pressure-support is not sufficient, the upper atmosphere should expand and undergo mass-loss. To quantify this, I model the atmosphere above the boundary condition with an isothermal layer, also set to $T_{\rm eq}$. As such, I can solve the fluid equations from Section 2.3 in Chapter 2 to determine the fluid flow structure and hence mass-

[2] In practise this mass-change is spread across a portion of the domain, predominantly in the upper atmosphere.

flux into/out of the MESA domain. The problem is set up as follows. For a given density, $\rho_{\rm BC}$, at $R_{\rm BC}$, which now acts as the inner boundary condition for the isothermal layer, one must solve the steady-state momentum equation in the case of non-hydrostatic equilibrium:

$$\left(u - \frac{c_{\rm s}^2}{u}\right)\frac{\partial u}{\partial r} = 2c_{\rm s}^2\left(\frac{1}{r} - \frac{R_{\rm B}}{r^2}\right), \tag{5.2.1}$$

subject to an outer boundary condition which is controlled by the local gas density of the disc, $\rho_{\rm disc}$. Strictly speaking, the location of this boundary condition should be $\min(R_{\rm B}, R_{\rm Hill})$ since these are the points where gas dynamics are controlled by the disc (and therefore the host star) and not the planet. Following Rafikov (2006); Ginzburg et al. (2016), one can show that $R_{\rm B} < R_{\rm Hill}$ for small mass planets close to their star. As such, the density at the outer boundary for this isothermal layer is set to $\rho_{\rm disc}$ at $R_{\rm B}$. One can then relate $\rho_{\rm disc}$ to the gas surface density $\Sigma_{\rm g}$ with Eq. 5.1.15, which will incorporate the physics of disc dispersal. As shown in Chapter 2, Eq. 5.2.1 can be solved with Lambert W functions:

$$u(r) = c_{\rm s}\sqrt{-W_k(-D(r))}, \tag{5.2.2}$$

where k is the function-branch and $D(r)$ is:

$$D(r) = \exp\left\{-4\ln r - 4\frac{R_{\rm B}}{r} - \ln c_{\rm s}^2 + C\right\}. \tag{5.2.3}$$

Here, $C = 4\ln R_{\rm B} + 2\ln u_{\rm crit} - (u_{\rm crit}/c_{\rm s})^2 + 4$ is a constant and controlled by the fluid velocity, $u_{\rm crit}$, at the critical point, $R_{\rm B}$. If $|u_{\rm crit}| < c_{\rm s}$ then the solution is a subsonic breeze, whereas if $|u_{\rm crit}| = c_{\rm s}$, then the solution corresponds to transonic Bondi accretion (Bondi, 1952) or mass-loss i.e. the classic Parker wind solution Parker (1958). In Figure 5.4, I show example solutions with varying critical velocity, $u_{\rm crit}$. Breeze solutions are those that remain subsonic for all radii, whilst transonic solutions are those that exceed the sound speed beyond/interior to $R_{\rm B}$. Note that for the transonic accretion solution, the velocity increases in magnitude as one approaches the planet and would eventually become supersonic and shock inside the MESA domain. As discussed, MESA is not capable of treating a supersonic fluid. However, I find that the accretion occurs via breeze solutions that are sufficiently subsonic to not be of concern in the accretion regime.

The difference between requiring a breeze solution or a transonic solution depends on the fluid pressure at the outer boundary. In Owen and Wu (2016), only the transonic solution was considered due to the vacuum boundary condition and the inherent assumption of an instantaneous disc dispersal. In reality, the flow will increase in launch velocity and transition from breeze solutions to the transonic wind as the disc gradually disperses. The faster this process happens, the quicker outflow will tend to the transonic solution. In order to model this self-consistently, consider the density profile $\rho_{\rm HSE}$ in the case of hydrostatic equilibrium from Eq. 2.2.10, if applied to the isothermal layer beyond the MESA domain:

$$\rho_{\rm HSE} = \rho_{\rm BC}\exp\left\{\frac{GM}{c_{\rm s}^2}\left(\frac{1}{r} - \frac{1}{R_{\rm BC}}\right)\right\}. \tag{5.2.4}$$

Strictly speaking, this is a solution to the fluid equations if $u_{\rm crit} = 0$. Now, if one manually decreases the density at the Bondi radius such that $\rho(R_{\rm B}) < \rho_{\rm HSE}(R_{\rm B})$ (which in this problem is akin to reducing the gas density of the disc) then $u_{\rm crit}$ will increase in an attempt to return to equilibrium. As shown in Figure 5.5, if $\exp(-0.5) \leq \rho(R_{\rm B})/\rho_{\rm HSE}(R_{\rm B}) < 1$, then the solution will remain as a subsonic breeze (Lamers and Cassinelli, 1999). If the density is reduced below this value however, the flow is no longer pressure-supported and the only viable solution is transonic wind/accretion.

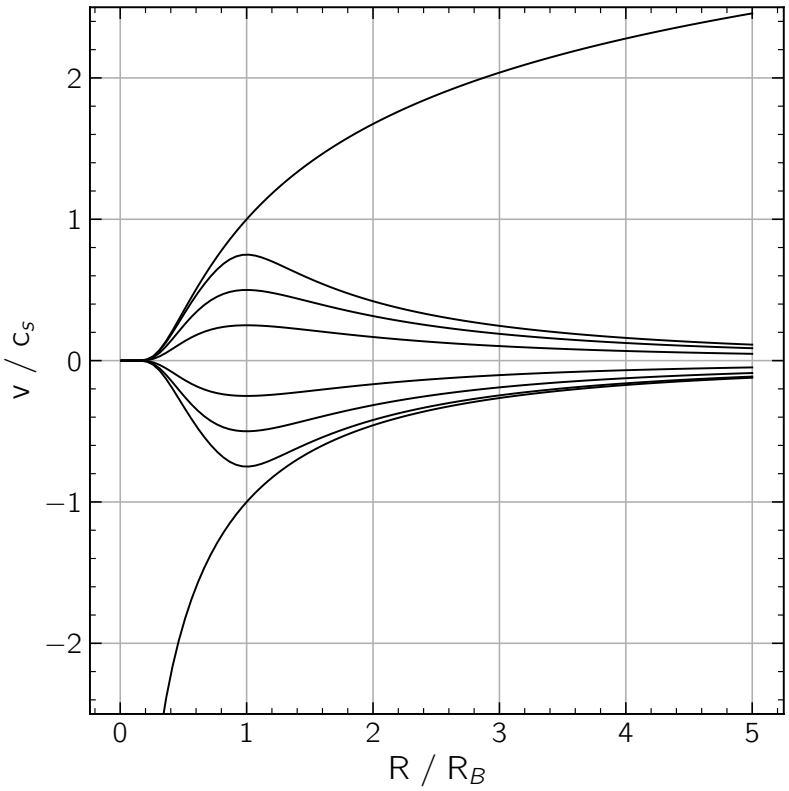


Figure 5.4: Velocity profiles for physical solutions to the isothermal fluid equations in non-hydrostatic equilibrium are shown with radius. Here, the six breeze solutions are those that remain subsonic at all positions i.e. $v/c_c < 1$ for all R/R_B. Positive and negative velocities correspond to mass-loss and accretion respectively. The two transonic solutions are those that exceed the sound speed above/below the Bondi radius R_B. In the simulations discussed in this Chapter, planets can accrete/lose mass via any of these solutions.

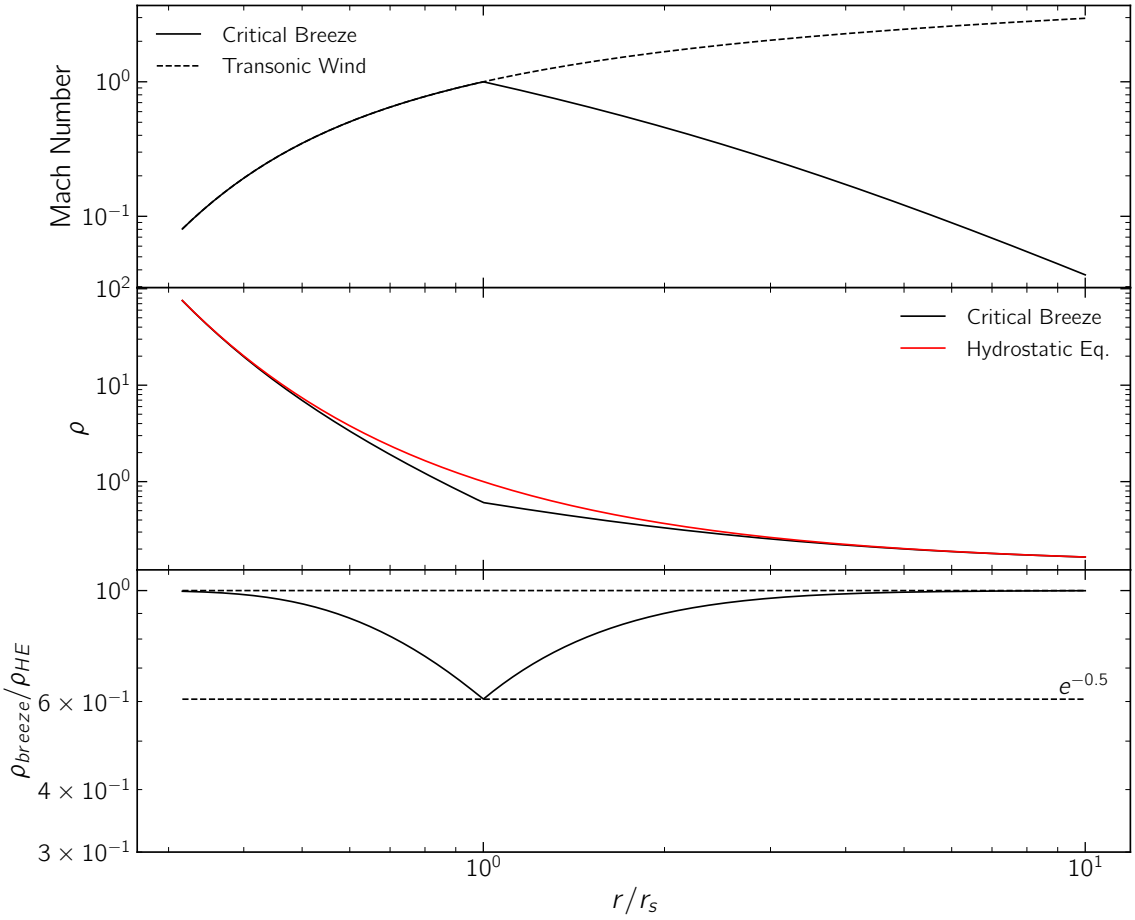


Figure 5.5: In the top panel, the dashed line shows Mach number for the transonic wind solution as a function of distance, relative to the sonic point ($r_s = R_B$). The solid line shows a critical breeze solution, for which the Mach number $\mathcal{M} = 0.99$. In the middle panel, the density profile for this critical breeze solution is shown in solid black, as well as the density profile for the case of hydrostatic equilibrium in red. One can see that density is reduced for the breeze solution. The bottom panel shows the ratio for these two quantities, which falls to a value of $\exp(-0.5)$ *at the Bondi radius. If the ratio drops below this point, the breeze solution will transition into the transonic solution and is no longer pressure-supported.*

For this problem, it follows that for a given ratio of $\rho_{\rm BC}/\rho_{\rm disc}$ there exists a solution to the fluid equations in non-hydrostatic equilibrium with a specific $u_{\rm crit} \in [-c_{\rm s}, +c_{\rm s}]$. From mass-continuity at the two boundaries we have that:

$$R_{\rm BC}^2\,\rho_{\rm BC}\,u_{\rm BC} = R_{\rm B}^2\,\rho_{\rm disc}\,u_{\rm crit}, \tag{5.2.5}$$

Recall that there is only one solution to the isothermal fluid equations for a given $u_{\rm crit}$ i.e. solutions are unique, hence one can calculate $u_{\rm BC}$ for a given $u_{\rm crit}$. Rearranging:

$$\frac{u_{\rm BC}}{u_{\rm crit}} = \frac{\rho_{\rm disc}}{\rho_{\rm BC}}\left(\frac{R_{\rm B}}{R_{\rm BC}}\right)^2 \tag{5.2.6}$$

we may now solve this equation to find $u_{\rm BC}$. Since $u_{\rm BC}$ is a function of $u_{\rm crit}$, this equation must be solved numerically, in this case with Brent's method (Press et al., 1992). Finally, the mass-flux $\dot{M}_{\rm P}$ due to Parker-type solutions (including subsonic breezes) at the `MESA` outer boundary is:

$$\dot{M}_{\rm P} = 4\pi r_{\rm RB}^2 \rho_{\rm BC} u_{\rm BC}. \tag{5.2.7}$$

Naively, one can adopt this value for each timestep of the simulation, however this results in large, unstable oscillations in planet mass. This is because the value calculated for $\dot{M}_{\rm P}$ in Eq. 5.2.7 is an *instantaneous* mass-loss rate that is only appropriate for a set of boundary conditions at a given time. Recall however that `MESA` is an implicit numerical scheme, meaning that it takes large steps in time. The value of $\dot{M}_{\rm P}$ at the beginning of this timestep may therefore be very different to that at the end, particularly when noting that it is very sensitive to $\rho_{\rm BC}/\rho_{\rm disc}$. In fact, this sensitivity means that the value is a dramatic over-estimate, meaning that the planet loses too much mass and then requires accretion to return to pressure support. This unstable oscillation continues, resulting in meaningless results. To remedy this, I adapt `MESA` with the following implicit mass-loss scheme.

Firstly, `MESA` takes a step forward in time with $\dot{M} = 0$. The state of the system at the end of this timestep is then used to calculate $\dot{M}_{\rm P}$, which represents the mass-loss that the Parker-type solutions can provide[3]. Then, `MESA` goes back and repeats the same timestep with different values of $\dot{M}$ until it converges on a value such that the system has returned to quasi-hydrostatic equilibrium i.e. $|(\rho(R_{\rm B})/\rho_{\rm disc}) - 1| < \epsilon$, where $\epsilon = 10^{-3}$ is the numerical tolerance. This mass-loss rate is labelled as $\dot{M}_{\rm HSE}$. Finally, `MESA` performs the timestep one more time adopting the following mass-loss rate:

$$|\dot{M}| = \min\{|\dot{M}_{\rm P}|, |\dot{M}_{\rm HSE}|\}. \tag{5.2.8}$$

In other words, if the Parker-type solutions can provide enough mass-flux to return to quasi-hydrostatic equilibrium, then `MESA` adopts $\dot{M}_{\rm HSE}$. If this is not the case, then the system cannot return to equilibrium and $\dot{M}_{\rm P}$ is adopted. Although being more computationally expensive i.e. multiple evaluations are required for a single timestep, this scheme provides a stable numerical approach to self-consistent accretion and mass-loss. Note that once the threshold of $|\dot{M}_{\rm P}| < |\dot{M}_{\rm HSE}|$ is crossed, typically in the latter stages of disc dispersal, then one need not calculate $\dot{M}_{\rm HSE}$ which reduces the computationally cost of each timestep.

Now that mass-loss rate is determined for each timestep, the density structure beyond the `MESA` domain can be calculated, allowing one to evaluate the optical depth to outgoing radiation as required in the `MESA` outer boundary condition (see Section 5.2.2). For each evaluation of the model at a given timestep, the adopted value of $\dot{M}$ provides $u_{\rm BC}$. Therefore, one can calculate the density profile in the isothermal layer from mass continuity:

$$\rho(r) = \rho_{\rm BC}\left(\frac{R_{\rm BC}}{r}\right)^2 \frac{u_{\rm BC}}{u(r)}, \tag{5.2.9}$$

[3]This value is in-fact averaged with that of the previous timestep so as to better estimate how $\dot{M}$ evolves with time.

where $u(r)$ is calculated using Lambert W function for a given $u_{\rm BC}$. The optical depth contribution from this isothermal layer is thus:

$$\tau_{\rm iso} = \int_{R_{\rm BC}}^{R_{\rm B}} \rho(r)\,\kappa(r)\,dr, \tag{5.2.10}$$

where I also adopt the dust-free opacities from Freedman et al. (2008) evaluated at $T_{\rm eq}$ and $\rho(r)$. The contribution to optical depth from the the disc is approximated to be $\tau_{\rm disc} \approx \kappa_{\rm disc}\Sigma_{\rm g}/2$ where $\kappa_{\rm disc} = 1$ cm^2 g^{-1} (e.g. D'Alessio et al., 2001). Note that this choice in disc opacity has little effect on the evolution. Finally, the total optical depth to outgoing radiation beyond the `MESA` boundary is:

$$\tau_{\rm BC} = \tau_{\rm disc} + \tau_{\rm iso} = \frac{\kappa_{\rm disc}\Sigma_{\rm g}}{2} + \int_{R_{\rm BC}}^{R_{\rm B}} \rho(r)\,\kappa(r)\,dr. \tag{5.2.11}$$

In addition to a contribution to optical depth, the isothermal layer also adds a contribution to the atmospheric mass. Recall that the outer boundary $R_{\rm BC}$ for the `MESA` domain is defined at a fixed temperature and pressure, rendering it meaningless in terms of a physical radius or mass of the planet. The true mass is calculated as the sum of the mass inside the `MESA` domain and the mass in the isothermal layer out to the Bondi radius. The planet radius is approximated by the photospheric radius, which is determined in the same way as in Section 2.6 i.e. locating the pressure such that $P = 2g/3\kappa$, as derived in Chapter 2. Note that whilst the planet is immersed in the protoplanetary disc, this radius is also physically meaningless since the disc is also optically thick to out-going radiation.

In all of this argument, I have assumed that the outflow beyond $R_{\rm BC}$ is isothermal and constant with time. As argued in Owen and Wu (2016), this is likely to be (at least approximately) true since the flow will be heated up to $T_{\rm eq}$ by the stellar heating if exposed, or to the disc mid-plane temperature if not. I comment on this approximation in Section 5.4.4

5.2.4 Building Planets

Constructing initial models in `MESA` for young planetary atmospheres is a non-trivial process. Since the original design of the code is for stellar structure, the routines for building models are designed for such a purpose. Constructing a low-mass, low-luminosity and yet highly inflated body with a solid core therefore requires multiple steps, which are outlined below and follows previous works of Owen and Wu (2013, 2016); Chen and Rogers (2016); Kubyshkina et al. (2020).

1. **Initial model**: To begin, the `create_initial_model` routine is used to produce an initial H/He dominated model of $2M_{\rm Jup}$ and $6R_{\rm Jup}$. I choose a hydrogen mass fraction of 0.74, helium mass-fraction of 0.24 and metal mass-fraction of 0.02 to reflect typical nebular conditions. This model is evolved for 10^{12} yrs for model stability and to ensure that mass can be easily removed in the Step 3 to produce a small mass planet.

2. **Add core**: Next, the `relax_core` routine is used to add an inert core of mass $M_{\rm p}$. I adopt a bulk core density for a $1M_\oplus$ core of 5.0 g cm^{-3} to reflect the findings of Chapter 4 and use the mass-radius relationships of Fortney et al. (2007) to determine the core radius $R_{\rm c}$. Note that `MESA` treats the core as inert and does not model its interior.

3. **Choose initial atmospheric mass fraction**: To choose an atmospheric mass fraction X, the `relax_mass` routine is used to reduce the initial $2M_{\rm Jup}$ model to a desired value. Note however, that this value is arbitrary at this point, since I later allow the system to come into equilibrium with the disc and then accrete material prior to disc dispersal. It is the atmospheric mass fraction at this point that is of physical interest to the problem.

4. **Impose outer boundary condition**: At this point, I switch from the standard photospheric boundary condition implemented in `MESA` (see Paxton et al., 2011) to one with a fixed pressure and temperature as described in Section 5.2.2. This is prescribed with the `other_surface_PT` routine. At this point I also perform tests to ensure this boundary is in an optically thick, subsonic and isothermal region (see Figure 5.3).

5. **Inflate and irradiate**: To simulate a young planet, one must now increase the model entropy from the value determined in Step 1. To do so, I use the `other_energy` routine to introduce a core luminosity that acts to inflate the planet. This can be interpreted as the accretion luminosity, $L_{\rm acc}$, that would have been present during core formation (see Section 5.1.2). Since the planet will come into equilibrium with the disc in the next step, I choose this luminosity such that $\rho(R_{\rm B})$ in the isothermal layer beyond the `MESA` domain is close to $\rho_{\rm disc}$. In addition, I also irradiate the planet at a temperature of $T_{\rm eq}$. To do so, I follow Owen and Wu (2016) and implement the `set_irradiation` $F_* - \Sigma$ routine which deposits a total irradiation of $F_*/4$ through a column density of $\Sigma = 250$ g cm^{-2} which is appropriate for an opacity of $\kappa = 4 \times 10^{-3}$ cm^2 g^{-1} to the incoming stellar flux, the value suggested by Guillot (2010) for sun-like stars. Note that, in reality, I need not irradiate the planet at the beginning of the evolution since the disc is optically thick to incoming stellar irradiation. Instead, the thermal energy from the disc is inherently modelled via the temperature imposed at the outer boundary. However, note that the majority of mass-loss occurs once the disc becomes optically thin. Hence, as a simplification of the model, I assume the planet is irradiated for the entire simulation.

6. **Turn on accretion/mass-loss**: The model must now come into equilibrium with the disc via accretion/mass-loss. The routine, as described in Section 5.2.3, is implemented in the `extras_check_model` since multiple evaluations of the model are required before continuing to the next timestep. I additionally switch on the hydrodynamic fluid terms with `change_v_flag` to allow advection. During this step, $\rho(R_{\rm B})$ in the isothermal layer beyond the `MESA` domain is equalised to $\rho_{\rm disc}$, which in turn is determined by the disc parameters from Eq. 5.1.15.

7. **Turn off core luminosity and accrete**: The final preparation step is to switch off the core luminosity from Step 5 and to allow gas accretion to commence. This is allowed to happen for at least one cooling timescale $t_{\rm KH}$ before disc dispersal begins such that the planet 'forgets' its initial conditions. The disc begins to disperse at $t_{\rm disp}$, at which point the model can self-consistently switch from accretion to mass-loss.

The resulting model is representative of a young planet that has formed and accreted a H/He dominated atmosphere whilst immersed in a protoplanetary disc. The planet parameters at this point are the core mass, $M_{\rm c}$, initial atmospheric mass-fraction, $X_{\rm init}$, cooling timescale $t_{\rm KH}$, equilibrium temperature, $T_{\rm eq}$, semi-major axis, a, and planet age $t_{\rm pl}$. Note that this latter variable is in-fact a crude estimate of the actual planet age, but instead represents the point at which the accretion process begins. In essence then, it is a variable that controls the initial atmospheric mass-fraction for a planet before dispersal begins. Finally, the disc parameters for each model are the initial gas surface density, $\Sigma_{\rm g,0}$, the time at which the disc starts to disperse, $t_{\rm disp}$, and the dispersal timescale, $\tau_{\rm disp}$. The gas surface density is modelled as follows:

$$\Sigma_{\rm g} = \begin{cases} \Sigma_{\rm g,0} & t < t_{\rm disp} \\ \Sigma_{\rm g,0}\ \exp\{\frac{1}{\tau_{\rm disp}}(t_{\rm disp} - t)\} & t \geq t_{\rm disp}, \end{cases} \tag{5.2.12}$$

which was chosen so as to approximately follow the disc dispersal models of Owen et al. (2010).

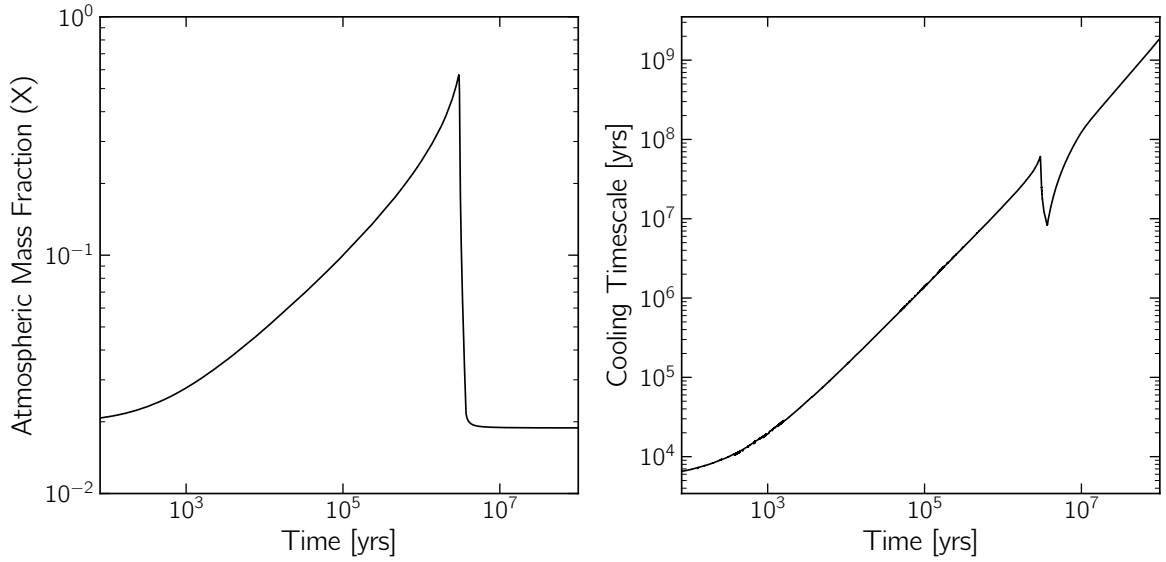


Figure 5.6: Accretion and boil-off evolution are shown for a planet with a $5M_{\oplus}$ core at $0.1AU$, with equilibrium temperature of $1000K$. The disc disperses at $t_{disp} = 3 \times 10^6$ yrs with a dispersal timescale of $\tau_{disp} = 10^5$ yrs. The planet is allowed to accrete for the entire disc lifetime i.e. $t_{pl} = 3 \times 10^6$ at the onset of disc dispersal. The left-hand panel shows the atmospheric mass fraction, whilst right-hand panel shows the cooling timescale. During disc dispersal, the cooling timescale drops, although is not an accurate measure of the planet's entropy at this time since it is not in equilibrium.

5.3 Results

Figure 5.6 demonstrates the effectiveness of the model, applied to a $5M_{\oplus}$ core at 0.1AU, with an equilibrium temperature of 1000K. The disc disperses at $t_{\rm disp} = 3 \times 10^6$ yrs with a dispersal timescale of $\tau_{\rm disp} = 10^5$ yrs. The planet is allowed to accrete for the entire disc lifetime i.e. $t_{\rm pl} = 3 \times 10^6$ at the onset of disc dispersal. One can see that the planet accretes an atmospheric mass fraction of $\sim 20\%$, before undergoing significant mass-loss as a result of the surrounding disc accreting onto the host star. The final atmospheric mass-fraction is $\sim 2\%$, which is in good agreement with the peak in Figure 4.16 from Chapter 4 i.e. the atmospheric masses inferred from the XUV photoevaporation model. I comment on this in Section 5.4.3.

As noted by Owen and Wu (2016); Owen (2020), the cooling timescale is a useful and intuitive measure of a planet's entropy: the longer this timescale, the more the planet has cooled and therefore implies a lower entropy. Figure 5.6 shows this timescale as a function of time. Overall, it increases as the planet cools, as one would expect. However, one can see that, during disc dispersal, this timescale in-fact decreases. Note that during this period, the atmosphere is not in equilibrium and undergoes dramatic advection, which changes the luminosity structure of the planet. Since there is an inherent assumption of hydrostatic equilibrium in the calculation of cooling timescale, its value at this time should not be trusted, since there is no way of calculating it at this time. To avoid this complication, I choose to interpret the thermodynamic state of the atmosphere based on the actual entropy at the base of the atmosphere for the remainder of this Chapter. Nonetheless, after this period, the cooling timescale is $\sim 10^8$ yrs within $\sim 10^6$ yrs after disc dispersal, which is in good agreement with Owen and Wu (2016), and confirms that planets will have prematurely cooled during disc-dispersal.

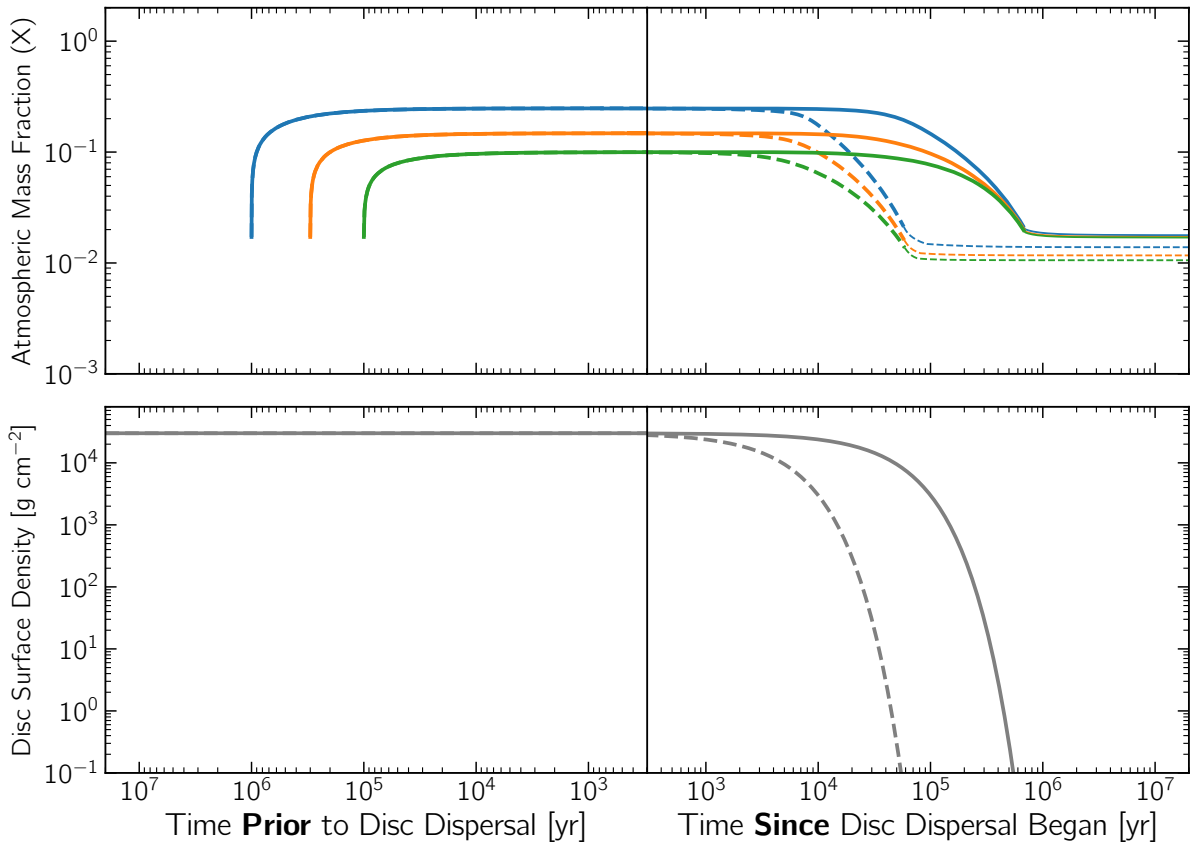


Figure 5.7: Accretion and boil-off are shown for a planet with core mass of $5M_\oplus$ core at $0.1AU$, with an equilibrium temperature of $1000K$. The disc disperses at $t_{disp} = 3 \times 10^6$ yrs, with time prior to this in left-hand panels, and time since dispersal begins in right-hand panels. Top panels show the atmospheric mass fraction, with planets forming at 10^6 yrs, 3×10^5 yrs and 10^5 yrs prior to disc dispersal shown in blue, orange and green respectively. Solid and dashed lines are for disc dispersal timescales of 10^5 yrs and 10^4 yrs respectively, with disc gas surface density shown in lower panels.

5.3.1 Formation time vs. disc dispersal timescale

Intuitively, it makes sense that the disc dispersal timescale will control the amount of atmospheric mass that will be lost during disc dispersal. To investigate this, Figure 5.7 shows the evolution of a $5M_\oplus$ core at 0.1AU, with an equilibrium temperature of 1000K. The disc disperses at $t_{\rm disp} = 3 \times 10^6$ yrs, but with two separate timescales: $\tau_{\rm disp} = 10^4$ yrs and 10^5 yrs. In addition, the planet is formed at different times: 10^6 yrs, 3×10^5 yrs and 10^5 yrs prior to disc dispersal. As a result, these planets have a range in initial atmospheric mass fractions since they have had different amount of time to cool and accrete. Figure 5.7 reveals that, despite having initial atmospheric mass-fractions of $\sim 10\%$, 15% and 20%, the final atmospheric mass fraction is determined by the disc dispersal timescale and not by the atmospheric mass prior to disc dispersal. Planets that undergo dispersal timescales of $\tau_{\rm disp} = 10^4$ yrs and 10^5 yrs are left with 1% and 2% atmospheric mass fractions respectively. Thus, for a given core mass and dispersal timescale, planets will all end with approximately the same atmospheric mass fraction. This in-turn will produce a peaked distribution in atmospheric mass fraction, such as that seen in Figure 4.16 from Chapter 4. Note however, that planets undergoing a quicker dispersal also present themselves with a larger dispersion in atmospheric mass fractions, which I comment on in Section 5.4.3.

The previous result suggests that formation history of the planet may be removed during disc dispersal, whereas the disc history, specifically its dispersal timescale, may be imprinted upon

the planet instead. To investigate this claim, Figure 5.8 shows the evolution for Mach number at the Bondi radius and entropy at the base of the atmosphere i.e. at the core radius, for the same planets in Figure 5.7. The Mach number evolution demonstrates how the disc dispersal timescale controls the boil-off process. Note that bold lines represent the case in which the planet can maintain quasi-hydrostatic equilibrium with the disc via mass-loss. As the disc begins to disperse, the outflow presents as low Mach number breeze solutions, with initial values dependent on dispersal timescale and, to a smaller degree, initial atmospheric mass fraction. However, as the disc density drops further, the Mach number must increase to maintain pressure support at a rate that is completely controlled by dispersal timescale i.e. the Mach number converges to values only dependent on disc dispersal timescales. Eventually, the Mach number must rise so high that equilibrium cannot be maintained, which is represented by thin lines in Figure 5.8. Note that this switch occurs close to, but not exactly, at the point when the the outflow changes to the transonic solution with unity Mach number at the Bondi radius.

Intuitively, the disc dispersal rate must control the outflow speed since the ratio of planet density and disc density controls the Mach number and therefore mass-loss rate (see Eq. 5.2.6). If a planet were to have too much mass removed at a given time, then mass-loss would slow until the disc disperses to a sufficient level. Mass-loss would then resume at the rate determined by disc dispersal. This therefore highlights an important result from this work, that the breeze solutions allow the planet to remain in quasi-hydrostatic equilibrium and also allow the disc to dictate the final atmospheric mass fraction. In essence this then builds upon the separate effects from Ikoma and Hori (2012) and Owen and Wu (2016), which I discuss in Section 5.4.2

Figure 5.8 also shows the entropy at the base of the atmosphere as a function of time after disc dispersal begins. Planets with an initially larger atmospheric mass fraction have been accreting, and therefore cooling for longer, meaning that their initial entropy is lower. However, planets that are subject to a quicker disc dispersal also cool quicker since the boil-off process requires advection to remove material upwards, which in turn removes energy from the atmosphere. As a result, planets undergoing a quicker disc dispersal converge to a smaller value in entropy. Although being a very small effect, it is also interesting to note that, for a given disc dispersal timescale, planets with the initially largest entropy (and therefore smallest initial atmospheric mass fractions) evolve to have the smallest entropy i.e. the ordering of entropy switches before and after boil-off. This can be understood since planets with the smallest atmospheric masses will be the most efficient at cooling, meaning that they will end up with a smaller entropy than those with a larger atmospheric mass. Note the absolute differences in entropy between models of different disc dispersal timescales and initial atmospheric mass fractions are very small, which is unsurprising since entropy is a logarithmic variable $S = k_{\rm B} \log \Omega$, where Ω is the number of available microstates of the system. Nonetheless, when considering the cooling timescale, which is a proxy for planet entropy, the final value for all models in Figure 5.8 is $t_{\rm KH} \sim 10^8$ yrs after 3×10^6 yrs after disc dispersal, consistent with Owen and Wu (2016).

To further understand how mass-loss ensues in the boil-off regime, Figure 5.9 shows the mass-loss rate as a function of planet size, specifically the location of the `MESA` outer boundary relative to the Bondi radius. Although this radius is defined at a fixed pressure and temperature, and does not represent a physical size of the planet (see Section 5.2.2 for details), it is a good measure of how the atmospheric structure and physical size changes with time. This is shown for the same fiducial models shown in Figures 5.2 and 5.8. As also found in Owen and Wu (2016), one can see that the mass-loss quickly shuts off once the planet shrinks to $\sim 0.1 R_{\rm B}$. As discussed, this is understood since the Mach number at $\sim 0.1 R_{\rm B}$ for any hydrodynamic outflow solution is very small, as shown in see Figure 5.4. One can also see that, as confirmed in Figure 5.8, planets with a faster disc dispersal undergo a faster mass-loss rate. In addition, while the planet is still large, planets with an initially larger atmospheric mass fraction experience a larger mass-loss rate, before converging to a rate determined by the disc dispersal rate. Note that there is a

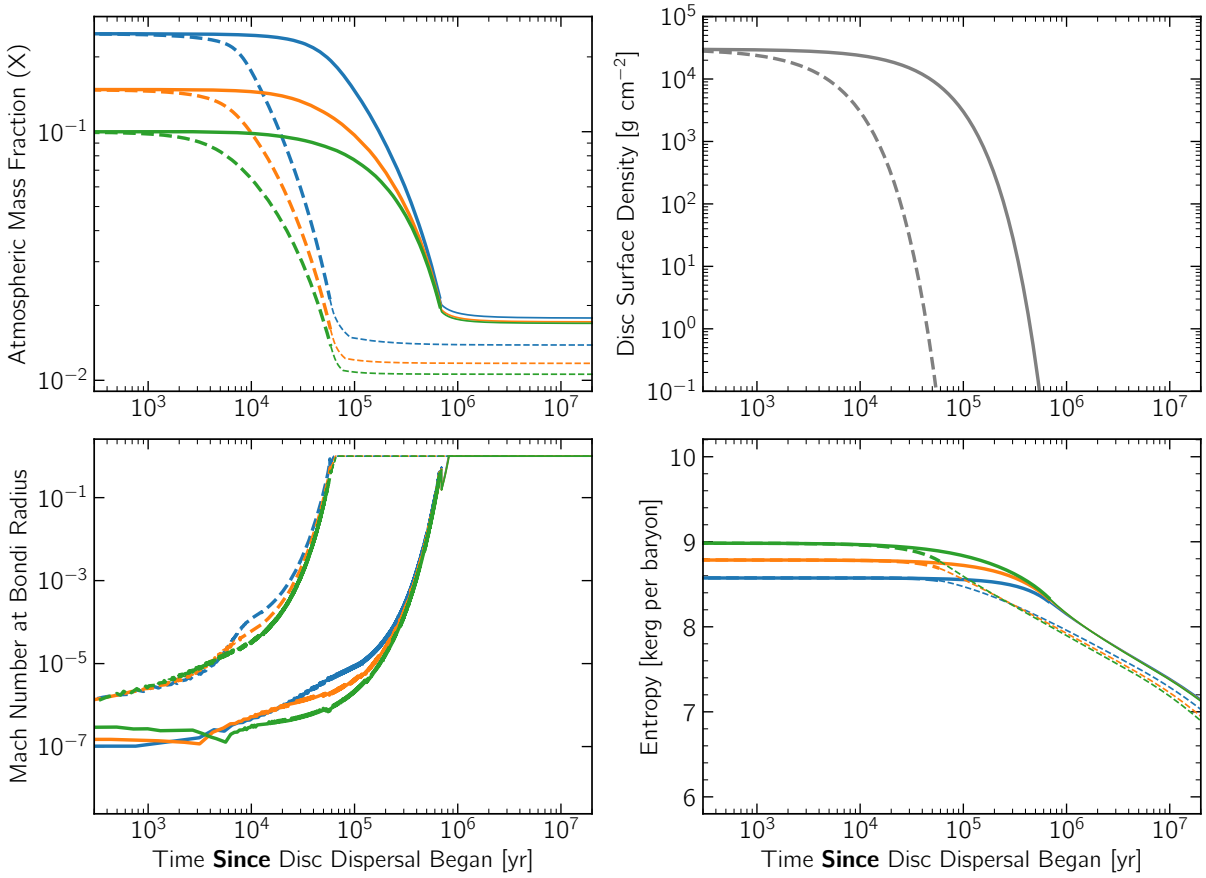


Figure 5.8: Accretion and boil-off are shown for a planet with core mass of $5M_\oplus$ core at $0.1AU$, with an equilibrium temperature of $1000K$. The disc disperses at $t_{disp} = 3 \times 10^6$ yrs, with time measured since this point. Upper-left and upper-right panels are the same as upper-right and lower-right from Figure 5.7. The lower-left panels shows the Mach number at the Bondi radius of the outflow as a function of time, whilst the lower-right shows the entropy at the base of the atmosphere i.e. at the core radius. Planets forming at 10^6 yrs, 3×10^5 yrs and 10^5 yrs prior to disc dispersal are shown in blue, orange and green respectively. Solid and dashed lines are for disc dispersal timescales of 10^5 yrs and 10^4 yrs respectively. In addition, thick lines show where the atmosphere is able to maintain quasi-hydrostatic equilibrium with the disc, whereas thin lines show the opposite.

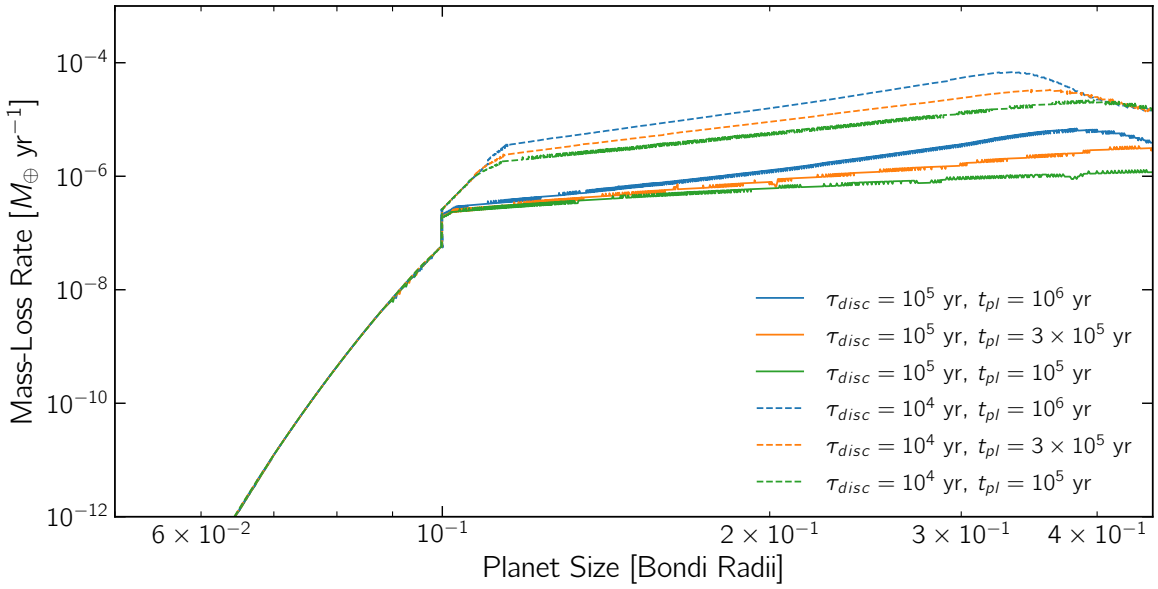


Figure 5.9: Mass-loss rate is shown as a function of planet size in units of Bondi radius, for a planet with core mass of $5M_\oplus$ core at $0.1AU$, with an equilibrium temperature of $1000K$. Planets forming at 10^6 yrs, 3×10^5 yrs and 10^5 yrs prior to disc dispersal are shown in blue, orange and green respectively. Solid and dashed lines are for disc dispersal timescales of 10^5 yrs and 10^4 yrs respectively.

small, yet sudden drop in mass-loss rate at exactly $0.1R_\mathrm{B}$. Although this has little effect on the overall evolution, this arises due to a numerical limitation, which is discussed in Section 5.4.4.

5.3.2 Orbital separation vs. Core mass

Thermodynamically speaking, planets further out in the disc will be at lower equilibrium temperatures and can therefore cool more efficiently, allowing accretion to occur at a quicker rate. In the same vain, a boil-off process will be less important further out in the disc since it is the irradiation from the host star, and therefore also the equilibrium temperature, that actually removes the material from the planet's influence. However, both accretion and boil-off are also strongly controlled by the core mass, since this sets the gravitational potential and thus the ease with which mass can enter/leave the planet. To investigate this, I perform a small parameter study for planets with core masses of $3M_\oplus$, $5M_\oplus$ and $10M_\oplus$, in two locations in the disc: 0.1 AU with equilibrium temperature of 1000 K and initial disc gas surface density of 3×10^4 g cm^{-2}; and 0.3 AU with equilibrium temperature of 500 K and initial disc gas surface density of 1×10^4 g cm^{-2}. Results are shown in Figure 5.10. Note that I have chosen to form the $3M_\oplus$, $5M_\oplus$ and $10M_\oplus$ core mass planets at 10^6 yrs, 10^5 yrs and 10^4 yrs prior to disc dispersal respectively, so that the initial atmospheric mass fractions are comparable. In reality, if all planets were formed at the same time, the planet with the largest core would accrete a significantly larger atmospheric mass fraction. A clear trend can be seen that planets in the outer, and therefore cooler, disc accrete an atmospheric mass fraction that is $\sim 50\%$ more massive than their hotter counterpart in the inner disc. Indeed, one can see the $10M_\oplus$ core at 0.3AU eventually undergoes runaway accretion before the disc has dispersed significantly, at which point the simulation is stopped. As summarised in Figure 5.11, higher mass cores lose a smaller fraction of their atmospheric mass due to their deep gravitational wells. In addition, planets at further orbital separations lose a smaller fraction of their atmospheric mass since they are subject to lower temperatures/irradiation at such distances. For comparison, I also include results from Owen and Wu (2016). Note that these models were performed at 500 K

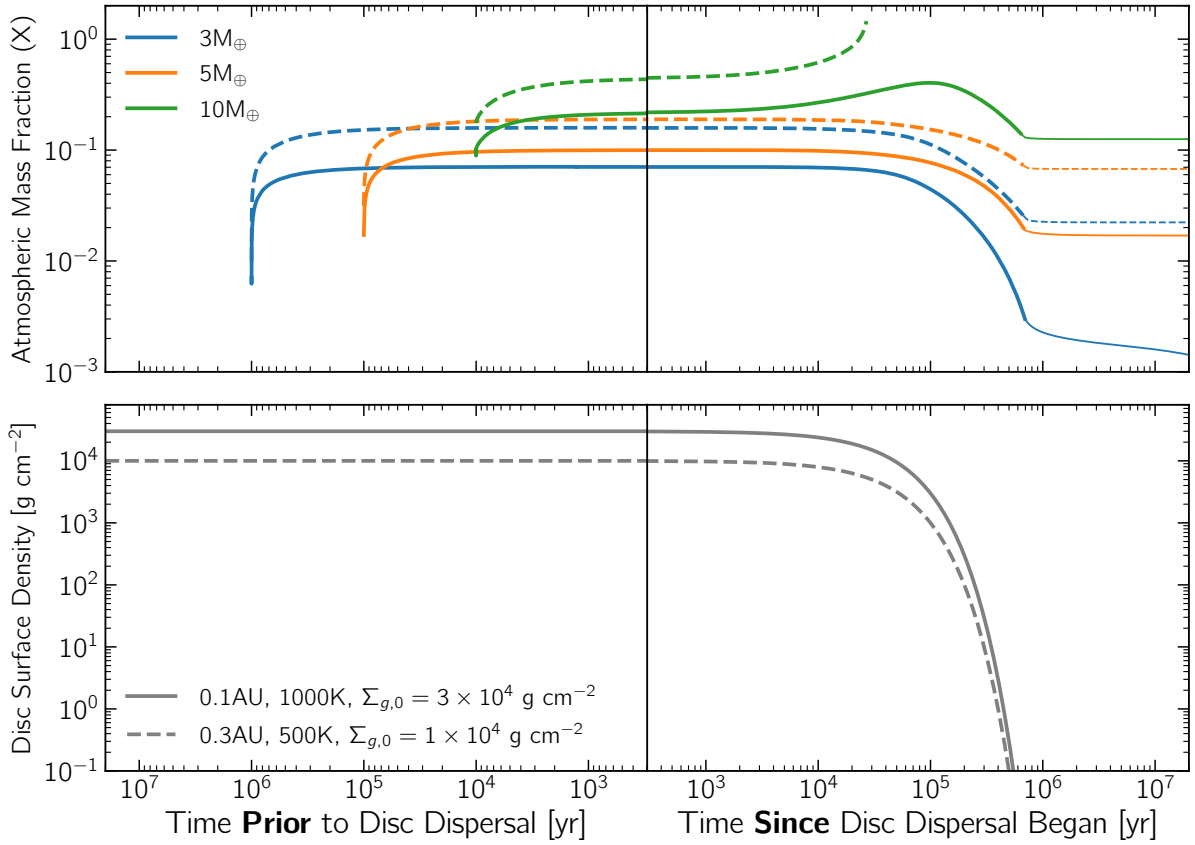


Figure 5.10: In the upper-panel, the evolution of atmospheric mass fraction is shown for planets with core masses of $3M_\oplus$, $5M_\oplus$ and $10M_\oplus$ in blue, orange and green respectively, which are formed at 10^6 yrs, 10^5 yrs and 10^4 yrs prior to disc disc dispersal respectively. Solid lines represent planets at 0.1 AU with equilibrium temperature of 1000 K and initial disc gas surface density of 3×10^4 g cm^{-2}, whereas dashed lines represent planets at 0.3 AU with equilibrium temperature of 500 K and initial disc gas surface density of 1×10^4 g cm^{-2}. The disc surface density evolution at these two locations is shown in the lower-panel.

and 900 K, as opposed to 500 K and 1000 K in this study. Nonetheless, one can see that the new models confirm the prediction that the mass-loss rates from Owen and Wu (2016) were indeed upper-limits, due to the assumptions of instantaneous disc dispersal and maximally entropic initial conditions.

5.4 Discussion

These models highlight the importance of disc dispersal in setting the initial conditions for planet evolution. They also demonstrate the need for self-consistent modelling: from gas accretion all the way to irradiation from the host star. I have shown that planets will typically accrete $\gtrsim 10\%$ in atmospheric mass before disc dispersal begins, although this exact value depends on core mass, orbital separation and the time at which the planet formed. This is consistent with previous works on gas accretion (e.g. Ikoma and Hori, 2012; Lee et al., 2014; Lee and Chiang, 2015a; Ginzburg et al., 2016). In addition, the time at which the planet forms also sets the initial entropy (and therefore cooling timescale) of the planet. However, one of the major results of this study is that, despite different initial entropies and atmospheric mass fractions, this information is almost completely removed during disc dispersal. In these models, it is the disc dispersal rate that is imprinted on the final atmospheric mass fraction of the planet, with

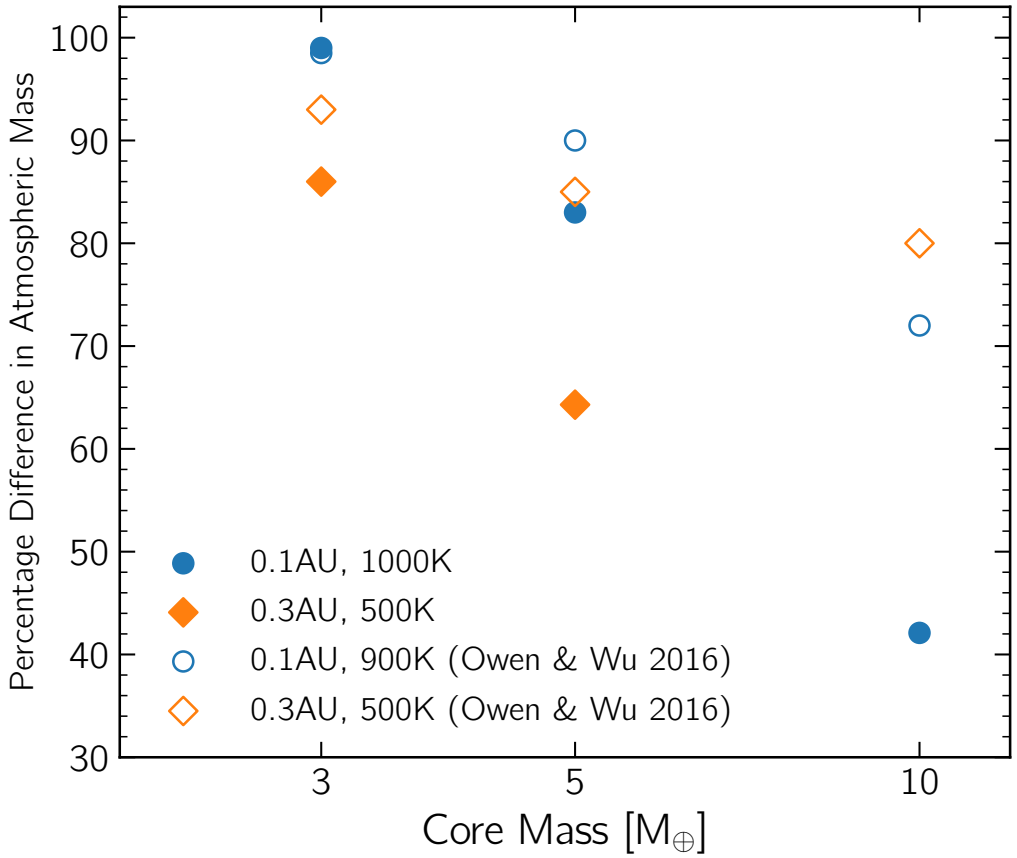


Figure 5.11: The percentage difference in atmospheric mass due to boil-off is shown for varying core masses. Filled makers refer to models performed in this study, whereas open markers refer to results from Owen and Wu (2016). Blue circles represent planets at $0.1AU$ *with equilibrium temperature of* $1000K$ *(*$900K$ *for Owen and Wu (2016)) and initial disc gas surface density of* 3×10^4 *g* cm^{-2}*, whereas orange diamonds represent planets at* $0.3AU$ *with equilibrium temperature of* $500K$ *and initial disc gas surface density of* 1×10^4 *g* cm^{-2}*. Note that there is no value for* $10M_\oplus$ *at* $500K$ *since this model underwent runaway accretion before disc dispersal.*

quicker dispersal times inducing a larger atmospheric mass change. I now discuss the implication of these findings.

5.4.1 Boil-off sets the initial conditions for photoevaporation

One of the major motivations of this work was the discrepancy between core accretion models (e.g. Lee and Chiang, 2015b; Lee et al., 2018; Fung and Lee, 2018) and photoevaporation, as shown in Figure 4.16 from Chapter 4. As initially found by Jankovic et al. (2019), core accretion predicts that planets should accrete $\gtrsim 10\%$ in atmospheric mass during the disc lifetime, whereas the amount of mass that XUV photoevaporation can strip in order to be consistent with the *Kepler* data is $\sim 2\%$. Note that this discrepancy is likely not an indication that XUV photoevaporation underestimates mass-loss rates, since planets beyond orbital periods of ~ 30 days (which are not strongly effected by photoevaporation) have mass and radii measurements consistent with hosting a few percent in H/He mass, instead of $\gtrsim 10\%$ (e.g. Weiss and Marcy, 2014; Jontof-Hutter et al., 2016; Wolfgang et al., 2016). As I have previously discussed, possible mechanisms to resolve this tension include forming the planets at the very end of the disc lifetime (Ikoma and Hori, 2012; Lee and Chiang, 2016), or alternatively improving the accuracy of gas accretion models to include the effects of giant mergers (Liu et al., 2015; Inamdar and Schlichting, 2016) which would result in potentially significant atmospheric mass-loss and the production of heat which would take typically kyrs to disperse. The inclusion of 3D simulations has additionally shown that recycling of high-entropy gas during the accretion phase can act to reduce the final atmospheric mass of the planet (Ormel et al., 2015; Fung et al., 2015; Cimerman et al., 2017; Ali-Dib et al., 2020; Chen et al., 2020).

In this Chapter however, I have taken a different approach, which is to consider the mass-loss that is initiated by the rapid dispersal of the protoplanetary disc. Figures 5.6, 5.7 and 5.10 highlight that boil-off can offer a resolution to this discrepancy. Whilst planets are seen to accrete $\gtrsim 10\%$ in atmospheric mass, they rapidly lose this mass during disc dispersal to a value of $\sim 1\%$, which is consistent with the results from Chapter 4. This offers a potential for future physics-based inference, in which populations of planets can be modelled through the boil-off process in order to find the disc dispersal timescales and core masses required to fit the initial population of planets inferred from XUV photoevaporation in Chapter 4. This would be a key step in connecting planet formation models to the exoplanet population that we observe today. For now however, this is left for future work.

5.4.2 A Gentle Breeze

Previous studies that have considered mass-loss during disc dispersal have taken different approaches which yielded different results. In Ikoma and Hori (2012), accretion and mass-loss were considered in a hydrostatic equilibrium regime. They found that planets will continue to lose their entire atmospheric masses as the disc disperses, in order to maintain this equilibrium. Owen and Wu (2016), on the other hand, took a different approach. By considering the timescales of the problem, they argued that the planet will not be able to stay in hydrostatic equilibrium with the disc, since dispersal occurs on timescales of $10^4 - 10^5$ yrs (Kenyon and Hartmann, 1995; Ercolano et al., 2011; Koepferl et al., 2013), whereas the cooling timescale of a planet has upper limits of $t_{\rm KH} \sim 10^6 - 10^7$ yrs. In the case that hydrostatic equilibrium cannot be maintained, mass-loss will ensue via a hydrodynamic wind. Owen and Wu (2016) modelled this process in `MESA` (Paxton et al., 2011, 2013, 2015, 2018) and assumed an isothermal Parker wind Parker (1958), as well as an instantaneous disc dispersal. They also assumed the planet began in a maximally entropic state i.e. fully convective. Unlike Ikoma and Hori (2012), they found that planets do not lose all of their mass, since mass-loss shuts-off once the planets shrink to small sizes $\sim 0.1 R_{\rm B}$. They also found the planet prematurely cooled due to advection

removing high entropy material from the depths of the atmosphere. However, the assumptions that they made implied that their mass-loss rates were likely upper limits.

In this study, I have built upon the models of Ikoma and Hori (2012) and Owen and Wu (2016), by self-consistently modelling the planet through gas accretion and disc dispersal. Importantly, the planet is allowed to remain in a quasi-hydrostatic equilibrium with the disc via isothermal breeze solutions, as well as non-hydrostatic equilibrium if this is not physically possible. I find that the planets require physics from both Ikoma and Hori (2012) and Owen and Wu (2016) when introducing this self-consistent approach. Planets begin by cooling and accreting, before losing mass via breeze solutions, which allow the planet to sustain pressure support with the disc as it begins to disperse. Eventually however, the mass-loss that a breeze can provide is not sufficient to remain in equilibrium, so the planet quickly transitions to mass-loss via the transonic wind from Parker (1958) as the disc continues to completely disperse. Similar to Owen and Wu (2016) however, I find this mass-loss quickly shuts off once the planet has shrunk to $\sim 0.1 R_{\rm B}$. However, the mass-loss rates that I calculate are significantly lower than that of Owen and Wu (2016). As shown in Figure 5.11, the change in atmospheric mass is consistently lower for the models presented in this study, when compared to those of Owen and Wu (2016). This is especially noticeable for higher mass cores. Nonetheless, I also find that planets undergo premature cooling as a result of contraction and advection. As also found by Owen and Wu (2016), the typical cooling timescales for planets are $t_{\rm KH} \sim 10^8$ yrs after 3×10^6 yrs after disc dispersal. As shown in Owen (2020), this means that observations of young planets with measured masses, radii and ages may allow us to constrain their initial entropies and therefore determine whether they did in-fact undergo a boil-off phase.

5.4.3 Planet formation history is removed

One of the key findings from this study has been the impact of disc dispersal timescale. Despite planets forming at different times in the disc lifetime and therefore accreting different atmospheric mass fractions, the final state of the planet is predominantly controlled by the speed at which the disc density drops. This is because the speed at which this happens determines how much mass needs to be removed via a breeze in order to stay maintain quasi-hydrostatic equilibrium. Nonetheless, planets that start with a larger initial atmospheric mass fraction do experience a higher initial mass-loss rate. I find that the quicker the dispersal timescale, the more impact this has on the final state of the atmosphere (see upper left-hand panel of Figure 5.8). In the limit that the dispersal becomes instantaneous, it follows that planets with an initially larger atmospheric mass fraction should evolve to a larger final atmospheric mass fraction. This is indeed the conclusion that Owen and Wu (2016) came to under their assumptions, suggesting that my models are consistent with theirs.

In the realistic case however of a non-instantaneous disc dispersal based on observed timescales of $10^4 - 10^5$ yrs (Kenyon and Hartmann, 1995; Ercolano et al., 2011; Koepferl et al., 2013), these results suggest that the history of the planet i.e. the point at which it formed in the disc lifetime is removed. Instead, information about the rate of disc dispersal is imprinted, which will affect the distribution of atmospheric mass fractions that planets host after disc dispersal. As discussed in , this opens up the possibility of using demographic surveys of exoplanets to perform physics-based inference with boil-off and XUV photoevaporation / core-powered mass-loss models to place constraints on disc dispersal timescales.

5.4.4 Model Improvements

I have shown that self-consistent modelling of gas accretion and mass-loss can shed light on the atmospheric evolution of exoplanets during disc dispersal. However, there is one major assumption that I have made which limits this analysis: an isothermal outflow. This has been

assumed since the flow will be heated to $T_{\rm eq}$ by the stellar heating if exposed, or to the disc mid-plane temperature if not. As shown in Figure 5.3, the assumption of isothermality is valid in the radiative zone of the planet, however this may not be the case, particularly in optically thin regions at large distances from the planet. In assuming an isothermal outflow, I am inherently assuming that energy is inserted into the flow to keep is heated to $T_{\rm eq}$ and thus remove it from the planet's Hill sphere. I have also assumed that the planet is irradiated by the host star for the entire of the planet lifetime. Since most of the mass-loss occurs at the point when the disc becomes optically thin, this assumption is not a bad one. However, in reality, irradiation will increase as the disc disperses which may affect the mass-loss rates. These effects combine to imply that I am not considering another mechanism by which mass-loss may shut-off: that the flow cannot be heated to sufficient temperatures to leave the planet's influence. As discussed in 5.4.2, mass-loss in the current models shuts off due to the planet cooling and contracting. However, this effect may be compounded when considering non-isothermal outflows, to further reduce the effectiveness of the boil-off process. To remedy this, one potential option is to use hydrodynamic models that consider the radiative transfer and subsequent mass-loss in the upper atmosphere. These models can then be used to construct a grid of mass-loss rates, which can be then be implemented in `MESA`. This was indeed the approach of Owen and Wu (2013), applied to XUV photoevaporation, in using the mass-loss rates from hydrodynamic models from Owen and Jackson (2012). For now however, this is beyond the scope of this study, but is planned for future work.

In addition, this latter approach would fix a current numerical limitation of the adopted numerical scheme, as highlighted in Figure 5.9. Currently, mass-loss rates are calculated with Lambert W functions (e.g. Cranmer, 2004). However, these non-analytic functions are calculated with a Taylor expansion for small values of $R_{\rm BC}/R_{\rm B} \leq 0.1$. As a result, a small drop in mass-loss rate is seen in Figure 5.9, that is caused by the numerical scheme transitioning to this approximation. If one were to calculate mass-loss rates in the non-isothermal case, this problem would be bypassed, since Lambert W functions would no longer be appropriate.

Another simplification in this model was the simple analytic expression for disc dispersal (see Eq. 5.2.12), which assumes a constant dispersal rate as a function of time for all distances from the host star. In reality this is not the case since dispersal in the inner disc begins near the gap that is opened at $\sim$ 1AU due to the balance between stellar accretion and photoevaporation (see Section 5.1.3). As such, another future improvement to the model is to use hydrodynamic simulations of photoevaporative driven disc dispersal from works such as Owen et al. (2010); Winter et al. (2018); Coleman and Haworth (2022) to accurately model this process as a function of time and orbital separation.

Finally, an important effect that has been neglected is that of dust. Thus far, only dust-free gas has been considered, by-way of the opacity tables from Freedman et al. (2008). In reality, dust is also present in the protoplanetary disc and therefore the atmosphere of the planet. One can argue that for late stages of disc evolution the dust is not important since it will rapidly drain onto the host star once the inner and outer disc have been disconnected (e.g. Alexander and Armitage, 2007; Owen and Kollmeier, 2019). However, if dust is present, it will have two important effects of the atmospheric evolution: firstly dust acts to increase the efficiency of cooling of an optically thin outflow, hence decreasing the mass-loss rate; on the other hand it also prevents cooling of the optically thick atmosphere, meaning that planet contraction is slowed and mass-loss rates will increase. What is not clear, is which of these two competing effects will dominate. Thus, future work will also incorporate dust opacities in the atmosphere and outflow to investigate its importance.

5.4.5 Future direction

Once model refinements have been made, a full grid of models is required to truly understand how variables such a core mass, orbital separation and disc dispersal rate control the atmospheric properties of the planets after disc dispersal. After this, there are many avenues to explore in future work.

One such example is to note that, as discussed in Section 5.1.4, the boil-off process is in fact the precursor to core-powered mass-loss Ginzburg et al. (2016, 2018). Whilst the disc is dispersing, mass-loss can ensue because rapid contraction of the core provides the luminosity to advect material into the upper atmosphere such that it can be removed via bolometric irradiation. After this period however, mass-loss can continue if a core luminosity can sufficiently inflate the atmosphere, albeit on a much longer timescale of order ∼Gyrs (see Chapter 2). As such, core-powered mass-loss can be self-consistently modelled under the presented framework. In addition, if one were to incorporate XUV heating into the mass-loss rates, one could self-consistently model boil-off, core-powered mass-loss and photoevaporation. The interplay between these mechanisms could then be investigated to place theoretical constraints as to which is dominant. Comparisons can then be made between models and demographic surveys, as highlighted in Chapters 3 and 4.

5.5 Conclusion

In this Chapter I have constructed simulations of atmospheric evolution that self-consistently models gas accretion and mass-loss during protoplanetary disc dispersal. These models build upon previous attempts to understand the so-called 'boil-off' phase from Ikoma and Hori (2012) and (Owen and Wu, 2016), which assumed hydrostatic equilibrium and non-hydrostatic equilibrium with the disc respectively. In the models presented here, I combine both of these effects to better understand the boil-off process. My main conclusions are as follows::

- The boil-off phase is a viable solution to the discrepancy between gas core accretion and XUV photoevaporation. Despite planets accreting $\gtrsim 10\%$ in atmospheric mass before disc dispersal begins, the ensuing mass-loss removes a significant fraction of this, to typical values of $\sim 1\%$, which is in agreement with the results of Chapter 4.

- Planets remain in a quasi-hydrostatic equilibrium with the disc for most of the disc dispersal process and lose mass via subsonic breeze solutions. However, they eventually cannot maintain this equilibrium and mass-loss ensues via a hydrodynamic wind. Eventually, mass-loss shuts off once the planet has cooled and contracted to $\sim 0.1 R_{\rm B}$.

- The final atmospheric mass fraction of a planet is strongly controlled by the rate of disc dispersal since this informs how much mass should be lost to maintain quasi-hydrostatic equilibrium. As such, information on the initial properties of the planet, such as initial entropy and atmospheric mass fraction are lost.

- Planets of larger core masses will lose less atmospheric mass, owing to their deeper gravitational wells. In addition, planets at larger orbital separations exist at a lower equilibrium temperature and therefore will not lose as much mass during the boil-off phase.

One of the major assumptions made in this work was that of an isothermal, dust-free outflow. In reality, this may not be the case and material may significantly cool as it leaves the planet's atmosphere, which will reduce mass-loss rates. Future work will therefore address this assumption by accurately modelling the heating in the outflow region using hydrodynamic simulations.

A large parameter study will then allow one to fully understand the physics that goes into the initial conditions of planets after disc dispersal. In addition, this will allow self-consistent models of boil-off, core-powered mass-loss and XUV photoevaporation to be performed, yielding theoretical constraints on which mechanisms dominate atmospheric evolution.

Chapter 6

Summary and Outlook

In this thesis I have investigated the atmospheric evolution of small, close-in exoplanets and its consequences on planet formation. One of the key features in the exoplanet demographics is the radius gap found at $\sim 1.8R_{\oplus}$ (e.g. Fulton et al., 2017; Van Eylen et al., 2018; Berger et al., 2020a). I have considered how one might observationally distinguish between different sculptors of the radius gap, namely XUV photoevaporation (e.g. Owen and Wu, 2013; Lopez and Fortney, 2013) and core-powered mass-loss (e.g. Ginzburg et al., 2018; Gupta and Schlichting, 2019), as well as shown that atmospheric escape can be used to break the degeneracies of planet structure in the mass-radius plane. In addition, I have also shown that atmospheric mass-loss during protoplanetary disc dispersal may play a pivotal role in setting up the conditions of planets before undergoing late-stage mass-loss (e.g. Owen and Wu, 2016; Ginzburg et al., 2016). In this Chapter, I provide a brief summary of the findings from this thesis, as well as an outlook on the field with areas that warrant future thought.

6.1 Summary

In Chapter 2, I discussed the structure of small mass planets that host large H/He dominated atmospheres. I showed that, in the optically thick case, the atmosphere is separated into an inner convective region and an outer radiative layer. Since these atmospheres are collisional and subject to extreme irradiation from their host stars, they can undergo hydrodynamic mass-loss resulting in dramatic changes in their size. I presented the analytic model for XUV driven escape from Owen and Wu (2017) and core-powered mass-loss from Gupta and Schlichting (2019). These have been shown to accurately reproduce the results of the radius gap and provide a computationally cheap way of generating synthetic populations of planets which can be compared with demographic surveys, such as *Kepler*.

Next, in Chapter 3, I used these analytic models of mass-loss to present a new method of differentiating between these mechanisms. This arises due the fact that, at fixed bolometric flux, planets will receive more X-rays around lower mass stars under the XUV-driven model and hence the radius gap will systematically shift to larger radii. In the core-powered mass-loss model, mass-loss relies on the bolometric flux from the host star, meaning that this shift should not be present as fixed incident flux. To resolve the presence, or absence, of this feature in the data, one must analyse the radius gap in 3D parameter space, specifically with incident flux, stellar mass and planet size. I used synthetic populations from both models to create mock observations from the *Kepler* survey to show that surveys of $\gtrsim 5000$ planets were required to differentiate the models to a high statistical significance. I also showed that the presented methodology is robust to moderate false positive contamination of $\lesssim 10\%$, whilst reducing the measurement uncertainty and widening the stellar mass range of the survey have positive effects

on the analysis. I also performed the experiment on three existing subsets of the *Kepler* survey: CKS (Fulton et al., 2017), GKS (Berger et al., 2020b) and CKS-Cool (Petigura et al., 2022), all of which are consistent with both models. With future demographic surveys such as *PLATO* (Rauer et al., 2014) on the horizon, this model comparison can be used to determine which of these mechanisms is dominant. As shown in Chapter 3, reducing measurement uncertainty is key to extracting information from planet distributions. This is all the more important for planets found near or within the radius gap. I would therefore recommend future surveys to perform asteroseismic follow-up on such planets to allow analyses, such as that presented in Chapter 3, to yield a higher statistical significance.

In Chapter 4 I showed that atmospheric escape can be used to break the degeneracies of internal composition of small, close-in exoplanets. Under the assumption that XUV driven atmospheric escape is the dominant driver of mass-loss, I constructed a sophisticated Bayesian hierarchical inference model to infer the core mass, core composition and initial atmospheric mass fraction distributions of small, close-in exoplanets from the CKS sample, all three of which are hidden from standard survey techniques. The core mass function was shown to be peaked at $\sim 4M_{\oplus}$, hinting towards a planet formation mechanism that is preferentially forming planets this mass. The core composition distribution was shown to be extremely narrow and centred around an Earth-like silicate/iron composition. This provides hints that such planets form in the inner disc and not exterior to the water snow-line as is suggested in migration models. Finally, the initial atmospheric mass-fraction distribution was shown to strongly peak at $\sim 2\%$, which is typical for sub-Neptunes that host atmospheres that typically double their core size. This suggests that the radius gap forms via evolution and that planets are ~ 5 times more likely to evolve to a terrestrial nature i.e. negligible atmospheric mass fraction, than form in such a way. Finally, when considering the atmospheric mass that is accreted by planets given the inferred core mass distribution, I find that this is ~ 5 times larger than the inferred initial atmospheric mass fraction from this work. This suggests that either models for core accretion or photoevaporation need adjusting, or there is another mass-loss mechanism that occurs after gas accretion but before XUV-driven mass-loss to bring these two distributions into alignment.

Finally, in Chapter 5, I consider such a mass-loss mechanism which occurs during disc dispersal, referred to as 'boil-off'. I built upon evolution models from Ikoma and Hori (2012); Owen and Wu (2016) to show that planets may remain in a quasi-hydrostatic equilibrium with the disc as it begins to disperse. Eventually, the disc density drops to a level where the planet leaves equilibrium and loses mass via a hydrodynamic wind. This mass-loss quickly shuts off however, once the planet has contracted to $0.1R_{\rm B}$. I showed that the final atmospheric mass fraction of the planet is strongly controlled by core mass, orbital separation and disc dispersal rate. It is not, however, controlled by the time at which the planet formed within the disc. This latter statement suggests that the history of planet formation is removed during disc dispersal. Nonetheless, I confirm the findings of Owen and Wu (2016) and show that a planet prematurely cools during the boil-off phase due to rapid contraction and subsequent advection of high entropy material to the escaping upper atmosphere. Typical cooling timescale of such planets are found to be $t_{\rm KH} \sim 10^8$ yrs after 3×10^6 yrs after disc dispersal.

6.2 Outlook

In this Section, I discuss future avenues of research, in-light of the findings in this thesis. I also explore areas which need addressing if we are to gain a better understanding of atmospheric evolution.

6.2.1 Rapid Evaluation of Exoplanet Evolution Models

One of the main results from this thesis is the potential of physics-based inference analyses of exoplanet demographics. In Chapter 4, I exploited XUV photoevaporation to place constraints on the internal structure and evolutionary history of small, close-in exoplanets. However, this was an extremely computationally challenging task due to the number of models that need to be evaluated. For example, the results in Chapter 4 required the evaluation of $\sim 10^{10}$ planetary evolution models, consuming ~ 1 million CPU-hrs. To make this computationally feasible, the simplified, analytic evolutionary model from Owen and Wu (2017) was used. However, even in this work, it could only be applied over a narrow range of stellar mass and expected correlations (for example between core-mass and initial atmospheric mass) were not taken into account.

Thus, if one desires to use evolutionary models for large-scale population level statistical inferences, a new, more computationally efficient approach is required. This is all the more pertinent if one wishes to use accurate evolution models that include, for example a realistic equation of state, radiative transfer and self-gravity. Fortunately, machine-learning may provide an answer. Instead of solving the planetary structure and evolution equations for each planetary model, one can use a machine-learnt model to emulate the planet's evolution. Here I present a proof-of-concept machine-learnt emulator, trained on analytic models of XUV photoevaporation from (Owen and Wu, 2017), which demonstrate high accuracy and considerable computational speed-up.

The overall goal here is to predict the photospheric radius and atmospheric mass fraction $X \equiv M_{\rm atm}/M_{\rm core}$ after billions of years of evolution, without computing evolutionary tracks. I wish to predict this "final" evolutionary state of a planet as a function of core mass $M_{\rm core}$, initial atmospheric mass fraction $X_{\rm init}$, orbital period P, host stellar mass M_* and core density $\rho_{\rm core}$.

The design of the emulator, as shown schematically in Figure 6.1, is a two pronged-approach with all adopted algorithms taken from `scikit-learn` (Pedregosa et al., 2011). In the photoevaporation model, the transition from a sub-Neptune to a super-Earth is quick since the mass-loss timescale drops rapidly for atmospheric mass fractions $\lesssim 1\%$ (see Chapter 2). Thus, the outcome of exoplanetary evolution is bimodal. This bimodality motivates the two-pronged approach: I begin by implementing a random forest classifier to determine whether a given planet will evolve to become a super-Earth or a sub-Neptune. Since the planets that become super-Earths have a negligible atmospheric mass fraction, the final radius for these is simply given by the mass-radius relation of the stripped cores. For the sub-Neptunes I employ a second machine learnt algorithm, in this case a random forest regressor, to predict the final radius and atmospheric mass fraction of a sub-Neptune. Note that random forests were chosen over other machine learning algorithms such as support-vector machines and gradient-boosted linear regression. These alternative algorithms yielded inferior performance or required more training to achieve the same accuracy as the random forest approach.

To train the machine learning models, 10^6 planets were simulated to 5 Gyr with the analytic photoevaporation model of Owen and Wu (2017), with input parameters of P, M_*, $M_{\rm core}$, $\rho_{\rm core}$ and $X_{\rm init}$ uniformly drawn such as to sample parameter space evenly. In this case, the bounds were as follows: $P \in [1, 100]$ days, $M_* \in [0.5, 1.5] M_\oplus$, $M_{\rm core} \in [0.6, 20.0] M_\oplus$, $\tilde{\rho} \in [-1.0, 1.0]$ and $\log X_{\rm init} \in [-4, 0]$. Random cross-validation was used to optimise the performance of the random forest classifier and regressor, which yielded highest accuracy when both contained 19 decision trees.

Swapping a physical model for a machine learning emulator will dramatically decrease the computational expense of calculations. In this case, I found that for an equivalent set of planets, the emulator has a computational speedup factor of ~ 400. This speed-up is all the more impressive as the standard evolutionary model was heavily optimised, whereas no such optimisation

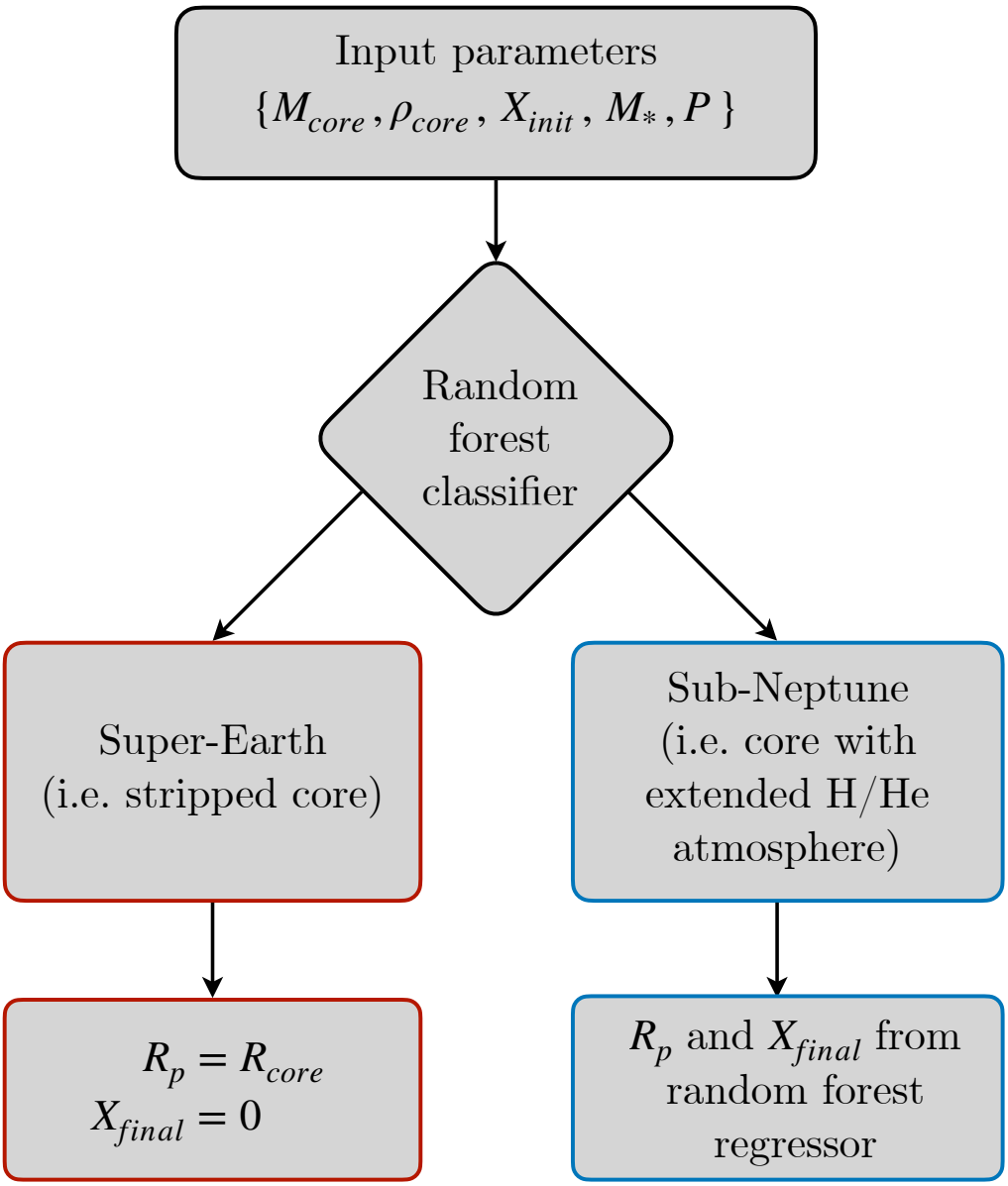


Figure 6.1: Flowchart to demonstrate the atmospheric evolution emulator. For a given set of parameters, an initial random forest classifier determines whether a planet will be stripped of its primordial H/He atmosphere. If so, the standard mass-radius relation of Fortney et al. (2007) is used to provide the final radius. If the initial classifier predicts an atmosphere is retained, a random forest regressor predicts its final radius and atmospheric mass fraction.

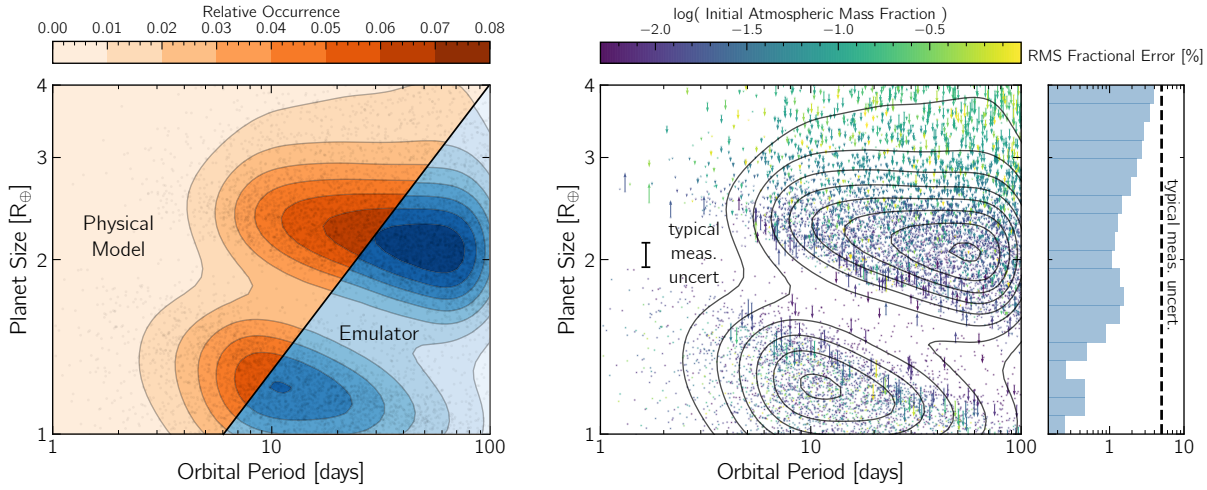


Figure 6.2: Left: The radius gap computed with the semi-analytic model of XUV photoevaporation (orange) and with the evolution emulator (blue). Contours represent relative occurrence and individual planets are plotted in black. Right: A population of planets are shown with their periods and final radii according to the analytic model, with arrows pointing to this position as predicted with evolution emulator. They hence represent the error between physical model and emulator for each planet. Note that the majority of planets do not have arrows since the error is negligible. Colours represent the initial atmospheric mass fraction. Whilst the average RMS error for the emulator is 1% *(as most planets have radii of either 1.3 or 2.5* $R_\oplus$*), the right-hand histogram shows this error as a function of radius. A typical radius measurement uncertainty is shown in black, taken from the value in the CKS catalogue (Fulton and Petigura, 2018).*

was undertaken on the evaluation of the emulator. This indicates that speedups of $\gtrsim 1000$ are potentially achievable.

However, the trade-off for an improved speed is the introduction of output error. Overall, the accuracy of the classifier was found to be 99.98%, with very occasional planets being misclassified as sub-Neptunes when they should be super-Earths. The RMS fractional error for the sub-Neptune regressor was 1%. This value is significantly smaller than typical radii measurement uncertainties of $\sim$ 5% (Fulton and Petigura, 2018), implying that it is suitable for use in comparing large number of evolutionary models to the observed properties of bulk exoplanets. The RMS fractional error in final atmospheric mass fraction was found to be 0.5% and since the mass of a planet is dominated by the core, this introduces a negligible error in final mass. In Figure 6.2, I show planets in the radius-period plane evolved using the analytic photoevaporation model in orange and evolution emulator in blue. One can see that the occurrence is near-indistinguishable. I also show the errors of the emulator in Figure 6.2, which are largest for larger planets with higher initial atmospheric mass-fraction. Despite constructing a training sample with uniformly drawn parameters, this does not evenly sample R_p or X_final space. Since the majority of planets evolve to smaller radii, this area of parameter space is more sparsely populated in the training data and hence accuracy is lower here.

In conclusion, this proof-of-concept machine-learnt emulator is exceptionally fast to evaluate at a very small penalty in accuracy. In the future, sophisticated numerical simulations with higher accuracy should be used for training, instead of simple evolutionary models adopted here. Furthermore, other parameters such as system age and initial cooling timescale may be used as inputs. This will allow physics-based inference, such as that presented in Chapter 4, to investigate the trends between core mass and initial atmospheric mass fraction distributions i.e. probing how gas accretion and the boil-off phase present in the data, as well correlations of core mass with stellar mass, such as that proposed Wu (2019).

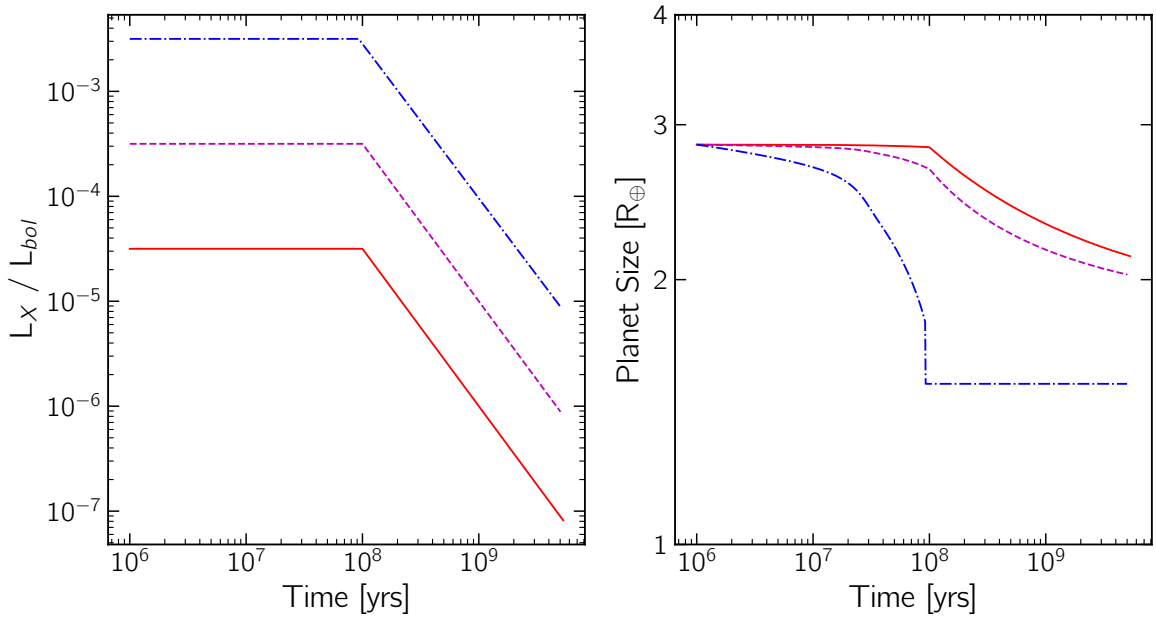


Figure 6.3: In the left-hand panel, the ratio of X-ray to bolometric luminosity is shown as a function of time for a Solar mass star in magenta dashed, with artificial offsets of ±1 dex in blue dashed and red solid lines, to model the potential intrinsic scatter in high energy flux from host stars. In the right-hand panel, the corresponding radius evolution is shown for a $5M_\oplus$ core with initial 1% atmospheric mass-fraction at a 10 day orbital period for the different X-ray evolutions. One can see that the outcomes can be dramatically different, with the planet experiencing the highest X-ray flux being completely stripped, whereas the others do not. Thus for a given planet, the X-ray history can determine which side of the radius gap the planet will sit. Note however, that an order of magnitude scatter in L_{XUV}/L_{bol} for a given stellar type may be an extreme interpretation of the results from Tu et al. (2015); Johnstone et al. (2021).

6.2.2 Variable Stellar X-ray Luminosity

One important point that has been overlooked in the work presented in this thesis, and the field in general, is the role of variability in stellar X-ray on the population of close-in exoplanets undergoing photoevaporation. As discussed in Chapters 1 and 2, surveys of stellar clusters at different ages reveal that solar-type stars emit a high energy luminosity with $(L_{\rm XUV}/L_{\rm bol})_{\rm sat} \sim 10^{-3.5}$, before spinning down at $\sim$ 100 Myr due to magnetic braking (e.g. Güdel et al., 1997; Ribas et al., 2005; Wright et al., 2011; Claire et al., 2012; Jackson et al., 2012). However, it has also been shown that there is up to an order of magnitude of scatter in exact value of $L_{\rm XUV}/L_{\rm bol}$, even for stars of the same mass (Johnstone et al., 2015). In Tu et al. (2015), simple energy-limited mass-loss calculations were performed to show that this scatter can have a dramatic effect on the outcome of atmospheric evolution of small, close-in exoplanets.

The source of this intrinsic scatter comes from the initial rotation period distribution of young stars. X-ray generation in the corona of stellar atmospheres is strongly coupled to their rotation period via a magnetic dynamo. Therefore stars of a given mass with an faster initial rotation period will generate a larger X-ray luminosity, before spinning down to a unique mass-dependent value (Soderblom et al., 1995). In Figure 6.3, I model this order-of-magnitude scatter in $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$ with the analytic XUV photoevaporation model of Owen and Wu (2017) to confirm the findings of Tu et al. (2015). It demonstrates that a $5M_\oplus$ core with initial atmospheric mass fraction of 1% around a solar-mass star at 10 day orbital period can have dramatically different evolutionary outcomes with different values of $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$. In the case where the scatter increases this ratio an order of magnitude, the planet is completely stripped and becomes

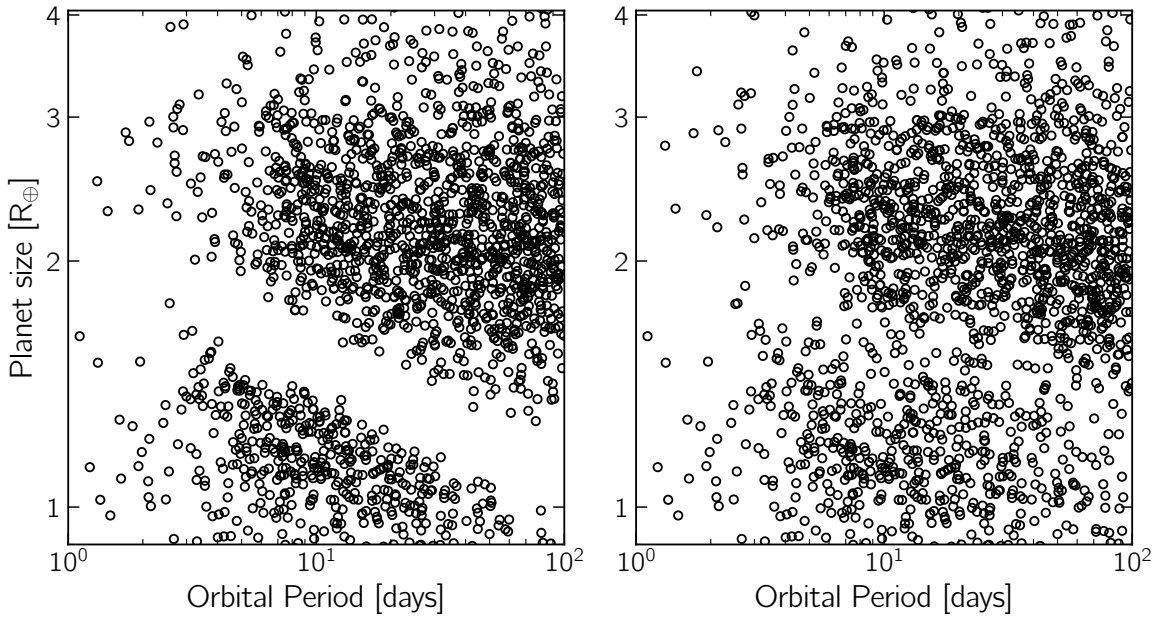


Figure 6.4: The radius gap is shown as a function of orbital period, as produced by the analytic model of Owen and Wu (2017). In the left-hand panel, no scatter is modelled in the X-ray luminosity of host stars, yielding a radius gap devoid of planets. In the right-hand panel, an order of magnitude scatter is added to $(L_{XUV}/L_{bol})_{sat}$ for a given stellar mass, which scatters planets into the gap, reducing its purity.

a super-Earth, which is not the case for lower luminosities.

Additionally, in Figure 6.4, I demonstrate the effect this scatter has on a population level, specifically the appearance of the radius gap. In the left-hand panel, I show a population of planets evolved through the analytic XUV photoevaporation model of Owen and Wu (2017) (specifically MODEL II from Chapter 4), which clearly shows a radius gap that is devoid of planets. However, in the right-hand panel, I add an order-of-magnitude scatter in the values of $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$ for a given mass star. As shown in Figure 6.3, this changes the evolutionary outcome of planets, meaning that the radius gap becomes significantly filled-in. Upon inspection, the planets that are shifted into the gap are super-Earths that would have remained as sub-Neptunes if they has received less XUV flux.

Upon first inspection, this result may cause concern as to the efficacy of the photoevaporation model. X-ray scatter shifts planets into the radius gap, meaning that it becomes degenerate with planet radii measurement uncertainty, destroying information that is stored in the radius gap distribution. This all implies that physics-based inference may become even more statistically challenging to perform. In addition, with regards to the model comparison performed in Chapter 3, it also removes information imprinted by any stellar mass trend, since stars of a given mass do not produce a unique value of $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$. However, recall that in the asteroseismic planet sample of Van Eylen et al. (2018), which has the smallest radii measurement uncertainties of any survey, the radius gap is completely devoid of planets. This implies that there is very little, or even no scatter in the X-ray luminosities of young host stars. How can this be the case?

As discussed in Johnstone et al. (2021), there are in-fact three sources of scatter that contribute to X-ray observations of host stars: the first is the aforementioned intrinsic scatter in $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$ for stars of a given mass; the second comes from observing stars at different points in their variability cycles; and finally there is the measurement uncertainty in the X-ray photon detection count from telescopes such as *Einstein* (Giacconi et al., 1979), *ROSAT* (Voges

et al., 1999), and *XMM-Newton* (Jansen et al., 2001). Johnstone et al. (2021) investigated the observed activity cycles of young stars to show that stellar variability on timescales of decades or longer is likely the dominant cause of the spread in $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$. If this is assumed to be true, they showed that stars have $L_{\rm X}$ values greater than twice the average 20% of the time, greater than three times the average 9% of the time, and greater than five times the average 3% of the time. With regards to exoplanet atmospheric evolution, the cyclic variability of host stars on these timescales will have little effect since it is the integrated high-energy flux that controls the outcome of photoevaporation. Hence, if stars of a given mass have similar activity cycles and similar $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$, the integrated XUV flux will produce a radius gap devoid of planets, as observed in Van Eylen et al. (2018). Nonetheless, it is interesting to note that physics-based inference may be used in the future to place constraints on the intrinsic scatter in $(L_{\rm XUV}/L_{\rm bol})_{\rm sat}$ for stars of a given mass. One could use a similar methodology to that in Chapter 4 to find the intrinsic X-ray scatter required to reproduce the radius gap 'filling-factor'.

In general, this area of research will benefit from future observational efforts to characterise the evolution of high-energy flux from host stars. In particular, more work is needed to understand how X-ray scatter may change as a function of stellar mass, as well as how different initial rotation periods may effect the length of the saturation regime (e.g. Tu et al., 2015; Johnstone et al., 2015).

6.2.3 X-ray vs. EUV dominated mass-loss

One of the main focuses of this thesis has been the atmospheric evolution of small, close-in exoplanets undergoing XUV photoevaporation. Based on the work of Owen and Jackson (2012), I have assumed throughout that it is the X-rays that dictate the atmospheric escape, as opposed to the EUV photons. However, a recent study from King and Wheatley (2021) suggested that the EUV flux may persist on Gyr timescales, potentially inducing mass-loss far beyond the canonical X-ray saturation time of ~ 100 Myr.

To quantify the difference between X-ray and EUV luminosities, King and Wheatley (2021) used archival EUV data from Woods et al. (2005) to derive an empirical relation for the evolution of $L_{\rm EUV}/L_{\rm bol}$, as well as the well-documented evolution of $L_{\rm X}/L_{\rm bol}$ (e.g. Ribas et al., 2005; Claire et al., 2012; Jackson et al., 2012):

$$\frac{L_x}{L_{\rm bol}} = \begin{cases} \left(\frac{L_x}{L_{\rm bol}}\right)_{\rm sat} & \text{for } t < t_{\rm sat}, \\ \left(\frac{L_x}{L_{\rm bol}}\right)_{\rm sat} \left(\frac{t}{t_{\rm sat}}\right)^{\alpha} & \text{for } t \geq t_{\rm sat}, \end{cases} \tag{6.2.1}$$

where L_x can refer to $L_{\rm EUV}$ or $L_{\rm X}$. King and Wheatley (2021) find that $(L_{\rm EUV}/L_{\rm bol})_{\rm sat} \sim 10^{-4.0}$ and $(L_{\rm X}/L_{\rm bol})_{\rm sat} \sim 10^{-3.7}$. However, their key result is that $\alpha \approx -0.65$ for EUV evolution and $\alpha \approx -1.30$ for X-ray, implying that EUV luminosity decays slower that X-ray. This difference in evolution is shown in the upper panel of Figure 6.5. Additionally, the bottom panel also shows the integrated cumulative energy of the two contributions. Note that, despite the EUV luminosity rising above X-ray at ~ 500 Myr, the cumulative energy is only dominated by EUV beyond ~ 4 Gyr.

To investigate the effects of this differing high-energy evolution, I perform simple calculations using the analytic model of photoevaporation from Owen and Wu (2017). To differentiate between X-ray dominated and EUV-dominated escape, one must consider the differences between the two regimes. X-rays will tend to penetrate further into an atmosphere, due to their smaller collisional cross-section. However, as discussed in Chapter 1 and demonstrated by Owen and Jackson (2012), an X-ray regime is one in which the outflow sonic point exists above the X-ray heated layer. In this case, material is heated by X-rays deep in the atmosphere, which then

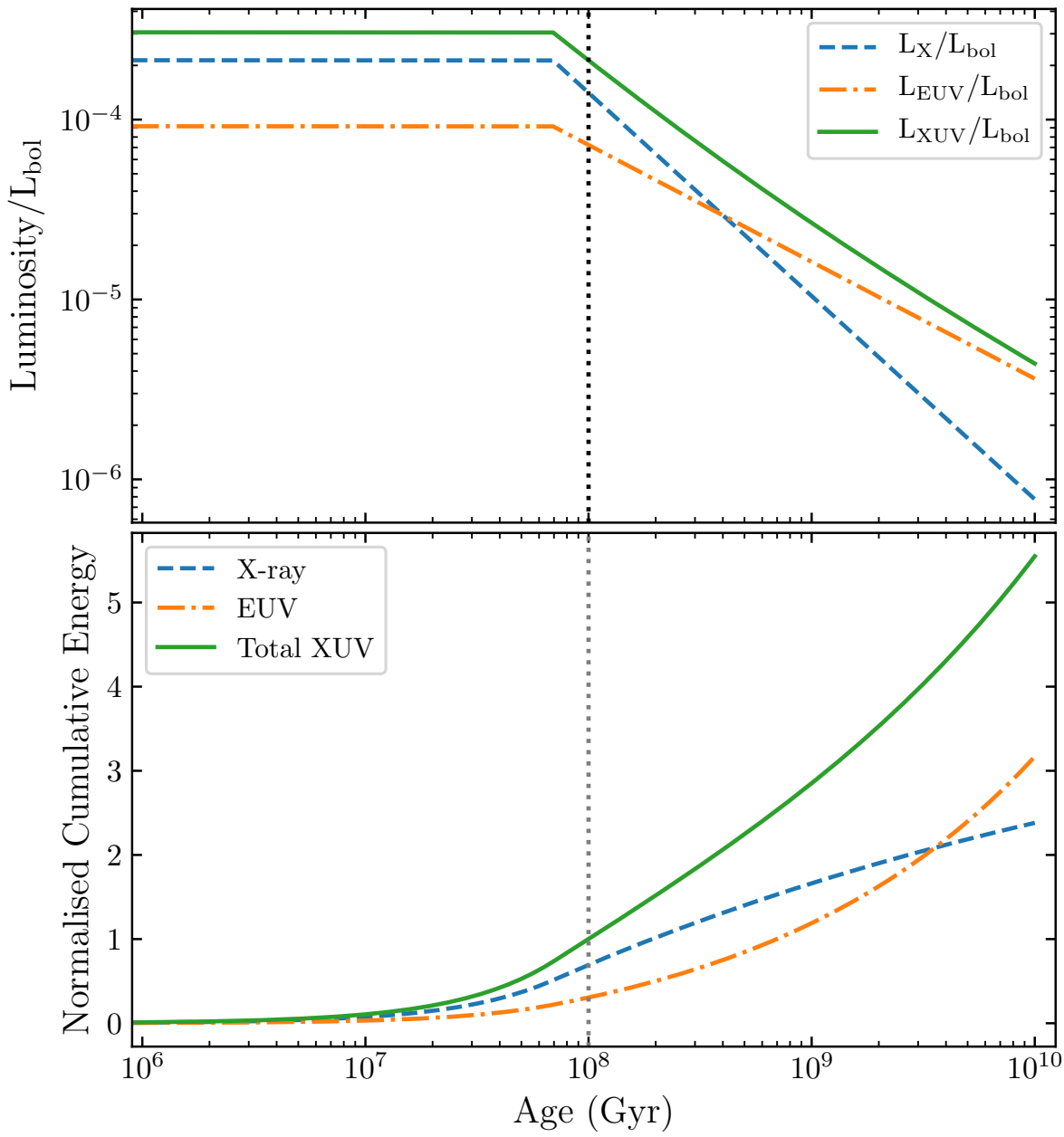


Figure 6.5: In the upper panel, evolution of X-ray (blue dashed) and EUV (orange dot-dashed) luminosities are shown relative to bolometric luminosity. The combined XUV is also shown in green. In the lower panel, the normalised cumulative energy is shown as a function of time. One can see that L_{EUV}/L_{bol} rises above that of X-ray at $\sim$ 500 Myr in the upper panel, and becomes energetically dominant at $\sim$ 4 Gyr in the lower panel. Figure taken from King and Wheatley (2021).

accelerate to up to sonic point. Beyond this, the material may shock before reaching an ionisation front which marks the boundary between X-ray and EUV heated material. Since the sonic point exists below this ionisation front, the outflow is controlled by the X-ray flux. In the EUV regime however, the X-ray heated material is still subsonic as it passes through the ionisation front. As a result, it is the EUV heating that dictates the outflow structure. The position of the ionisation front is set by the balance of hydrogen recombination and ionisation. Owen and Jackson (2012) derived an analytic expression for the critical EUV luminosity, $L_{\rm EUV,crit}$, above which the atmospheric escape is EUV dominated[1]:

$$L_{\rm EUV,crit} \approx 2 \times 10^{29} {\rm erg\ s^{-1}} \left(\frac{a}{0.1\ {\rm AU}}\right)^2 \left(\frac{\dot{M}_{\rm X}}{10^{12}\ {\rm g\ s^{-1}}}\right)^2 \left(\frac{R_{\rm p}}{10 R_\oplus}\right)^{-3}, \tag{6.2.2}$$

where $\dot{M}_{\rm X}$ is the mass-loss rate due to X-ray heating, in this case given by the energy-limited formalism presented in Chapters 1 and 2:

$$\dot{M}_{\rm X} = \eta \frac{\pi R_{\rm p}^3 L_{\rm X}}{G M_{\rm p} a^2}. \tag{6.2.3}$$

As a result of these scalings, one would expect planets in the EUV-dominated regime to be at large orbital distances and small in size. Thus, I incorporate both EUV and X-ray dominated regimes by performing this check, given $L_{\rm EUV}$ and $L_{\rm X}$ at a given time. If the flow is found to be EUV-dominated, I recalculate the mass-loss rate from Eq. 6.2.3 using $L_{\rm EUV}$. In Figure 6.6, I show the regimes in which a planet around a solar-mass star with an initial atmospheric mass fraction of 1% would be observed undergoing X-ray dominated escape or EUV dominated escape at 5 Gyr. Since planets that are highly irradiated and low in mass would be completely stripped, I also highlight this area of parameter space. One can see that planets at large separations and with large core masses (and therefore small, gravitationally contracted sizes) are in the EUV-dominated regime at 5 Gyr, as predicted by Eq. 6.2.2. It is interesting to note that planets observed with transit spectroscopy in different areas of such parameter space may be found in different escape regimes, which will change their outflow structure. To probe this difference, one would need to use an atomic line that probes the outflow near the ionisation front. As a result, Ly-α is not suitable since interstellar medium absorption removes line-core information, which stores the detail of low velocity flow and thus the location at which the escape in launched. On the other hand, the He 1083 nm line (e.g. Spake et al., 2018; Ninan et al., 2020) may be appropriate since it is populated by recombinations and therefore dominated near the EUV ionisation depth. However, transit spectroscopy is difficult for planets as large orbital distances, meaning that observing planets that are in EUV dominated regime would be technically challenging with current technology. Note also that planets at such large orbital separations at 5 Gyr may have cooled to such an extend that mass-loss no longer occurs via a hydrodynamic regime, but instead by Jeans escape. In this case, the mass-loss rate would be significantly lower and even harder (if not impossible) to observe.

These results suggest that some planets may evolve into an EUV-dominated escape regime, due to its prolonged luminosity found by King et al. (2018). In spite of this however, I find that this introduces a negligible effect on the atmospheric evolution for all planets of varying masses and orbital separations. For planets of sufficient mass and orbital separation, the transition to EUV dominated escape happens on Gyr timescales (see Figure 6.5), by which time the planet has cooled and contracted significantly. Despite being greater than its X-ray counterpart, the EUV luminosity has also dramatically fallen from its saturated value at this point, implying that any mass-loss is negligible. Even for a $10 M_\oplus$ planet at 1000 day orbital period i.e. upper-right hand corner of Figure 6.6, I calculate a difference in planet size at 10 Gyr of $\sim 0.1\%$

[1]Note that I have neglected order unity factors that account for the density drop-off in ionised flow, and the relation between the planet radius and sonic point (see Owen and Jackson, 2012, for details).

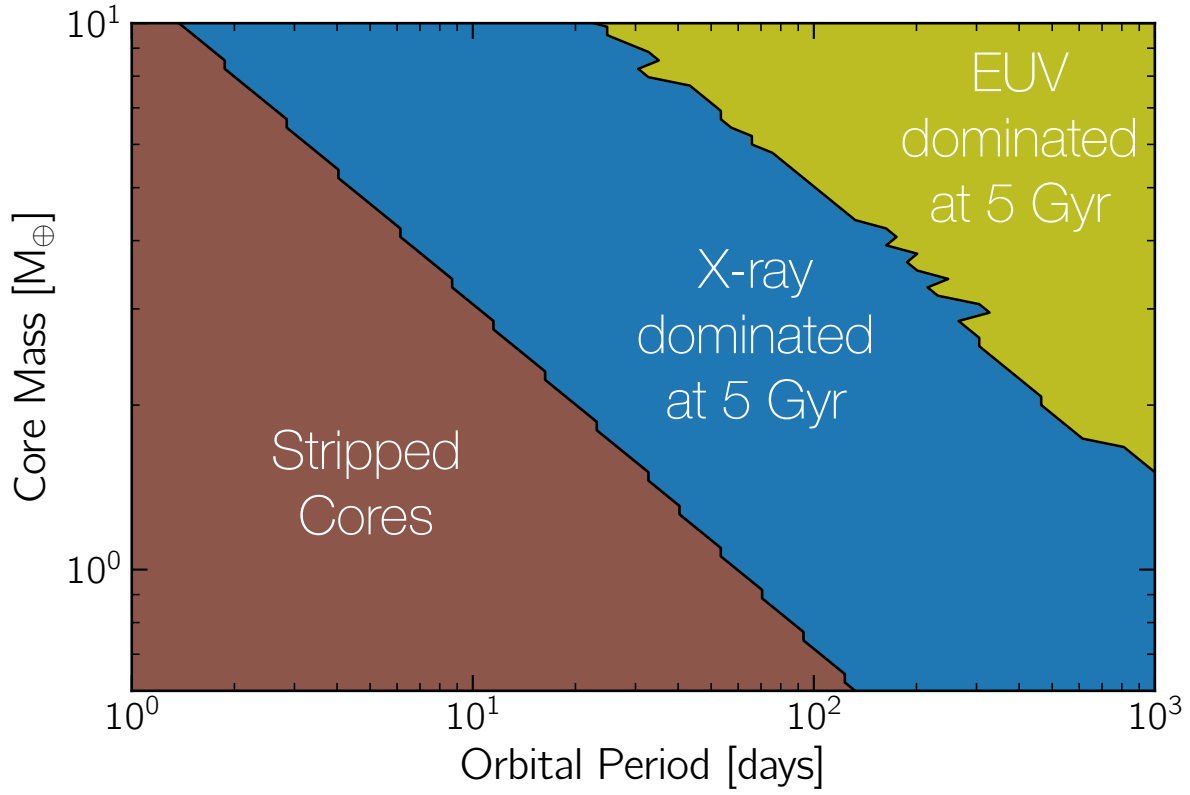


Figure 6.6: Different regimes of mass-loss are shown as a function of orbital period and core mass for planets around a solar-mass star and initial atmospheric mass-fraction of 1%*. Planets with low-mass and small orbital separation are completely stripped (left-hand region), planets with sufficiently large mass and orbital separation transition to the EUV dominated escape regime (right-hand region). In between, there is a population of planets that remain in the X-ray dominated regime, even after 5 Gyr of evolution. Despite this difference, I find a negligible effect on the evolution of all planets, since the transition to EUV dominated escape happens on* $\sim$ *Gyr timescales, by which time the planets have cooled and contracted, meaning that photoevaporation becomes inefficient.*

when only considering the X-ray dominated phase instead of both regimes. In conclusion, these simple calculations have shown that, although some high mass and large separation planets may evolve into the EUV dominated escape regime, this has a negligible effect on their evolution. As a result, the radius gap is unaffected and one can safely assume that the majority of mass-loss under the XUV photoevaporation model does indeed occur in the first ~ 100 Myr.

6.2.4 Closing Remarks

In this thesis I have focused on using our knowledge of exoplanet evolution to better understand exoplanet formation. It is therefore no surprise that, in order for this field to progress, further work is needed on evolution theory and observations.

A key aspect that is missing from the majority of theoretical studies is the evolution of planetary cores and how it might effect atmospheric structure and subsequent mass-loss. Super-Earths and sub-Neptunes are likely to have molten silicate surfaces for their early evolution, which can introduce silicate vapour atmospheres. These high mean molecular weight atmosphere components may inhibit convection and early studies of this suggest that it may significantly effect atmospheric structure and evolution (e.g. Ormel et al., 2021; Schlichting and Young, 2022; Misener and Schlichting, 2022). Indeed, as with all the calculations in this thesis, complex chemistry has not been considered in the atmospheres of close-in exoplanets. As the field progresses, I believe that we must consider the important effects of chemistry in escaping atmospheres.

Observationally speaking, a common idiom used in the field is '*know thy star, know thy planet*'. With regards to atmospheric escape, this is especially pertinent since without constraints on the high-energy luminosity of host stars, we would not be able to understand XUV photoevaporation. Nevertheless, further observations of stellar evolution is paramount to understand planet evolution. In particular, observational efforts are required to understand how X-ray/EUV emission changes with stellar mass. In addition, constraints are required as to how intrinsic scatter of high energy emission may effect the evolution of exoplanet atmospheres for a given stellar type. An exciting prospect that is recently becoming a reality is detecting transiting planets around young stars (e.g. Newton et al., 2019; Rizzuto et al., 2020). Although observationally challenging due to their high-levels of stellar activity, such observations will be key to constraining exoplanet evolution at early time. Recent demographic results for small planets around young stars show tentative evidence that sub-Neptunes reduce in size with time, providing further evidence of thermal cooling and contraction in the presence of a H/He dominated atmospheres Fernandes et al. (2022).

As future large-scale surveys come online, such as *PLATO* and *Nancy Grace Roman*, we will be able to probe the distributions of exoplanets to smaller sizes and larger orbital distances. With a solid understanding of exoplanet evolution, we will then be able to rewind the clock, to understand the connection between planet formation and the population we observe today.

Bibliography

Adams, E. R., Seager, S., and Elkins-Tanton, L. (2008). Ocean Planet or Thick Atmosphere: On the Mass-Radius Relationship for Solid Exoplanets with Massive Atmospheres. ApJ, 673(2):1160–1164.

Adams, F. C. (2011). Magnetically Controlled Outflows from Hot Jupiters. ApJ, 730(1):27.

Agol, E., Steffen, J., Sari, R., and Clarkson, W. (2005). On detecting terrestrial planets with timing of giant planet transits. MNRAS, 359(2):567–579.

Alexander, R. D. and Armitage, P. J. (2007). Dust dynamics during protoplanetary disc clearing. MNRAS, 375(2):500–512.

Ali-Dib, M., Cumming, A., and Lin, D. N. C. (2020). The imprint of the protoplanetary disc in the accretion of super-Earth envelopes. MNRAS, 494(2):2440–2448.

Ali-Dib, M., Cumming, A., and Lin, D. N. C. (2022). The effect of late giant collisions on the atmospheres of protoplanets and the formation of cold sub-Saturns. MNRAS, 509(1):1413–1431.

Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., Rosenfeld, K. A., Öberg, K. I., Birnstiel, T., Espaillat, C., Cieza, L. A., Williams, J. P., Lin, S.-Y., and Ho, P. T. P. (2012). The TW Hya Disk at 870 μm: Comparison of CO and Dust Radial Structures. ApJ, 744:162.

Arakcheev, A. S., Zhilkin, A. G., Kaigorodov, P. V., Bisikalo, D. V., and Kosovichev, A. G. (2017). Reduction of mass loss by the hot Jupiter WASP-12b due to its magnetic field. *Astronomy Reports*, 61(11):932–941.

Armitage, P. J. (2015). Physical processes in protoplanetary disks. *arXiv e-prints*, page arXiv:1509.06382.

Armstrong, D. J., Lopez, T. A., Adibekyan, V., Booth, R. A., Bryant, E. M., Collins, K. A., Emsenhuber, A., Huang, C. X., King, G. W., Lillo-box, J., Lissauer, J. J., Matthews, E. C., Mousis, O., Nielsen, L. D., Osborn, H., Otegi, J., Santos, N. C., Sousa, S. G., Stassun, K. G., Veras, D., Ziegler, C., Acton, J. S., Almenara, J. M., Anderson, D. R., Barrado, D., Barros, S. C. C., Bayliss, D., Belardi, C., Bouchy, F., Briceno, C., Brogi, M., Brown, D. J. A., Burleigh, M. R., Casewell, S. L., Chaushev, A., Ciardi, D. R., Collins, K. I., Colón, K. D., Cooke, B. F., Crossfield, I. J. M., Díaz, R. F., Deleuil, M., Delgado Mena, E., Demangeon, O. D. S., Dorn, C., Dumusque, X., Eigmuller, P., Fausnaugh, M., Figueira, P., Gan, T., Gand hi, S., Gill, S., Goad, M. R., Guenther, M. N., Helled, R., Hojjatpanah, S., Howell, S. B., Jackman, J., Jenkins, J. S., Jenkins, J. M., Jensen, E. L. N., Kennedy, G. M., Latham, D. W., Law, N., Lendl, M., Lozovsky, M., Mann, A. W., Moyano, M., McCormac, J., Meru, F., Mordasini, C., Osborn, A., Pollacco, D., Queloz, D., Raynard, L., Ricker, G. R., Rowden, P., Santerne, A., Schlieder, J. E., Seager, S., Sha, L., Tan, T.-G., Tilbrook, R. H., Ting, E., Udry, S., Vanderspek, R., Watson, C. A., West, R. G., Wilson, P. A., Winn, J. N., Wheatley,

P., Villasenor, J. N., Vines, J. I., and Zhan, Z. (2020). A remnant planetary core in the hot-Neptune desert. *arXiv e-prints*, page arXiv:2003.10314.

Auvergne, M., Bodin, P., Boisnard, L., Buey, J. T., Chaintreuil, S., Epstein, G., Jouret, M., Lam-Trong, T., Levacher, P., Magnan, A., Perez, R., Plasson, P., Plesseria, J., Peter, G., Steller, M., Tiphène, D., Baglin, A., Agogué, P., Appourchaux, T., Barbet, D., Beaufort, T., Bellenger, R., Berlin, R., Bernardi, P., Blouin, D., Boumier, P., Bonneau, F., Briet, R., Butler, B., Cautain, R., Chiavassa, F., Costes, V., Cuvilho, J., Cunha-Parro, V., de Oliveira Fialho, F., Decaudin, M., Defise, J. M., Djalal, S., Docclo, A., Drummond, R., Dupuis, O., Exil, G., Fauré, C., Gaboriaud, A., Gamet, P., Gavalda, P., Grolleau, E., Gueguen, L., Guivarc'h, V., Guterman, P., Hasiba, J., Huntzinger, G., Hustaix, H., Imbert, C., Jeanville, G., Johlander, B., Jorda, L., Journoud, P., Karioty, F., Kerjean, L., Lafond, L., Lapeyrere, V., Landiech, P., Larqué, T., Laudet, P., Le Merrer, J., Leporati, L., Leruyet, B., Levieuge, B., Llebaria, A., Martin, L., Mazy, E., Mesnager, J. M., Michel, J. P., Moalic, J. P., Monjoin, W., Naudet, D., Neukirchner, S., Nguyen-Kim, K., Ollivier, M., Orcesi, J. L., Ottacher, H., Oulali, A., Parisot, J., Perruchot, S., Piacentino, A., Pinheiro da Silva, L., Platzer, J., Pontet, B., Pradines, A., Quentin, C., Rohbeck, U., Rolland, G., Rollenhagen, F., Romagnan, R., Russ, N., Samadi, R., Schmidt, R., Schwartz, N., Sebbag, I., Smit, H., Sunter, W., Tello, M., Toulouse, P., Ulmer, B., Vandermarcq, O., Vergnault, E., Wallner, R., Waultier, G., and Zanatta, P. (2009). The CoRoT satellite in flight: description and performance. A&A, 506(1):411–424.

Bakos, G., Noyes, R. W., Kovács, G., Stanek, K. Z., Sasselov, D. D., and Domsa, I. (2004). Wide-Field Millimagnitude Photometry with the HAT: A Tool for Extrasolar Planet Detection. PASP, 116(817):266–277.

Baraffe, I., Chabrier, G., Allard, F., and Hauschildt, P. H. (1998). Evolutionary models for solar metallicity low-mass stars: mass-magnitude relationships and color-magnitude diagrams. A&A, 337:403–412.

Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., and Hauschildt, P. H. (2003). Evolutionary models for cool brown dwarfs and extrasolar giant planets. The case of HD 209458. A&A, 402:701–712.

Batalha, N. M., Rowe, J. F., Bryson, S. T., Barclay, T., Burke, C. J., Caldwell, D. A., Christiansen, J. L., Mullally, F., Thompson, S. E., Brown, T. M., Dupree, A. K., Fabrycky, D. C., Ford, E. B., Fortney, J. J., Gilliland, R. L., Isaacson, H., Latham, D. W., Marcy, G. W., Quinn, S. N., Ragozzine, D., Shporer, A., Borucki, W. J., Ciardi, D. R., Gautier, Thomas N., I., Haas, M. R., Jenkins, J. M., Koch, D. G., Lissauer, J. J., Rapin, W., Basri, G. S., Boss, A. P., Buchhave, L. A., Carter, J. A., Charbonneau, D., Christensen-Dalsgaard, J., Clarke, B. D., Cochran, W. D., Demory, B.-O., Desert, J.-M., Devore, E., Doyle, L. R., Esquerdo, G. A., Everett, M., Fressin, F., Geary, J. C., Girouard, F. R., Gould, A., Hall, J. R., Holman, M. J., Howard, A. W., Howell, S. B., Ibrahim, K. A., Kinemuchi, K., Kjeldsen, H., Klaus, T. C., Li, J., Lucas, P. W., Meibom, S., Morris, R. L., Prša, A., Quintana, E., Sanderfer, D. T., Sasselov, D., Seader, S. E., Smith, J. C., Steffen, J. H., Still, M., Stumpe, M. C., Tarter, J. C., Tenenbaum, P., Torres, G., Twicken, J. D., Uddin, K., Van Cleve, J., Walkowicz, L., and Welsh, W. F. (2013). Planetary Candidates Observed by Kepler. III. Analysis of the First 16 Months of Data. ApJS, 204(2):24.

Batygin, K. and Stevenson, D. J. (2010). Inflating Hot Jupiters with Ohmic Dissipation. ApJ, 714(2):L238–L243.

Bear, E. and Soker, N. (2011). Evaporation of Jupiter-like planets orbiting extreme horizontal branch stars. MNRAS, 414(2):1788–1792.

Beaulieu, J. P., Bennett, D. P., Fouqué, P., Williams, A., Dominik, M., Jørgensen, U. G., Kubas, D., Cassan, A., Coutures, C., Greenhill, J., Hill, K., Menzies, J., Sackett, P. D., Albrow, M., Brillant, S., Caldwell, J. A. R., Calitz, J. J., Cook, K. H., Corrales, E., Desort, M., Dieters, S., Dominis, D., Donatowicz, J., Hoffman, M., Kane, S., Marquette, J. B., Martin, R., Meintjes, P., Pollard, K., Sahu, K., Vinter, C., Wambsganss, J., Woller, K., Horne, K., Steele, I., Bramich, D. M., Burgdorf, M., Snodgrass, C., Bode, M., Udalski, A., Szymański, M. K., Kubiak, M., Więckowski, T., Pietrzyński, G., Soszyński, I., Szewczyk, O., Wyrzykowski, Ł., Paczyński, B., Abe, F., Bond, I. A., Britton, T. R., Gilmore, A. C., Hearnshaw, J. B., Itow, Y., Kamiya, K., Kilmartin, P. M., Korpela, A. V., Masuda, K., Matsubara, Y., Motomura, M., Muraki, Y., Nakamura, S., Okada, C., Ohnishi, K., Rattenbury, N. J., Sako, T., Sato, S., Sasaki, M., Sekiguchi, T., Sullivan, D. J., Tristram, P. J., Yock, P. C. M., and Yoshioka, T. (2006). Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. Nature, 439(7075):437–440.

Bennett, D. P., Anderson, J., Bond, I. A., Udalski, A., and Gould, A. (2006). Identification of the OGLE-2003-BLG-235/MOA-2003-BLG-53 Planetary Host Star. ApJ, 647(2):L171–L174.

Berger, T. A., Huber, D., Gaidos, E., van Saders, J. L., and Weiss, L. M. (2020a). The Gaia-Kepler Stellar Properties Catalog. II. Planet Radius Demographics as a Function of Stellar Mass and Age. *arXiv e-prints*, page arXiv:2005.14671.

Berger, T. A., Huber, D., van Saders, J. L., Gaidos, E., Tayar, J., and Kraus, A. L. (2020b). The Gaia-Kepler Stellar Properties Catalog. I. Homogeneous Fundamental Properties for 186,301 Kepler Stars. AJ, 159(6):280.

Bitsch, B., Forsberg, R., Liu, F., and Johansen, A. (2018). Stellar abundance of binary stars: their role in determining the formation location of super-Earths and ice giants. MNRAS, 479(3):3690–3707.

Bodenheimer, P. and Lissauer, J. J. (2014). Accretion and Evolution of ˜2.5 M $_\oplus$ Planets with Voluminous H/He Envelopes. ApJ, 791(2):103.

Böhm-Vitense, E. (1958). Über die Wasserstoffkonvektionszone in Sternen verschiedener Effektivtemperaturen und Leuchtkräfte. Mit 5 Textabbildungen. ZAp, 46:108.

Bond, I. A., Udalski, A., Jaroszyński, M., Rattenbury, N. J., Paczyński, B., Soszyński, I., Wyrzykowski, L., Szymański, M. K., Kubiak, M., Szewczyk, O., Żebruń, K., Pietrzyński, G., Abe, F., Bennett, D. P., Eguchi, S., Furuta, Y., Hearnshaw, J. B., Kamiya, K., Kilmartin, P. M., Kurata, Y., Masuda, K., Matsubara, Y., Muraki, Y., Noda, S., Okajima, K., Sako, T., Sekiguchi, T., Sullivan, D. J., Sumi, T., Tristram, P. J., Yanagisawa, T., Yock, P. C. M., and OGLE Collaboration (2004). OGLE 2003-BLG-235/MOA 2003-BLG-53: A Planetary Microlensing Event. ApJ, 606(2):L155–L158.

Bondi, H. (1952). On spherically symmetrical accretion. MNRAS, 112:195.

Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E. W., Dupree, A. K., Gautier, T. N., Geary, J. C., Gilliland, R., Gould, A., Howell, S. B., Jenkins, J. M., Kondo, Y., Latham, D. W., Marcy, G. W., Meibom, S., Kjeldsen, H., Lissauer, J. J., Monet, D. G., Morrison, D., Sasselov, D., Tarter, J., Boss, A., Brownlee, D., Owen, T., Buzasi, D., Charbonneau, D., Doyle, L., Fortney, J., Ford, E. B., Holman, M. J., Seager, S., Steffen, J. H., Welsh, W. F., Rowe, J., Anderson, H., Buchhave, L., Ciardi, D., Walkowicz, L., Sherry, W., Horch, E., Isaacson, H., Everett, M. E., Fischer, D., Torres, G., Johnson, J. A., Endl, M., MacQueen, P., Bryson, S. T., Dotson, J., Haas, M., Kolodziejczak, J., Van Cleve, J., Chandrasekaran, H., Twicken, J. D., Quintana, E. V., Clarke, B. D., Allen, C., Li, J., Wu, H., Tenenbaum, P.,

Verner, E., Bruhweiler, F., Barnes, J., and Prsa, A. (2010). Kepler Planet-Detection Mission: Introduction and First Results. *Science*, 327(5968):977.

Borucki, W. J., Koch, D. G., Basri, G., Batalha, N., Brown, T. M., Bryson, S. T., Caldwell, D., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E. W., Gautier, III, T. N., Geary, J. C., Gilliland, R., Gould, A., Howell, S. B., Jenkins, J. M., Latham, D. W., Lissauer, J. J., Marcy, G. W., Rowe, J., Sasselov, D., Boss, A., Charbonneau, D., Ciardi, D., Doyle, L., Dupree, A. K., Ford, E. B., Fortney, J., Holman, M. J., Seager, S., Steffen, J. H., Tarter, J., Welsh, W. F., Allen, C., Buchhave, L. A., Christiansen, J. L., Clarke, B. D., Das, S., Désert, J.-M., Endl, M., Fabrycky, D., Fressin, F., Haas, M., Horch, E., Howard, A., Isaacson, H., Kjeldsen, H., Kolodziejczak, J., Kulesa, C., Li, J., Lucas, P. W., Machalek, P., McCarthy, D., MacQueen, P., Meibom, S., Miquel, T., Prsa, A., Quinn, S. N., Quintana, E. V., Ragozzine, D., Sherry, W., Shporer, A., Tenenbaum, P., Torres, G., Twicken, J. D., Van Cleve, J., Walkowicz, L., Witteborn, F. C., and Still, M. (2011). Characteristics of Planetary Candidates Observed by Kepler. II. Analysis of the First Four Months of Data. ApJ, 736:19.

Boss, A. P. (1995). Proximity of Jupiter-Like Planets to Low-Mass Stars. *Science*, 267(5196):360–362.

Bourrier, V., Lecavelier des Etangs, A., Ehrenreich, D., Sanz-Forcada, J., Allart, R., Ballester, G. E., Buchhave, L. A., Cohen, O., Deming, D., Evans, T. M., García Muñoz, A., Henry, G. W., Kataria, T., Lavvas, P., Lewis, N., López-Morales, M., Marley, M., Sing, D. K., and Wakeford, H. R. (2018). Hubble PanCET: an extended upper atmosphere of neutral hydrogen around the warm Neptune GJ 3470b. A&A, 620:A147.

Bouvier, J., Alencar, S. H. P., Harries, T. J., Johns-Krull, C. M., and Romanova, M. M. (2007). Magnetospheric Accretion in Classical T Tauri Stars. In Reipurth, B., Jewitt, D., and Keil, K., editors, *Protostars and Planets V*, page 479.

Brown, T. M., Latham, D. W., Everett, M. E., and Esquerdo, G. A. (2011). Kepler Input Catalog: Photometric Calibration and Stellar Classification. AJ, 142(4):112.

Bryson, S., Coughlin, J., Batalha, N. M., Berger, T., Huber, D., Burke, C., Dotson, J., and Mullally, S. E. (2020). A Probabilistic Approach to Kepler Completeness and Reliability for Exoplanet Occurrence Rates. AJ, 159(6):279.

Caldiroli, A., Haardt, F., Gallo, E., Spinelli, R., Malsky, I., and Rauscher, E. (2021). Irradiation-driven escape of primordial planetary atmospheres II. Evaporation efficiency of sub-Neptunes through hot Jupiters. *arXiv e-prints*, page arXiv:2112.00744.

Carroll, B. W. and Ostlie, D. A. (2006). *An introduction to modern astrophysics and cosmology*.

Carter, J. A., Agol, E., Chaplin, W. J., Basu, S., Bedding, T. R., Buchhave, L. A., Christensen-Dalsgaard, J., Deck, K. M., Elsworth, Y., Fabrycky, D. C., Ford, E. B., Fortney, J. J., Hale, S. J., Handberg, R., Hekker, S., Holman, M. J., Huber, D., Karoff, C., Kawaler, S. D., Kjeldsen, H., Lissauer, J. J., Lopez, E. D., Lund, M. N., Lundkvist, M., Metcalfe, T. S., Miglio, A., Rogers, L. A., Stello, D., Borucki, W. J., Bryson, S., Christiansen, J. L., Cochran, W. D., Geary, J. C., Gilliland, R. L., Haas, M. R., Hall, J., Howard, A. W., Jenkins, J. M., Klaus, T., Koch, D. G., Latham, D. W., MacQueen, P. J., Sasselov, D., Steffen, J. H., Twicken, J. D., and Winn, J. N. (2012). Kepler-36: A Pair of Planets with Neighboring Orbits and Dissimilar Densities. *Science*, 337(6094):556.

Chabrier, G. and Baraffe, I. (1997). Structure and evolution of low-mass stars. A&A, 327:1039–1053.

Chaplin, W. J. and Miglio, A. (2013). Asteroseismology of Solar-Type and Red-Giant Stars. ARA&A, 51(1):353–392.

Charbonneau, D., Brown, T. M., Latham, D. W., and Mayor, M. (2000). Detection of Planetary Transits Across a Sun-like Star. ApJ, 529(1):L45–L48.

Chatterjee, S. and Tan, J. C. (2014). Inside-out Planet Formation. ApJ, 780(1):53.

Chauvin, G., Desidera, S., Lagrange, A. M., Vigan, A., Gratton, R., Langlois, M., Bonnefoy, M., Beuzit, J. L., Feldt, M., Mouillet, D., Meyer, M., Cheetham, A., Biller, B., Boccaletti, A., D'Orazi, V., Galicher, R., Hagelberg, J., Maire, A. L., Mesa, D., Olofsson, J., Samland, M., Schmidt, T. O. B., Sissa, E., Bonavita, M., Charnay, B., Cudel, M., Daemgen, S., Delorme, P., Janin-Potiron, P., Janson, M., Keppler, M., Le Coroller, H., Ligi, R., Marleau, G. D., Messina, S., Mollière, P., Mordasini, C., Müller, A., Peretti, S., Perrot, C., Rodet, L., Rouan, D., Zurlo, A., Dominik, C., Henning, T., Menard, F., Schmid, H. M., Turatto, M., Udry, S., Vakili, F., Abe, L., Antichi, J., Baruffolo, A., Baudoz, P., Baudrand, J., Blanchard, P., Bazzon, A., Buey, T., Carbillet, M., Carle, M., Charton, J., Cascone, E., Claudi, R., Costille, A., Deboulbe, A., De Caprio, V., Dohlen, K., Fantinel, D., Feautrier, P., Fusco, T., Gigan, P., Giro, E., Gisler, D., Gluck, L., Hubin, N., Hugot, E., Jaquet, M., Kasper, M., Madec, F., Magnard, Y., Martinez, P., Maurel, D., Le Mignant, D., Möller-Nilsson, O., Llored, M., Moulin, T., Origné, A., Pavlov, A., Perret, D., Petit, C., Pragt, J., Puget, P., Rabou, P., Ramos, J., Rigal, R., Rochat, S., Roelfsema, R., Rousset, G., Roux, A., Salasnich, B., Sauvage, J. F., Sevin, A., Soenke, C., Stadler, E., Suarez, M., Weber, L., Wildi, F., Antoniucci, S., Augereau, J. C., Baudino, J. L., Brandner, W., Engler, N., Girard, J., Gry, C., Kral, Q., Kopytova, T., Lagadec, E., Milli, J., Moutou, C., Schlieder, J., Szulágyi, J., Thalmann, C., and Wahhaj, Z. (2017). Discovery of a warm, dusty giant planet around HIP 65426. A&A, 605:L9.

Chauvin, G., Lagrange, A. M., Dumas, C., Zuckerman, B., Mouillet, D., Song, I., Beuzit, J. L., and Lowrance, P. (2004). A giant planet candidate near a young brown dwarf. Direct VLT/NACO observations using IR wavefront sensing. A&A, 425:L29–L32.

Chen, H. and Rogers, L. A. (2016). Evolutionary Analysis of Gaseous Sub-Neptune-mass Planets with MESA. ApJ, 831(2):180.

Chen, J. and Kipping, D. (2017). Probabilistic Forecasting of the Masses and Radii of Other Worlds. ApJ, 834(1):17.

Chen, Y.-X., Li, Y.-P., Li, H., and Lin, D. N. C. (2020). The Preservation of Super-Earths and the Emergence of Gas Giants after Their Progenitor Cores Have Entered the Pebble-isolation Phase. ApJ, 896(2):135.

Chiang, E. and Laughlin, G. (2013). The minimum-mass extrasolar nebula: in situ formation of close-in super-Earths. MNRAS, 431(4):3444–3455.

Choi, J., Dotter, A., Conroy, C., Cantiello, M., Paxton, B., and Johnson, B. D. (2016). Mesa Isochrones and Stellar Tracks (MIST). I. Solar-scaled Models. ApJ, 823(2):102.

Christian, R. and George, C. (2010). *Introducing Monte Carlo Methods with R*. Springer-Verlag New York, 1 edition.

Christiansen, J. L., Clarke, B. D., Burke, C. J., Jenkins, J. M., Bryson, S. T., Coughlin, J. L., Mullally, F., Thompson, S. E., Twicken, J. D., Batalha, N. M., Haas, M. R., Catanzarite, J., Campbell, J. R., Kamal Uddin, A., Zamudio, K., Smith, J. C., and Henze, C. E. (2016). Measuring Transit Signal Recovery in the Kepler Pipeline. III. Completeness of the Q1-Q17 DR24 Planet Candidate Catalogue with Important Caveats for Occurrence Rate Calculations. ApJ, 828(2):99.

Christiansen, J. L., Clarke, B. D., Burke, C. J., Seader, S., Jenkins, J. M., Twicken, J. D., Catanzarite, J. D., Smith, J. C., Batalha, N. M., Haas, M. R., Thompson, S. E., Campbell, J. R., Sabale, A., and Kamal Uddin, A. (2015). Measuring Transit Signal Recovery in the Kepler Pipeline II: Detection Efficiency as Calculated in One Year of Data. ApJ, 810(2):95.

Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., Burke, C. J., Tenenbaum, P., Seader, S., Thompson, S. E., Barclay, T. S., Clarke, B. D., Li, J., Smith, J. C., Stumpe, M. C., Twicken, J. D., and Van Cleve, J. (2012). The Derivation, Properties, and Value of Kepler's Combined Differential Photometric Precision. PASP, 124(922):1279.

Ciardi, D. R., Fabrycky, D. C., Ford, E. B., Gautier, T. N., I., Howell, S. B., Lissauer, J. J., Ragozzine, D., and Rowe, J. F. (2013). On the Relative Sizes of Planets within Kepler Multiple-candidate Systems. ApJ, 763(1):41.

Cimerman, N. P., Kuiper, R., and Ormel, C. W. (2017). Hydrodynamics of embedded planets' first atmospheres - III. The role of radiation transport for super-Earth planets. MNRAS, 471(4):4662–4676.

Claire, M. W., Sheets, J., Cohen, M., Ribas, I., Meadows, V. S., and Catling, D. C. (2012). The Evolution of Solar Flux from 0.1 nm to 160 μm: Quantitative Estimates for Planetary Studies. ApJ, 757(1):95.

Clarke, C. J., Gendrin, A., and Sotomayor, M. (2001). The dispersal of circumstellar discs: the role of the ultraviolet switch. MNRAS, 328(2):485–491.

Coleman, G. A. L. and Haworth, T. J. (2022). Dispersal of protoplanetary discs: how stellar properties and the local environment determine the pathway of evolution. MNRAS, 514(2):2315–2332.

Cox, A. N. (2000). *Allen's astrophysical quantities*.

Cox, J. P. and Giuli, R. T. (1968). *Principles of stellar structure*.

Cranmer, S. R. (2004). New views of the solar wind with the Lambert W function. *American Journal of Physics*, 72:1397–1403.

Currie, T., Lawson, K., Schneider, G., Lyra, W., Wisniewski, J., Grady, C., Guyon, O., Tamura, M., Kotani, T., Kawahara, H., Brandt, T., Uyama, T., Muto, T., Dong, R., Kudo, T., Hashimoto, J., Fukagawa, M., Wagner, K., Lozi, J., Chilcote, J., Tobin, T., Groff, T., Ward-Duong, K., Januszewski, W., Norris, B., Tuthill, P., van der Marel, N., Sitko, M., Deo, V., Vievard, S., Jovanovic, N., Martinache, F., and Skaf, N. (2022). Images of embedded Jovian planet formation at a wide separation around AB Aurigae. *Nature Astronomy*, 6:751–759.

D'Alessio, P., Calvet, N., and Hartmann, L. (2001). Accretion Disks around Young Objects. III. Grain Growth. ApJ, 553(1):321–334.

Dorn, C., Harrison, J. H. D., Bonsor, A., and Hands, T. O. (2019). A new class of Super-Earths formed from high-temperature condensates: HD219134 b, 55 Cnc e, WASP-47 e. MNRAS, 484(1):712–727.

dos Santos, L. A., Ehrenreich, D., Bourrier, V., Astudillo-Defru, N., Bonfils, X., Forget, F., Lovis, C., Pepe, F., and Udry, S. (2020). The high-energy environment and atmospheric escape of the mini-Neptune K2-18 b. A&A, 634:L4.

dos Santos, L. A., Ehrenreich, D., Bourrier, V., Lecavelier des Etangs, A., López-Morales, M., Sing, D. K., Ballester, G., Ben-Jaffel, L., Buchhave, L. A., García Muñoz, A., Henry, G. W., Kataria, T., Lavie, B., Lavvas, P., Lewis, N. K., Mikal-Evans, T., Sanz-Forcada, J., and

Wakeford, H. (2019). The Hubble PanCET program: an extensive search for metallic ions in the exosphere of GJ 436 b. A&A, 629:A47.

Dotter, A. (2016). MESA Isochrones and Stellar Tracks (MIST) 0: Methods for the Construction of Stellar Isochrones. ApJS, 222(1):8.

Dressing, C. D. and Charbonneau, D. (2015). The Occurrence of Potentially Habitable Planets Orbiting M Dwarfs Estimated from the Full Kepler Dataset and an Empirical Measurement of the Detection Sensitivity. ApJ, 807(1):45.

Dressing, C. D., Charbonneau, D., Dumusque, X., Gettel, S., Pepe, F., Collier Cameron, A., Latham, D. W., Molinari, E., Udry, S., Affer, L., Bonomo, A. S., Buchhave, L. A., Cosentino, R., Figueira, P., Fiorenzano, A. F. M., Harutyunyan, A., Haywood, R. D., Johnson, J. A., Lopez-Morales, M., Lovis, C., Malavolta, L., Mayor, M., Micela, G., Motalebi, F., Nascimbeni, V., Phillips, D. F., Piotto, G., Pollacco, D., Queloz, D., Rice, K., Sasselov, D., Ségransan, D., Sozzetti, A., Szentgyorgyi, A., and Watson, C. (2015). The Mass of Kepler-93b and The Composition of Terrestrial Planets. ApJ, 800(2):135.

Ehrenreich, D., Bourrier, V., Wheatley, P. J., Lecavelier des Etangs, A., Hébrard, G., Udry, S., Bonfils, X., Delfosse, X., Désert, J.-M., Sing, D. K., and Vidal-Madjar, A. (2015). A giant comet-like cloud of hydrogen escaping the warm Neptune-mass exoplanet GJ 436b. Nature, 522(7557):459–461.

Eker, Z., Bakış, V., Bilir, S., Soydugan, F., Steer, I., Soydugan, E., Bakış, H., Aliçavuş, F., Aslan, G., and Alpsoy, M. (2018). Interrelated main-sequence mass-luminosity, mass-radius, and mass-effective temperature relations. MNRAS, 479(4):5491–5511.

Ercolano, B., Clarke, C. J., and Hall, A. C. (2011). The clearing of discs around late-type T Tauri stars: constraints from the infrared two-colour plane. MNRAS, 410:671–678.

Erkaev, N. V., Kulikov, Y. N., Lammer, H., Selsis, F., Langmayr, D., Jaritz, G. F., and Biernat, H. K. (2007). Roche lobe effects on the atmospheric loss from "Hot Jupiters". A&A, 472(1):329–334.

Erkaev, N. V., Lammer, H., Odert, P., Kislyakova, K. G., Johnstone, C. P., Güdel, M., and Khodachenko, M. L. (2016). EUV-driven mass-loss of protoplanetary cores with hydrogen-dominated atmospheres: the influences of ionization and orbital distance. MNRAS, 460(2):1300–1309.

Erkaev, N. V., Lammer, H., Odert, P., Kulikov, Y. N., Kislyakova, K. G., Khodachenko, M. L., Güdel, M., Hanslmeier, A., and Biernat, H. (2013). XUV-Exposed, Non-Hydrostatic Hydrogen-Rich Upper Atmospheres of Terrestrial Planets. Part I: Atmospheric Expansion and Thermal Escape. *Astrobiology*, 13(11):1011–1029.

Farouki, R. T. (2012). The bernstein polynomial basis: A centennial retrospective. *Computer Aided Geometric Design*, 29(6):379 – 419.

Ferguson, J. W., Alexander, D. R., Allard, F., Barman, T., Bodnarik, J. G., Hauschildt, P. H., Heffner-Wong, A., and Tamanai, A. (2005). Low-Temperature Opacities. ApJ, 623(1):585–596.

Fernandes, R. B., Mulders, G. D., Pascucci, I., Bergsten, G. J., Koskinen, T. T., Hardegree-Ullman, K. K., Pearson, K. A., Giacalone, S., Zink, J., Ciardi, D. R., and O'Brien, P. (2022). Pterodactyls: A Tool to Uniformly Search and Vet for Young Transiting Planets In TESS Primary Mission Photometry. *arXiv e-prints*, page arXiv:2206.03989.

Fischer, D. A. and Valenti, J. (2005). The Planet-Metallicity Correlation. ApJ, 622(2):1102–1117.

Ford, E. B. and Rasio, F. A. (2006). On the Relation between Hot Jupiters and the Roche Limit. ApJ, 638(1):L45–L48.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. (2013). emcee: The mcmc hammer. *PASP*, 125:306–312.

Foreman-Mackey, D., Hogg, D. W., and Morton, T. D. (2014). Exoplanet Population Inference and the Abundance of Earth Analogs from Noisy, Incomplete Catalogs. ApJ, 795(1):64.

Fortney, J. J., Marley, M. S., and Barnes, J. W. (2007). Planetary Radii across Five Orders of Magnitude in Mass and Stellar Insolation: Application to Transits. ApJ, 659(2):1661–1672.

Freedman, R. S., Marley, M. S., and Lodders, K. (2008). Line and Mean Opacities for Ultracool Dwarfs and Extrasolar Planets. ApJS, 174(2):504–513.

Fressin, F., Torres, G., Charbonneau, D., Bryson, S. T., Christiansen, J., Dressing, C. D., Jenkins, J. M., Walkowicz, L. M., and Batalha, N. M. (2013). The False Positive Rate of Kepler and the Occurrence of Planets. ApJ, 766(2):81.

Fulton, B. J. and Petigura, E. A. (2018). The California-Kepler Survey. VII. Precise Planet Radii Leveraging Gaia DR2 Reveal the Stellar Mass Dependence of the Planet Radius Gap. AJ, 156(6):264.

Fulton, B. J., Petigura, E. A., Howard, A. W., Isaacson, H., Marcy, G. W., Cargile, P. A., Hebb, L., Weiss, L. M., Johnson, J. A., Morton, T. D., Sinukoff, E., Crossfield, I. J. M., and Hirsch, L. A. (2017). The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets. AJ, 154(3):109.

Fung, J., Artymowicz, P., and Wu, Y. (2015). The 3D Flow Field Around an Embedded Planet. ApJ, 811(2):101.

Fung, J. and Lee, E. J. (2018). Inner Super-Earths, Outer Gas Giants: How Pebble Isolation and Migration Feedback Keep Jupiters Cold. ApJ, 859(2):126.

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Babusiaux, C., Bailer-Jones, C. A. L., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., Jordi, C., Klioner, S. A., Lammers, U., Lindegren, L., Luri, X., Mignard, F., Panem, C., Pourbaix, D., Randich, S., Sartoretti, P., Siddiqui, H. I., Soubiran, C., van Leeuwen, F., Walton, N. A., Arenou, F., Bastian, U., Cropper, M., Drimmel, R., Katz, D., Lattanzi, M. G., Bakker, J., Cacciari, C., Castañeda, J., Chaoul, L., Cheek, N., De Angeli, F., Fabricius, C., Guerra, R., Holl, B., Masana, E., Messineo, R., Mowlavi, N., Nienartowicz, K., Panuzzo, P., Portell, J., Riello, M., Seabroke, G. M., Tanga, P., Thévenin, F., Gracia-Abril, G., Comoretto, G., Garcia-Reinaldos, M., Teyssier, D., Altmann, M., Andrae, R., Audard, M., Bellas-Velidis, I., Benson, K., Berthier, J., Blomme, R., Burgess, P., Busso, G., Carry, B., Cellino, A., Clementini, G., Clotet, M., Creevey, O., Davidson, M., De Ridder, J., Delchambre, L., Dell'Oro, A., Ducourant, C., Fernández-Hernández, J., Fouesneau, M., Frémat, Y., Galluccio, L., García-Torres, M., González-Núñez, J., González-Vidal, J. J., Gosset, E., Guy, L. P., Halbwachs, J. L., Hambly, N. C., Harrison, D. L., Hernández, J., Hestroffer, D., Hodgkin, S. T., Hutton, A., Jasniewicz, G., Jean-Antoine-Piccolo, A., Jordan, S., Korn, A. J., Krone-Martins, A., Lanzafame, A. C., Lebzelter, T., Löffler, W., Manteiga, M., Marrese, P. M., Martín-Fleitas, J. M., Moitinho, A., Mora, A., Muinonen, K., Osinde, J., Pancino, E., Pauwels, T., Petit, J. M., Recio-Blanco, A., Richards, P. J., Rimoldini, L., Robin, A. C., Sarro, L. M., Siopis, C., Smith, M., Sozzetti, A., Süveges, M., Torra, J., van Reeven, W., Abbas, U., Abreu Aramburu, A., Accart, S., Aerts, C., Altavilla, G., Álvarez, M. A., Alvarez, R., Alves, J., Anderson, R. I., Andrei, A. H., Anglada Varela, E., Antiche, E., Antoja, T., Arcay, B., Astraatmadja, T. L., Bach, N., Baker, S. G., Balaguer-Núñez, L., Balm, P., Barache, C., Barata, C., Barbato, D.,

Barblan, F., Barklem, P. S., Barrado, D., Barros, M., Barstow, M. A., Bartholomé Muñoz, S., Bassilana, J. L., Becciani, U., Bellazzini, M., Berihuete, A., Bertone, S., Bianchi, L., Bienaymé, O., Blanco-Cuaresma, S., Boch, T., Boeche, C., Bombrun, A., Borrachero, R., Bossini, D., Bouquillon, S., Bourda, G., Bragaglia, A., Bramante, L., Breddels, M. A., Bressan, A., Brouillet, N., Brüsemeister, T., Brugaletta, E., Bucciarelli, B., Burlacu, A., Busonero, D., Butkevich, A. G., Buzzi, R., Caffau, E., Cancelliere, R., Cannizzaro, G., Cantat-Gaudin, T., Carballo, R., Carlucci, T., Carrasco, J. M., Casamiquela, L., Castellani, M., Castro-Ginard, A., Charlot, P., Chemin, L., Chiavassa, A., Cocozza, G., Costigan, G., Cowell, S., Crifo, F., Crosta, M., Crowley, C., Cuypers, J., Dafonte, C., Damerdji, Y., Dapergolas, A., David, P., David, M., de Laverny, P., De Luise, F., De March, R., de Martino, D., de Souza, R., de Torres, A., Debosscher, J., del Pozo, E., Delbo, M., Delgado, A., Delgado, H. E., Di Matteo, P., Diakite, S., Diener, C., Distefano, E., Dolding, C., Drazinos, P., Durán, J., Edvardsson, B., Enke, H., Eriksson, K., Esquej, P., Eynard Bontemps, G., Fabre, C., Fabrizio, M., Faigler, S., Falcão, A. J., Farràs Casas, M., Federici, L., Fedorets, G., Fernique, P., Figueras, F., Filippi, F., Findeisen, K., Fonti, A., Fraile, E., Fraser, M., Frézouls, B., Gai, M., Galleti, S., Garabato, D., García-Sedano, F., Garofalo, A., Garralda, N., Gavel, A., Gavras, P., Gerssen, J., Geyer, R., Giacobbe, P., Gilmore, G., Girona, S., Giuffrida, G., Glass, F., Gomes, M., Granvik, M., Gueguen, A., Guerrier, A., Guiraud, J., Gutiérrez-Sánchez, R., Haigron, R., Hatzidimitriou, D., Hauser, M., Haywood, M., Heiter, U., Helmi, A., Heu, J., Hilger, T., Hobbs, D., Hofmann, W., Holland, G., Huckle, H. E., Hypki, A., Icardi, V., Janßen, K., Jevardat de Fombelle, G., Jonker, P. G., Juhász, Á. L., Julbe, F., Karampelas, A., Kewley, A., Klar, J., Kochoska, A., Kohley, R., Kolenberg, K., Kontizas, M., Kontizas, E., Koposov, S. E., Kordopatis, G., Kostrzewa-Rutkowska, Z., Koubsky, P., Lambert, S., Lanza, A. F., Lasne, Y., Lavigne, J. B., Le Fustec, Y., Le Poncin-Lafitte, C., Lebreton, Y., Leccia, S., Leclerc, N., Lecoeur-Taibi, I., Lenhardt, H., Leroux, F., Liao, S., Licata, E., Lindstrøm, H. E. P., Lister, T. A., Livanou, E., Lobel, A., López, M., Managau, S., Mann, R. G., Mantelet, G., Marchal, O., Marchant, J. M., Marconi, M., Marinoni, S., Marschalkó, G., Marshall, D. J., Martino, M., Marton, G., Mary, N., Massari, D., Matijevič, G., Mazeh, T., McMillan, P. J., Messina, S., Michalik, D., Millar, N. R., Molina, D., Molinaro, R., Molnár, L., Montegriffo, P., Mor, R., Morbidelli, R., Morel, T., Morris, D., Mulone, A. F., Muraveva, T., Musella, I., Nelemans, G., Nicastro, L., Noval, L., O'Mullane, W., Ordénovic, C., Ordóñez-Blanco, D., Osborne, P., Pagani, C., Pagano, I., Pailler, F., Palacin, H., Palaversa, L., Panahi, A., Pawlak, M., Piersimoni, A. M., Pineau, F. X., Plachy, E., Plum, G., Poggio, E., Poujoulet, E., Prša, A., Pulone, L., Racero, E., Ragaini, S., Rambaux, N., Ramos-Lerate, M., Regibo, S., Reylé, C., Riclet, F., Ripepi, V., Riva, A., Rivard, A., Rixon, G., Roegiers, T., Roelens, M., Romero-Gómez, M., Rowell, N., Royer, F., Ruiz-Dern, L., Sadowski, G., Sagristà Sellés, T., Sahlmann, J., Salgado, J., Salguero, E., Sanna, N., Santana-Ros, T., Sarasso, M., Savietto, H., Schultheis, M., Sciacca, E., Segol, M., Segovia, J. C., Ségransan, D., Shih, I. C., Siltala, L., Silva, A. F., Smart, R. L., Smith, K. W., Solano, E., Solitro, F., Sordo, R., Soria Nieto, S., Souchay, J., Spagna, A., Spoto, F., Stampa, U., Steele, I. A., Steidelmüller, H., Stephenson, C. A., Stoev, H., Suess, F. F., Surdej, J., Szabados, L., Szegedi-Elek, E., Tapiador, D., Taris, F., Tauran, G., Taylor, M. B., Teixeira, R., Terrett, D., Teyssand ier, P., Thuillot, W., Titarenko, A., Torra Clotet, F., Turon, C., Ulla, A., Utrilla, E., Uzzi, S., Vaillant, M., Valentini, G., Valette, V., van Elteren, A., Van Hemelryck, E., van Leeuwen, M., Vaschetto, M., Vecchiato, A., Veljanoski, J., Viala, Y., Vicente, D., Vogt, S., von Essen, C., Voss, H., Votruba, V., Voutsinas, S., Walmsley, G., Weiler, M., Wertz, O., Wevers, T., Wyrzykowski, Ł., Yoldas, A., Žerjal, M., Ziaeepour, H., Zorec, J., Zschocke, S., Zucker, S., Zurbach, C., and Zwitter, T. (2018). Gaia Data Release 2. Summary of the contents and survey properties. A&A, 616:A1.

García Muñoz, A. (2007). Physical and chemical aeronomy of HD 209458b. Planet. Space Sci., 55(10):1426–1455.

Gaudi, B. S., Bennett, D. P., Udalski, A., Gould, A., Christie, G. W., Maoz, D., Dong, S., McCormick, J., Szymański, M. K., Tristram, P. J., Nikolaev, S., Paczyński, B., Kubiak, M., Pietrzyński, G., Soszyński, I., Szewczyk, O., Ulaczyk, K., Wyrzykowski, Ł., OGLE Collaboration, DePoy, D. L., Han, C., Kaspi, S., Lee, C. U., Mallia, F., Natusch, T., Pogge, R. W., Park, B. G., μ-Fun Collabortion, Abe, F., Bond, I. A., Botzler, C. S., Fukui, A., Hearnshaw, J. B., Itow, Y., Kamiya, K., Korpela, A. V., Kilmartin, P. M., Lin, W., Masuda, K., Matsubara, Y., Motomura, M., Muraki, Y., Nakamura, S., Okumura, T., Ohnishi, K., Rattenbury, N. J., Sako, T., Saito, T., Sato, S., Skuljan, L., Sullivan, D. J., Sumi, T., Sweatman, W. L., Yock, P. C. M., MOA Collaboration, Albrow, M. D., Allan, A., Beaulieu, J. P., Burgdorf, M. J., Cook, K. H., Coutures, C., Dominik, M., Dieters, S., Fouqué, P., Greenhill, J., Horne, K., Steele, I., Tsapras, Y., Planet Collaboration, RoboNet Collaborations, Chaboyer, B., Crocker, A., Frank, S., and Macintosh, B. (2008). Discovery of a Jupiter/Saturn Analog with Gravitational Microlensing. *Science*, 319(5865):927.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.*, 7(4):457–472.

Genda, H. and Abe, Y. (2003). Survival of a proto-atmosphere through the stage of giant impacts: the mechanical aspects. Icarus, 164(1):149–162.

Genda, H. and Abe, Y. (2005). Enhanced atmospheric loss on protoplanets at the giant impact phase in the presence of oceans. Nature, 433(7028):842–844.

Giacconi, R., Branduardi, G., Briel, U., Epstein, A., Fabricant, D., Feigelson, E., Forman, W., Gorenstein, P., Grindlay, J., Gursky, H., Harnden, F. R., Henry, J. P., Jones, C., Kellogg, E., Koch, D., Murray, S., Schreier, E., Seward, F., Tananbaum, H., Topka, K., Van Speybroeck, L., Holt, S. S., Becker, R. H., Boldt, E. A., Serlemitsos, P. J., Clark, G., Canizares, C., Markert, T., Novick, R., Helfand, D., and Long, K. (1979). The Einstein (HEAO 2) X-ray Observatory. ApJ, 230:540–550.

Ginzburg, S., Schlichting, H. E., and Sari, R. (2016). Super-Earth Atmospheres: Self-consistent Gas Accretion and Retention. ApJ, 825:29.

Ginzburg, S., Schlichting, H. E., and Sari, R. (2018). Core-powered mass-loss and the radius distribution of small exoplanets. MNRAS, 476(1):759–765.

Glassgold, A. E., Najita, J., and Igea, J. (2004). Heating Protoplanetary Disk Atmospheres. ApJ, 615(2):972–990.

Goodman, J. and Weare, J. (2010). Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80.

Green, J., Schechter, P., Baltay, C., Bean, R., Bennett, D., Brown, R., Conselice, C., Donahue, M., Fan, X., Gaudi, B. S., Hirata, C., Kalirai, J., Lauer, T., Nichol, B., Padmanabhan, N., Perlmutter, S., Rauscher, B., Rhodes, J., Roellig, T., Stern, D., Sumi, T., Tanner, A., Wang, Y., Weinberg, D., Wright, E., Gehrels, N., Sambruna, R., Traub, W., Anderson, J., Cook, K., Garnavich, P., Hillenbrand, L., Ivezic, Z., Kerins, E., Lunine, J., McDonald, P., Penny, M., Phillips, M., Rieke, G., Riess, A., van der Marel, R., Barry, R. K., Cheng, E., Content, D., Cutri, R., Goullioud, R., Grady, K., Helou, G., Jackson, C., Kruk, J., Melton, M., Peddie, C., Rioux, N., and Seiffert, M. (2012). Wide-Field InfraRed Survey Telescope (WFIRST) Final Report. *arXiv e-prints*, page arXiv:1208.4012.

Güdel, M. (2004). X-ray astronomy of stellar coronae. A&A Rev., 12(2-3):71–237.

Güdel, M., Guinan, E. F., and Skinner, S. L. (1997). The X-Ray Sun in Time: A Study of the Long-Term Evolution of Coronae of Solar-Type Stars. ApJ, 483(2):947–960.

Guillot, T. (2010). On the radiative equilibrium of irradiated planetary atmospheres. A&A, 520:A27.

Gupta, A. and Schlichting, H. E. (2019). Sculpting the valley in the radius distribution of small exoplanets as a by-product of planet formation: the core-powered mass-loss mechanism. MNRAS, 487(1):24–33.

Gupta, A. and Schlichting, H. E. (2020). Signatures of the core-powered mass-loss mechanism in the exoplanet population: dependence on stellar properties and observational predictions. MNRAS, 493(1):792–806.

Gupta, A. and Schlichting, H. E. (2021). Caught in the Act: Core-powered Mass-loss Predictions for Observing Atmospheric Escape. *arXiv e-prints*, page arXiv:2103.08785.

Hadden, S. and Lithwick, Y. (2014). Densities and Eccentricities of 139 Kepler Planets from Transit Time Variations. ApJ, 787(1):80.

Haisch, Karl E., J., Lada, E. A., and Lada, C. J. (2001). Circumstellar Disks in the IC 348 Cluster. AJ, 121(4):2065–2074.

Hansen, B. M. S. and Murray, N. (2012). Migration Then Assembly: Formation of Neptune-mass Planets inside 1 AU. ApJ, 751(2):158.

Hansen, M. and Oh, S. P. (2006). Lyman α radiative transfer in a multiphase medium. MNRAS, 367(3):979–1002.

Hayashi, C. (1981). Structure of the Solar Nebula, Growth and Decay of Magnetic Fields and Effects of Magnetic and Turbulent Viscosities on the Nebula. *Progress of Theoretical Physics Supplement*, 70:35–53.

Hempelmann, A., Schmitt, J. H. M. M., Schultz, M., Ruediger, G., and Stepien, K. (1995). Coronal X-ray emission and rotation of cool main-sequence stars. A&A, 294:515–524.

Henyey, L., Vardya, M. S., and Bodenheimer, P. (1965). Studies in Stellar Evolution. III. The Calculation of Model Envelopes. ApJ, 142:841.

Hernández, J., Calvet, N., Briceño, C., Hartmann, L., Vivas, A. K., Muzerolle, J., Downes, J., Allen, L., and Gutermuth, R. (2007). Spitzer Observations of the Orion OB1 Association: Disk Census in the Low-Mass Stars. ApJ, 671(2):1784–1799.

Holman, M. J. and Murray, N. W. (2005). The Use of Transit Timing to Detect Terrestrial-Mass Extrasolar Planets. *Science*, 307(5713):1288–1291.

Howard, A. W., Marcy, G. W., Bryson, S. T., Jenkins, J. M., Rowe, J. F., Batalha, N. M., Borucki, W. J., Koch, D. G., Dunham, E. W., Gautier, Thomas N., I., Van Cleve, J., Cochran, W. D., Latham, D. W., Lissauer, J. J., Torres, G., Brown, T. M., Gilliland, R. L., Buchhave, L. A., Caldwell, D. A., Christensen-Dalsgaard, J., Ciardi, D., Fressin, F., Haas, M. R., Howell, S. B., Kjeldsen, H., Seager, S., Rogers, L., Sasselov, D. D., Steffen, J. H., Basri, G. S., Charbonneau, D., Christiansen, J., Clarke, B., Dupree, A., Fabrycky, D. C., Fischer, D. A., Ford, E. B., Fortney, J. J., Tarter, J., Girouard, F. R., Holman, M. J., Johnson, J. A., Klaus, T. C., Machalek, P., Moorhead, A. V., Morehead, R. C., Ragozzine, D., Tenenbaum, P., Twicken, J. D., Quinn, S. N., Isaacson, H., Shporer, A., Lucas, P. W., Walkowicz, L. M., Welsh, W. F., Boss, A., Devore, E., Gould, A., Smith, J. C., Morris, R. L., Prsa, A., Morton, T. D., Still, M., Thompson, S. E., Mullally, F., Endl, M., and MacQueen, P. J. (2012). Planet Occurrence within 0.25 AU of Solar-type Stars from Kepler. ApJS, 201(2):15.

Howard, A. W., Marcy, G. W., Johnson, J. A., Fischer, D. A., Wright, J. T., Isaacson, H.,

Valenti, J. A., Anderson, J., Lin, D. N. C., and Ida, S. (2010). The Occurrence and Mass Distribution of Close-in Super-Earths, Neptunes, and Jupiters. *Science*, 330(6004):653.

Howe, A. R. and Burrows, A. (2015). Evolutionary Models of Super-Earths and Mini-Neptunes Incorporating Cooling and Mass Loss. ApJ, 808(2):150.

Howell, S. B., Sobeck, C., Haas, M., Still, M., Barclay, T., Mullally, F., Troeltzsch, J., Aigrain, S., Bryson, S. T., Caldwell, D., Chaplin, W. J., Cochran, W. D., Huber, D., Marcy, G. W., Miglio, A., Najita, J. R., Smith, M., Twicken, J. D., and Fortney, J. J. (2014). The K2 Mission: Characterization and Early Results. PASP, 126(938):398.

Hsu, D. C., Ford, E. B., Ragozzine, D., and Ashby, K. (2019). Occurrence Rates of Planets Orbiting FGK Stars: Combining Kepler DR25, Gaia DR2, and Bayesian Inference. AJ, 158(3):109.

Hsu, D. C., Ford, E. B., Ragozzine, D., and Morehead, R. C. (2018). Improving the Accuracy of Planet Occurrence Rates from Kepler Using Approximate Bayesian Computation. AJ, 155(5):205.

Ida, S. and Lin, D. N. C. (2004). Toward a Deterministic Model of Planetary Formation. I. A Desert in the Mass and Semimajor Axis Distributions of Extrasolar Planets. ApJ, 604(1):388–413.

Ida, S. and Lin, D. N. C. (2005). Toward a Deterministic Model of Planetary Formation. III. Mass Distribution of Short-Period Planets around Stars of Various Masses. ApJ, 626(2):1045–1060.

Ida, S. and Lin, D. N. C. (2010). Toward a Deterministic Model of Planetary Formation. VI. Dynamical Interaction and Coagulation of Multiple Rocky Embryos and Super-Earth Systems around Solar-type Stars. ApJ, 719(1):810–830.

Ikoma, M. and Hori, Y. (2012). In Situ Accretion of Hydrogen-rich Atmospheres on Short-period Super-Earths: Implications for the Kepler-11 Planets. ApJ, 753(1):66.

Inamdar, N. K. and Schlichting, H. E. (2016). Stealing the Gas: Giant Impacts and the Large Diversity in Exoplanet Densities. ApJ, 817:L13.

Isella, A., Pérez, L. M., and Carpenter, J. M. (2012). On the Nature of the Transition Disk around LkCa 15. ApJ, 747:136.

Jackson, A. P., Davis, T. A., and Wheatley, P. J. (2012). The coronal X-ray-age relation and its implications for the evaporation of exoplanets. MNRAS, 422(3):2024–2043.

Jackson, B., Jensen, E., Peacock, S., Arras, P., and Penev, K. (2016). Tidal decay and stable Roche-lobe overflow of short-period gaseous exoplanets. *Celestial Mechanics and Dynamical Astronomy*, 126(1-3):227–248.

Jankovic, M. R., Owen, J. E., and Mohanty, S. (2019). Close-in Super-Earths: The first and the last stages of planet formation in an MRI-accreting disc. *Monthly Notices of the Royal Astronomical Society*, 484(2):2296–2308.

Jansen, F., Lumb, D., Altieri, B., Clavel, J., Ehle, M., Erd, C., Gabriel, C., Guainazzi, M., Gondoin, P., Much, R., Munoz, R., Santos, M., Schartel, N., Texier, D., and Vacanti, G. (2001). XMM-Newton observatory. I. The spacecraft and operations. A&A, 365:L1–L6.

Jeans, J. (1925). *The Dynamical Theory of Gases*. Cambridge Library Collection - Physical Sciences. Cambridge University Press, 4 edition.

Jenkins, J. (2019). TESS Discovery of the First Ultra Hot Neptune, LTT9779b. In *AAS/Division for Extreme Solar Systems Abstracts*, volume 51 of *AAS/Division for Extreme Solar Systems Abstracts*, page 103.07.

Jensen, A. G., Redfield, S., Endl, M., Cochran, W. D., Koesterke, L., and Barman, T. (2012). A Detection of Hα in an Exoplanetary Exosphere. ApJ, 751(2):86.

Jin, S. and Mordasini, C. (2018). Compositional Imprints in Density-Distance-Time: A Rocky Composition for Close-in Low-mass Exoplanets from the Location of the Valley of Evaporation. ApJ, 853(2):163.

Jin, S., Mordasini, C., Parmentier, V., van Boekel, R., Henning, T., and Ji, J. (2014). Planetary Population Synthesis Coupled with Atmospheric Escape: A Statistical View of Evaporation. ApJ, 795(1):65.

Johnson, J. A., Petigura, E. A., Fulton, B. J., Marcy, G. W., Howard, A. W., Isaacson, H., Hebb, L., Cargile, P. A., Morton, T. D., Weiss, L. M., Winn, J. N., Rogers, L. A., Sinukoff, E., and Hirsch, L. A. (2017a). The California-Kepler Survey. II. Precise Physical Properties of 2025 Kepler Planets and Their Host Stars. AJ, 154(3):108.

Johnson, J. A., Petigura, E. A., Fulton, B. J., Marcy, G. W., Howard, A. W., Isaacson, H., Hebb, L., Cargile, P. A., Morton, T. D., Weiss, L. M., Winn, J. N., Rogers, L. A., Sinukoff, E., and Hirsch, L. A. (2017b). The California-Kepler Survey. II. Precise Physical Properties of 2025 Kepler Planets and Their Host Stars. AJ, 154(3):108.

Johnstone, C. P., Bartel, M., and Güdel, M. (2021). The active lives of stars: A complete description of the rotation and XUV evolution of F, G, K, and M dwarfs. A&A, 649:A96.

Johnstone, C. P., Güdel, M., Brott, I., and Lüftinger, T. (2015). Stellar winds on the main-sequence. II. The evolution of rotation and winds. A&A, 577:A28.

Jontof-Hutter, D., Ford, E. B., Rowe, J. F., Lissauer, J. J., Fabrycky, D. C., Van Laerhoven, C., Agol, E., Deck, K. M., Holczer, T., and Mazeh, T. (2016). Secure Mass Measurements from Transit Timing: 10 Kepler Exoplanets between 3 and 8 $M_\oplus$ with Diverse Densities and Incident Fluxes. ApJ, 820(1):39.

Kalas, P., Graham, J. R., Chiang, E., Fitzgerald, M. P., Clampin, M., Kite, E. S., Stapelfeldt, K., Marois, C., and Krist, J. (2008). Optical Images of an Exosolar Planet 25 Light-Years from Earth. *Science*, 322(5906):1345.

Kenyon, S. J. and Hartmann, L. (1995). Pre-Main-Sequence Evolution in the Taurus-Auriga Molecular Cloud. ApJS, 101:117.

Khodachenko, M. L., Shaikhislamov, I. F., Lammer, H., and Prokopov, P. A. (2015). Atmosphere Expansion and Mass Loss of Close-orbit Giant Exoplanets Heated by Stellar XUV. II. Effects of Planetary Magnetic Field; Structuring of Inner Magnetosphere. ApJ, 813(1):50.

King, G. W. and Wheatley, P. J. (2021). EUV irradiation of exoplanet atmospheres occurs on Gyr time-scales. MNRAS, 501(1):L28–L32.

King, G. W., Wheatley, P. J., Salz, M., Bourrier, V., Czesla, S., Ehrenreich, D., Kirk, J., Lecavelier des Etangs, A., Louden, T., Schmitt, J., and Schneider, P. C. (2018). The XUV environments of exoplanets from Jupiter-size to super-Earth. MNRAS, 478(1):1193–1208.

Kippenhahn, R. and Weigert, A. (1994). *Stellar Structure and Evolution*.

Kipping, D. M. and Sandford, E. (2016). Observational biases for transiting planets. MNRAS, 463(2):1323–1331.

Koepferl, C. M., Ercolano, B., Dale, J., Teixeira, P. S., Ratzka, T., and Spezzi, L. (2013). Disc clearing of young stellar objects: evidence for fast inside-out dispersal. MNRAS, 428:3327–3354.

Kubyshkina, D., Vidotto, A. A., Fossati, L., and Farrell, E. (2020). Coupling thermal evolution of planets and hydrodynamic atmospheric escape in MESA. MNRAS, 499(1):77–88.

Kulow, J. R., France, K., Linsky, J., and Loyd, R. O. P. (2014). Lyα Transit Spectroscopy and the Neutral Hydrogen Tail of the Hot Neptune GJ 436b. ApJ, 786(2):132.

Lamers, H. J. G. L. M. and Cassinelli, J. P. (1999). *Introduction to Stellar Winds*.

Lammer, H., Erkaev, N. V., Fossati, L., Juvan, I., Odert, P., Cubillos, P. E., Guenther, E., Kislyakova, K. G., Johnstone, C. P., Lüftinger, T., and Güdel, M. (2016). Identifying the 'true' radius of the hot sub-Neptune CoRoT-24b by mass-loss modelling. MNRAS, 461(1):L62–L66.

Lammer, H., Erkaev, N. V., Odert, P., Kislyakova, K. G., Leitzinger, M., and Khodachenko, M. L. (2013). Probing the blow-off criteria of hydrogen-rich 'super-Earths'. MNRAS, 430(2):1247–1256.

Lammer, H., Selsis, F., Ribas, I., Guinan, E. F., Bauer, S. J., and Weiss, W. W. (2003). Atmospheric Loss of Exoplanets Resulting from Stellar X-Ray and Extreme-Ultraviolet Heating. ApJ, 598(2):L121–L124.

Lammer, H., Stökl, A., Erkaev, N. V., Dorfi, E. A., Odert, P., Güdel, M., Kulikov, Y. N., Kislyakova, K. G., and Leitzinger, M. (2014). Origin and loss of nebula-captured hydrogen envelopes from 'sub'- to 'super-Earths' in the habitable zone of Sun-like stars. MNRAS, 439(4):3225–3238.

Laughlin, G., Crismani, M., and Adams, F. C. (2011). On the Anomalous Radii of the Transiting Extrasolar Planets. ApJ, 729(1):L7.

Lavie, B., Ehrenreich, D., Bourrier, V., Lecavelier des Etangs, A., Vidal-Madjar, A., Delfosse, X., Gracia Berna, A., Heng, K., Thomas, N., Udry, S., and Wheatley, P. J. (2017). The long egress of GJ 436b's giant exosphere. A&A, 605:L7.

Lecavelier Des Etangs, A. (2007). A diagram to determine the evaporation status of extrasolar planets. A&A, 461(3):1185–1193.

Lee, E. J. and Chiang, E. (2015a). To Cool is to Accrete: Analytic Scalings for Nebular Accretion of Planetary Atmospheres. ApJ, 811:41.

Lee, E. J. and Chiang, E. (2015b). To Cool is to Accrete: Analytic Scalings for Nebular Accretion of Planetary Atmospheres. ApJ, 811(1):41.

Lee, E. J. and Chiang, E. (2016). Breeding Super-Earths and Birthing Super-puffs in Transitional Disks. ApJ, 817(2):90.

Lee, E. J. and Chiang, E. (2017). Magnetospheric Truncation, Tidal Inspiral, and the Creation of Short-period and Ultra-short-period Planets. ApJ, 842(1):40.

Lee, E. J., Chiang, E., and Ferguson, J. W. (2018). Optically thin core accretion: how planets get their gas in nearly gas-free discs. MNRAS, 476(2):2199–2208.

Lee, E. J., Chiang, E., and Ormel, C. W. (2014). Make Super-Earths, Not Jupiters: Accreting Nebular Gas onto Solid Cores at 0.1 AU and Beyond. ApJ, 797:95.

Li, Z. Y., Banerjee, R., Pudritz, R. E., Jørgensen, J. K., Shang, H., Krasnopolsky, R., and Maury, A. (2014). The Earliest Stages of Star and Planet Formation: Core Collapse, and

the Formation of Disks and Outflows. In Beuther, H., Klessen, R. S., Dullemond, C. P., and Henning, T., editors, *Protostars and Planets VI*, page 173.

Liu, S.-F., Hori, Y., Lin, D. N. C., and Asphaug, E. (2015). Giant Impact: An Efficient Mechanism for the Devolatilization of Super-Earths. ApJ, 812(2):164.

Lopez, E. D. and Fortney, J. J. (2013). The Role of Core Mass in Controlling Evaporation: The Kepler Radius Distribution and the Kepler-36 Density Dichotomy. ApJ, 776(1):2.

Lopez, E. D. and Fortney, J. J. (2014). Understanding the Mass-Radius Relation for Sub-neptunes: Radius as a Proxy for Composition. ApJ, 792:1.

Lopez, E. D., Fortney, J. J., and Miller, N. (2012). How Thermal Evolution and Mass-loss Sculpt Populations of Super-Earths and Sub-Neptunes: Application to the Kepler-11 System and Beyond. ApJ, 761(1):59.

Lopez, E. D. and Rice, K. (2018). How formation time-scales affect the period dependence of the transition between rocky super-Earths and gaseous sub-Neptunesand implications for $\eta_{\oplus}$. MNRAS, 479(4):5303–5311.

Loyd, R. O. P. and France, K. (2014). Fluctuations and Flares in the Ultraviolet Line Emission of Cool Stars: Implications for Exoplanet Transit Observations. ApJS, 211(1):9.

Loyd, R. O. P., Shkolnik, E. L., Schneider, A. C., Richey-Yowell, T., Barman, T. S., Peacock, S., and Pagano, I. (2020). Current Population Statistics Do Not Favor Photoevaporation over Core-powered Mass Loss as the Dominant Cause of the Exoplanet Radius Gap. ApJ, 890(1):23.

Luhman, K. L., Allen, P. R., Espaillat, C., Hartmann, L., and Calvet, N. (2010). The Disk Population of the Taurus Star-Forming Region. ApJS, 186(1):111–174.

Mamajek, E. E. (2009). Initial Conditions of Planet Formation: Lifetimes of Primordial Disks. In Usuda, T., Tamura, M., and Ishii, M., editors, *American Institute of Physics Conference Series*, volume 1158 of *American Institute of Physics Conference Series*, pages 3–10.

Mandel, K. and Agol, E. (2002). Analytic Light Curves for Planetary Transit Searches. ApJ, 580(2):L171–L175.

Mann, A. W., Johnson, M. C., Vanderburg, A., Kraus, A. L., Rizzuto, A. C., Wood, M. L., Bush, J. L., Rockcliffe, K., Newton, E. R., Latham, D. W., Mamajek, E. E., Zhou, G., Quinn, S. N., Thao, P. C., Benatti, S., Cosentino, R., Desidera, S., Harutyunyan, A., Lovis, C., Mortier, A., Pepe, F. A., Poretti, E., Wilson, T. G., Kristiansen, M. H., Gagliano, R., Jacobs, T., LaCourse, D. M., Omohundro, M., Schwengeler, H. M., Terentev, I. A., Kane, S. R., Hill, M. L., Rabus, M., Esquerdo, G. A., Berlind, P., Collins, K. A., Murawski, G., Sallam, N. H., Aitken, M. M., Massey, B., Ricker, G. R., Vanderspek, R., Seager, S., Winn, J. N., Jenkins, J. M., Barclay, T., Caldwell, D. A., Dragomir, D., Doty, J. P., Glidden, A., Tenenbaum, P., Torres, G., Twicken, J. D., and Villanueva, Steven, J. (2020). TESS Hunt for Young and Maturing Exoplanets (THYME). III. A Two-planet System in the 400 Myr Ursa Major Group. AJ, 160(4):179.

Marois, C., Macintosh, B., Barman, T., Zuckerman, B., Song, I., Patience, J., Lafrenière, D., and Doyon, R. (2008). Direct Imaging of Multiple Planets Orbiting the Star HR 8799. *Science*, 322(5906):1348.

Massol, H., Hamano, K., Tian, F., Ikoma, M., Abe, Y., Chassefière, E., Davaille, A., Genda, H., Güdel, M., Hori, Y., Leblanc, F., Marcq, E., Sarda, P., Shematovich, V. I., Stökl, A., and Lammer, H. (2016). Formation and Evolution of Protoatmospheres. Space Sci. Rev., 205(1-4):153–211.

Mathur, S., Huber, D., Batalha, N. M., Ciardi, D. R., Bastien, F. A., Bieryla, A., Buchhave, L. A., Cochran, W. D., Endl, M., Esquerdo, G. A., Furlan, E., Howard, A., Howell, S. B., Isaacson, H., Latham, D. W., MacQueen, P. J., and Silva, D. R. (2017). Revised Stellar Properties of Kepler Targets for the Q1-17 (DR25) Transit Detection Run. ApJS, 229(2):30.

Mayor, M., Marmier, M., Lovis, C., Udry, S., Ségransan, D., Pepe, F., Benz, W., Bertaux, J. L., Bouchy, F., Dumusque, X., Lo Curto, G., Mordasini, C., Queloz, D., and Santos, N. C. (2011). The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets. *arXiv e-prints*, page arXiv:1109.2497.

Mayor, M., Pepe, F., Queloz, D., Bouchy, F., Rupprecht, G., Lo Curto, G., Avila, G., Benz, W., Bertaux, J. L., Bonfils, X., Dall, T., Dekker, H., Delabre, B., Eckert, W., Fleury, M., Gilliotte, A., Gojak, D., Guzman, J. C., Kohler, D., Lizon, J. L., Longinotti, A., Lovis, C., Megevand, D., Pasquini, L., Reyes, J., Sivan, J. P., Sosnowska, D., Soto, R., Udry, S., van Kesteren, A., Weber, L., and Weilenmann, U. (2003). Setting New Standards with HARPS. *The Messenger*, 114:20–24.

Mayor, M. and Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. Nature, 378(6555):355–359.

Mayor, M., Udry, S., Lovis, C., Pepe, F., Queloz, D., Benz, W., Bertaux, J. L., Bouchy, F., Mordasini, C., and Segransan, D. (2009). The HARPS search for southern extra-solar planets. XIII. A planetary system with 3 super-Earths (4.2, 6.9, and 9.2 $M_{\oplus}$). A&A, 493(2):639–644.

McDonald, G. D., Kreidberg, L., and Lopez, E. (2019). The Sub-Neptune Desert and Its Dependence on Stellar Type: Controlled by Lifetime X-Ray Irradiation. ApJ, 876(1):22.

Michel, A., Haldemann, J., Mordasini, C., and Alibert, Y. (2020). Planetary mass-radius relations across the galaxy. *arXiv e-prints*, page arXiv:2006.03601.

Miller, N. and Fortney, J. J. (2011). The Heavy-element Masses of Extrasolar Giant Planets, Revealed. ApJ, 736(2):L29.

Misener, W. and Schlichting, H. E. (2022). The importance of silicate vapour in determining the structure, radii, and envelope mass fractions of sub-Neptunes. MNRAS, 514(4):6025–6037.

Mizuno, H. (1980). Formation of the Giant Planets. *Progress of Theoretical Physics*, 64(2):544–557.

Mizuno, H., Nakazawa, K., and Hayashi, C. (1978). Instability of a gaseous envelope surrounding a planetary core and formation of giant planets. *Progress of Theoretical Physics*, 60:699–710.

Mordasini, C. (2020). Planetary evolution with atmospheric photoevaporation I. Analytical derivation and numerical study of the evaporation valley and transition from super-Earths to sub-Neptunes. *arXiv e-prints*, page arXiv:2002.02455.

Mróz, P., Poleski, R., Gould, A., Udalski, A., Sumi, T., Szymański, M. K., Soszyński, I., Pietrukowicz, P., Kozłowski, S., Skowron, J., Ulaczyk, K., OGLE Collaboration, Albrow, M. D., Chung, S.-J., Han, C., Hwang, K.-H., Jung, Y. K., Kim, H.-W., Ryu, Y.-H., Shin, I.-G., Shvartzvald, Y., Yee, J. C., Zang, W., Cha, S.-M., Kim, D.-J., Kim, S.-L., Lee, C.-U., Lee, D.-J., Lee, Y., Park, B.-G., Pogge, R. W., and KMT Collaboration (2020). A Terrestrial-mass Rogue Planet Candidate Detected in the Shortest-timescale Microlensing Event. ApJ, 903(1):L11.

Mulders, G. D., Pascucci, I., and Apai, D. (2015). A Stellar-mass-dependent Drop in Planet Occurrence Rates. ApJ, 798(2):112.

Mulders, G. D., Pascucci, I., Apai, D., and Ciesla, F. J. (2018). The Exoplanet Population Observation Simulator. I. The Inner Edges of Planetary Systems. AJ, 156(1):24.

Murray-Clay, R. A., Chiang, E. I., and Murray, N. (2009). Atmospheric Escape From Hot Jupiters. ApJ, 693(1):23–42.

Neil, A. R., Liston, J., and Rogers, L. A. (2022). Evaluating the Evidence for Water World Populations using Mixture Models. *arXiv e-prints*, page arXiv:2205.00006.

Neil, A. R. and Rogers, L. A. (2020). A Joint Mass Radius Period Distribution of Exoplanets. ApJ, 891(1):12.

Newton, E. R., Mann, A. W., Kraus, A. L., Livingston, J. H., Vanderburg, A., Curtis, J. L., Thao, P. C., Hawkins, K., Wood, M. L., Rizzuto, A. C., Soubkiou, A., Tofflemire, B. M., Zhou, G., Crossfield, I. J. M., Pearce, L. A., Collins, K. A., Conti, D. M., Tan, T.-G., Villeneuva, S., Spencer, A., Dragomir, D., Quinn, S. N., Jensen, E. L. N., Collins, K. I., Stockdale, C., Cloutier, R., Hellier, C., Benkhaldoun, Z., Ziegler, C., Briceño, C., Law, N., Benneke, B., Christiansen, J. L., Gorjian, V., Kane, S. R., Kreidberg, L., Morales, F. Y., Werner, M. W., Twicken, J. D., Levine, A. M., Ciardi, D. R., Guerrero, N. M., Hesse, K., Quintana, E. V., Shiao, B., Smith, J. C., Torres, G., Ricker, G. R., Vanderspek, R., Seager, S., Winn, J. N., Jenkins, J. M., and Latham, D. W. (2021). TESS Hunt for Young and Maturing Exoplanets (THYME). IV. Three Small Planets Orbiting a 120 Myr Old Star in the Pisces-Eridanus Stream. AJ, 161(2):65.

Newton, E. R., Mann, A. W., Tofflemire, B. M., Pearce, L., Rizzuto, A. C., Vanderburg, A., Martinez, R. A., Wang, J. J., Ruffio, J.-B., Kraus, A. L., Johnson, M. C., Thao, P. C., Wood, M. L., Rampalli, R., Nielsen, E. L., Collins, K. A., Dragomir, D., Hellier, C., Anderson, D. R., Barclay, T., Brown, C., Feiden, G., Hart, R., Isopi, G., Kielkopf, J. F., Mallia, F., Nelson, P., Rodriguez, J. E., Stockdale, C., Waite, I. A., Wright, D. J., Lissauer, J. J., Ricker, G. R., Vanderspek, R., Latham, D. W., Seager, S., Winn, J. N., Jenkins, J. M., Bouma, L. G., Burke, C. J., Davies, M., Fausnaugh, M., Li, J., Morris, R. L., Mukai, K., Villaseñor, J., Villeneuva, S., De Rosa, R. J., Macintosh, B., Mengel, M. W., Okumura, J., and Wittenmyer, R. A. (2019). TESS Hunt for Young and Maturing Exoplanets (THYME): A Planet in the 45 Myr Tucana-Horologium Association. ApJ, 880(1):L17.

Ninan, J. P., Stefansson, G., Mahadevan, S., Bender, C., Robertson, P., Ramsey, L., Terrien, R., Wright, J., Diddams, S. A., Kanodia, S., Cochran, W., Endl, M., Ford, E. B., Fredrick, C., Halverson, S., Hearty, F., Jennings, J., Kaplan, K., Lubar, E., Metcalf, A. J., Monson, A., Nitroy, C., Roy, A., and Schwab, C. (2020). Evidence for He I 10830 Å Absorption during the Transit of a Warm Neptune around the M-dwarf GJ 3470 with the Habitable-zone Planet Finder. ApJ, 894(2):97.

Ning, B., Wolfgang, A., and Ghosh, S. (2018). Predicting Exoplanet Masses and Radii: A Nonparametric Approach. ApJ, 869(1):5.

O'Brien, T. A., Collins, W. D., Rauscher, S. A., and Ringler, T. D. (2014). Reducing the computational cost of the ECF using a nuFFT: A fast and objective probability density estimation method. *"Computational Statistics and Data Analysis"*, 79:222 – 234.

O'Brien, T. A., Kashinath, K., Cavanaugh, N. R., Collins, W. D., and O'Brien, J. P. (2016). A fast and objective multidimensional kernel density estimation method: fastKDE. *"Computational Statistics and Data Analysis"*, 101:148 – 160.

Ogilvie, G. (2005). Lecture course on accretion discs.

Ormel, C. W., Shi, J.-M., and Kuiper, R. (2015). Hydrodynamics of embedded planets' first atmospheres - II. A rapid recycling of atmospheric gas. MNRAS, 447(4):3512–3525.

Ormel, C. W., Vazan, A., and Brouwers, M. G. (2021). How planets grow by pebble accretion. III. Emergence of an interior composition gradient. A&A, 647:A175.

Owen, J. E. (2019). Atmospheric Escape and the Evolution of Close-In Exoplanets. *Annual Review of Earth and Planetary Sciences*, 47:67–90.

Owen, J. E. (2020). Constraining the entropy of formation from young transiting planet. MNRAS, 498(4):5030–5040.

Owen, J. E. and Adams, F. C. (2014). Magnetically controlled mass-loss from extrasolar planets in close orbits. MNRAS, 444(4):3761–3779.

Owen, J. E. and Adams, F. C. (2019). Effects of magnetic fields on the location of the evaporation valley for low-mass exoplanets. MNRAS, 490(1):15–20.

Owen, J. E. and Alvarez, M. A. (2016). UV Driven Evaporation of Close-in Planets: Energy-limited, Recombination-limited, and Photon-limited Flows. ApJ, 816(1):34.

Owen, J. E. and Campos Estrada, B. (2020). Testing exoplanet evaporation with multitransiting systems. MNRAS, 491(4):5287–5297.

Owen, J. E., Ercolano, B., Clarke, C. J., and Alexander, R. D. (2010). Radiation-hydrodynamic models of X-ray and EUV photoevaporating protoplanetary discs. MNRAS, 401(3):1415–1428.

Owen, J. E. and Jackson, A. P. (2012). Planetary evaporation by UV & X-ray radiation: basic hydrodynamics. MNRAS, 425(4):2931–2947.

Owen, J. E. and Kollmeier, J. A. (2019). Radiation pressure clear-out of dusty photoevaporating discs. MNRAS, 487(3):3702–3714.

Owen, J. E. and Lai, D. (2018). Photoevaporation and high-eccentricity migration created the sub-Jovian desert. MNRAS, 479(4):5012–5021.

Owen, J. E. and Morton, T. D. (2016). The Initial Physical Conditions of Kepler-36 b and c. ApJ, 819(1):L10.

Owen, J. E. and Murray-Clay, R. (2018). Metallicity-dependent signatures in the Kepler planets. MNRAS, 480:2206–2216.

Owen, J. E. and Wu, Y. (2013). Kepler Planets: A Tale of Evaporation. ApJ, 775(2):105.

Owen, J. E. and Wu, Y. (2016). Atmospheres of Low-mass Planets: The "Boil-off". ApJ, 817(2):107.

Owen, J. E. and Wu, Y. (2017). The Evaporation Valley in the Kepler Planets. ApJ, 847(1):29.

Panić, O., Hogerheijde, M. R., Wilner, D., and Qi, C. (2009). A break in the gas and dust surface density of the disc around the T Tauri star IM Lupi. A&A, 501:269–278.

Parker, E. N. (1958). Dynamics of the Interplanetary Gas and Magnetic Fields. ApJ, 128:664.

Paxton, B., Bildsten, L., Dotter, A., Herwig, F., Lesaffre, P., and Timmes, F. (2011). Modules for Experiments in Stellar Astrophysics (MESA). ApJS, 192(1):3.

Paxton, B., Cantiello, M., Arras, P., Bildsten, L., Brown, E. F., Dotter, A., Mankovich, C., Montgomery, M. H., Stello, D., Timmes, F. X., and Townsend, R. (2013). Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars. ApJS, 208(1):4.

Paxton, B., Marchant, P., Schwab, J., Bauer, E. B., Bildsten, L., Cantiello, M., Dessart, L., Farmer, R., Hu, H., Langer, N., Townsend, R. H. D., Townsley, D. M., and Timmes, F. X. (2015). Modules for Experiments in Stellar Astrophysics (MESA): Binaries, Pulsations, and Explosions. ApJS, 220(1):15.

Paxton, B., Schwab, J., Bauer, E. B., Bildsten, L., Blinnikov, S., Duffell, P., Farmer, R., Goldberg, J. A., Marchant, P., Sorokina, E., Thoul, A., Townsend, R. H. D., and Timmes, F. X. (2018). Modules for Experiments in Stellar Astrophysics (MESA): Convective Boundaries, Element Diffusion, and Massive Star Explosions. ApJS, 234(2):34.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Pepper, J., Pogge, R. W., DePoy, D. L., Marshall, J. L., Stanek, K. Z., Stutz, A. M., Poindexter, S., Siverd, R., O'Brien, T. P., Trueblood, M., and Trueblood, P. (2007). The Kilodegree Extremely Little Telescope (KELT): A Small Robotic Telescope for Large-Area Synoptic Surveys. PASP, 119(858):923–935.

Perri, F. and Cameron, A. G. W. (1974). Hydrodynamic Instability of the Solar Nebula in the Presence of a Planetary Core. Icarus, 22(4):416–425.

Petigura, E. A. (2020). Two Views of the Radius Gap and the Role of Light Curve Fitting. *arXiv e-prints*, page arXiv:2006.15649.

Petigura, E. A., Howard, A. W., and Marcy, G. W. (2013). Prevalence of Earth-size planets orbiting Sun-like stars. *Proceedings of the National Academy of Science*, 110(48):19273–19278.

Petigura, E. A., Howard, A. W., Marcy, G. W., Johnson, J. A., Isaacson, H., Cargile, P. A., Hebb, L., Fulton, B. J., Weiss, L. M., Morton, T. D., Winn, J. N., Rogers, L. A., Sinukoff, E., Hirsch, L. A., and Crossfield, I. J. M. (2017). The California-Kepler Survey. I. High-resolution Spectroscopy of 1305 Stars Hosting Kepler Transiting Planets. AJ, 154(3):107.

Petigura, E. A., Marcy, G. W., Winn, J. N., Weiss, L. M., Fulton, B. J., Howard, A. W., Sinukoff, E., Isaacson, H., Morton, T. D., and Johnson, J. A. (2018). The California-Kepler Survey. IV. Metal-rich Stars Host a Greater Diversity of Planets. AJ, 155(2):89.

Petigura, E. A., Rogers, J. G., Isaacson, H., Owen, J. E., Kraus, A. L., Winn, J. N., MacDougall, M. G., Howard, A. W., Fulton, B., Kosiarek, M. R., Weiss, L. M., Behmard, A., and Blunt, S. (2022). The California-Kepler Survey. X. The Radius Gap as a Function of Stellar Mass, Metallicity, and Age. AJ, 163(4):179.

Piso, A.-M. A. and Youdin, A. N. (2014). On the Minimum Core Mass for Giant Planet Formation at Wide Separations. ApJ, 786(1):21.

Piso, A.-M. A., Youdin, A. N., and Murray-Clay, R. A. (2015). Minimum Core Masses for Giant Planet Formation with Realistic Equations of State and Opacities. ApJ, 800(2):82.

Pollacco, D. L., Skillen, I., Collier Cameron, A., Christian, D. J., Hellier, C., Irwin, J., Lister, T. A., Street, R. A., West, R. G., Anderson, D. R., Clarkson, W. I., Deeg, H., Enoch, B., Evans, A., Fitzsimmons, A., Haswell, C. A., Hodgkin, S., Horne, K., Kane, S. R., Keenan, F. P., Maxted, P. F. L., Norton, A. J., Osborne, J., Parley, N. R., Ryans, R. S. I., Smalley, B., Wheatley, P. J., and Wilson, D. M. (2006). The WASP Project and the SuperWASP Cameras. PASP, 118(848):1407–1418.

Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., and Greenzweig, Y.

(1996). Formation of the Giant Planets by Concurrent Accretion of Solids and Gas. Icarus, 124(1):62–85.

Powell, D., Murray-Clay, R., and Schlichting, H. E. (2017). Using Ice and Dust Lines to Constrain the Surface Densities of Protoplanetary Disks. ApJ, 840:93.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in FORTRAN. The art of scientific computing.*

Pringle, J. E. (1981). Accretion discs in astrophysics. ARA&A, 19:137–162.

Rafikov, R. R. (2006). Atmospheres of Protoplanetary Cores: Critical Mass for Nucleated Instability. ApJ, 648(1):666–682.

Rauer, H., Catala, C., Aerts, C., Appourchaux, T., Benz, W., Brandeker, A., Christensen-Dalsgaard, J., Deleuil, M., Gizon, L., Goupil, M. J., Güdel, M., Janot-Pacheco, E., Mas-Hesse, M., Pagano, I., Piotto, G., Pollacco, D., Santos, Ċ., Smith, A., Suárez, J. C., Szabó, R., Udry, S., Adibekyan, V., Alibert, Y., Almenara, J. M., Amaro-Seoane, P., Eiff, M. A.-v., Asplund, M., Antonello, E., Barnes, S., Baudin, F., Belkacem, K., Bergemann, M., Bihain, G., Birch, A. C., Bonfils, X., Boisse, I., Bonomo, A. S., Borsa, F., Brandão, I. M., Brocato, E., Brun, S., Burleigh, M., Burston, R., Cabrera, J., Cassisi, S., Chaplin, W., Charpinet, S., Chiappini, C., Church, R. P., Csizmadia, S., Cunha, M., Damasso, M., Davies, M. B., Deeg, H. J., Díaz, R. F., Dreizler, S., Dreyer, C., Eggenberger, P., Ehrenreich, D., Eigmüller, P., Erikson, A., Farmer, R., Feltzing, S., de Oliveira Fialho, F., Figueira, P., Forveille, T., Fridlund, M., García, R. A., Giommi, P., Giuffrida, G., Godolt, M., Gomes da Silva, J., Granzer, T., Grenfell, J. L., Grotsch-Noels, A., Günther, E., Haswell, C. A., Hatzes, A. P., Hébrard, G., Hekker, S., Helled, R., Heng, K., Jenkins, J. M., Johansen, A., Khodachenko, M. L., Kislyakova, K. G., Kley, W., Kolb, U., Krivova, N., Kupka, F., Lammer, H., Lanza, A. F., Lebreton, Y., Magrin, D., Marcos-Arenal, P., Marrese, P. M., Marques, J. P., Martins, J., Mathis, S., Mathur, S., Messina, S., Miglio, A., Montalban, J., Montalto, M., Monteiro, M. J. P. F. G., Moradi, H., Moravveji, E., Mordasini, C., Morel, T., Mortier, A., Nascimbeni, V., Nelson, R. P., Nielsen, M. B., Noack, L., Norton, A. J., Ofir, A., Oshagh, M., Ouazzani, R. M., Pápics, P., Parro, V. C., Petit, P., Plez, B., Poretti, E., Quirrenbach, A., Ragazzoni, R., Raimondo, G., Rainer, M., Reese, D. R., Redmer, R., Reffert, S., Rojas-Ayala, B., Roxburgh, I. W., Salmon, S., Santerne, A., Schneider, J., Schou, J., Schuh, S., Schunker, H., Silva-Valio, A., Silvotti, R., Skillen, I., Snellen, I., Sohl, F., Sousa, S. G., Sozzetti, A., Stello, D., Strassmeier, K. G., Švanda, M., Szabó, G. M., Tkachenko, A., Valencia, D., Van Grootel, V., Vauclair, S. D., Ventura, P., Wagner, F. W., Walton, N. A., Weingrill, J., Werner, S. C., Wheatley, P. J., and Zwintz, K. (2014). The PLATO 2.0 mission. *Experimental Astronomy*, 38(1-2):249–330.

Raymond, S. N., Boulet, T., Izidoro, A., Esteves, L., and Bitsch, B. (2018). Migration-driven diversity of super-Earth compositions. MNRAS, 479(1):L81–L85.

Raymond, S. N. and Cossou, C. (2014). No universal minimum-mass extrasolar nebula: evidence against in situ accretion of systems of hot super-Earths. MNRAS, 440:L11–L15.

Ribas, I., Guinan, E. F., Güdel, M., and Audard, M. (2005). Evolution of the Solar Activity over Time and Effects on Planetary Atmospheres. I. High-Energy Irradiances (1-1700 Å). ApJ, 622(1):680–694.

Richert, A. J. W., Getman, K. V., Feigelson, E. D., Kuhn, M. A., Broos, P. S., Povich, M. S., Bate, M. R., and Garmire, G. P. (2018). Circumstellar disc lifetimes in numerous galactic young stellar clusters. MNRAS, 477(4):5191–5206.

Ricker, G. R., Winn, J. N., Vanderspek, R., Latham, D. W., Bakos, G. Á., Bean, J. L., Berta-

Thompson, Z. K., Brown, T. M., Buchhave, L., Butler, N. R., Butler, R. P., Chaplin, W. J., Charbonneau, D., Christensen-Dalsgaard, J., Clampin, M., Deming, D., Doty, J., De Lee, N., Dressing, C., Dunham, E. W., Endl, M., Fressin, F., Ge, J., Henning, T., Holman, M. J., Howard, A. W., Ida, S., Jenkins, J. M., Jernigan, G., Johnson, J. A., Kaltenegger, L., Kawai, N., Kjeldsen, H., Laughlin, G., Levine, A. M., Lin, D., Lissauer, J. J., MacQueen, P., Marcy, G., McCullough, P. R., Morton, T. D., Narita, N., Paegert, M., Palle, E., Pepe, F., Pepper, J., Quirrenbach, A., Rinehart, S. A., Sasselov, D., Sato, B., Seager, S., Sozzetti, A., Stassun, K. G., Sullivan, P., Szentgyorgyi, A., Torres, G., Udry, S., and Villasenor, J. (2015). Transiting Exoplanet Survey Satellite (TESS). *Journal of Astronomical Telescopes, Instruments, and Systems*, 1:014003.

Rivera, E. J., Lissauer, J. J., Butler, R. P., Marcy, G. W., Vogt, S. S., Fischer, D. A., Brown, T. M., Laughlin, G., and Henry, G. W. (2005). A ~7.5 $M_\oplus$ Planet Orbiting the Nearby Star, GJ 876. ApJ, 634(1):625–640.

Rizzuto, A. C., Newton, E. R., Mann, A. W., Tofflemire, B. M., Vanderburg, A., Kraus, A. L., Wood, M. L., Quinn, S. N., Zhou, G., Thao, P. C., Law, N. M., Ziegler, C., and Briceño, C. (2020). TESS Hunt for Young and Maturing Exoplanets (THYME). II. A 17 Myr Old Transiting Hot Jupiter in the Sco-Cen Association. AJ, 160(1):33.

Rogers, J. G., Gupta, A., Owen, J. E., and Schlichting, H. E. (2021). Photoevaporation versus core-powered mass-loss: model comparison with the 3D radius gap. MNRAS, 508(4):5886–5902.

Rogers, J. G. and Owen, J. E. (2021). Unveiling the planet population at birth. MNRAS, 503(1):1526–1542.

Rogers, L. A. (2015). Most 1.6 Earth-radius Planets are Not Rocky. ApJ, 801(1):41.

Rogers, L. A., Bodenheimer, P., Lissauer, J. J., and Seager, S. (2011). Formation and Structure of Low-density exo-Neptunes. ApJ, 738:59.

Rogers, L. A. and Seager, S. (2010). A Framework for Quantifying the Degeneracies of Exoplanet Interior Compositions. ApJ, 712(2):974–991.

Salz, M., Schneider, P. C., Czesla, S., and Schmitt, J. H. M. M. (2016). Energy-limited escape revised. The transition from strong planetary winds to stable thermospheres. A&A, 585:L2.

Sandoval, A., Contardo, G., and David, T. J. (2020). The Influence of Age on the Relative Frequency of Super-Earths and Sub-Neptunes. *arXiv e-prints*, page arXiv:2012.09239.

Saumon, D., Chabrier, G., and van Horn, H. M. (1995). An Equation of State for Low-Mass Stars and Giant Planets. ApJS, 99:713.

Schlichting, H. E., Sari, R., and Yalinewich, A. (2015). Atmospheric mass loss during planet formation: The importance of planetesimal impacts. Icarus, 247:81–94.

Schlichting, H. E. and Young, E. D. (2022). Chemical Equilibrium between Cores, Mantles, and Atmospheres of Super-Earths and Sub-Neptunes and Implications for Their Compositions, Interiors, and Evolution. *PSJ*, 3(5):127.

Shakura, N. I. and Sunyaev, R. A. (1973). Black holes in binary systems. Observational appearance. A&A, 24:337–355.

Shkolnik, E. L. and Barman, T. S. (2014). HAZMAT. I. The Evolution of Far-UV and Near-UV Emission from Early M Stars. AJ, 148(4):64.

Soderblom, D. R., Jones, B. F., Stauffer, J. R., and Chaboyer, B. (1995). The Evolution of the Lithium Abundances of Solar-Type Stars.V. K Dwarfs in the Hyades. AJ, 110:729.

Spake, J. J., Sing, D. K., Evans, T. M., Oklopčić, , A., Bourrier, V., Kreidberg, L., Rackham, B. V., Irwin, J., Ehrenreich, D., Wyttenbach, A., Wakeford, H. R., Zhou, Y., Chubb, K. L., Nikolov, N., Goyal, J. M., Henry, G. W., Williamson, M. H., Blumenthal, S., Anderson, D. R., Hellier, C., Charbonneau, D., Udry, S., and Madhusudhan, N. (2018). Helium in the eroding atmosphere of an exoplanet. Nature, 557(7703):68–70.

Stone, J. M. and Proga, D. (2009). Anisotropic Winds from Close-In Extrasolar Planets. ApJ, 694(1):205–213.

Thorngren, D. P. and Fortney, J. J. (2018). Bayesian Analysis of Hot-Jupiter Radius Anomalies: Evidence for Ohmic Dissipation? AJ, 155(5):214.

Tian, F., Toon, O. B., Pavlov, A. A., and De Sterck, H. (2005). Transonic Hydrodynamic Escape of Hydrogen from Extrasolar Planetary Atmospheres. ApJ, 621(2):1049–1060.

Trammell, G. B., Arras, P., and Li, Z.-Y. (2011). Hot Jupiter Magnetospheres. ApJ, 728(2):152.

Tripathi, A., Kratter, K. M., Murray-Clay, R. A., and Krumholz, M. R. (2015). Simulated Photoevaporative Mass Loss from Hot Jupiters in 3D. ApJ, 808(2):173.

Tu, L., Johnstone, C. P., Güdel, M., and Lammer, H. (2015). The extreme ultraviolet and X-ray Sun in Time: High-energy evolutionary tracks of a solar-like star. A&A, 577:L3.

Turner, N. J., Fromang, S., Gammie, C., Klahr, H., Lesur, G., Wardle, M., and Bai, X. N. (2014). Transport and Accretion in Planet-Forming Disks. In Beuther, H., Klessen, R. S., Dullemond, C. P., and Henning, T., editors, *Protostars and Planets VI*, page 411.

Valencia, D., O'Connell, R. J., and Sasselov, D. (2006). Internal structure of massive terrestrial planets. Icarus, 181(2):545–554.

Valencia, D., Sasselov, D. D., and O'Connell, R. J. (2007). Detailed Models of Super-Earths: How Well Can We Infer Bulk Properties? ApJ, 665(2):1413–1420.

Valsecchi, F., Rasio, F. A., and Steffen, J. H. (2014). From Hot Jupiters to Super-Earths via Roche Lobe Overflow. ApJ, 793(1):L3.

Van Eylen, V., Agentoft, C., Lundkvist, M. S., Kjeldsen, H., Owen, J. E., Fulton, B. J., Petigura, E., and Snellen, I. (2018). An asteroseismic view of the radius valley: stripped cores, not born rocky. MNRAS, 479(4):4786–4795.

Van Eylen, V., Astudillo-Defru, N., Bonfils, X., Livingston, J., Hirano, T., Luque, R., Lam, K. W. F., Justesen, A. B., Winn, J. N., Gandolfi, D., Nowak, G., Palle, E., Albrecht, S., Dai, F., Campos Estrada, B., Owen, J. E., Foreman-Mackey, D., Fridlund, M., Korth, J., Mathur, S., Forveille, T., Mikal-Evans, T., Osborne, H. L. M., Ho, C. S. K., Almenara, J. M., Artigau, E., Barragán, O., Barros, S. C. C., Bouchy, F., Cabrera, J., Caldwell, D. A., Charbonneau, D., Chaturvedi, P., Cochran, W. D., Csizmadia, S., Damasso, M., Delfosse, X., De Medeiros, J. R., Díaz, R. F., Doyon, R., Esposito, M., Fűrész, G., Figueira, P., Georgieva, I., Goffo, E., Grziwa, S., Guenther, E., Hatzes, A. P., Jenkins, J. M., Kabath, P., Knudstrup, E., Latham, D. W., Lavie, B., Lovis, C., Mennickent, R. E., Mullally, S. E., Murgas, F., Narita, N., Pepe, F. A., Persson, C. M., Redfield, S., Ricker, G. R., Santos, N. C., Seager, S., Serrano, L. M., Smith, A. M. S., Suárez Mascareño, A., Subjak, J., Twicken, J. D., Udry, S., Vanderspek, R., and Zapatero Osorio, M. R. (2021). Masses and compositions of three small planets orbiting the nearby M dwarf L231-32 (TOI-270) and the M dwarf radius valley. MNRAS, 507(2):2154–2173.

Vardya, M. S. (1962). Planck and Rosseland Mean of Rayleigh Scattering by H_2. ApJ, 135:303.

Vidal-Madjar, A., Désert, J. M., Lecavelier des Etangs, A., Hébrard, G., Ballester, G. E., Ehrenreich, D., Ferlet, R., McConnell, J. C., Mayor, M., and Parkinson, C. D. (2004). Detection of Oxygen and Carbon in the Hydrodynamically Escaping Atmosphere of the Extrasolar Planet HD 209458b. ApJ, 604(1):L69–L72.

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J. M., Ballester, G. E., Ferlet, R., Hébrard, G., and Mayor, M. (2003). An extended upper atmosphere around the extrasolar planet HD209458b. Nature, 422(6928):143–146.

Voges, W., Aschenbach, B., Boller, T., Bräuninger, H., Briel, U., Burkert, W., Dennerl, K., Englhauser, J., Gruber, R., Haberl, F., Hartner, G., Hasinger, G., Kürster, M., Pfeffermann, E., Pietsch, W., Predehl, P., Rosso, C., Schmitt, J. H. M. M., Trümper, J., and Zimmermann, H. U. (1999). The ROSAT all-sky survey bright source catalogue. A&A, 349:389–405.

Volkov, A. N., Johnson, R. E., Tucker, O. J., and Erwin, J. T. (2011). Thermally Driven Atmospheric Escape: Transition from Hydrodynamic to Jeans Escape. ApJ, 729(2):L24.

Waite, J. H., Cravens, T. E., Kozyra, J., Nagy, A. F., Atreya, S. K., and Chen, R. H. (1983). Electron precipitation and related aeronomy of the Jovian thermosphere and ionosphere. J. Geophys. Res., 88(A8):6143–6164.

Watson, A. J., Donahue, T. M., and Walker, J. C. G. (1981). The dynamics of a rapidly escaping atmosphere: Applications to the evolution of Earth and Venus. Icarus, 48(2):150–166.

Weidenschilling, S. J. (1977). Aerodynamics of solid bodies in the solar nebula. MNRAS, 180:57–70.

Weiss, L. M. and Marcy, G. W. (2014). The Mass-Radius Relation for 65 Exoplanets Smaller than 4 Earth Radii. ApJ, 783(1):L6.

West, R. G., Gillen, E., Bayliss, D., Burleigh, M. R., Delrez, L., Günther, M. N., Hodgkin, S. T., Jackman, J. A. G., Jenkins, J. S., King, G., McCormac, J., Nielsen, L. D., Raynard, L., Smith, A. M. S., Soto, M., Turner, O., Wheatley, P. J., Almleaky, Y., Armstrong, D. J., Belardi, C., Bouchy, F., Briegal, J. T., Burdanov, A., Cabrera, J., Casewell, S. L., Chaushev, A., Chazelas, B., Chote, P., Cooke, B. F., Csizmadia, S., Ducrot, E., Eigmüller, P., Erikson, A., Foxell, E., Gänsicke, B. T., Gillon, M., Goad, M. R., Jehin, E., Lambert, G., Longstaff, E. S., Louden, T., Moyano, M., Murray, C., Pollacco, D., Queloz, D., Rauer, H., Sohy, S., Thompson, S. J., Udry, S., Walker, S. R., and Watson, C. A. (2019). NGTS-4b: A sub-Neptune transiting in the desert. MNRAS, 486(4):5094–5103.

Winn, J. N., Fabrycky, D., Albrecht, S., and Johnson, J. A. (2010). Hot Stars with Hot Jupiters Have High Obliquities. ApJ, 718(2):L145–L149.

Winter, A. J., Clarke, C. J., Rosotti, G., Ih, J., Facchini, S., and Haworth, T. J. (2018). Protoplanetary disc truncation mechanisms in stellar clusters: comparing external photoevaporation and tidal encounters. MNRAS, 478(2):2700–2722.

Wolfgang, A. and Lopez, E. (2015). How Rocky Are They? The Composition Distribution of Kepler's Sub-Neptune Planet Candidates within 0.15 AU. ApJ, 806(2):183.

Wolfgang, A., Rogers, L. A., and Ford, E. B. (2016). Probabilistic Mass-Radius Relationship for Sub-Neptune-Sized Planets. ApJ, 825(1):19.

Woods, T. N., Eparvier, F. G., Bailey, S. M., Chamberlin, P. C., Lean, J., Rottman, G. J., Solomon, S. C., Tobiska, W. K., and Woodraska, D. L. (2005). Solar EUV Experiment (SEE): Mission overview and first results. *Journal of Geophysical Research (Space Physics)*, 110(A1):A01312.

Wright, J. T., Marcy, G. W., Howard, A. W., Johnson, J. A., Morton, T. D., and Fischer, D. A. (2012). The Frequency of Hot Jupiters Orbiting nearby Solar-type Stars. ApJ, 753(2):160.

Wright, N. J., Drake, J. J., Mamajek, E. E., and Henry, G. W. (2011). The Stellar-activity-Rotation Relationship and the Evolution of Stellar Dynamos. ApJ, 743(1):48.

Wu, Y. (2019). Mass and Mass Scalings of Super-Earths. ApJ, 874(1):91.

Wu, Y. and Lithwick, Y. (2013). Density and Eccentricity of Kepler Planets. ApJ, 772(1):74.

Yan, F. and Henning, T. (2018). An extended hydrogen envelope of the extremely hot giant exoplanet KELT-9b. *Nature Astronomy*, 2:714–718.

Yelle, R. V. (2004). Aeronomy of extra-solar giant planets at small orbital distances. Icarus, 170(1):167–179.

Zeng, L., Jacobsen, S. B., Sasselov, D. D., Petaev, M. I., Vanderburg, A., Lopez-Morales, M., Perez-Mercader, J., Mattsson, T. R., Li, G., Heising, M. Z., Bonomo, A. S., Damasso, M., Berger, T. A., Cao, H., Levi, A., and Wordsworth, R. D. (2019). Growth model interpretation of planet size distribution. *Proceedings of the National Academy of Science*, 116(20):9723–9728.

Zeng, L. and Seager, S. (2008). A Computational Tool to Interpret the Bulk Composition of Solid Exoplanets based on Mass and Radius Measurements. PASP, 120(871):983.

Zhang, M., Knutson, H. A., Wang, L., Dai, F., dos Santos, L. A., Fossati, L., Henry, G. W., Ehrenreich, D., Alibert, Y., Hoyer, S., Wilson, T. G., and Bonfanti, A. (2022). Detection of Ongoing Mass Loss from HD 63433c, a Young Mini-Neptune. AJ, 163(2):68.

Zhu, W., Petrovich, C., Wu, Y., Dong, S., and Xie, J. (2018). About 30% of Sun-like Stars Have Kepler-like Planetary Systems: A Study of Their Intrinsic Architecture. ApJ, 860(2):101.

Zink, J. K., Christiansen, J. L., and Hansen, B. M. S. (2019). Accounting for incompleteness due to transit multiplicity in Kepler planet occurrence rates. MNRAS, 483(4):4479–4494.

Zink, J. K., Hardegree-Ullman, K. K., Christiansen, J. L., Petigura, E. A., Dressing, C. D., Schlieder, J. E., Ciardi, D. R., and Crossfield, I. J. M. (2020). Scaling K2. III. Comparable Planet Occurrence in the FGK Samples of Campaign 5 and Kepler. AJ, 160(2):94.

Appendices

Permissions

All figures in this thesis taken from journal articles were reproduced with permission, where appropriate. Details for various journals are listed below.

MNRAS

From the MNRAS website `https://academic.oup.com/mnras/pages/rights_and_new_business_development`: 'You are not required to obtain permission to reuse figures (including tables) for academic, educational or non-commercial research for the following material provided full acknowledgment is given to the original source including Figure Number, Title of the Article, Author, Journal Title and Volume or Issue.' This applies to the following figures:

1. Figure 1.9, reproduces Figure 8 from Jankovic et al. (2019).
2. Figure 1.10, reproduces Figure 2 from Gupta and Schlichting (2019).
3. Figure 6.5, reproduces Figure 1 from King and Wheatley (2021).

AAS

From the AAS website `https://journals.aas.org/article-charges-and-copyright/#AAS_material`: 'Note that, effective 1 January 2022, all AAS journals are fully open access.' Under the Creative Commons License (BY): 'This licence lets others distribute, remix, tweak and build upon your work, even commercially, as long as they credit you for the original creation.' This applies to the following figures:

1. Figures 1.7 and 1.1, reproduce Figures 5 and 7 from Fulton et al. (2017).
2. Figure 1.3, reproduces Figure 25 from Jontof-Hutter et al. (2016).
3. Figure 1.5, reproduces Figure 1 from McDonald et al. (2019).
4. Figure 1.6, reproduces Figure 8 from Owen and Wu (2013) and Figure 9 from Lopez and Fortney (2013).
5. Figure 1.8, reproduces Figure 7 from Petigura et al. (2022).
6. Figure 2.3, reproduces Figure 1 from Piso et al. (2015).
7. Figure 5.1, reproduces Figure 18 from Luhman et al. (2010).

Nature

Permission was granted to reproduce Figure 1 from Ehrenreich et al. (2015), shown in Figure 1.4. The license is printed below:

SPRINGER NATURE LICENSE

TERMS AND CONDITIONS: Jul 17, 2022

This Agreement between Mr. James Rogers ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number: 5351270735012

License date: Jul 17, 2022

Licensed Content Publisher: Springer Nature

Licensed Content Publication: Nature

Licensed Content Title: A giant comet-like cloud of hydrogen escaping the warm Neptune-mass exoplanet GJ 436b

Licensed Content Author: David Ehrenreich et al

Licensed Content Date: Jun 24, 2015

Type of Use: Thesis/Dissertation

Requestor type: academic/university or research institute

Format: electronic

Portion: figures/tables/illustrations

Number of figures/tables/illustrations: 1