Ensemble Asteroseismology and Hierarchical Bayesian Models: New Inferences of Astrophysics with Oscillating Stars

by

Oliver James Hall

A thesis submitted to the

University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Sun, Stars and Exoplanets

School of Physics and Astronomy

University of Birmingham

Birmingham, B15 2TT

July 2020

Abstract

Asteroseismology – the study of stellar oscillations – is one of the key tools used to study stars in modern astronomy. This thesis focuses on asteroseismology of solar-like oscillators, which exhibit variability driven by the same mechanism as the Sun. By studying variations in the brightness of stars on their surface, asteroseismology probes the full stellar interior, giving insight into stars' internal processes and fundamental properties.

In this thesis, I provide an introduction to asteroseismology and the state of the field. This is followed by a description of asteroseismic analysis tools I developed for the open source Python package 'Lightkurve', and a presentation of two studies that use asteroseismology to probe different aspects of astrophysics.

The first study uses an asteroseismic ensemble of 5576 evolved Red Clump stars to calibrate data from the *Gaia* mission, and quantify systematic differences in asteroseismic modelling techniques. The second focuses on 91 main sequence stars like the Sun, using asteroseismology to measure their rotation rates. Along with asteroseismic ages, these new rotation rates are used to quantitatively state that stars experience a change in how they lose angular momentum half-way through their main sequence lifetimes. A common thread throughout both these studies is the use of Bayesian statistics, which allows us to leverage large asteroseismic ensembles to make inferences about adjacent fields of astronomy.

To my family

Thanks and Acknowledgements

First and foremost, endless thanks to my supervisor Guy Davies. Not just for excellent supervision and support, but for being a friend, mentor and teacher. If I do half as good a job with my first PhD student, I'll be happy.

Second I would like to thank my secondary supervisor, Bill Chaplin, for expert advice during my PhD, and for getting me interested in asteroseismology back when I was an undergraduate at Birmingham.

The HiROS group at the University of Birmingham – now called Sun, Stars and Exoplanets – has grown tremendously since I started my PhD, and I've seen many colleagues come and go. I'd like to extend a huge thanks to all faculty, staff and visitors to the group during my time here for their mentorship and support. Most of all I would like to thank my fellow PhD students (and some postdocs): Caitlin, James, Thomas, Ben, Mat, Warrick, Alex D, Saniya, Eddie, Vedad, Martin, Walter, Matt, George, Ted, Tanda, Emma, Alex L and Lindsey, for all the games, ciders, cuppas and friendships that will hopefully last much longer than my time at Birmingham. Special thanks also to the SASP group secretary Lou, who's appearance half way through my PhD hugely increased both helpful student support and the number of gluten-free biscuits available.

This thesis would not have been possible without the large number of group, departmental and support staff at the University of Birmingham. I can't possibly name them all, so in particular I would like to thank: the Physics and Astronomy stores (for signing for everything I bought online in the last 3 years), the Teaching Support

Office, the University Library staff and the Physics Astronomy postgraduate reps. Thanks to the faculty and staff in the ASR group and beyond for organising journal clubs, seminars, colloquia and other events I enjoyed and benefited from. Thanks to the IT Service Desk and BlueBEAR team (and in particular to Ed and Luke for helping me with my carbon footprint calculations at the end of this thesis).

During my research I have had the privilege to visit and be supported by scientists and groups beyond Birmingham, for which I am grateful.

I would like to thank the Stellar Astrophysics Center (SAC) at the University of Aarhus and all its staff, who have supported my research both directly and indirectly. In particular I would like to thank Mikkel and Rasmus for their support and guidance, and Brigitte and Louise for helping me get set up on my many visits.

A personal highlight of my PhD was the three weeks I spent with the Kepler Guest Observer office at NASA Ames in California. Thank you so much to Geert, Christina, Jessie, Ann-Marie, Gully and Nick, not just for helping me develop skills that launched most of my success as an early career researcher, but also for the mentorship and friendship you've shown me well beyond my visit.

I'd like to extend a huge thank you all my external collaborators and in particular Derek, Jen and Keith for working with me not as a student, but as a peer as I started forming collaborations outside my group.

Finally, I'd like to thank the anonymous reviewers of my research and my thesis committee, consisting of Ian Stevens, Chris Moore and Jérôme Ballot, for their careful and thoughtful scrutiny of my work and an excellent viva.

Before moving on, I'd like to acknowledge the funding and support that enabled this PhD (and outreach activities) both on personal and institutional grants. Thanks are due to Marc Pinsonneault and Dan Huber for funding great visits to their respective departments. I'd also like acknowledge funding and support from SAC (again),

the Royal Society, Bishop Challoner Catholic College, the Ogden Trust, the Science Technology and Facilities Council (STFC), the Institute of Physics (IoP) and NASA.

Last but not least, I want to give some more personal thank yous.

Thanks to the Astrobites collaboration for its meaningful, impactful and progressive outreach. Being an author has improved me both as a scientist as a member of the astronomy community.

Thanks to the astronomers on Twitter, who made me feel closer to the astronomy community at large, and who continue to improve my life with insightful research and memes in equal measure.

Thank you to all my new and amazing friends in the astronomy community (on- and offline). The work you do is incredible and I am your biggest fan. You are easily the best parts of this job, and I am excited to meet you again at conferences once this pandemic is over.

Endless thanks to the St Martin's Centre for Health and Healing and its incredible therapists. They gave me the tools to understand my anxiety, and I would not be finishing this PhD on as solid ground as I am without them.

Thank you to the late Randy Pausch for showing me a vision of the academic I want to be.

Thank you to my family, and in particular to Mum and Dad for always supporting me in everything I've ever attempted with a clear head and a full heart. This PhD is an equal (if not greater) amount of your effort as it is mine, built on a lifetime of care, love and support.

Thank you to my close friends from my time at the University of Birmingham. My life wouldn't be the same without you. You know who you are!

Finally thank you to my wife Kristina. My biggest supporter, my sternest reviewer, my collaborator for life. I love you so much.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Thesis Plan

Astronomy and astrophysics aim to understand how the Universe works, and to provide context to our place within it. When looking to understand the Universe around us, the first port of call is stars, starting with our own Sun. Due to our proximity to the Sun, we have a deep understanding of how it works. For other stars, which only appear in our sky as point sources (single pinpricks of light), obtaining this understanding is more challenging.

Over the course of history, astronomers have become increasingly adept at maximising the use of information available to us about stars other than the Sun. Stars emit light over a broad range of wavelengths, revealing details of their current elemental compositions, temperatures, and surface gravities. They (appear to) move on the night sky, revealing their distance to us and movement relative to other objects in the Milky Way. By observing a star's brightness over time, astronomers can detect flares of magnetic activity, planets and stellar companions passing between us and their host star, and long- and short- timescale changes to a star's brightness. In many cases, these changes to their brightness repeat, revealing that the star is oscillating. The study of these oscillations in stars, called *asteroseismology*, is the core topic of this thesis.

Asteroseismology is one of the key tools used to study stars in modern astron-

omy. Almost all stars exhibit oscillations of some kind, caused by their internal processes. Because of this, the form, strength, and frequency of stellar brightness variability tells us *directly* about the conditions inside a star. In early astronomy, this revelation was used to construct relationships between the true brightness and period of oscillations of so-called Cepheid-type stars, revealing the distance to these objects throughout the Milky Way as well as other galaxies (Leavitt & Pickering, 1912). Much like musical instruments, different types of star make different 'sounds', as their internal structure and internal processes change due to size, mass or evolutionary state. By observing the 'music' of stars, asteroseismology allows us to access information about the stellar interior just using observations of the stellar surface.

This thesis focuses closely on so-called solar-like oscillators. What makes this class of stars both unique and interesting is that their oscillations are driven by the exact same processes that drive oscillations in our Sun: a convective outer layer. Convective outer layers are found both in cool Sun-like dwarf stars on the main sequence as well as in evolved sub-giant and red giant stars. Due to our proximity to the Sun, we have a strong theoretical framework for how these types of stars oscillate. Solar-like oscillators are both abundant in the universe (main sequence dwarfs) and relatively luminous (red giant stars). Because of this, asteroseismology of solar-like oscillators can be used to study the past, present, and future of both individual stars and the Milky-Way at large.

Astrophysics as a field of study is currently experiencing a transition into the era of 'large data astronomy', with increasingly expansive datasets providing a comprehensive view of astrophysical objects. The past two decades have seen releases of major catalogues of photometry such as 2MASS (Skrutskie et al., 2006), spectroscopy such as SDSS (Eisenstein et al., 2011), APOGEE (Majewski et al., 2017) and LAMOST (Deng et al., 2012; Zhao et al., 2012) among many others. The ongoing *Gaia* mission (Gaia Collaboration et al., 2016, 2018) has provided astrometric

distances and proper motions to millions of stars, with spectroscopic information planned for future data releases. Combining these large scale surveys, most astrophysical research now incorporates information from multiple analysis techniques, in one form or another. For example, a stellar physics study could use *Gaia* astrometry to confirm if stars belong to a cluster, and use APOGEE metallicities and temperatures to model oscillating stars found in the cluster.

Asteroseismology in particular has benefited from large catalogues of data. Recent space-based planet finding missions such as CoRoT (Baglin et al., 2006), *Kepler* (Borucki et al., 2010), *K2* (Howell et al., 2014) and *TESS* (Ricker et al., 2015) have provided long, uninterrupted photometric time series, hugely improving the precision of asteroseismic observations. Combined with spectroscopic and astrometric measurements of temperature, metallicity, surface gravity and luminosity, theoretical models of oscillating stars have allowed for new insights into the physics of individual stars as well as those of larger stellar populations across the Milky Way. As the range of data on a star broadens, so too does our understanding of them.

In order to leverage the large quantities of data available in this era of astronomy, many astrophysicists (including myself) use *Bayesian statistics*. The fundamental of Bayesian statistics is that it doesn't describe an observation as a datum with an uncertainty, but as a probability distribution. These probability distributions can (or even should!) be informed by our *prior knowledge*. For example, imagine fitting a model to observations of a stellar cluster to find the cluster's age. The measured age is expressed as a probability distribution, with a most probable value (commonly the published result), and some spread of probability around that value (commonly the published uncertainty). To incorporate their prior knowledge, the astronomer could say that for ages older than the Universe (13.8 Gyr, Planck Collaboration et al., 2018) – the probability is zero.

Most of the research presented in this thesis is done using Bayesian statistics,

and so some common terminology will surface repeatedly. Given a model M and some data D, the fundamentals of Bayesian statistics are (Bayes & Price, 1763)

$$p(M|D) \propto p(D|M) \times p(M)\,, \tag{1.1}$$

called Bayes' Theorem. Here, the lower-case p indicates a probability distribution function, i.e.: $p(M|D)$ is the probability of our model (M) being true given our data (D), and is called the *posterior*. Similarly, $p(D|M)$, is the probability of obtaining our data given our model. This is commonly referred to as the *likelihood function*. Next, $p(M)$ is the probability of obtaining our model in the first place, which encodes our prior information (such as the age of the Universe). If we understand how our measurements were made, and have a prior knowledge of our situation, we can calculate how likely our model is to be true.

Integrating over all values a model parameters can take, a process called *marginalization*, is computationally expensive. Instead, advances in so-called sampling techniques allow us to estimate the marginalised posterior probability by making small changes in model parameters, and seeing how that affects the posterior probability. For small numbers ($\lesssim 100$) of model parameters, Markov Chain Monte Carlo (MCMC, see e.g. Goodman & Weare, 2010; Foreman-Mackey et al., 2017) does so by essentially making parameters go on an random-walk. For models with more parameters, Hamiltonian Monte Carlo (HMC, see e.g. Betancourt & Girolami, 2013) takes only informed steps, but requires fine-tuning by a user. Many modern implementations of HMC (Salvatier et al., 2016; Carpenter et al., 2017) apply a No-U-Turns Sampler (NUTS, Hoffman & Gelman, 2014), an extension to HMC that removes the need for manual fine-tuning. Different techniques are applicable to different data analysis problems (see the two models used in Chapter 5), and being able to know what to apply and when is a big part of modern astronomy.

In this thesis, I present two separate studies of asteroseismic ensembles of red giant and main sequence stars respectively. Both studies apply modern Bayesian statistical techniques in order to make new inferences of astrophysics. With these studies, this thesis aims to show how combining multiple fields of research using a robust Bayesian statistical framework allows for inferences that would otherwise be inaccessible.

This thesis is structured as follows: Chapter 2 provides an introduction to asteroseismology, giving historical context and the theoretical foundation required to follow the rest of the thesis. Chapter 3 provides a closer look into the signal-processing analysis of asteroseismic data with the space-based telescopes, and presents asteroseismic analysis software I developed for the Lightkurve project (Lightkurve Collaboration et al., 2018). Chapter 4 gives a walk-through example of building a hierarchical Bayesian statistical model. This is followed by the first study: using an ensemble of asteroseismic observations of core-helium burning giant stars to test asteroseismic theory and systematics of the *Gaia* mission (Hall et al., 2019). Chapter 5 presents the second study, which has two key components: the use of modern statistical techniques to make new asteroseismic measurements in *Kepler* data of main sequence stars, and the use of this new asteroseismic ensemble to draw conclusions about stellar rotational evolution. Finally, in Chapter 6 I will draw conclusions and provide a future outlook for these and similar studies. My individual contributions are detailed at the heads of Chapters 3, 4 and 5.

Chapter 2

Asteroseismology of Solar-Like Oscillators

2.1 Introduction

The study of stellar oscillations, known as *asteroseismology*, has grown over the course of the last century into one of the most important tools used to study stellar physics. Stars' internal processes create standing waves inside them, which create global fluctuations on their surface that we can observe. By studying this oscillating surface variability, asteroseismology provides a window into stars' internal structure and rotation.

While the term 'asteroseismology' refers to the study of all oscillating stars, the primary focus of this thesis is stars that oscillate in the same manner as the Sun: so-called solar-like oscillators (see Figure 2.1)[1]. The variability of solar-like oscillators is driven by their common property of a convective outer envelope. This property is shared by cool dwarf stars like the Sun, sub-giant stars, and more evolved red giants. Despite the different evolutionary stages, the process by which these stars oscillate is the same. Herein lies the unique power of this type of variable star. Despite not being exactly like the Sun, they exhibit behaviour analogous to the Sun. As we can observe, and therefore model, the Sun in much greater detail than any other star, we have the possibility to extend this analysis to stars throughout the Milky Way.

[1]Throughout this thesis, the word 'asteroseismology' can be taken colloquially to refer exclusively to the study of solar-like oscillators.

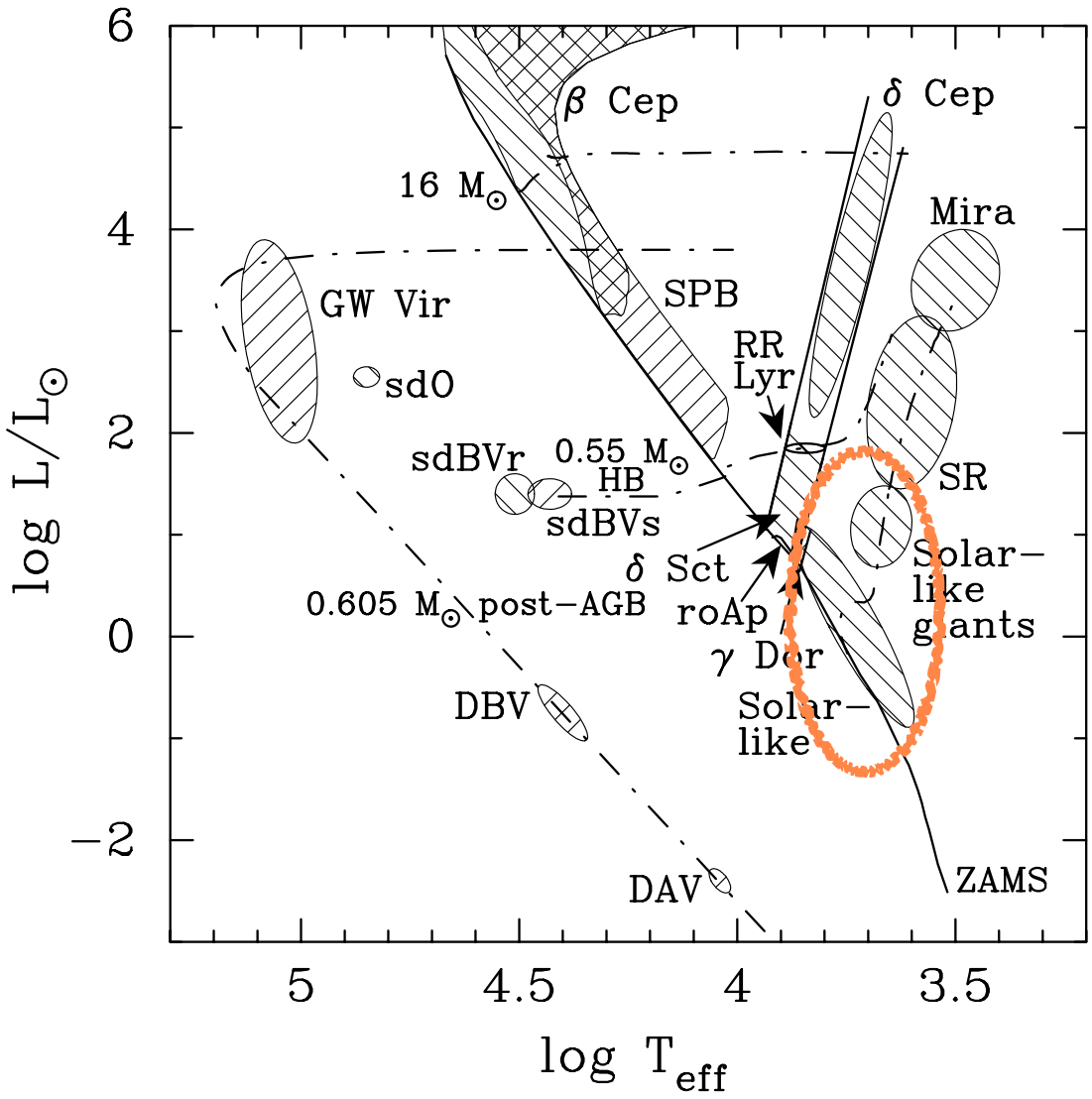

Figure 2.1: A Hertzsprung-Russell diagram (HRD), adapted from Figure 1 of Handler (2009), showing regions of the HRD occupied by different variable types. The orange ellipse indicates the approximate region occupied by solar-like oscillators, the main topic of this thesis. Solar-like giant oscillators may be included under this label. Areas hatched from lower-left to upper-right indicate the domain of g-mode pulsators, and vice-versa for p-mode pulsators (see text). Overlapping areas contain 'hybrid' pulsators. The dash-dotted lines indicate sections of evolutionary tracks for stars at different masses (given in units of solar mass, $M_{\odot}$). The solid line indicates the Zero-Age Main Sequence (ZAMS). For other acronyms, see Handler (2013). It should be noted that recent work has found Semiregular Variable (SR) stars at luminosities of $\log(L) \lesssim 2.6\,L_{\odot}$ to exhibit solar-like oscillation patterns (Yu et al., 2020).

This Chapter is structured as follows: Section 2.2 will provide a brief history of asteroseismology and the state of the field today. Section 2.3 will introduce the key principles of asteroseismic theory required to follow the research presented in the

remaining Chapters. Finally, Section 2.4 will summarise and provide an outlook for the field.

2.2 A brief history of Helio- and Asteroseismology

The solar-stellar connection has been the key to the success of modern asteroseismology. Unlike for distant stars, we can directly resolve small-scale changes on the solar surface. The asteroseismic study of the Sun, called *helioseismology*, provides the foundation upon which all asteroseismic theories are built, tested, and calibrated.

2.2.1 Early Helioseismology

The field of helioseismology has a rich, observation-driven history (see Christensen-Dalsgaard, 2002, for a more comprehensive review). While the first detection of solar variability may be attributed to Plaskett (1916) (with further confirmation in Hart, 1954, 1956), the first definitive measurements of solar oscillation frequencies were made by Leighton et al. (1962), by studying Doppler velocities of the solar surface. The observed 5 minute period, which we now know to be the frequency of maximum oscillation amplitude in the Sun, was first thought to be a local phenomenon (Bahng & Schwarzschild, 1963). Not until the 1970s did theoretical breakthroughs by Ulrich (1970) and Leibacher & Stein (1971) describe the observed 5 minute oscillations as being attributed to acoustic standing waves inside the Sun. These theories were further confirmed by both Deubner (1975), who observed individual oscillation frequencies around the 5 minute period, and found them to be consistent with what we would expect of acoustic waves in the solar cavity. Shortly after, high resolution spectroscopic observations by Claverie et al. (1979) and Grec et al. (1980) confirmed that these modes of oscillation were not just a surface phenomenon, but probed the full solar interior[2].

[2]The Grec et al. (1980) observations were made on the South Pole, and the Claverie et al. (1979) observations done (in part) in Birmingham!

Helioseismology (as well as asteroseismology) works best with long, uninterrupted observations. This has been made possible since 1976 by ground based continuous surveys such as GONG (Global Oscillations Network Group, Kennedy & GONG Team, 1994) and BiSON (Birmingham solar Oscillations Network, see e.g. Hale et al., 2016). The latter provides the longest time-baseline observations of the Sun as a point source, analogous to observations of distant stars. While so-called 'Sun-as-a-star' observations on the ground can only measure oscillations that create Doppler shifts on large scales, space based observatories such as ESA/NASA's SOHO (solar Heliospheric Observatory, Domingo et al., 1995) can resolve individual sections of the solar surface. These measurements, while on a shorter baseline, allow for a much richer array of mode frequencies to be observed.

The combination of ground- and space-based observations has allowed helioseismologists to paint a detailed theoretical picture of how the Sun operates, including but not limited to mapping the Sun's internal sound-speed (Basu et al., 1997; Turck-Chièze et al., 1997) and density profiles (Basu et al., 2009), measuring differential rotation rates throughout the Sun (e.g. Thompson et al., 1996; Elsworth et al., 1995; Chaplin et al., 1999; Couvidat et al., 2003a; García et al., 2008), measuring the base of the convection zone (Christensen-Dalsgaard et al., 1985; Ballot et al., 2004), understanding the Sun's core processes (e.g. Turck-Chièze et al., 2001; Couvidat et al., 2003b; Basu et al., 2009) and the effect that the Helium abundance in the core has on modes of oscillation (Gough, 1983; Vorontsov et al., 1991). These studies built the theoretical and observational framework required to expand asteroseismology to stars outside our Solar System.

2.2.2 Early Asteroseismology

Helioseismology was first expanded to other stars using ground based surveys. These detections were limited to bright stars visible with the naked eye, which exhibited variability at amplitudes large enough to be visible through the Earth's atmospheric

interference[3]. The first asteroseismic detections were of the star Procyon, the eighth brightest star in the sky, including a tentative measurement of the spacing between individual oscillation frequencies (Brown et al., 1991), although this was controversial until later independent confirmations (Martić et al., 1999). The first clear detection and identification of modes of oscillation in a solar-like star was not until almost a decade later, for the star α Cen A (Bouchy & Carrier, 2001).

During the efforts to characterise asteroseismic signals using ground-based data in the '90s (e.g. Kjeldsen et al., 1995), the field had to grapple with both low resolution observations and aliased signals, due to the short baselines and lack of continuous observations. The jump to space, the next step in asteroseismology, was not planned, but instead happened serendipitously. At the turn of the millennium Buzasi (2000) presented a case to repurpose the NASA Wide field Infra Red Explorer (WIRE, Hacking et al., 1999) satellite, which failed its intended purpose, for asteroseismology. Using its onboard tracker telescope, WIRE was able to monitor variability of various bright stars uninterrupted. These data showed small changes in variation (at a 'parts-per-thousand' level), allowing for characterisation of both solar-like oscillators and other variable stars (Buzasi et al., 2005; Stello et al., 2008). WIRE highlighted the wealth of asteroseismic data that could be obtained with a dedicated space telescope.

2.2.3 Modern Asteroseismology

The first major breakthrough in space-based asteroseismology came with the French-led CoRoT mission (Convection, Rotation and planetary Transits, Baglin et al., 2006), launched in 2006. The first space-mission dedicated to the search for exoplanets, it also provided the uninterrupted long time-series photometry required to perform asteroseismic analyses.

CoRoT observed in two cadences: a 32 second short cadence for 10 targets per

[3]For example, the star η Boo with a magnitude of $V = 2.68$ exhibits a maximum oscillation amplitude of roughly 22 parts-per-million (Kjeldsen et al., 1995)

field, and a long cadence of 8.5 minutes for 12,000 targets per field, switching field every 150 days. This observing strategy allowed asteroseismologists to go beyond global measures of stellar oscillations, and measure individual oscillation frequencies. The ability to measure multiple stars at once without a dedicated ground-based campaign meant that, for the first time, *ensemble asteroseismology* was possible. In many ways, CoRoT set the stage for asteroseismology of the following decade.

The advance of modern asteroseismology was continued by the *Kepler* mission, launched in 2009 (Borucki et al., 2010). Similarly to CoRoT, *Kepler* monitored 170,000 stars in Long Cadence (LC, 29.4 min) and 512 stars in Short Cadence (SC, 58.85 s) at a given time. Unlike CoRoT, *Kepler* observed a single field for four continuous years. Years of uninterrupted photometry allowed not only for more long period exoplanets to be found through the transit method, but also provided asteroseismic spectra of never-before-seen quality and resolution[4].

Besides its impact on the field of exoplanetology (see e.g. Batalha, 2014), *Kepler* carried on the momentum from CoRoT to make asteroseismology the major field it is today (see e.g. Chaplin & Miglio, 2013, for a comprehensive review). The LC observations probed oscillation frequencies below 284 μHz, for a large number of stars. As red giant stars oscillate at frequencies lower than this, Kepler's LC observations provided masses and radii of these evolved stars (e.g. Hekker et al., 2011; Pinsonneault et al., 2014; Pinsonneault et al., 2018; Yu et al., 2018). Asteroseismic ages of these stars, obtained through comparisons to stellar models, ushered in a new era of *galactic archaeology*, the study of the Milky Way's history (e.g. Mathur et al., 2016; Miglio et al., 2009, 2013; Stello et al., 2015; Davies & Miglio, 2016). The SC observations, which probed up to 8496 μHz, provided the same for main sequence stars (Chaplin et al., 2010, 2011, 2014) oscillating at higher frequencies, and allowed for synergy with exoplanetology (e.g. Christensen-Dalsgaard et al., 2010; Huber et al., 2013a; Silva Aguirre et al., 2015; Van Eylen et al., 2014, 2018). As

[4]In fact, it will likely be many years until a mission again provides this kind of data.

asteroseismic oscillations probe the stellar interior, *Kepler* and CoRoT data allowed for inferences of rotation of the near surface layers of main sequence stars (e.g. Davies et al., 2015; Nielsen et al., 2015; Benomar et al., 2018) and rotation of the *cores* of red giant stars (e.g. Beck et al., 2012; Deheuvels et al., 2012, 2014; Mosser et al., 2012c; Gehan et al., 2017). With these insights into the cores of evolved stars, asteroseismic modes of oscillation could be used to distinguish clearly between whether a star was burning hydrogen in a shell around its core, or whether it had started burning helium inside its core (Bedding et al., 2011; Mosser et al., 2012b, 2015; Stello et al., 2013; Vrard et al., 2016; Elsworth et al., 2017). Finally, these large quantities of data allowed for ensemble asteroseismology, providing new inference of stellar physics based on stellar populations (e.g. Khan et al., 2018; Yu et al., 2018; Hall et al., 2019, among many others) and careful checks of theoretical asteroseismic relations (e.g. Huber et al., 2011b, 2017; Miglio et al., 2012, see section 2.3).

The *Kepler* mission ended in 2013 due to the loss of two reaction wheels, leaving the spacecraft unable to maintain its fixed field of view. To salvage the telescope, the *Kepler* team were able to repurpose it as the ecliptic-plane mission, *K2* (Howell et al., 2014).

Over the course of the *K2* mission's lifespan, it performed 19 observing campaigns of approximately 80 days each, in different fields around the ecliptic. While *K2* campaigns were shorter than *Kepler*'s, the new focus on the ecliptic enabled asteroseismic observations of new stellar clusters and star-planet systems (Chaplin et al., 2015; Lund et al., 2016b,a), and new regions of the Milky Way for galactic archaeology (Stello et al., 2015; Miglio et al., 2016; Rendle et al., 2019b). Asteroseismic analysis of *K2* data (as well as that of *Kepler* data) is still ongoing.

The *Kepler* mission has since been succeeded by *TESS* (Transiting Exoplanet Survey Satellite, Ricker et al., 2015), launched in 2018. Similar in design (30 minute

long cadence, 20 and 120 second short cadences), this exoplanet and asteroseismology mission, nearing the end of its 2 year nominal run, observes $> 90\%$ of the night sky, from near the ecliptic to the ecliptic pole. It does so by observing sectors of the sky for 27 days. The overlap between these sectors means that near the ecliptic stars will have 27 days of data, while near the ecliptic poles they will have 351 days of data. While the shorter baselines and brighter focus of *TESS* will provide fewer asteroseismic observations of main sequence stars than *Kepler*, the long cadence observations of every star in its field of view will provide an unprecedented number of asteroseismic detections in red giant stars. Asteroseismology with *TESS* has already been used to study exoplanet hosts (Huber et al., 2019), red giant stars (Silva Aguirre et al., 2020), and the history of the Milky Way (Chaplin et al., 2020), among others. Soon to enter into its confirmed extended mission, the asteroseismic success of *TESS* will be ongoing for the foreseeable future.

2.3 Theory of solar-Like Oscillations

2.3.1 Introduction

Stars cool enough to have a convective outer layer will exhibit solar-like oscillations. In the outer layers of a star, material will rise and fall as heat is transported to the surface. This convective process is turbulent and stochastic in nature, and therefore convection both *drives* and *damps* waves that propagate throughout a star's interior (Goldreich & Keeley, 1977b; Goldreich & Kumar, 1988; Gough et al., 1996; Houdek et al., 1999). Where these waves interact they form standing waves which probe the full stellar cavity.

As in a musical instrument, each star will have certain unique resonant frequencies at which standing waves are formed. Due to the stochastic nature of the excitation mechanism, these frequencies will be relatively low amplitude and have short lifetimes (Goldreich & Keeley, 1977a). This is stark in comparison to other oscillators (shown in Figure 2.1), which are typically coherent (long lifetimes).

A commonly used analogy for solar-like oscillators is a bell in a sandstorm. In this scenario, the bell buffeted by sand will ring out at resonant frequencies, but be dampened by the stochastic nature of the sandstorm. For both the bell and the star, the resonant oscillation frequencies of the ringing cavity carry information about the physical properties of the cavity. In the same way you can tell musical instruments apart by their sound, an analysis of a star's modes of oscillation tells us about the properties of the star, providing a window into its internal processes.

Unlike most instruments, stars have a complex interior structure, giving rise to different types of oscillation modes (Cox, 1980). There are two distinct types of modes in solar-like oscillators, which are differentiated by their restoring force. Acoustic – or pressure – modes (p-modes), are restored by pressure gradients, and propagate in the convective outer layers. Gravity modes (g-modes) are instead restored by buoyancy forces, and typically propagate in the radiative interior and stellar core (see Figure 2.2).

In main sequence solar-like oscillators, p-modes are the main topic of interest, as they cause observable displacement on the surface of the star. G-modes on the other hand, have not yet been observed directly in these stars due to their extremely low amplitudes near the surface (see e.g. Turck-Chièze et al., 2004; Appourchaux et al., 2010), although potential signatures have been reported for the Sun (García et al., 2007; Fossat et al., 2017)[5]. In evolved giant stars, which oscillate at low frequencies, g-modes interfere with p-modes at the boundary between the convective and radiative regions of the star, affecting how p-modes appear at the surface (Scuflaire, 1974; Kjeldsen et al., 2003; Bedding et al., 2010). Through these so called 'mixed modes', it becomes possible to probe the rotation and conditions of the inner regions of the star, including the core (see e.g. Bedding et al., 2011; Deheuvels et al., 2012; Gehan et al., 2017).

[5]These results are still controversial, and refuted in the literature (Appourchaux et al., 2018; Schunker et al., 2018; Appourchaux & Corbard, 2019).

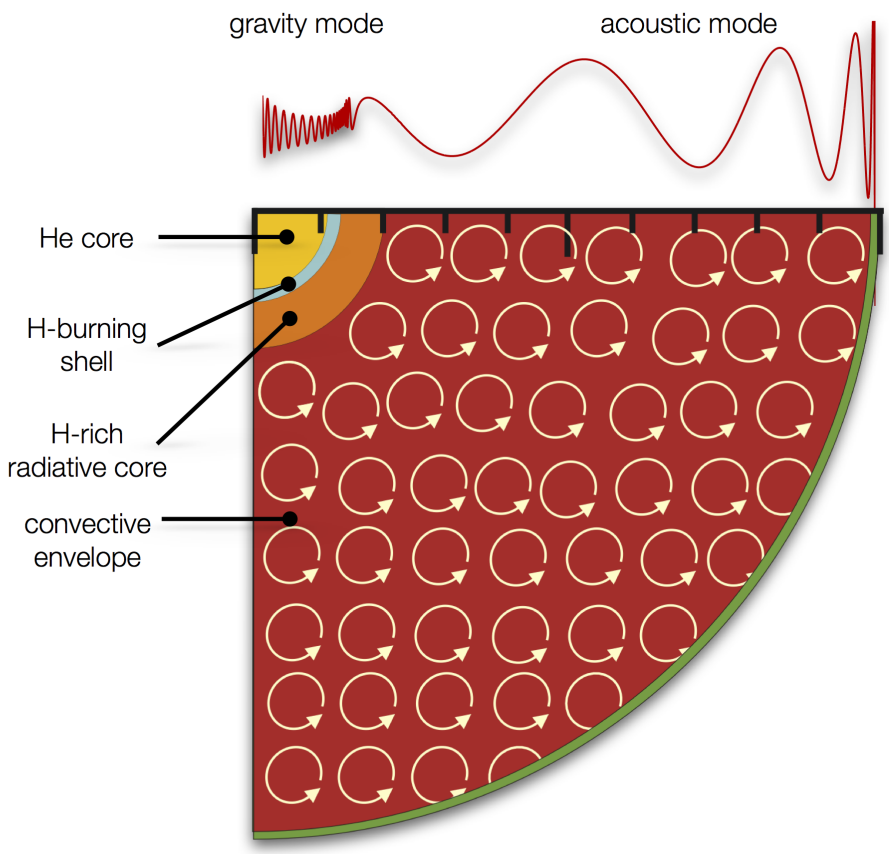

Figure 2.2: An infographic showing the regimes of different mode types in a Red Giant Branch star exhibiting solar-like oscillations. The acoustic modes (p-modes) propagate in the outer convective region, where the gravity modes (g-modes) propagate in the inner radiative region and the core. Image created by Andrea Miglio, available at asterostep.eu.

2.3.2 Modes of Oscillation

Solar-like oscillations are commonly described by spherical harmonic functions. While an in-depth introduction to spherical harmonics is beyond the scope of this thesis (see e.g. Christensen-Dalsgaard, 2002; Aerts et al., 2010), it is important to have an intuition for the parameters that describe a mode of oscillation, namely: n, the radial order, ℓ, the angular degree and m, the azimuthal order.

To gain a physical intuition for what the parameters n, ℓ and m represent, let us

start by considering a standing wave on a string. The fundamental mode, where all particles on the string are perturbed from their equilibrium position, is described by $n = 0$. The first harmonic, $n = 1$, has a node in the centre of the string where no particles are perturbed. The second harmonic, $n = 2$ has two nodes, and so forth.

In the example of a string, n describes the number of nodes in the string's 1D line. To describe standing waves in stars, we need to generalise this to three dimensions, which requires two more parameters, one for each additional dimension.

In a star, the radial order n describes the number of nodes in the radial direction, outward from the centre of the star. You can imagine the star as being composed of 'shells' of oscillating material, separated by nodes where no movement occurs.

The angular degree ℓ, describes the number of nodes present on the surface of the star. For example, a radial oscillation, where a star increases and decreases in radius, has no nodes on the surface, and is described by $\ell = 0$. A dipole oscillation will see the northern hemisphere shrink while the southern hemisphere expands (and vice versa), with a node along the equator, and is described by $\ell = 1$.

Finally the azimuthal order m describes the orientation of the oscillations, and can be seen as the number of nodes crossing the equator. In spherical harmonics, the number of available azimuthal orders present per angular degree is equal to $2\ell + 1$. For example, for a radial oscillation, there are no nodes, and so $\ell = 0$ and $m = 0$. For a dipole oscillation ($\ell = 1$) there are three available azimuthal orders: $m = 0$, which has a node along the equator, and $|m| = 1$, which is the same oscillation but rotated by 90°, placing the node along the poles. These oscillations represented by $m = -1$ and $m = 1$ are the same but opposite on the star, and therefore degenerate with one another. However if a star is rotating, their observed oscillation frequencies change, breaking the degeneracy and making all three oscillations visible to an observer (Ledoux, 1951; Hansen et al., 1977). A representation of how modes of different ℓ and m appear on a sphere can bee seen in Figure 2.3.

All the modes of oscillation occur at different frequencies and are present simul-

taneously, resulting in a rich spectrum of visible modes. Typically, main sequence and sub-giant stars oscillate at high frequencies and high radial orders ($n \gtrsim 20$), whereas evolved Red Giant stars oscillate at low frequencies and low overtone numbers ($n \lesssim 20$). When observing a solar-like oscillator with a telescope such as *Kepler*, we only view a point source, and are therefore only capable of detecting those modes that cause large-scale changes to the stellar surface intensity (Dziembowski, 1977). For $\ell = (0, 1, 2)$, and occasionally $\ell = 3$, we can resolve modes of oscillation in distant stars. However for higher values of ℓ, the number of nodes on the stellar surface cancel out when integrated across the stellar surface, making the modes invisible to us.

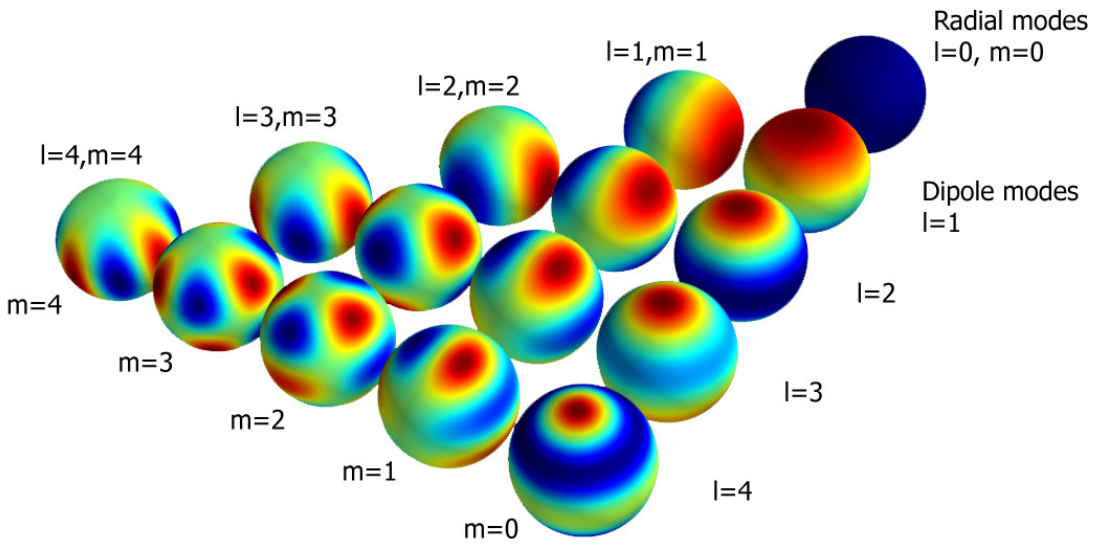

Figure 2.3: A model representation of spherical harmonic modes of oscillation at different combinations of angular degree (ℓ) and azimuthal order (m). The three examples in the top right corner are described in the text. The opposing red and blue colours indicate the oscillating regions at a minima or maxima, and the green regions between represent the stationary nodes. Note how at larger ℓ, the perturbed regions increasingly cancel out across the stellar disk. Figure 1.2 in Beck (2013).

2.3.3 The Asymptotic Expression

As described above, the common formalism of solar-like oscillations is to use spherical harmonics, which are analogous to a 1D example of a standing wave on a string. By describing how stellar oscillations arise from disturbances inside the star, we can

relate observed properties of solar-like oscillations to the physical structure of stars. While a complete derivation of the equations that describe solar-like oscillations is outside the scope of this thesis (see e.g. Chaplin & Basu, 2014, Chapter 3 instead), we can briefly discuss the relevant approximations and physical processes that give rise to modes of oscillation.

To good approximation, modes of oscillation are first and foremost considered linear (modes have velocities much slower than the sound speed) and adiabatic (the oscillation timescale is much smaller than the heat transport timescale)[6]. Under this approximation, we can consider oscillating perturbations occurring inside a star that impact four properties: the displacement of material from the equilibrium position, the pressure, the density and the gravitational potential. Next, we consider the boundary conditions of the star, which assume that at the surface pressure is zero and the gravitational potential is continuous. Finally we can apply the Cowling approximation (Cowling, 1941), which states that it is appropriate to ignore perturbations to the gravitational potential in the limit that the radial order (n) is large. By considering the remaining perturbations, we can arrive at a second-order differential equation that describes an oscillatory displacement of material inside a star as a function of pressure, density, radial position and the frequency of the oscillation. The harmonic frequencies of the star form the set of solutions to this differential equation, and are referred to as its *eigenfrequencies* (Christensen-Dalsgaard, 2002; Chaplin & Basu, 2014).

Extracting the values of the eigenfrequencies at which modes of of oscillation appear is a mathematically rigorous process. This may be simplified through the approximation that the radial order is larger than the angular degree, so that $n \gg \ell$[7]. Main sequence stars typically oscillate strongest for $20 < n < 30$, and for red giants this is typically $n < 15$. Therefore the frequencies of the modes of oscillation

[6]This approximation does not fully hold true near the surface, which must be accounted for in stellar models (Ball & Gizon, 2017).

[7]This is referred to as the JWKB approximation, which is described for the purposes of helio- and asteroseismology in Unno et al. (1989).

visible in most solar-like oscillators will follow an asymptotic expression, that goes as (Vandakurov, 1967; Tassoul, 1980; Deubner & Gough, 1984; Lopes & Turck-Chieze, 1994)

$$\nu_{n,\ell} \approx \left(n + \frac{\ell}{2} + \epsilon\right) \Delta\nu - \delta\nu_{0,\ell}\,, \tag{2.1}$$

where $\nu_{n,\ell}$ is a frequency at a given radial order n and angular degree ℓ and ϵ is a phase offset. The variable $\Delta\nu$ is the spacing between two consecutive radial orders of equal angular degree (also called the *large frequency spacing*). On the other hand $\delta\nu_{0,\ell}$ is the so-called *small separation*, describing the spacing between two modes of different angular degree at the same radial order. In main sequence stars, $\Delta\nu$ is typically of the order of tens to hundreds of μHz, and is typically below 20 μHz for red giants. The small separation is usually less than 10% of $\Delta\nu$, and ϵ tends to lie around 1 (Lund et al., 2017). It is important to note that the asymptotic expression is in reasonable agreement with observations of high-n modes of main sequence stars, but breaks down for more evolved stars, due to the presence of mixed modes.

From Equation 2.1 we can see that modes of oscillation will appear regularly spaced. Radial modes ($\ell = 0$) will appear evenly spaced by $\Delta\nu$, and dipole modes ($\ell = 1$) will appear half-way between the two consecutive radial modes, shifted by an offset $\delta\nu_{0,1}$. Finally, quadrupole and octopole modes ($\ell = (2, 3)$) will appear near the $\ell = (0, 1)$ modes (respectively) of the *following* radial order, shifted by their respective small separations.

Figure 2.4 shows a frequency-power spectrum of the Sun-like star 16 Cyg B, where the repeating pattern described by Equation 2.1 can be seen. What is not described by Equation 2.1 is the heights of the modes of oscillation, which rise to a peak of maximum oscillation after which it falls off again. The frequency of this peak, called $\nu_{\rm max}$ or the *frequency of maximum oscillation*, lies at the centre of what is colloquially referred to as the 'mode envelope' or 'mode hump'. In asteroseismic analyses, the mode hump is usually described as Gaussian in shape, but this does

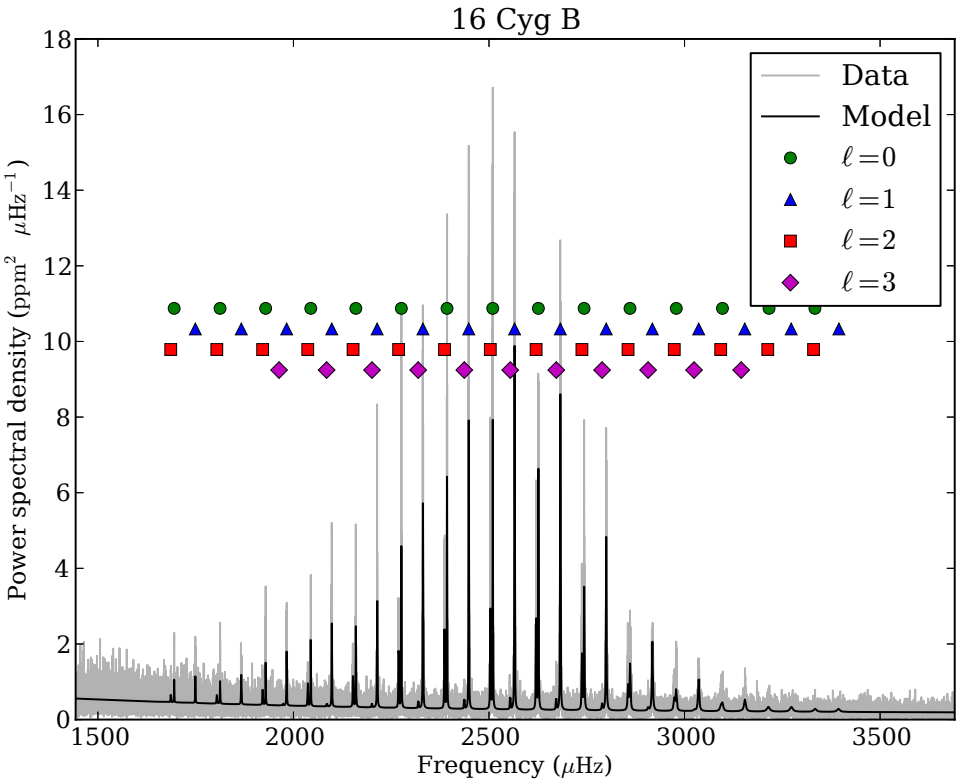

Figure 2.4: An example of a frequency model (black) fit to asteroseismic data (grey) of the solar-like star 16 Cyg B. The 'envelope' of frequencies rises to a maximum near the $\nu_{\rm max}$ of this star, and then reduces in power. Circles represent radial modes ($\ell = 0$), triangles dipole modes ($\ell = 1$), squares quadrupole ($\ell = 2$) and diamonds octopole ($\ell = 3$). The regular pattern closely follows that described by the asymptotic expression in equations 2.1 and 2.2. Adapted from Figure 1 of Davies et al. (2015).

not necessarily have to be the case. A fit of two Lorentzian profiles has been shown to fit just as well for solar data (Lefebvre et al., 2008), and the height and width of the envelope changes significantly with temperature (Lund et al., 2017). While a full theoretical explanation of $\nu_{\rm max}$ is an ongoing field of study, it is known that the location of the power excess is expected to be tied to the balance between the damping and excitation rates (Houdek et al., 1999; Chaplin et al., 2008), as well as to the acoustic cut-off frequency, which describes the maximum frequency for which modes of oscillation are visible on the stellar surface (Belkacem et al., 2011, see Section 2.3.4 below).

Many observations of solar-like oscillators report a departure from the asymptotic expression of Equation 2.1 around $\nu_{\rm max}$ (Mosser et al., 2011). When describing modes of oscillation (e.g. to fit a model to data), it is useful to include this 'curvature' of the large frequency separation in the asymptotic expression. Accounting for this, the expanded expression has the form (Vrard et al., 2016)

$$\nu_{n,\ell} \approx \left[n + \frac{\ell}{2} + \epsilon + \frac{\alpha}{2}\left(n - \frac{\nu_{\rm max}}{\Delta\nu}\right)^2\right]\Delta\nu - \delta\nu_{0,\ell}\,, \tag{2.2}$$

where we have added the frequency of maximum oscillation, $\nu_{\rm max}$, and α, a curvature term. It is important to note this is no longer an asymptotic expression, and that the presence of a curvature term is an *observational* addition. What this tells us in practice is that $\Delta\nu$ is *not constant* in frequency. By adding a curvature term, we can treat $\Delta\nu$ as if it is constant. For this reason, it is sometime referred to (more properly) as the average frequency spacing, and denoted as $\langle\Delta\nu\rangle$.

The cases described above are for a non-rotating star only, which is why azimuthal orders are not considered, as they are degenerate in that scenario. The inclusion of magnetic fields alongside stellar rotation can also introduce further perturbations to mode frequencies (see e.g. Gough & Thompson, 1990; Goode & Thompson, 1992; Kiefer & Roth, 2018; Thomas et al., 2019). The modelling of asteroseismic oscillation frequencies of rotating stars is described in more detail in

Chapter 5.

Finally, not included here are the effects of perturbations due to rapid changes in the density of the stellar interior, called 'glitches' (see e.g. Miglio et al., 2010; Mazumdar et al., 2014). While I account for their presence as slight perturbations to the asymptotic expression in later Chapters, I will not explicitly discuss them in this work.

2.3.4 Global Seismic Observables

In asteroseismic analyses, two key parameters describing the frequencies of modes of oscillation can be observed even for low time-resolution data. These are $\nu_{\rm max}$, the frequency at which the power of the oscillations is highest, and $\Delta\nu$, the large frequency spacing.

The large frequency spacing, as shown in Equation 2.1, describes the even distribution of overtones of a given angular degree in frequency space. Oscillations of different radial orders probe different regions of a star, and so the spacing between them is related to the sound travel time within the star. Formally, $\Delta\nu$ is defined as (Christensen-Dalsgaard, 2002)

$$\Delta\nu = \left[2\int_0^R \frac{\mathrm{d}\,r}{c_{\rm s}}\right]^{-1}, \tag{2.3}$$

where R is the stellar radius, and $c_{\rm s}$ is the adiabatic sound speed, which is defined as $c_{\rm s}^2 = \sqrt{\Gamma_1 P/\rho}$, where Γ_1 is an adiabatic index, and P and ρ are the pressure and stellar density respectively.

Without having to solve the integral in Equation 2.3, a dimensional analysis can show that a measurement of $\Delta\nu$ yields a useful scaling relation. In the limit that $r = R$, the sound speed scales as $c_{\rm s}^2 \propto P/\rho$. If we assume that the density of a star is uniform throughout and that it is in hydrostatic equilibrium, P in turn scales as $P \propto M^2/R^4$ where M and R are stellar mass and radius (Chaplin & Basu, 2008). Substituting this into a dimensional analysis of Equation 2.3 yields

$$\Delta\nu^2 \propto \frac{c_s^2}{R^2} \propto \frac{M^2}{\bar{\rho} \times R^6} \propto \bar{\rho}\,, \tag{2.4}$$

where $\bar{\rho}$ is the average density of the star. By dividing through the solar values of $\Delta\nu$ and density, we find a scaling relation:

$$\Delta\nu \simeq \Delta\nu_\odot \sqrt{\frac{\rho}{\rho_\odot}}\,, \tag{2.5}$$

where the '$\odot$' symbol denotes a solar value.

The second global asteroseismic observable is the frequency of maximum oscillation, $\nu_{\rm max}$. Unlike $\Delta\nu$, which has a clear theoretical basis, $\nu_{\rm max}$ is only an observed property of solar-like oscillations. However oscillations at and around $\nu_{\rm max}$ are most clearly seen on the surface, and it is therefore expected to be affected by the conditions in the near-surface layers of a star. As a result, $\nu_{\rm max}$ is argued to be closely related to the acoustic cut-off frequency ($\nu_{\rm ac}$, Deubner & Gough, 1984), the frequency above which waves no longer fully reflect at the stellar surface, and cease to form standing waves (Belkacem et al., 2011).

The acoustic cut off frequency is known to scale with the surface gravity (g) and effective temperature ($T_{\rm eff}$) of a star as

$$\nu_{\rm ac} \propto \frac{g}{\sqrt{T_{\rm eff}}}\,, \tag{2.6}$$

where surface gravity scales as $g \propto M/R^2$. If we take $\nu_{\rm max}$ to be directly proportional to the acoustic cut-off frequency, we recover the scaling relation

$$\nu_{\rm max} \simeq \nu_{\rm max,\odot} \frac{M/M_\odot}{(R/R_\odot)^2} \sqrt{\frac{T_{\rm eff,\odot}}{T_{\rm eff}}}\,, \tag{2.7}$$

where we have substituted in mass and radius for surface gravity, and divided through by solar properties. Both the scaling relations in Equations 2.5 and 2.7 have been repeatedly tested through comparison to observations, and found good

agreement (see e.g. Bedding, 2014).

Using dimensional analysis and theoretical stellar structure, we have constructed two scaling relations between the most easily observed asteroseismic properties and a star's mass, radius and effective temperature. Substituting Equations 2.5 and 2.7 together and re-arranging for mass and radius we arrive at the two so-called 'seismic scaling relations' (Brown et al., 1991; Kjeldsen & Bedding, 1995):

$$\begin{aligned} \frac{M}{M_\odot} &\simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^3 \left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-4} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{3/2} \quad \text{and} \\ \frac{R}{R_\odot} &\simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right) \left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-2} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{1/2}, \end{aligned} \tag{2.8}$$

where all symbols are as defined above. Throughout the work presented in this thesis, I have used values of $\nu_{\rm max,\odot} = 3090 \pm 30\,\mu$Hz, $\Delta\nu_\odot = 135.1 \pm 0.1\,\mu$Hz (Huber et al., 2011b) and $T_{\rm eff,\odot} = 5777.2 \pm 0.8\,$K (Prša et al., 2016).

The use of the seismic scaling relations is commonly referred to as the 'direct method'. If we have an external measure of $T_{\rm eff}$, typically through photometry or spectroscopy, this top-level asteroseismic analysis can reveal a star's mass and radius in a distance-independent manner. The real power of the direct method lies in its simplicity, allowing for bulk-estimation of fundamental stellar properties of hundreds of main sequence stars, and thousands of red giants observed by *Kepler*, *K2* and *TESS*.

2.3.5 Limitations of the direct method

The seismic scaling relations used in the direct method are scaled to solar values, but not all stars that host solar-like oscillations are solar-like in nature. This is an important distinction that impacts the uncertainty on masses and radii obtained from the seismic scaling relations for red giant stars especially.

Stellar models can alleviate discrepancies in the seismic scaling relations for more

evolved stars by modelling correction terms that account for stellar properties such as temperature, metallicity, and evolutionary state. Including these terms, we can rewrite Equations 2.8 as (Sharma et al., 2016)

$$\begin{aligned} \frac{M}{M_\odot} &\simeq \left(\frac{\nu_{\rm max}}{f_{\nu_{\rm max}}\nu_{{\rm max},\odot}}\right)^3 \left(\frac{\Delta\nu}{f_{\Delta\nu}\Delta\nu_\odot}\right)^{-4} \left(\frac{T_{\rm eff}}{T_{{\rm eff},\odot}}\right)^{3/2} \text{ and} \\ \frac{R}{R_\odot} &\simeq \left(\frac{\nu_{\rm max}}{f_{\nu_{\rm max}}\nu_{{\rm max},\odot}}\right) \left(\frac{\Delta\nu}{f_{\Delta\nu}\Delta\nu_\odot}\right)^{-2} \left(\frac{T_{\rm eff}}{T_{{\rm eff},\odot}}\right)^{1/2}, \end{aligned} \tag{2.9}$$

where we have introduced the terms $f_{\Delta\nu}$ and $f_{\nu_{\rm max}}$ as the model motivated corrections to the scaling relations. While the correction to the $\nu_{\rm max}$ term is likely negligible (Viani et al., 2017), the $\Delta\nu$ term has been found to be critical when studying evolved stars (Brogaard et al., 2018). The exact value of $f_{\Delta\nu}$ however is still an active area of research, some of which is presented in Chapter 4, and the choice of stellar models when calculating $f_{\Delta\nu}$ can lead to significant changes in the inference of stellar properties (Brogaard et al., 2018; Hall et al., 2019).

2.3.6 Asteroseismic Modelling

The direct method of estimating stellar mass and radius through seismic scaling relations is powerful but fundamentally flawed, as it does not take into account stellar characteristics that impact on mass, radius and temperature (such as metallicity). Instead, observations can be compared with grids of stellar models to obtain more reliable measurements of stellar properties, called 'grid-based' modelling (see e.g. Stello et al., 2009b; Gai et al., 2011).

In grid-based modelling, observables are compared to a multi-dimensional grid of stellar models in different parameter spaces. Commonly, evolutionary tracks at different masses and metallicities are used for this purpose (for a review, see Lebreton et al., 2014a,b). By interpolating between the evolutionary tracks closest to the observed stellar parameters for a star, we can find an asteroseismic mass and

radius at a higher precision and without the biases inherent to the direct method. While grid-based modelling is commonly done using $\Delta\nu$, $\nu_{\rm max}$, [Fe/H] and $T_{\rm eff}$, in the era of *Gaia* it is sometimes expanded to include precise measures of stellar luminosity (Rendle et al., 2019a; Huber et al., 2019). Inclusion of the small frequency separation, $\delta\nu_{0,2}$ (which probes the Helium abundance in the core, Vorontsov et al., 1991), can also provide precise ($\lesssim$ 25%, Serenelli et al., 2017) measurements of asteroseismic age through these techniques.

Grid-based modelling efforts can be expanded to use more detailed asteroseismic information for well observed stars, by directly fitting observed seismic mode frequencies to grids of stellar models (see e.g. Metcalfe et al., 2012; Silva Aguirre et al., 2013, 2015, 2017). This so-called 'detailed grid-based modelling' (or just detailed modelling) more comprehensively leverages individual frequency information, eliminating typical simplifications in asteroseismic observations, such as the use of $\nu_{\rm max}$ and an average value for $\Delta\nu$, which in practice varies with radial order. For comparison, state of the art detailed grid-based modelling approaches find ages as precise as $\lesssim$ 15% (Silva Aguirre et al., 2017).

Finally, it should be noted that as with the model-motivated corrections to the direct method, grid-based and detailed modelling introduces a dependence on the choice of stellar model prescriptions (see e.g. Brogaard et al., 2018).

2.4 Summary

The brightness of most stars varies periodically in time. These oscillations are borne from internal processes in the star, and so by combining stellar physical theory with observations, we can determine fundamental stellar properties in a distance independent manner.

The study of solar-like oscillations, which this thesis focuses on, is possible for both Sun-like stars on the main sequence and more evolved (sub-)giant stars, so long as they host a convective outer layer. One of the powers of this class of oscillations

is its direct comparison to the Sun, for which we have much richer theoretical and observational understanding than is possible for distant stars.

Describing a star as a three-dimensional spherical cavity[8] hosting standing waves, we can use the mathematical framework of spherical harmonics, combined with stellar structural theory, to describe the oscillations we see on a star's surface. Combining these observations with theory, I have shown the derivation of the so-called 'seismic scaling relations'.

By combining spectroscopic observations of effective temperature with fundamental seismic observables, seismic scaling relations allow for bulk-estimates of mass and radius for thousands of stars. Measuring global asteroseismic observables with *Kepler*, *K2* and *TESS* is explored further in Chapter 3. For stars more evolved than the Sun, the direct method is subject to ongoing corrections, improvements, and verifications, some of which are presented in Chapter 4.

By combining global asteroseismic observables with grids of stellar evolutionary models, asteroseismology can hugely expand the precision of its mass and radius estimates, while accounting for biases and shortcomings the direct method is subject to. By comparing observations of individual frequencies to modelled stellar frequencies, so-called 'detailed modelling' can improve the precision of asteroseismic mass and radius estimates to the per-cent level. One of the major benefits from working with stellar models over the direct method is that it can provide stellar age to $\lesssim 15\%$ precision, which is otherwise difficult to obtain for field stars[9].

In this Chapter, I have shown that asteroseismology is a powerful stellar laboratory in modern astronomy. Leaning on a rich history of helioseismic observations, asteroseismology with *Kepler* and *K2* has allowed for new insights for both individual stars and larger stellar populations. One such study is presented in Chapter 5, combining detailed modelling, asteroseismic frequency fitting, and stellar evolution-

[8]While we refer to it as a cavity, since the stellar surface is not fixed, it's more like a spherical organ pipe.

[9]Efforts to develop scaling relations for age are currently in their infancy (Bellinger, 2019; Bellinger, 2020)

ary theory.

The future of asteroseismology is promising, with observations with *TESS* ongoing, and the *PLATO* mission (Rauer et al., 2014) on the horizon ensuring a steady flow of new asteroseismic observations. At the same pace, asteroseismic analysis techniques are becoming increasingly robust, accessible and automated. New large-scale spectroscopic surveys are providing auxiliary data needed to improve the precision and scope of asteroseismic techniques even further. Hopefully, as the scale and accessibility of asteroseismic ensembles grow, so too will its use as an cross-disciplinary tool and laboratory of stellar physics.

Chapter 3

Asteroseismology with Lightkurve

This Chapter discusses the asteroseismic capabilities of the Python package Lightkurve, an open-source tool. I developed the majority of Lightkurve's `periodogram` *and* `seismology` *modules, discussed in this Chapter. This Chapter discusses Lightkurve's capabilities at time of writing, which may have changed since. The code used to plot figures 3.4 and 3.5 is adapted from code written by Dr. Christina Hedges, and the full list of Lightkurve collaborators can be found here.*

3.1 Introduction to Lightkurve

The open-source Python package 'Lightkurve' (Lightkurve Collaboration et al., 2018) aims to make the use of space-based photometry consistent and accessible to astronomers and civilian scientists alike. Developed primarily for use with the NASA *Kepler* (Borucki et al., 2010) and *K2* (Howell et al., 2014) missions, and now extensible to the *TESS* mission (Ricker et al., 2015), its built-in methods help users to perform multiple levels of data reduction and analysis. For example, a user could download a snapshot of a star (called a target-pixel-file, or TPF), extract the stellar flux, remove long term trends, and investigate the resulting time series, without having to learn these techniques from first principles.

Lightkurve's suite of tools includes software that I developed for asteroseismic analysis in frequency space. Given a time series from *Kepler*, *K2* or *TESS*, the user

can convert a time series into the frequency-domain to study stellar oscillations (or other periodic signals, such as eclipsing binary transits). For solar-like oscillators, the asteroseismic tools perform a basic extraction of the two fundamental seismic parameters, $\nu_{\rm max}$ and $\Delta\nu$, using tried and tested methods (see Viani et al., 2019, and references therein), and basic stellar parameters for main sequence stars from the seismic scaling relations (Brown et al., 1991; Kjeldsen & Bedding, 1995, and Chapter 2).

With the backdrop of explaining Lightkurve's features for asteroseismology, this Chapter will provide an introduction to time series analysis and asteroseismic measurement techniques. The Chapter is laid out as follows: Section 3.2 will give a thorough overview of how to convert an astronomical time series to the frequency domain and the properties of the frequency-domain. Section 3.3 will discuss the frequency-domain tools included in Lightkurve's `periodogram` module. Section 3.4 will work through how Lightkurve's `seismology` sub-package obtains the fundamental asteroseismic parameters $\nu_{\rm max}$ and $\Delta\nu$. Finally, 3.5 will discuss where Lightkurve's frequency-domain tools can be extended, and briefly address the strengths and weaknesses of the discussed methods. While this Chapter is centred around Lightkurve's functionality, it should provide an understanding of frequency domain and asteroseismic analysis for those unfamiliar with both topics.

3.2 The Frequency-Domain

When observing an oscillating star, a time series of its brightness will show periodic signals, of varying amplitude and frequency. For some pulsators, such as δ-Scuti stars, these will be coherent (non-damped) throughout the time series. For solar-like oscillators, oscillations are both driven and damped by turbulent convection, and so their phase does not remain consistent over long time-scales (Goldreich & Keeley, 1977a,b). Both of these types of oscillations – coherent and stochastic – are commonly studied in the frequency-domain.

Due to their inconsistent phase, the modes of solar-like oscillations in the frequency-domain are commonly modeled as Lorentzian peaks (Garcia & Ballot, 2019), which will have a given width and power indicative of their amplitudes in the time-domain. The range and resolution of the time series in the frequency-domain depends on the length and cadence of observations, and can be found using back-of-the-envelope estimation techniques.

Asteroseismology of solar-like oscillators is most commonly done in the frequency-domain[1], where the stochastic oscillations are reduced to distinct, independent peaks. Before understanding how Lightkurve's `periodogram` tool performs analysis in the frequency-domain, it is important to understand the fundamentals of Fourier Transforms.

3.2.1 Fourier Transforms

The conversion from an observed time series to the frequency-domain can be done through the mathematical framework of a Fourier transform (FT). A FT decomposes a function into the constituent frequencies of its signal. In observational asteroseismology, we use the FT to find the constituent oscillations of a time series.

If we denote the observed flux as a function of time as a $x(t)$, it will have a FT $X(\nu)$. This conversion is written as

$$X(\nu) = \int_{-\infty}^{+\infty} x(t) e^{-2\pi j \nu t} \mathrm{d}t \,, \tag{3.1}$$

where j is the imaginary unit, and ν is frequency, as is the convention in asteroseismology. This function describes a combination of sinusoids at different frequencies ν[2], integrated for the full range t of the time series function. As seen in Figure 3.1, the FT of a perfect sinusoid at frequency ν is a pair of Dirac delta functions located at $\pm\nu$, with amplitudes $X(\nu)$. Functions that are not regular sinusoids, such as

[1] Although time-domain asteroseismology is up-and-coming (see e.g. Pereira et al., 2019; Foreman-Mackey et al., 2017; Farr et al., 2018).

[2] Since $e^{j\theta} = cos(\theta) + j\sin(\theta)$.

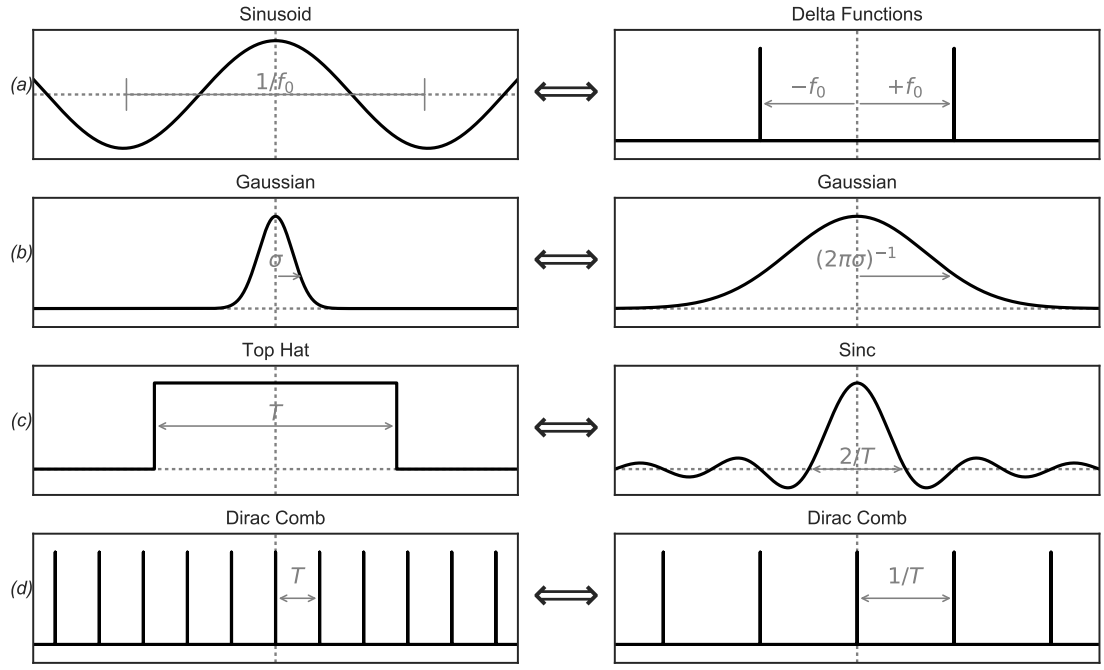

Figure 3.1: A visualisation of various continuous time series signals (left) and their Fourier Transformed counterparts (right). The central dotted lines indicate the zeroth frequency, across which the signal in the frequency domain is reflected. Negative frequencies to the left of this line are imaginary, and typically not considered in asteroseismic work. (Figure 3 of VanderPlas, 2018)

the Top Hat function in Figure 3.1 are described by sinusoids at a larger range of frequencies, and have a less straightforward structure in Fourier space.

3.2.2 Discrete Fourier Transforms

In practice, we do not deal with a continuous function $x(t)$. An astronomical observation is composed of many individual measurements of flux, making the time series a *discrete set* of data. If we instead consider N flux measurements $x[n]$ with indices n, we can compute the Discrete Fourier Transform (DFT) at N discrete frequencies, where each frequency index is given by k. The DFT is given as

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-\frac{2\pi j}{N} kn}, \tag{3.2}$$

where $X[k]$ is the DFT at a frequency bin k, with a frequency bin-width Δ_ν. Equation 3.2 effectively calculates the amplitude of constituent sinusoids of frequency

$k \times \Delta_\nu$ summed across each observation $x[n]$. $X[k]$ is sometimes written as a function of frequency as $X(\nu)$ for simplicity, although it is important to bear in mind that it remains a discrete function, not a continuous one, when calculated through a DFT.

If we compute the squared amplitude of the DFT, we remove the imaginary component and phase information of $X[k]$ and instead calculate the *power spectrum*, as

$$P[k] = |X[k]|^2 \,, \tag{3.3}$$

where $P[k]$ is referred to as the *power*.

3.2.3 Frequency Resolution

Given a DFT of a time series, we may want to find precise details of the resulting power spectrum. Consider a telescope with an observing cadence Δt. If N observations are made, the total length of the time series is

$$T = N\Delta t \,. \tag{3.4}$$

Now consider an observed signal with a period $P = T$. This signal would complete one full cycle within the observation. For oscillations at longer periods, a full observation of the signal would not be possible. This sets the *natural frequency resolution* of a time series to $1/T$, so that

$$\Delta_\nu \equiv T^{-1} = (N\Delta t)^{-1} \,, \tag{3.5}$$

where Δ_ν is the frequency bin-width of our DFT. Equation 3.5 highlights the importance of long-baseline observations for asteroseismology; increasing the observational baseline T increases the frequency resolution. For stars with a full 4 years of uninterrupted data such as that obtained by *Kepler*, the frequency resolution can be as

low as 7.9 nHz.

3.2.4 Nyquist Frequency

Now that we know the frequency resolution of our time series, the next question to ask is what its maximum observable frequency is. We again consider our observation of length T and cadence Δt. If we want to reliably measure the frequency of a periodic signal in our observations, we must have at least two measurements per period (one on the 'rise' and one on the 'fall'). This sets the maximum frequency we can reliably observe as

$$\nu_{\mathrm{Nyq}} \equiv (2\Delta t)^{-1}\,, \tag{3.6}$$

where ν_{Nyq} is the so-called Nyquist frequency (Grenander, 1959). Following Equation 3.6, the shorter our observational cadence, the higher the maximum frequency we can observe (i.e. faster oscillations). For an oscillation with a frequency $\nu \leq \nu_{\mathrm{Nyq}}$, a peak will appear in the frequency spectrum at frequency ν. For an oscillation with $\nu > \nu_{\mathrm{Nyq}}$ a peak will appear beyond the Nyquist frequency in the so-called *super-Nyquist* regime. This peak will be reflected (aliased) back across ν_{Nyq} to the sub-Nyquist regime. However these frequencies can not be implicitly trusted, as it is impossible to know whether the peak represents a true frequency below ν_{Nyq}, or an aliased peak from the super-Nyquist regime (for examples of super-Nyquist asteroseismology, see e.g. Murphy et al., 2013; Chaplin et al., 2014).

For example: for the *Kepler* long cadence ($\Delta t \sim 30\,\mathrm{min}$), $\nu_{\mathrm{Nyq}} \sim 278\,\mu\mathrm{Hz}$, allowing for observations of large (slowly oscillating) red giant stars. For the *Kepler* short cadence ($\Delta t \sim 1\,\mathrm{min}$), $\nu_{\mathrm{Nyq}} \sim 8333\,\mu\mathrm{Hz}$, allowing for observations of the smaller (fast oscillating) main sequence dwarfs. However it is important to note that in the case of uneven observations, Equation 3.6 breaks down and ν_{Nyq} becomes ambiguous. For consistent space-based observations, such as *Kepler*, this poses little issue, but must be considered especially when fitting frequency peaks close to the

Nyquist frequency, or when using the Nyquist frequency as part of a forward-model.

3.2.5 Features of the Power Spectrum

Parseval's Theorem

The amount of power a signal produces in the frequency-domain is tied to its amplitude in the time-domain. This relationship is described by Parseval's Theorem, which states that the integral of the square of a function is equal to the integral of the square of its FT. This is expressed as

$$\int_{-\infty}^{+\infty} \mathrm{d}\, t|x(t)|^2 = \int_{-\infty}^{+\infty} \mathrm{d}\, \nu|X(\nu)|^2\,, \tag{3.7}$$

for a function $x(t)$ with a FT $X(\nu)$. This equation is especially important in Fourier analysis, as it can be used to normalise FTs of astronomical data to ensure all power measurements are consistent (see e.g. Chaplin & Basu, 2017).

Parseval's theorem as described in Equation 3.7 is for a continuous FT (Eq. 3.1). It can be rewritten for the DFT (Eq. 3.2) as (Appourchaux, 2014)

$$\sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2\,, \tag{3.8}$$

where conventions are as above.

It is important to note that Parseval's theorem *must* be true for any FT, and therefore can be used to calibrate a calculated DFT. For different definitions of the DFT (such as those employed by different Python packages and methods), the normalisation of Parseval's theorem may be different. Getting this normalisation correct is, understandably, an important part of keeping methods consistent within the field of asteroseismology.

Noise Properties

Following a FT, data with Gaussian noise properties in the time-domain will appear in a power spectrum[3] as having the properties of a chi-squared (χ^2_k) distribution with two degrees of freedom (Woodard, 1984; Appourchaux et al., 1998). The generalised χ^2_k probability distribution goes as

$$\chi^2_k : p(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}\,, \tag{3.9}$$

where k is the degrees of freedom, Γ is the so-called Gamma function[4], x is the value the noise property is being applied to, and $p(x)$ represents the probability of finding a data point at a value x. Reduced for $k = 2$ and given a signal in power P, the noise properties of the power spectrum follow

$$\chi^2_2 : p(P) = \frac{1}{2} e^{-P/2}\,. \tag{3.10}$$

Unlike Gaussian noise, which is additive (i.e. added to the time series data), χ^2_k noise is *multiplicative*. This means that the mean and median of a power spectrum signal are affected by the noise properties. For two degrees of freedom ($k = 2$), Gaussian noise with a standard deviation of 1 will appear in the frequency domain with a mean power of 2 and a median power of $2(8/9)^3 \approx 1.4$. Taking these non-trivial noise properties into account is especially important when re-binning asteroseismic data (in which case the properties change), dividing out low-frequency background noise, and fitting an asteroseismic forward-model (see Chapter 5).

3.2.6 Conventions in Asteroseismology

There are a number of conventions in the study of solar-like oscillators that help contextualise power spectra found in the literature and throughout this thesis, discussed below. For a more in-depth text on contemporary asteroseismic data analysis, see

[3] Meaning, the modulus squared of its Fourier Transform, see Eq. 3.3.
[4] $\Gamma(n) = (n-1)!$ for any positive integer n.

e.g. Appourchaux (2014); Chaplin & Basu (2017).

Renormalising Parseval's theorem

In asteroseismology of solar-like oscillators, it is common practice to normalise the output of a DFT by the length of the time series, N. This impacts the definition of power described in Equation 3.3, so that

$$P_{\rm seis}[k] = \frac{|X[k]|^2}{N^2} = \frac{1}{N^2}\left|\sum_{n=0}^{N-1} x[n] e^{-\frac{2\pi j}{N}kn}\right|^2 , \tag{3.11}$$

where conventions are as above.

With this renormalisation of power, we need to renormalise Parseval's theorem so that it continues to hold true. By substituting $|X[k]|^2 = N^2 P_{\rm seis}[k]$ into Equation 3.8, we find, as in Chaplin & Basu (2017),

$$\frac{1}{N}\sum_{n=0}^{N-1}|x[n]|^2 = \sum_{k=1}^{N/2} P_{\rm seis}[k] \, , \tag{3.12}$$

where we have now expressed Parseval's theorem as a function of our defined power directly, making it more practically useful. On the left-hand-side, we have an expression of the mean square amplitude of a signal.

There is another change of convention in Equation 3.12: note the change of limits on the frequency side of the equation. This is because in asteroseismology, we only care about the positive-frequency components of the FT, as we are dealing with a real signal, for which negative frequencies hold no physical meaning. In order to ensure agreement with Parseval's theorem, this often introduces factors of 2 when normalising power spectra, as seen below.

Signal and Variance

In astronomy, time domain observations typically have Gaussian (normal) noise properties (and often if they don't, they are assumed to). Gaussian noise in the time-domain has a particular effect on the frequency-domain signal. To start, let's

consider a discrete time series $x[n]$ consisting of normally distributed noise, centred on zero, and with a standard deviation σ. The mean squared power of this noise is equivalent to the variance, σ^2, and can be written as

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N} x[n]^2 \,. \tag{3.13}$$

Notice how this equation is identical to the time-domain side of our re-normalised Parseval's theorem (Eq. 3.12). From Parseval's theorem then, we know that the summed power in the frequency-domain across $N/2$ bins must be equal to σ^2. In other words, the time-domain power from the noise is distributed across all frequencies. This means that the average power in the frequency domain is

$$\langle P_{\rm seis} \rangle = \frac{2}{N} \times \sigma^2 \,, \tag{3.14}$$

appearing as a 'noise floor' in our power spectrum.

Equation 3.14 shows us that by reducing the uncertainty on our astronomical observations, we can see peaks of lower amplitude in the frequency domain. For example, let's consider a single sinusoid with an amplitude A.[5] The mean squared power of this signal (and thus the time-domain side of Parseval's theorem), is equal to $A^2/2$. The majority of this power will be found in the frequency bin containing the frequency of the time-domain signal, with power 'leaking' into surrounding frequency bins and those containing the harmonic frequencies of the signal. These additional peaks are too small to be visible when considering the DFT of a perfect sinusoid with no additional noise, but are exacerbated when noise and poor sampling are present. Ensuring that these properties of the frequency-domain are retained by using appropriate normalisation factors (see below), measurements of power spectrum properties can be more accurately mapped back to physical properties of the star. For examples of how both a sinusoid and white noise appear in the frequency

[5]Here and throughout this thesis, I colloquially use 'amplitude' to mean what is sometimes referred to as the semi-amplitude: half of the peak-to-peak amplitude of a sinusoid.

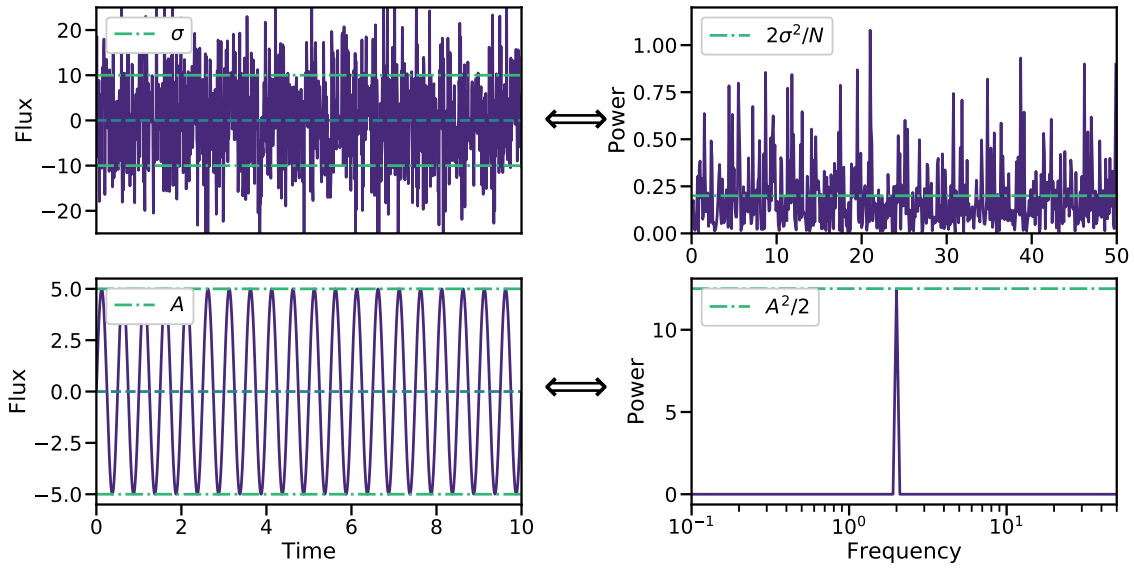

Figure 3.2: Two different time series (left) and their periodograms (right). *Top*: Gaussian noise with a variance σ^2 produces noise in the frequency-domain with mean value $2\sigma^2/N$, which is distributed following a chi-squared distribution with two degrees of freedom (Eq. 3.10). *Bottom:* A noiseless sinusoid of amplitude A and a frequency of 2, which produces a peak of height $A^2/2$ in power. See text for details.

domain, see Figure 3.2.

Power Spectral Density

As discussed above, our seismic power is proportional to the modulus squared of the DFT (Eq. 3.11). This yields units of $[x]^2$, where $[x]$ is the unit of the time series. In contemporary asteroseismology of solar-like oscillators using space-based photometry, this is parts-per-million (ppm). For other types of oscillators, magnitude is also used, or ms^{-1} for radial-velocity measurements.

In order to effectively compare power spectra for different stars, or different observations of the same star, it becomes necessary to normalise the power to *power spectral density*, by dividing through by the bin-width as

$$\mathrm{PSD}[k] = P_{\mathrm{seis}}[k]/\Delta_\nu \,. \tag{3.15}$$

In the current convention, units of μHz are often used for Δ_ν, resulting in units

of ppm$^2/\mu$Hz. Colloquially the phrase 'power spectrum' may sometimes refer to a power spectrum that has been normalised in this manner.

Apodization

Applying a FT – something that is by design continuous – to discrete observational data, has some detrimental effects to the resulting power spectrum. In most modern telescopes all flux between instrument read-outs is contained to a single data-point. This means that the *integration time* as well as the observing cadence is also equal to Δt, and that our signal, which we can consider continuous, is convolved with a window function when observed. In this case we can consider our window function to be a top-hat function of width Δt.

To calculate the effect of applying a window function to the time series, we rely on two known properties of the FT: that a convolution in the time-domain is equivalent to a multiplication in the frequency-domain, and that the FT of a top-hat function is a sinc function (see Figure 3.1, VanderPlas, 2018). Our top-hat window function in the frequency-domain is then given as

$$\eta(\nu) = \mathrm{sinc}[\pi\nu\Delta\mathrm{t}] = \mathrm{sinc}\left[\frac{\pi}{2}\left(\frac{\nu}{\nu_{\mathrm{Nyq}}}\right)\right], \tag{3.16}$$

where sinc = sin(x)/x, and we have used the fact that $\nu_{\mathrm{Nyq}} = \frac{1}{2\Delta t}$. The expansion on the right-hand-side of the equation relies on the assumption that the integration time is equal to the observing cadence, which is true for *Kepler*, but not necessarily for all observations. Equation 5.1 is referred to as the *apodization*, and attenuates the signal in the frequency spectrum. Note that when working in power the signal is attenuated as η^2, so that our *observed* power in the frequency domain is

$$P_{\mathrm{obs}}(\nu) = P_{\mathrm{seis}}(\nu) \times \eta(\nu)^2\,, \tag{3.17}$$

where we are expressing power as a continuous function for the sake of clarity. Finally, it is worth noting that this does not have a major impact on asteroseismology

in the sub-Nyquist regime, as $\eta^2(\nu = \nu_{\rm Nyq}) \approx 0.4$, and does not reach its first zero until $\nu = 2\nu_{\rm Nyq}$. However it is important to take into account when forward-modeling solar-like oscillations, particularly for smaller stars oscillating at high frequencies.

3.3 Lightkurve's `periodogram` functionality

The Lightkurve Python package (Lightkurve Collaboration et al., 2018) aims to make the use of time series photometry from the *Kepler*, *K2* and *TESS* missions accessible and consistent for users not necessarily familiar with the theory described above. The `periodogram` module of Lightkurve contains tools used to transform time series photometry into the frequency-domain, and perform further manipulation and data reduction of the resulting power spectrum.

3.3.1 Lomb-Scargle periodograms

At this stage, it becomes important to point out the conceptual difference between a 'power spectrum' and a 'periodogram'. The power spectrum, expressed *analytically* in Equations 3.3 and 3.11, is a continuous function representing the mathematical Fourier transform of a time series. A periodogram is the *estimation* of that power spectrum, usually obtained through *numerical* techniques.

One such method is the classical periodogram, which has the form (Schuster, 1898)

$$P_{\rm classic}[k] = \frac{1}{N}\left|\sum_{n=0}^{N-1} x[n] e^{-\frac{2\pi j}{N}kn}\right|^2 , \tag{3.18}$$

which takes a similar form to our previous expressions of power, but with a change in factors of N.

As discussed above, a power spectrum of a time series with normally distributed noise will be distributed following a chi-squared two degrees of freedom distribution. When using the classical periodogram to estimate a power spectrum, this holds true. However, if data are non-uniformly sampled, this property no longer holds.

In astronomy, where there can be long breaks between observations, accounting for non-uniform sampling is important. If we're being precise, even space-based photometry is subject to small-timescale changes in sampling rates[6]. To account for this, we can instead use the so-called Lomb-Scargle periodogram.

The Lomb-Scargle periodogram differs only slightly from the classical periodogram, and takes the form (Scargle, 1982)

$$\begin{aligned} P_{\mathrm{LS}}[\nu] =& \frac{A^2}{2}\left(\sum_{n=0}^{N-1} x[n]\cos(2\pi\nu[t[n]-\tau])\right)^2 \\ &+\frac{B^2}{2}\left(\sum_{n=0}^{N-1} x[n]\sin(2\pi\nu[t[n]-\tau])\right)^2 . \end{aligned} \tag{3.19}$$

The key changes here from the classical periodogram are the additional normalisation factors A and B and the phase component τ. The sine and cosine terms are expanded from the exponential term using Euler's method. The discrete function $t[n]$ represents the time at observations $x[n]$. The factors A, B and τ are arbitrary functions of frequency and time, the forms of which are chosen so that the Lomb-Scargle periodogram reduces to the classical periodogram in the case of equally-spaced observations, while not being affected by time-shifts in the data (see VanderPlas, 2018, Section 5).

The modified periodogram in Equation 3.19 is functionally identical to a result obtained by fitting sinusoids of given frequencies ν to the data $x[n]$, and constructing a periodogram based on the 'goodness-of-fit' of a χ^2_2 probability function described in Equation 3.10 (Lomb, 1976). This means that even for non-uniformly sampled data the Lomb-Scargle periodogram will have chi-squared distributed power, while still accurately estimating the power spectrum (VanderPlas, 2018).

[6]For an example of how small these can be, *Kepler*'s Long Cadence (29.42 min) has inconsistencies on the order of milliseconds.

3.3.2 The Lomb-Scargle periodogram implementation in Lightkurve

Lightkurve's `periodogram` module acts as a wrapper for the `astropy` (Astropy Collaboration et al., 2013, 2018) `LombScargle` module (VanderPlas et al., 2012; VanderPlas & Ivezić, 2015). This allows for custom default inputs that are more typical for asteroseismology, while still providing freedom of use.

By default, a periodogram will be created for a regular grid of frequencies, spanning from one bin-width ($(N\Delta t)^{-1}$, Eq. 3.5) to the Nyquist frequency ($(2\Delta t)^{-1}$, Eq. 3.6). The minimum and/or maximum frequencies can be provided instead to set limits to the frequency grid, in which case the Lomb-Scargle periodogram is only computed for this discrete grid of frequencies. A custom grid of frequencies can also be provided, evenly spaced or otherwise. While a regular grid of frequencies will most closely approximate the true power spectrum, the use of a Lomb-Scargle periodogram means that choices of irregular spacing will not affect the noise properties of the periodogram.

The periodogram is by default created for a grid with a bin-width of $(N\Delta t)^{-1}$, set by the length of the time series. This is commonly referred to as 'critical sampling', and provides a periodogram where bins of frequency are independent from one another. While critical sampling is preferred for fitting a forward-model to data, there are cases where an asteroseismologist may prefer to over- or undersample the spectrum. This is useful in order to measure an amplitude directly from a plot, with no fitting required, in which case an oversample factor of 5 to 10 is typically preferred (i.e. the bin-width is 5 to 10 times smaller than at critical sampling).

3.3.3 Periodogram Normalisation

In order to cater to asteroseismologists studying both stochastic and coherent oscillators, Lightkurve has multiple normalisation options, based on the criteria set forth in Kjeldsen & Bedding (1995) and Chaplin & Basu (2017).

For asteroseismologists studying solar-like oscillators, the y-axis units are com-

monly expressed in Power Spectral Density (PSD, Eq. 3.15). Before rescaling by the bin-width, the power must obey the renormalised Parseval's theorem shown in Equation 3.12. In the case of evenly-spaced data, the Lomb-Scargle periodogram approximates a classical periodogram. Comparing Parseval's theorem in Equation 3.12 and the classical periodogram in Equation 3.18, we can see that the output of the Lomb-Scargle method must be renormalised as

$$\mathrm{PSD}[k] = \frac{P_{\mathrm{seis}}[k]}{\Delta_\nu} = \frac{2 \times P_{\mathrm{LS}}[k]}{N \times \Delta_\nu}\,, \tag{3.20}$$

where N is the length of observations, and P_{LS} is the Lomb-Scargle power. The factor of two here is introduced to satisfy Parseval's theorem, due to only considering real components of the DFT.

For asteroseismology of coherent pulsators, looking at the periodogram solely in amplitude is preferred for its increased dynamic range, making it easier to study power excess by eye. Lightkurve's amplitude normalisation does not renormalise the Lomb-Scargle power to comply with Parseval's theorem. Instead, it normalises so that a signal with amplitude A in the time domain produces an amplitude of A in the amplitude spectrum, where the amplitude spectrum ($A[k]$) is the square root of the power spectrum ($P[k]$).

In Section 3.2.6 I showed how for a power spectrum adhering to Parseval's theorem (Eq. 3.12), a signal with amplitude A and frequency ν_0 will appear with a power $P(\nu_0) = A^2/2$. In order to obtain a periodogram with mode amplitudes that match those in the time-domain signal, we must renormalise the Lomb-Scargle periodogram by an additional factor of two. The amplitude spectrum[7] then becomes (Kjeldsen & Frandsen, 1992; Kjeldsen & Bedding, 1995)

$$A[k] = \sqrt{\frac{4 \times P_{\mathrm{LS}}[k]}{N}}\,. \tag{3.21}$$

[7]Or more properly, amplitude periodogram.

When creating a periodogram in Lightkurve, the 'psd' and 'amplitude' normalization options provide the periodograms described by Equations 3.20 and 3.21 respectively.

3.3.4 Smoothing and Binning

Smoothing a periodogram is useful for plotting and investigating data, as well as removing background noise. Lightkurve implements two smoothing techniques: a box kernel method, and a moving log-median method.

The box kernel method smooths the periodogram by convolving it with the `astropy.convolution.Box1DKernel`, which is a top hat function. The width of the 1D box kernel filter is customisable, with a default of 10 times the periodogram bin-width This simple approach is best for filtering out noise while retaining the seismic mode peaks, and does not work if the periodogram is not built on an evenly spaced grid of frequencies.

The moving log-median method smooths the power spectrum by calculating the median for increasingly large intervals of data. The filter moves in steps of $0.5\times$ the filter width, but is defined in log space. This means that for equal steps in log-space, the range of bins included in the filter in linear space increases at higher frequencies. This is better for estimating the noise background, as it uses fewer bins near the seismic modes, and more bins near the low-frequency granulation background.

The box kernel method convolves the periodogram with a kernel, smoothing the data but not altering the median. The log-median method on the other hand replaces the data in each frequency bin with the median of the smoothed region, effectively binning the data and changing the median. As mentioned above, the median of data with chi-squared two-degrees-of-freedom noise properties will have a median of $2(8/9)^3 \approx 1.4$. In order to maintain consistency between the two smoothing methods, the medians calculated in the log-median method are divided through by this correction factor. An example of the two different smoothing techniques applied in PSD and amplitude are shown in Figure 3.3.

Finally, you can also bin the power spectrum. This is technically also a smoothing technique, taking the mean of each bin. The difference with the other smoothing techniques is that re-binning changes the number of frequency bins, whereas this is maintained for the box kernel and log-median methods. The bin size, set to a default of 10, determines the number of frequency bins included in a single new bin (e.g. for a bin size of 5, the re-binned periodogram will have $N/5$ frequency bins).

While it is not applied in this thesis, it is useful to note that re-binning a power spectrum changes its noise properties. Instead, its noise properties become that of a chi-square distribution with $2 \times b$ degrees of freedom (see Eq. 3.9), where b is the number of frequency bins included in each new bin (i.e. the binsize parameter, Appourchaux, 2004). If fitting a forward model to data re-binned using a mean method, as is used in Lightkurve, extra care must be taken to re-normalise the binned data so that the median is equal to $k\left(1-\frac{2}{9k}\right)^2$, where k is the new degree of freedom.

3.3.5 Flattening

When studying asteroseismic modes, it is helpful to divide out the granulation background. When fitting a forward model, this is done by fitting the background components to the data (see Chapter 5), but for surface-level investigation of the data, an approximation of the background can provide the same effect.

Lightkurve's `periodogram.flatten` function first estimates the background using the log-median filter, with a filter width of 0.1 in log space. It then divides the periodogram through by the background, resulting in a periodogram with a mean of 1 in the y-axis units. This is commonly referred to as a Signal-to-Noise (SNR) periodogram. While this is technically a misnomer, as the only noise source removed was the background, we will continue to use this term throughout this Chapter.

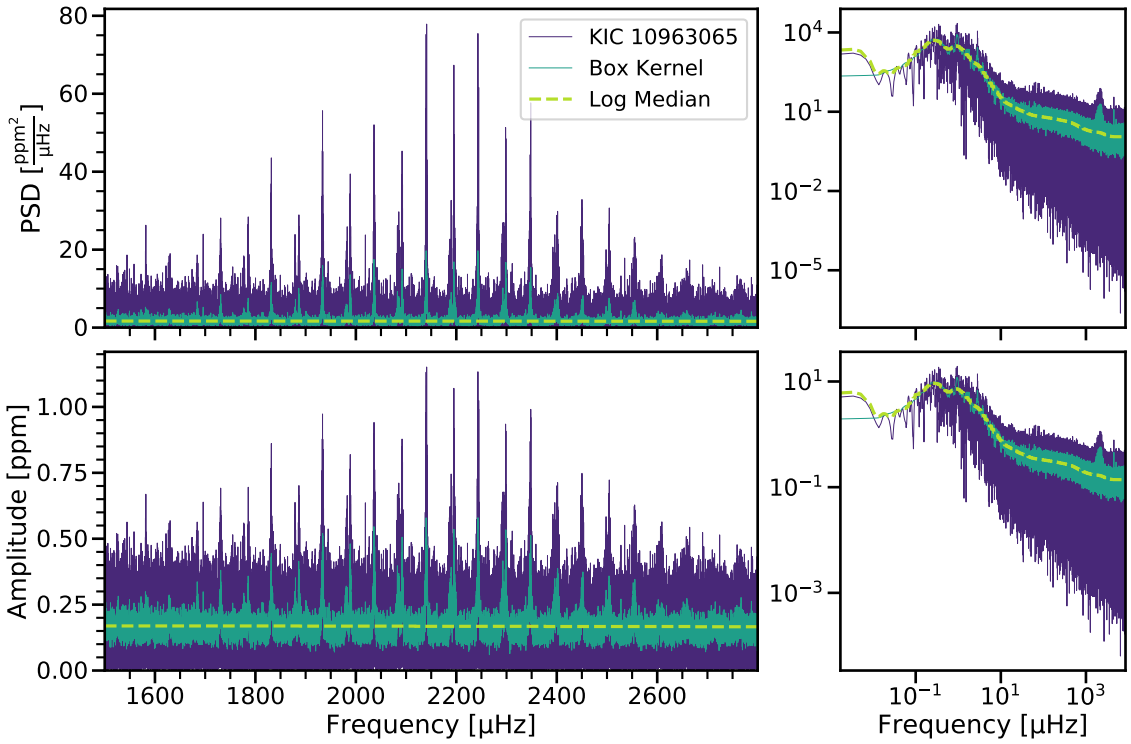

Figure 3.3: Example of different detrending and smoothing outcomes on the main sequence star KIC 10963065. *Left*: Modes of oscillation in power spectral density (top) and amplitude (bottom). The periodogram smoothed by a 1D Box Kernel is shown in front of the full periodogram. The dashed line indicates the periodogram smoothed using a moving log-median, which does not preserve the mode frequencies. Note the higher dynamic range in amplitude space. *Right*: The same, but for the full frequency range, shown in log space. The 1D Box Kernel preserves the modes of oscillation, while the log-median filter preserves the low-frequency noise and smooths over the modes. See text for details.

3.4 Obtaining stellar properties with asteroseismology

In Chapter 2 I showed how the two global asteroseismic observables, $\nu_{\rm max}$ and $\Delta\nu$, can be used to recover a rudimentary estimate of mass and radius using the seismic scaling relations

$$\begin{aligned} \frac{M}{M_\odot} &\simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^3 \left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-4} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{3/2} \text{ and} \\ \frac{R}{R_\odot} &\simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right) \left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-2} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{1/2}, \end{aligned} \tag{3.22}$$

as shown in Equation 2.8. Lightkurve's `seismology` sub-package takes in an SNR periodogram, and uses tried-and-tested techniques to make a rough estimate of $\nu_{\rm max}$ and $\Delta\nu$.

3.4.1 Estimating $\nu_{\rm max}$

One of the most robust automated methods of estimating the frequency of maximum oscillation, $\nu_{\rm max}$, is the use of an autocorrelation function (2D ACF method, see e.g. Mosser & Appourchaux, 2009; Huber et al., 2009; Verner & Roxburgh, 2011; Viani et al., 2019)[8]. This technique leverages the fact that solar-like modes of oscillation appear as a regularly spaced 'comb' in the periodogram. By correlating a spectrum with itself, an ACF will show both the spacing between repeating patterns as well as find power excess indicating the location of $\nu_{\rm max}$.

Lightkurve uses the 2D ACF method presented in Viani et al. (2019) and references therein. In this technique, a window of fixed width is moved along the periodogram, evaluating the correlation of the periodogram contained within the window with itself at each step. Lightkurve uses `numpy`'s correlation function (`numpy.correlate`,

[8]Other ACF techniques are also commonplace in asteroseismology, see e.g. Roxburgh & Vorontsov (2006) and Roxburgh (2009).

van der Walt et al., 2011), where the correlation between two real sets $x[n]$ and $y[n]$ of length N can be defined as

$$C_{(x,y)}[m] = \sum_{n=0}^{N-1}(x[n+m] \times y[n])\,, \text{ where} \\ x[n+m \geq N] = 0\,, \tag{3.23}$$

and where the index m indicates the shift in the spectrum, also called a 'lag'. For a range of $m = 0$ to $m = N$, the function $C_{(x,y)}$ will show the correlation between the two sets overlapping to different degrees. To calculate the correlation of a spectrum with itself is the autocorrelation function, which is then given as

$$\text{ACF}[m] \equiv C_{(x,x)}[m] = \sum_{n=0}^{N-1}(x[n+m] \times x[n])\,, \tag{3.24}$$

where $x[n+m \geq N] = 0$, as above. In Lightkurve, we take x to be a window of the SNR spectrum. As Equation 3.24 is a sum operation, the SNR is also rescaled by subtracting its mean, decreasing the overall noise levels in the ACF. An ACF calculated from $m = 0$ to $m = N$ will be mirrored across the central lag $m = N/2$ (the correlation of the window fully overlapped with itself), and so we consider only the latter half of the ACF when calculating seismic properties.

The ACF is calculated for multiple windows of the SNR spectrum. For example, for main sequence stars (with ν_{max} typically in the thousands of μHz), Lightkurve defaults to a window width of 250 μHz, taking steps of 10 μHz. For a periodogram with a Nyquist frequency at 8000 μHz, the window starts at 125 μHz and ends at 7875 μHz, resulting in 775 separate ACFs each of a length equal to the number of frequency bins contained in the 250 μHz window. This is the 2D aspect of this technique, with one dimension being the ACFs and the other the central frequencies of the windows used to calculate the ACFs. For red giant stars (with ν_{max} in the hundreds of μHz), Lightkurve uses defaults of a window width of 25 μHz and a step

size of 1 μHz.

In order to determine $\nu_{\rm max}$, we want to find the central frequency of the window that resulted in the highest ACF power, indicating a maximum power excess. To do so we calculate the Mean Collapsed Correlation (MCC, see Viani et al., 2019), as

$$\mathrm{MCC} = (\sum_{m=0}^{N/2} |\mathrm{ACF}[m]| - 1)/n_{\rm lags} \tag{3.25}$$

where $n_{\rm lags}$ is the number of lags (i.e. windows) that the ACF is calculated for, in this case $N/2$ (where N here is the length of the *periodogram*). In Lightkurve, the MCC is further convolved with an `astropy` Gaussian 1D kernel with a standard deviation of $1/5^{\rm th}$ of the window size to smooth out any spurious peaks. The value for $\nu_{\rm max}$ is then taken to be the frequency bin corresponding to the highest peak on the smoothed MCC. A representation of the different steps of the $\nu_{\rm max}$ estimating process are shown in Figure 3.4.

While this method works relatively well unsupervised, it is not robust against large peaks in the spectrum which are not of an asteroseismic origin, and does not work well in the case of low signal-to-noise data, where the asteroseismic modes do not clearly stand out from the background.

3.4.2 Estimating $\Delta\nu$

Given an estimate of $\nu_{\rm max}$, we can now estimate $\Delta\nu$. In fact, the regular spacing of the mode frequencies should already be visible in the 2D ACF shown in Figure 3.4. To estimate $\Delta\nu$ we only need to calculate one ACF, centred on our $\nu_{\rm max}$ estimate with a width large enough to include as many mode frequencies as possible. Around $\nu_{\rm max}$, the heights of the mode frequencies roughly follow a Gaussian function. For main sequence stars, the Full Width Half Maximum (FWHM) of this Gaussian 'hump' is approximately (Lund et al., 2017)

$$\mathrm{FWHM} \approx 0.25 \times \nu_{\rm max}\,, \tag{3.26}$$

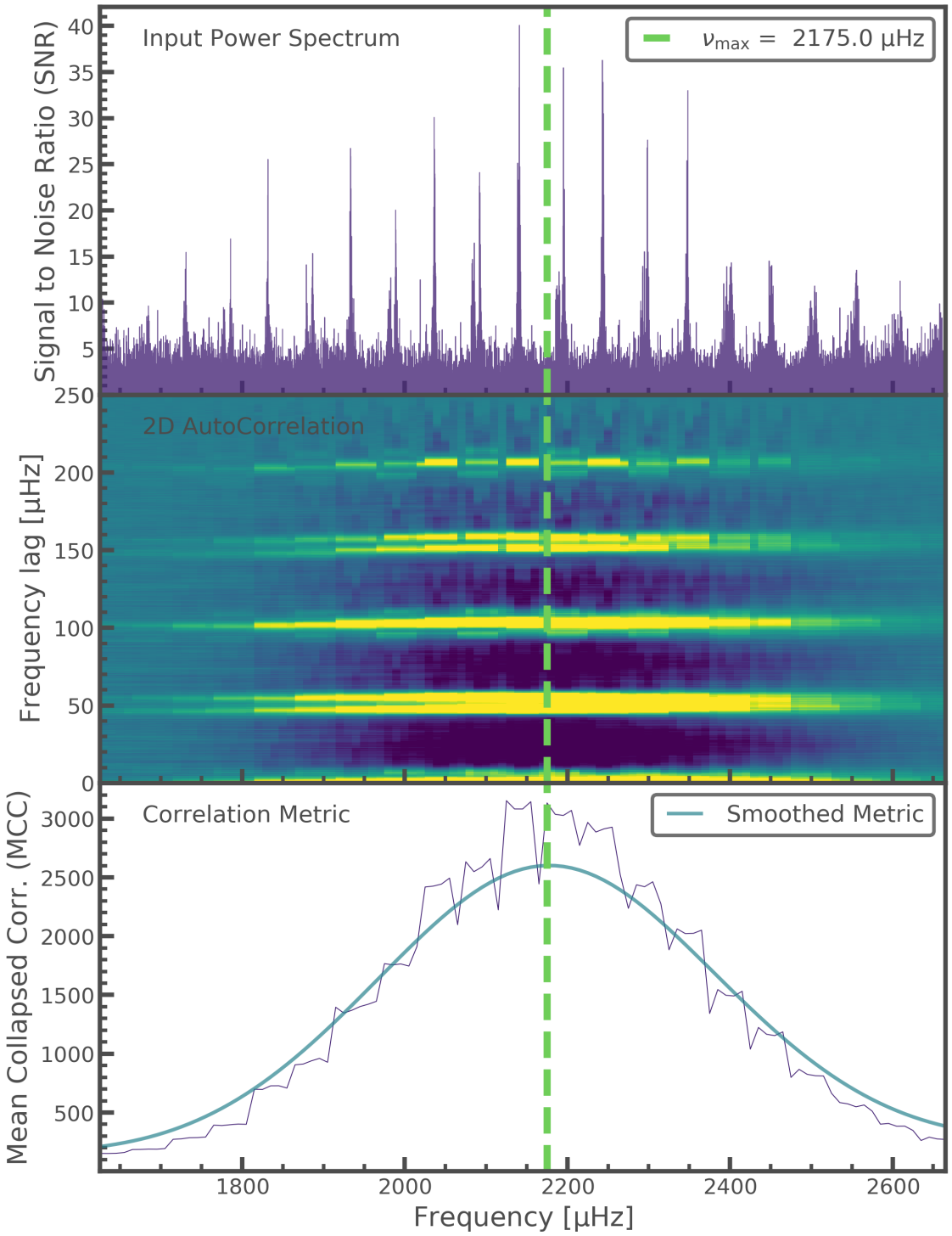

Figure 3.4: The ACF method of finding $\nu_{\max}$. *Top:* the input SNR periodogram, obtained by flattening a Lomb-Scargle periodogram. *Middle*: the 2D ACF. The y-axis shows the frequency lag of the ACFs, and the x-axis shows the central frequencies of the window for which the ACF was calculated. Brighter regions indicate relatively increased autocorrelation (and vice versa). *Bottom*: The Mean Collapsed Correlation (MCC, see Eq. 3.25), and a smoothed version of the MCC overlaid. The determined $\nu_{\max}$, indicated by the dotted line throughout all three plots, corresponds to the maximum of the smoothed MCC. See text for details.

whereas for red giant stars it is approximately (Mosser et al., 2010)

$$\text{FWHM} \approx 0.66 \times {\nu_{\text{max}}}^{0.88} \,. \tag{3.27}$$

The ACF is then calculated for a window centred on ν_{max} with a width of one FWHM either side of ν_{max}, encompassing all visible modes of oscillation. As done in Mosser & Appourchaux (2009), the ACF is then rescaled in terms of the mean noise level, as

$$A = \frac{|\text{ACF}[m]|^2}{|\text{ACF}[0]|^2} \times \frac{2n_{\text{lags}}}{3} \,, \tag{3.28}$$

where ACF[0] is the zeroth bin of the ACF (i.e. the sum of the window of the periodogram multiplied with itself). This renormalisation results in peaks in the ACF which are more distinct from the background noise[9].

The regular spacing of the modes of oscillation will cause repeating peaks as modes of one radial order overlap with the following radial order as the spectrum is shifted over itself. As such, peaks will appear at multiples of the large frequency spacing (and of the small frequency spacing, in high signal-to-noise spectra). In order to automate the process of identifying the peak corresponding to $\Delta\nu$, Lightkurve makes an estimate of $\Delta\nu$ based on the empirical relation (Stello et al., 2009a)

$$\Delta\nu \approx 0.294 \times {\nu_{\text{max}}}^{0.772} \,. \tag{3.29}$$

Lightkurve them runs a peak-finding algorithm (`scipy.signal.find_peaks`) within a range of $0.25 \times \Delta\nu$ around the empirical $\Delta\nu$. Lightkurve's estimate for $\Delta\nu$ is then the value of the shift in the ACF corresponding to the peak closest to the empirical value. A representation of the different steps of the ν_{max} estimating process are shown in Figure 3.5.

[9] As the overlap of the windows increases, noise levels in the ACF will increase as more background noise is included in the calculation

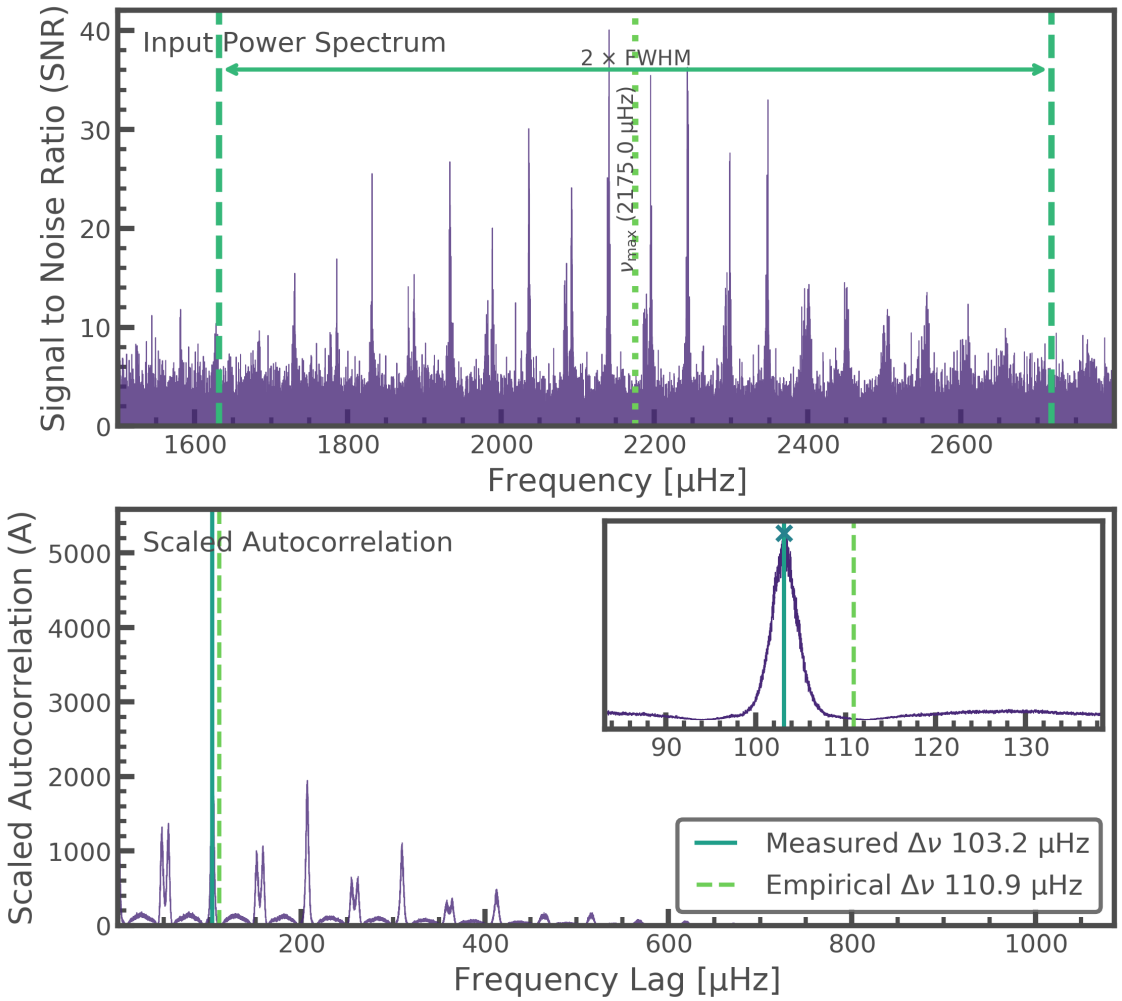

Figure 3.5: The ACF method of finding $\Delta\nu$ *Top*: the input SNR periodogram, obtained by flattening a Lomb-Scargle periodogram. The dotted line near the centre represents the $\nu_{\rm max}$ obtained from the 2D ACF method on the same data. The dashed lines either side of the plot show the limits for which the ACF is calculated to estimate $\Delta\nu$, which is one FWHM either side of the $\nu_{\rm max}$ estimate. *Bottom*: the scaled autocorrelation function (see Eq. 3.28). The sharp peaks show where the repeating pattern of the modes overlap. *Inset*: A region of 25% around the empirically determined $\Delta\nu$ (see Eq. 3.29), shown by the dashed line. The solid line shows the measured $\Delta\nu$, and the cross shows the exact location of the peak. Both lines are replicated in the lower plot. See text for details.

3.5 Discussion

Lightkurve provides accessible tools for first-look investigations into the frequency domain, based on best-practices of time series analysis within astronomy. The tools in Lightkurve are built upon the foundation laid by asteroseismic analysis techniques developed for- and over the course of- the *Kepler* mission.

Overall, the techniques implemented are rudimentary compared to the state-of-the art in asteroseismology. ACF approaches have been improved beyond the implementation presented in this thesis in other works (improved automation, see e.g. Yu et al., 2018), and can not be applied to all stellar types indiscriminately.

Lightkurve does not implement uncertainties at this stage, functioning primarily as a first-look tool. If it is to truly become a data analysis tool, it must provide uncertainties as part of its techniques (e.g. through fitting the ACF spectrum, Huber et al., 2009), and providing multiple estimation techniques for comparison (as in Viani et al., 2019).

However, Lightkurve is reaching the end of its core development, following the completion of the *K2* mission and the retirement of the *Kepler* telescope. The ease-of-access of Lightkurve has sparked development of new automated and open-source projects, such as PBJam (Nielsen et al. in prep.), DIAMONDS (Corsaro & De Ridder, 2015) and Pyriod (Bell, 2020). In the era of large-data astronomy, the more typical workflows are automated, the more time is saved re-inventing the wheel. Building well-documented and frequently used tools for common processes (such as the normalisation of the power spectrum) also comes with the additional benefit of consistency across multiple scientific papers. With the advance of similar projects, more time will be left to work on advancing the field of asteroseismology into the era of the *TESS* and *PLATO* missions.

Chapter 4

Hierarchical Bayesian Modeling

From Section 4.3 onwards, this chapter is taken almost verbatim from Hall et al. (2019), of which I was first author in the corresponding journal article. I completed the majority of the work, with exception of the detailed stellar classification, which was performed by Prof. Yvonne Elsworth. In order to retain the integrity of this study as an independent piece of work, there is some introductory material repeated from Chapter 2.

4.1 Introduction

In astronomy, the number of observations we receive are relatively few and far between. As a result, exact specifications of the models we fit to our data are especially important. If we want accurate inference of astrophysical properties from our data, we need to make sure that our model fully describes the data (and at least some degree of physics). One way of doing so is by building a *generative model* for the data.

A generative model quantitatively describes the process from which some data is generated (Hogg et al., 2010). For an astrophysical example, let's consider that we have observations of a population of stars in properties x and y, which are linearly related as

$$y(x) = mx + b\,, \tag{4.1}$$

where m and b are unknown parameters. Our observations y are subject to some random Gaussian noise, which is described by corresponding uncertainties σ_y. For simplicity, we'll assume that there are no uncertainties on x. We can then express a generative model for the data y through a likelihood function, as

$$p(y|m, b, x, \sigma_y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{[y - (mx + b)]^2}{2\sigma_y^2}\right)\,, \tag{4.2}$$

a normal distribution where m and b are the parameters in our model that we want to estimate. If we recall Bayes theorem (Eq. 1.1), we can express the posterior distribution on our parameters as

$$p(\theta|y, x, \sigma_y) \propto \prod_{i=0}^{N} p(y_i|\theta, x_i, \sigma_{y_i})\, p(\theta) \equiv \prod_{i=0}^{N} \mathcal{N}(y_i|mx_i + b, \sigma_{y_i})\, p(\theta)\,, \tag{4.3}$$

where our parameters are $\theta = \{m, b\}$, $p(\theta)$ is the prior on those parameters, and $\mathcal{N}$ indicates a normal distribution of the form $\mathcal{N}(y|\mu, \sigma)$ where μ is the mean and σ is the standard deviation. The posterior represents the product of the likelihoods and the prior for the full length of N data, where i indicates a single datum. Our understanding of our measurement process (Gaussian noise) and what we are observing (a linear relationship between two properties) has allowed us to construct a meaningful generative model for these data.

However in astrophysics there is rarely a clear relation between two physical parameters. Let's consider the more realistic case that our property y varies linearly with x, but that our observations are scattered in our observed parameter space by an additional physical property (usually, metallicity is the culprit), which affects the property y in a star. In practice, this will mean that the observations of y are spread out beyond the scatter on our observations, and that our relation $y(x)$

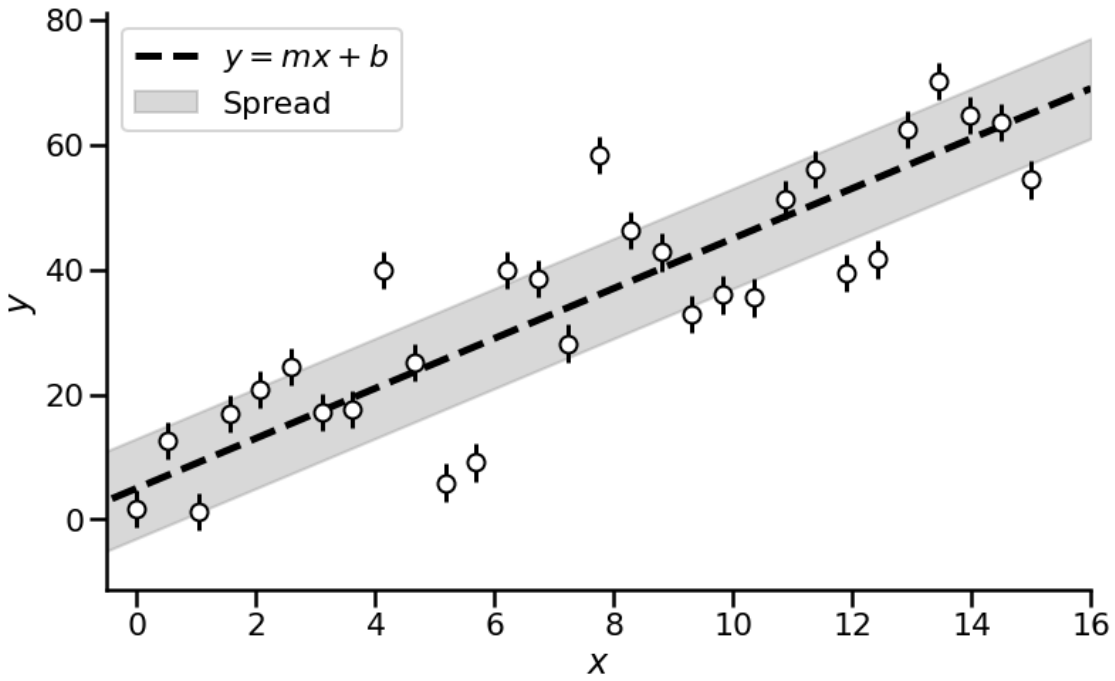

Figure 4.1: A sample of 30 simulated observations drawn from Equation 4.1, with some uncertainty on the y observations σ_y, scattered by an additional intrinsic spread that is significantly larger than the uncertainty, shown by the shaded area. The dashed black line shows the underlying relationship between x and y.

actually has some intrinsic spread with a physical origin. An example of this can be seen in Figure 4.1, where additional scatter on y is generated using Gaussian noise, representing the impact of another physical property.

Hierarchical models are an effective way of encoding these complicated relationships into a statistical model. They create associations within a group of parameters that all draw prior information from a shared distribution described by *hyperparameters*, creating a parameter hierarchy (Betancourt & Girolami, 2013). Hierarchical models are useful in a broad range of problems, and especially so in astrophysics, where it is important to take full advantage of sparse data. A situation where hierarchical models are naturally applicable is in population studies, where an observed population may have an intrinsic spread that is affected by multiple physical properties. A hierarchical approach can model population properties we are interested in as hyperparameters, from which other parameters draw their prior information.

In this Chapter, I will provide an introductory explanation of hierarchical models

and their applications to latent variable models. This Chapter is laid out as follows: Section 4.2 will provide an introduction to the concept of hierarchical latent variable modelling using an accessible example. Section 4.3 contains the published work Hall et al. (2019), where I applied a hierarchical latent variable model to the Red Clump – an overdensity of stars on the Hertzsprung-Russell diagram – to estimate the distribution of luminosities in Clump stars. Finally, Section 4.4 will draw conclusions on the applicability of hierarchical models to physical problems.

4.2 Hierarchical Latent Variable Modelling

4.2.1 A straight-line fit

Let us consider the sample described above and shown in Figure 4.1: observed data y (with uncertainty σ_y) and x (with no uncertainty), which are linearly related and subject to an intrinsic spread of physical origin. Let's say that we want to obtain constraints for the physical parameters m and b that describe the relationship between the properties x and y. We can set up a model as follows, using probabilistic notation:

$$\begin{aligned} m &\sim \mathcal{U}(0, 20)\,, \\ b &\sim \mathcal{U}(0, 20)\,, \\ y &\sim \mathcal{N}(y|mx + b, \sigma_y)\,, \end{aligned} \tag{4.4}$$

where the $\mathcal{U}$ indicates a uniform probability distribution (so our values for m and b are equally likely to be anywhere between 0 and 20). This can also be represented as a probabilistic graphical model (PGM), shown in Figure 4.2.

Fitting this model to the data shown in Figure 4.1 is unlikely to consistently find accurate values for m and b, because the intrinsic physical scatter on the data is larger than the uncertainties. We can see this by looking at the outcome of a fit

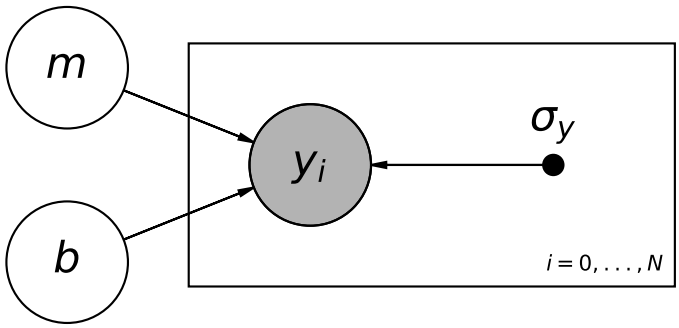

Figure 4.2: A probabilistic graphical model showing the model defined in Equation 4.4. Shaded circles show observed quantities. Filled in circles show fixed parameters that inform probability calculations, such as the uncertainty on the observations. The remaining circles indicate free parameters. The parameters m and b define the line $y = mx+b$, y_i represents y at a datum i for a total of N data, and σ_y is the uncertainty on the observations.

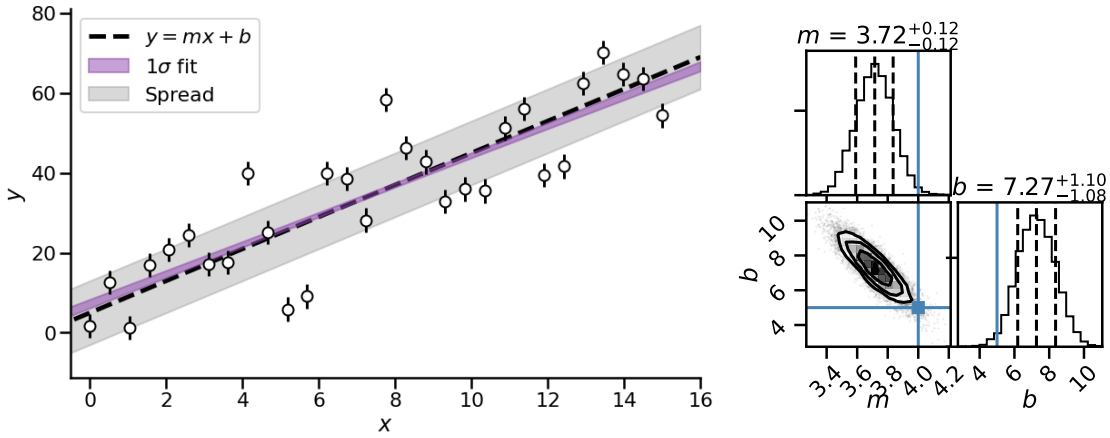

Figure 4.3: Results of fitting the model described in Equation 4.4 to the data in Figure 4.1. Left: The 1σ confidence interval on the fit relation is shown by the purple shaded area, with other conventions being the same as for Figure 4.1. Right: A 'corner plot' showing the samples of the posterior probabilities on our parameters m and b given the data. The solid blue lines indicate the truth values. From left to right, the dashed lines represent the 15.9^{th}, 50^{th} and 84.1^{st} percentiles respectively (i.e. the 1σ confidence interval and the median).

using this model, run using Hamiltonian Monte Carlo (HMC) in the Python package PyMC3, using 4 chains for 2,500 iterations each. The results are shown in Figure 4.3.

In Figure 4.3, the purple shaded area represents the 1σ confidence interval on

the fitted relation, $y = mx + b$. While it lies close to the dashed line representing the truth, it diverges from the truth at low and high x. Looking at the posterior distributions on the fit, we can see that our posterior estimates for both m and b do not fall within 1σ of the truth values, represented by the blue lines. Were we to repeat this analysis many times, the posterior results would average to the correct values, but with a vastly overestimated uncertainty due to the presence of the additional physical spread not accounted for in our model.

4.2.2 Hierarchical Latent Parameters

The differences between our fit and the truth in Figure 4.3 are due to us not accounting for the physical spread in our observations y. In other words, our model is not fully generative. One way of dealing with this is by building a model that more accurately reproduces the data, and using a *hierarchical latent variable model*.

A latent variable model relies on a change in our perspective of the problem: instead of comparing our model to data, as we did previously, we want to *draw variables from our model*, and compare *those* to data. This means that for N observations y, we want to generate N so-called latent variables Y from our model, one for each observation. Each of these latent variables Y is informed by our previous model (Eq. 4.1) defined by m and b, but also by a new parameter ρ_y, which represents the physical spread on the data. Because all the latent parameters Y are drawing common prior information from m, b and ρ_y, our model has become *hierarchical*. We can write the posterior distribution of our parameters as

$$p(\theta, Y|y, x, \sigma_y) \propto p(\theta) \prod_{i=0}^{N} \mathcal{N}(y_i|Y_i, \sigma_{y_i})\, p(Y_i|\theta)\,, \tag{4.5}$$

where our hyperparameters are now $\theta = \{m, b, \rho_y\}$, and Y are our latent variables[1]. $p(Y|\theta)$ then is the probability to obtain latent parameters Y given our hyperparameters θ, and $p(\theta)$ represents the priors on m, b and ρ_y, our hyperparam-

[1]The latent variables Y are included on the left hand side of the posterior, but since we typically only really care about θ the inclusion of Y in these notations is often omitted.

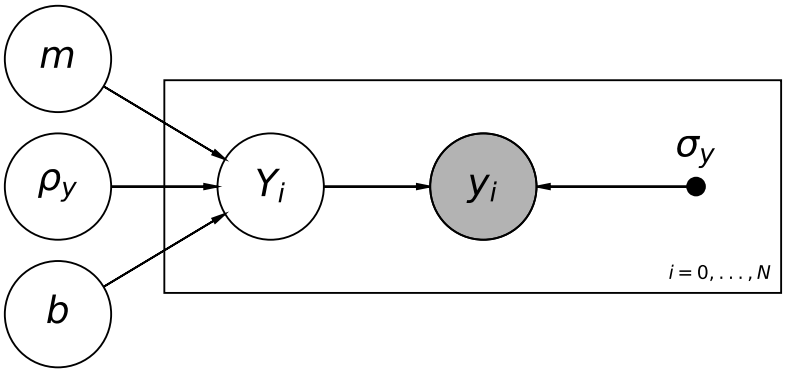

Figure 4.4: A probabilistic graphical model showing the model defined in Equation 4.6. The conventions used are the same as for Figure 4.2.

eters.

We can test this new model by performing a fit to the same data and seeing if we come closer to inferring the truth. In the same style as Equation 4.4, we can express our new model as

$$\begin{aligned} m &\sim \mathcal{U}(0, 20)\,, \\ b &\sim \mathcal{U}(0, 20)\,, \\ \rho_y &\sim \mathcal{U}(0, 20)\,, \\ Y &\sim \mathcal{N}(mx + b, \rho_y)\,, \\ y &\sim \mathcal{N}(y|Y, \sigma_y)\,, \end{aligned} \tag{4.6}$$

where we have placed an uninformative prior over ρ_y. This model is represented in a PGM shown in Figure 4.4, showing the hierarchical layout of the model.

A fit of this new model to our data, using 10,000 iterations[2] with PyMC3, is shown in Figure 4.5. The 1σ credible region has a spread that contains the truth across the full range our data occupies. Looking at the posterior distributions of the fit we can see that the truth values for m and b are contained within the 1σ credible region of the posteriors. The inferred posterior for ρ_y lies close to the truth, which

[2]Increased here from 2,500 to account for the more complicated model.

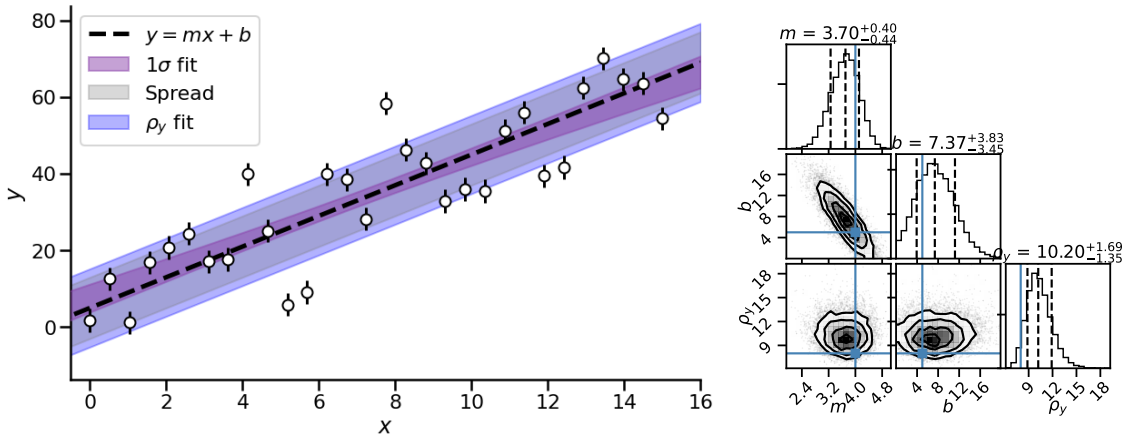

Figure 4.5: Results of fitting the model described in Equation 4.6 to the data in Figure 4.1. Conventions are the same as for Figure 4.3. The new blue shaded region indicates the measured spread μ_y, compared to the true spread (grey shaded area).

we can consider a success given the low number of data we are using to constrain this spread.

4.2.3 Increasing the data

Our hierarchical latent variable model, when applied to our set of 30 observations, does a good job at recovering m and b, but struggles to recover ρ_y. For only 30 observations, this is expected; the data y are given a physical spread following a normal distribution with a standard deviation ρ_y, and constraining the shape of a normal distribution with only 30 observations is difficult. To see how more observations would improve our inference, we can instead simulate $N = 1000$ observations, and fit our hierarchical latent variable model to them.

As seen from the posterior distributions in Figure 4.6, the inferences of m, b and ρ_y have all become significantly smaller (and therefore more precise). The spread ρ_y is still slightly offset from the truth, but within the realms of credibility, given the small spread on its posterior. While increasing the number of observations improved our inference on one physical parameter, the thought we put into our model prescription meant that we could obtain accurate inference for almost a factor of 30 fewer observations.

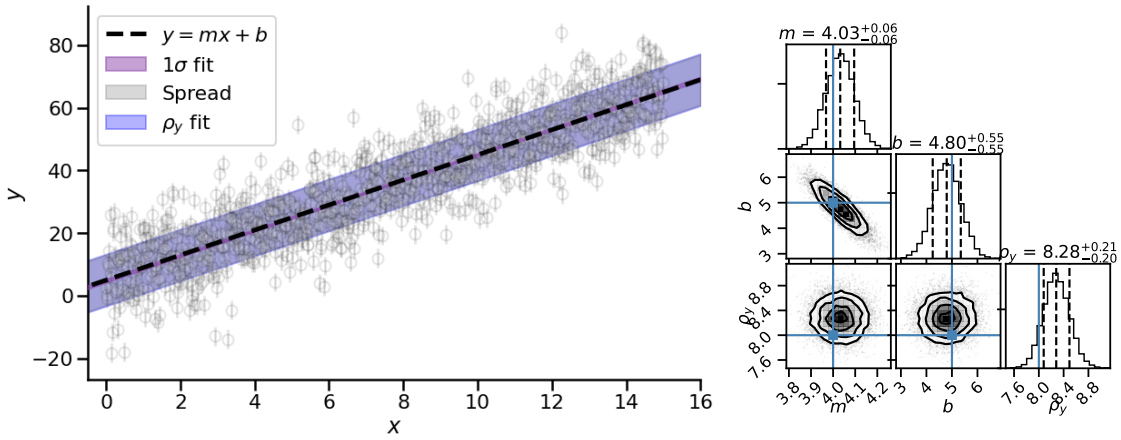

Figure 4.6: Results of fitting the model described in Equation 4.6 to a sample of 1000 simulated observations. Conventions are the same as for Figure 4.5.

4.3 Characterising the Red Clump with Asteroseismology and *Gaia*

4.3.1 Introduction

Since the launch of CoRoT (Baglin et al., 2006) and *Kepler* (Borucki et al., 2010), the use of asteroseismology — the study of stars' internal physics by observing their modes of oscillation — has become a crucial tool for testing fundamental stellar properties. The large quantity of long time series photometry from these missions (Chaplin & Miglio, 2013), and its distance independent nature, have allowed for measures of precise stellar radii of red giant stars (Hekker et al., 2011; Huber et al., 2011a, 2014; Mathur et al., 2016; Pinsonneault et al., 2014; Pinsonneault et al., 2018; Yu et al., 2018) and main sequence stars (Chaplin et al., 2010, 2011, 2014), studies of exoplanets and exoplanet hosts (Christensen-Dalsgaard et al., 2010; Batalha et al., 2011; Huber et al., 2013b,a; Chaplin et al., 2013; Silva Aguirre et al., 2015), internal and external stellar rotation (Beck et al., 2012; Deheuvels et al., 2012, 2014; Mosser et al., 2012a; Davies et al., 2015), ages of stellar populations (Miglio et al., 2009, 2013; Casagrande et al., 2014, 2016; Stello et al., 2015), and classifications of stellar types (Bedding et al., 2011; Mosser et al., 2012b, 2015; Stello et al., 2013; Vrard

et al., 2016; Elsworth et al., 2017), among others.

Many of these works rely on the so-called 'direct method' (see Chapter 2): the use of seismic scaling relations related to the two fundamental oscillation parameters, $\nu_{\rm max}$, the frequency of maximum power of the oscillation mode envelope, and $\Delta\nu$, the spacing between two oscillation modes of equal radial degree. These properties are individually proportional to mass, radius and temperature, and when combined and scaled with solar values, can provide measures of stellar mass, radius and surface gravity (Kjeldsen & Bedding, 1995). As such, stellar properties obtained through seismology depend on temperature as well as on the seismic parameters. Besides the direct method, results from seismology can also be obtained by comparing global seismic properties with a grid of models, referred to as 'grid modelling', and can be expanded to 'detailed modelling', which directly fits observed seismic mode frequencies to the grids (Metcalfe et al., 2012; Metcalfe et al., 2014; Silva Aguirre et al., 2013, 2015, 2017).

The seismic scaling relations have been thoroughly tested through interferometry (White et al., 2013), astrometry (Huber et al., 2017), eclipsing binaries (Gaulme et al., 2016), and open clusters (Miglio et al., 2012). Theoretically motivated corrections to the $\Delta\nu$ and $\nu_{\rm max}$ scaling relations have been proposed to depend on $T_{\rm eff}$, metallicity, and evolutionary state (Miglio et al., 2012; Sharma et al., 2016).

Recently, Brogaard et al. (2018) used precise measurements of masses and radii of eclipsing binary stars to compare to the same properties obtained through the seismic scaling relations as well as grid modelling, and found that not including a correction to the scaling relation for $\Delta\nu$ overestimated stellar masses by between $\sim$ 11 and 19%. Brogaard et al. (2018) also found that using corrections obtained through grid modelling by Rodrigues et al. (2017) provided stellar masses and radii on average slightly smaller than those using corrections by Sharma et al. (2016) obtained similarly. Crucially, no need was found for a correction to the $\nu_{\rm max}$ scaling relation, although it is known that a small correction for the mean molecular weight

could be needed for this relation (Belkacem et al., 2013; Viani et al., 2017).

When using the direct method, effective temperatures from spectroscopic analysis are often used (e.g. the APOKASC catalogue, Pinsonneault et al., 2014; Pinsonneault et al., 2018). However depending on the atmospheric models and temperature scales applied in spectroscopic analysis, inferred values for $T_{\rm eff}$ can vary up to $\sim 170\,K$ for Core Helium-Burning (CHeB) stars (Slumstrup et al., 2019). While Bellinger et al. (2019) have recently shown that these systematic uncertainties can be mitigated through the use of grid modelling for main-sequence and sub-giant stars, the question of which temperature scale for spectroscopy obtains the best value for $T_{\rm eff}$ remains open.

Seismic observations can be combined with distance dependent observations, such as astrometry, to improve and calibrate results. The second data release (DR2) of the astrometric *Gaia* mission (Gaia Collaboration et al., 2018) recently has provided data for a sample of over one billion targets, with uncertainties largely improved from the first data release (DR1, TGAS Gaia Collaboration et al., 2016), allowing for a broader range of science and calibrations to be done (Zinn et al., 2019a). With DR2 Lindegren et al. (2018) suggested a mean global parallax zero-point offset of $-29\,\mu$as, in the sense that *Gaia* parallaxes are too small, using a quasar sample, although it should be noted that the offset varies as a function of colour, magnitude and position on the sky. Arenou et al. (2018) computed the parallax difference between DR2 and existing catalogues, as well as prior data for individual targets, and found these on average to be the same order of magnitude as the Lindegren et al. (2018) zero-point. Riess et al. (2018) used Cepheid variables to derive a zero-point offset of $-46 \pm 13\,\mu$as, Stassun & Torres (2018) used Eclipsing Binaries to find a zero-point of $-83 \pm 33\,\mu$as, and Zinn et al. (2019a) compared parallaxes to seismic radii to identify a colour- and magnitude-dependent offset of $-52.8 \pm 2.4(stat.) \pm 1(syst.)\,\mu$as for red giant branch stars in the *Kepler* field. Finally,

using analysis of individual seismic mode frequencies for 93 dwarf stars, Sahlholdt & Silva Aguirre (2018) reported an offset in estimated stellar radii equal to a parallax offset of $-35 \pm 16\,\mu$as. As the parallax zero-point offset is known to vary with magnitude, colour, and position in the sky, the differences between these values for the zero-point are expected. Understanding how we quantify the offset is crucial if we want to use *Gaia* to calibrate asteroseismology and other methods.

One method of testing independent sets of measurements is calculating a well known astronomical property. An example of such a property is the luminosity of the 'Red Clump' (RC), an overdensity of red giant stars on the HR-diagram, in bands of absolute magnitude. Because they form a clear feature on the HR-diagram, so-called 'Red Clump' stars are commonly studied in astronomy, both inside and outside the Milky Way (for a historical overview see Girardi, 2016).

When a star runs out of hydrogen to burn in its core, marking the end of the main sequence, the core will begin to shrink. At the same time, the shell surrounding the core will begin hydrogen fusion, expanding the outer envelope and evolving the star on to the red giant branch. For stars of masses around $0.7 \lesssim M/M_{\odot} \lesssim 1.9$[3], the inactive core will compress enough to become degenerate. As the core accretes 'ash' from the hydrogen burning shell and compresses further, its mass and temperature will gradually increase. At a core mass of $\approx 0.33 M_{\odot}$, its temperature becomes high enough to ignite helium fusion, and the degeneracy is lifted due to the increased thermal pressure. This event is referred to as the Helium flash, and initiates the core helium-burning phase of a star's life (Sweigart et al., 1990; Girardi et al., 2013).

Because helium ignition begins at roughly the same core mass for all red giant stars with fully degenerate cores, the masses of their helium burning cores are extremely similar. As their luminosity is mainly determined by the core mass, they will all have similar luminosities, creating the RC overdensity of stars on the HR-diagram

[3]For [M/H] $\simeq 0.07$, upper limit subject to change with metallicity.

(Girardi, 2016, and references therein).

For RC stars, further differences in luminosity and temperature are mainly effects of metallicity and envelope mass, and thus the Clump has a relatively small spread. Stars at lower masses and low metallicities form a horizontal branch at a luminosity similar to the RC. Stars of slightly higher masses will form partially degenerate cores on the main sequence, and ignite helium fusion at slightly lower luminosities, forming a Secondary Red Clump (2CL, Girardi, 1999). At even higher masses the cores do not experience any degeneracy before reaching the temperature required for helium fusion. In the core helium-burning phases of these massive stars, their luminosity becomes a function of stellar mass, and so they form a vertical structure in the HR-diagram during their CHeB phase.

The luminosity of the RC overdensity may be used as a standard candle given constraints on mass and metallicity (Cannon, 1970), and has recently been used to calibrate *Gaia* DR1 parallaxes (Davies et al., 2017). Also using *Gaia* DR1 parallaxes, Hawkins et al. (2017) (hereafter H17) found precise measurements for the RC luminosity in various passbands, including the 2MASS (Skrutskie et al., 2006) K band, which is expected to minimise the spread in luminosity due to mass and metallicity (Salaris & Girardi, 2002). With *Gaia* DR2's improved parallax uncertainties and reduced systematic offset, now is a good time to revisit the RC as a calibrator.

In this work we investigate systematics in both asteroseismology and *Gaia simultaneously*, to see how differences in assumptions for one influence inferences of the other. Using a sample of over 5500 *Kepler* Red Giant stars in the RC for which parallaxes and seismology are available, we measure the position of the RC population in absolute magnitude in the 2MASS K band, and the *Gaia* G band. We do this using seismology and parallax (with photometry) independently. Since the distribution of RC stars should be the same for this population, independent of

method, a (dis)agreement of the measured positions and spreads of the RC using two independent methods sheds light on systematics in both. For the seismic method, we test the influence of the temperature scale used to obtain the values of $T_{\rm eff}$ fed into seismic scaling relations, as well as the impact of corrections to the $\Delta\nu$ scaling relation. For the *Gaia* method, we study how changes in the parallax zero-point offset for *Gaia* DR2 impact the inferred luminosity of the RC.

The rest of this Chapter is laid out as follows: Section 4.3.2 discusses how the data were obtained, and the theory used to calculate our observables. Section 4.3.3 discusses how we use hierarchical Bayesian modelling to study the RC. We present our results in Section 4.3.4 and discuss them in context of similar work in Section 4.3.5, and present our conclusions, both for this paper and the chapter, in Section 4.4.

4.3.2 Data

Our aim is to find the intrinsic position and spread of the Red Clump in absolute magnitude for various passbands using two approaches: one using a distance-independent luminosity calculated from asteroseismology, and the other using a magnitude inferred from photometry and *Gaia* DR2 parallaxes. Since the number of stars with asteroseismic data is significantly lower than those with data in *Gaia* DR2, this limits our sample.

For our asteroseismic sample, we used the catalogue of 16,094 oscillating *Kepler* red giants by Yu et al. (2018) (hereafter Y18), which contains global oscillation parameters $\nu_{\rm max}$ and $\Delta\nu$, as well as broad evolutionary state classifications, effective temperatures $T_{\rm eff}$ and metallicities [Fe/H] taken from Mathur et al. (2017).

We re-considered the classification of all stars labelled as CHeB in the Y18 catalogue using the method presented in Elsworth et al. (2017). This uses the structure of dipole-mode oscillations in the power spectra to classify stars as belonging to the 2CL, the Red Giant Branch (RGB), or the RC. We obtained light curves for 7437

stars labelled as CHeB in Y18, from two sources: the so-called KASOC light curves (Handberg & Lund, 2014)[4] and the KEPSEISMIC light curves (García et al., 2011)[5]. The latter have been produced with larger photometric masks to ensure a better stability at low frequencies, and have been gap-filled using in-painting techniques (García et al., 2014; Jofré et al., 2015).

Of these 7437 stars, we found that 5668 are RC, 737 are 2CL, and that no classification could be found for 499 stars. Notably, 533 stars were found to be RGB, disagreeing with the classification listed in Y18. This should be discussed in future work, but for the sake of internal consistency we have chosen to adopt Elsworth et al. (2017) classification in this work.

It should be noted that our classification does not specifically account for low-mass, low-metallicity horizontal branch stars, which are therefore expected to be retained in our sample, but are not expected to significantly affect the result as they have similar luminosities to the RC, and no extensive horizontal structure is present on the HR-diagram of the Y18 catalogue, or our subsample thereof (see Figure 4.7). A fraction of the newly classified stars had masses reported in Y18 as much higher than we would expect for a RC star. In order to exclude these from our sample, we apply a liberal cut for clump-corrected seismic masses of over $2.2\,M_{\odot}$, excluding 92 stars from our sample.

To obtain our astrometric sample, we cross-matched the RC stars we selected from the Y18 sample with the *Gaia* DR2 sample[6] (Gaia Collaboration et al., 2016, 2018). In cases of duplicate sources for a given KIC, we selected the star with the lowest angular separation to the target. We did not apply any truncation of the sample based on parallax uncertainty or negative parallax, since this is known to introduce a parallax dependent bias (Luri et al., 2018).

[4]Freely distributed at the KASOC webpage (http://kasoc.phys.au.dk)

[5]Freely distributed at the MAST website (https://archive.stsci.edu/prepds/kepseismic/)

[6]We make use of the of the https://www.gaia-kepler.fun crossmatch database created by Megan Bedell for this purpose.

The parallaxes ($\hat{\varpi}$) and parallax uncertainties ($\sigma_{\hat{\varpi}}$) make up our astrometric set of observables. We obtained the apparent magnitudes ($\hat{m}$) and their uncertainties ($\sigma_{\hat{m}}$) from the 2MASS survey for the K band and *Gaia* DR2 for the *Gaia* G band, and removed stars that do not have photometry or uncertainties on magnitude in 2MASS. Comparing the magnitude zero-points for the *Gaia* G, $G_{\rm BP}$ and $G_{\rm RP}$ bands, Casagrande & VandenBerg (2018b) found indication of a magnitude-dependent zero-point offset in the *Gaia* G band magnitudes in the range of $6\,{\rm mag} \lesssim {\rm G} \lesssim 16.5\,{\rm mag}$, corrected as

$$G^{\rm corr} = 0.0505 + 0.9966\,G\,, \tag{4.7}$$

where G is our uncorrected *Gaia* G band magnitude. This correction is small, and corresponds to 30 mmag over 10 magnitudes. We gave all our G band magnitudes a generous uncertainty of 10 mmag, the typical uncertainty quoted in Gaia Collaboration et al. (2018) for $G = 20$, in order to account for any additional uncertainty incurred by the above correction. It should be noted that a similar relation for the correction of G band magnitudes is presented in Maíz Apellániz & Weiler (2018). This correction places magnitudes about 30 mmag higher than when using the Casagrande & VandenBerg (2018b) correction in the applicable magnitude range. We expect the scale of this systematic offset to have a negligible impact on our results, and therefore adopt the Casagrande & VandenBerg (2018b) correction in this work for consistency with our chosen G band extinction coefficients and bolometric corrections (see below).

Our model also uses an extinction for each star in each band. Reddening values are taken from the Green et al. (2018) three-dimensional dustmap under the assumption that the distance to the object is that given by Bailer-Jones et al. (2018). We note that this is not expected to bias our results towards a previous measure of distance, because the spread in the obtained reddening values, regardless of choice of distance value, falls well within the spread of the prior set on these values in our

model (see Section 4.3.3). We converted reddening to the band-specific extinction $\hat{A}_\lambda$ using extinction coefficients unique to the Green et al. (2018) map for the K band[7]. For the *Gaia* G-band we calculated our band-specific extinction using the mean extinction coefficient presented in Casagrande & VandenBerg (2018b), after converting our reddening values to a measure of $E(B-V)$ following the conventions presented in Green et al. (2018).

The final sample contains 5576 RC stars, with minimal contamination from the 2CL or the RGB, and covers a magnitude range of ~ 8 to ~ 16 mag in G and ~ 6 to ~ 14 mag in K. Note that for this magnitude range we expect the *Gaia* DR2 catalogue to be practically complete, and do not need to apply any selection functions in magnitude. The data are shown in Figure 4.7 in a HR-diagram overlaid on the full Y18 sample.

The APOKASC-2 subsample

We used temperatures from Mathur et al. (2017), a catalogue compiling temperatures from a diverse set of papers including work with spectroscopy, photometry, and some asteroseismology. In order to investigate the impact of using differing temperature sources on our results, we also included runs on a subsample of 1637 stars that had $T_{\rm eff}$ values reported in the APOKASC-2 catalogue (Pinsonneault et al., 2014; Pinsonneault et al., 2018). When calculating seismic properties from these data, we only changed the values for $T_{\rm eff}$ to our new APOKASC-2 values. In Figure 4.8 we compare the distributions in $T_{\rm eff}$, mass, radius and [Fe/H] of the Y18 RC sample and the APOKASC-2 subsample. Also shown is the distribution of the APOKASC-2 temperatures, which are overall lower than the Y18 temperatures, and the distributions in mass and radius calculated through the direct method for these temperatures. Overall the APOKASC-2 subsample represents a lower temperature population, with its most distinct difference being in $T_{\rm eff}$ and [Fe/H].

[7]These coefficients can be found with the Green et al. (2018) usage notes.

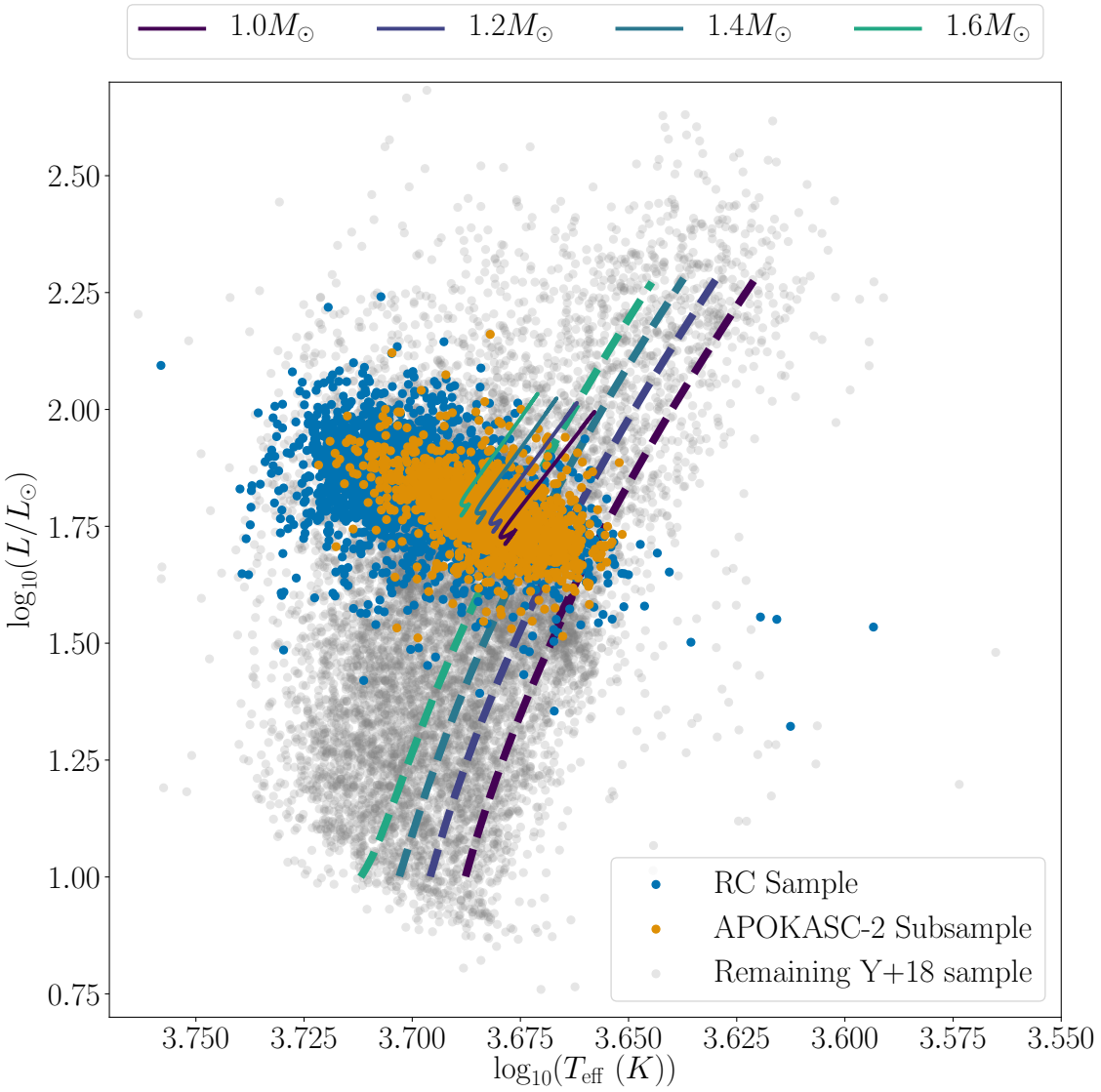

Figure 4.7: HR-diagram illustrating the data in our final set of 5576 stars overlaid on the Y18 sample, along with evolutionary tracks from MESA (Paxton et al., 2011, 2013, 2015)(for details about the physical inputs of the models see Khan et al., 2018). The stars in the Y18 sample not in our final selection are in grey. Plotted on top in blue are the stars that in our final sample, and the subsample of stars with temperatures reported in APOKASC-2 (Pinsonneault et al., 2018) is shown in orange. Evolutionary tracks are plotted for for masses ranging between (from right to left) 1.0 and 1.6 solar masses for a metallicity of $Z = 0.01108$ and helium content of $Y = 0.25971$. The dashed lines indicate the Red Giant Branch, whereas the solid lines indicate the main Core Helium Burning stage of the tracks (the Helium flash and sub-flashes are not included).

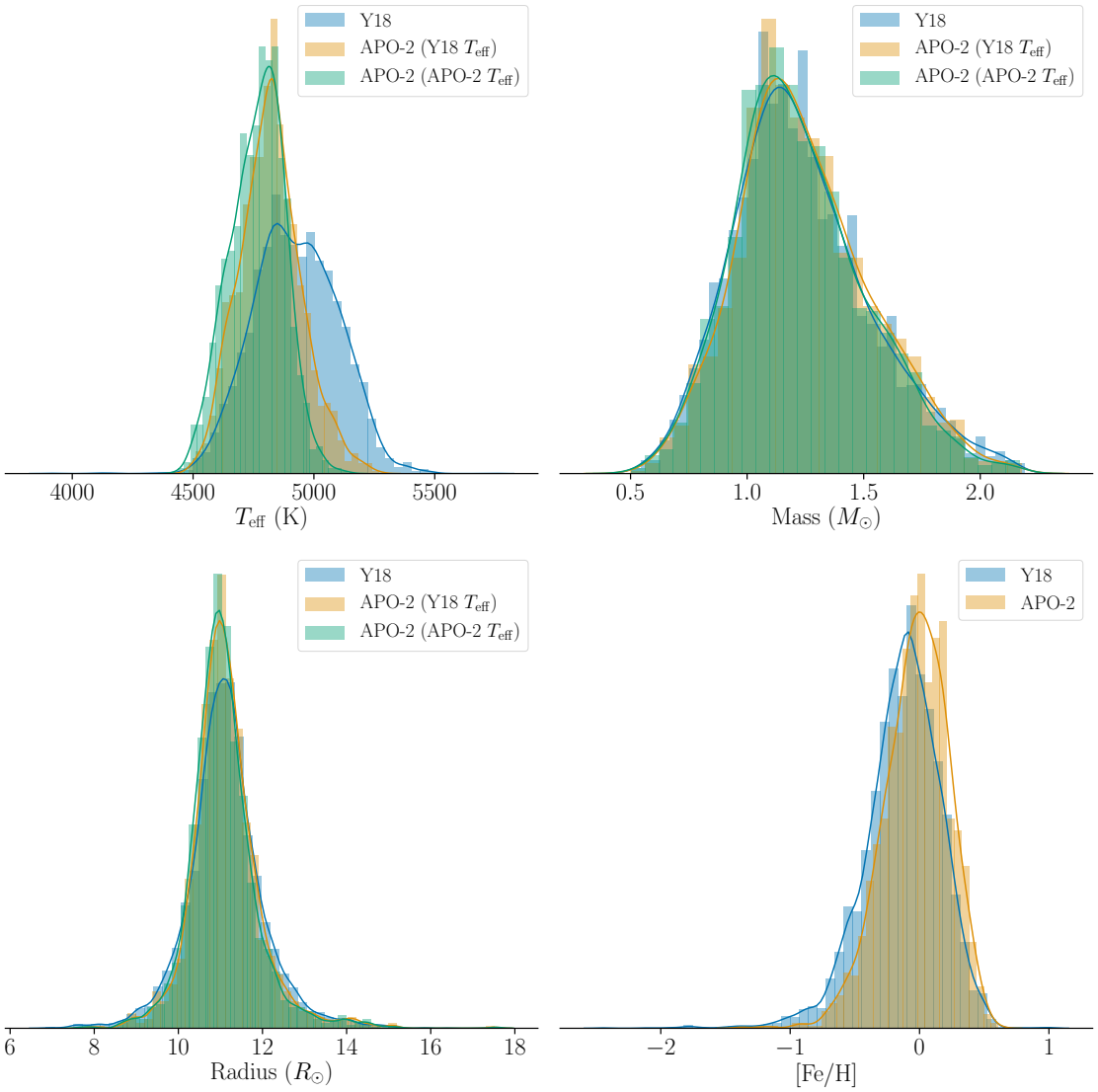

Figure 4.8: Distributions in $T_{\rm eff}$, mass, radius and [Fe/H] of the RC sample (Yu et al., 2018) and the APOKASC-2 subsample (Pinsonneault et al., 2014; Pinsonneault et al., 2018). In green are the distributions of the APOKASC-2 temperatures, which are overall lower, and the distributions in mass and radius calculated through the direct method for these temperatures. In the labels, 'APO-2' is a shorthand for APOKASC-2.

Obtaining the seismic sample

The two global observable seismic parameters, $\nu_{\rm max}$ and $\Delta\nu$, scale with fundamental stellar properties as (Brown et al., 1991; Kjeldsen & Bedding, 1995):

$$\frac{\nu_{\rm max}}{\nu_{\rm max\odot}} \simeq \left(\frac{M}{M_\odot}\right)\left(\frac{R}{R_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{-1/2} \text{ and} \tag{4.8}$$

$$\frac{\Delta\nu}{\Delta\nu_\odot} \simeq \left(\frac{M}{M_\odot}\right)^{1/2}\left(\frac{R}{R_\odot}\right)^{-3/2}, \tag{4.9}$$

where M is the stellar mass, R is the radius, $T_{\rm eff}$ is the effective temperature, and $\odot$ indicates a solar value. In this work we used $\nu_{\rm max\odot} = 3090 \pm 30\,\mu$Hz, $\Delta\nu_\odot = 135.1 \pm 0.1\,\mu$Hz and $T_{\rm eff\odot} = 5777$ K (Huber et al., 2011a). By rearranging these scaling relations, we can obtain stellar surface gravity and radius as

$$\frac{g}{g_\odot} \simeq \frac{\nu_{\rm max}}{\nu_{\rm max\odot}}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{1/2} \text{ and} \tag{4.10}$$

$$\frac{R}{R_\odot} \simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max\odot}}\right)\left(\frac{\Delta\nu}{f_{\Delta\nu}\Delta\nu_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{1/2}, \tag{4.11}$$

where the new term $f_{\Delta\nu}$ is a correction to the $\Delta\nu$ scaling relation in the notation of Sharma et al. (2016).

We calculated $f_{\Delta\nu}$ as a function of [Fe/H], $T_{\rm eff}$, $\nu_{\rm max}$, $\Delta\nu$ and evolutionary state using interpolation in a grid of models (Sharma & Stello, 2016). As part of our analysis, perturbations were made to values of effective temperature (see Section 4.3.3). For each perturbation of $T_{\rm eff}$ we recalculated $f_{\Delta\nu}$, changing no other parameters. We only extracted the correction values $f_{\Delta\nu}$ from the models, and used the seismic parameters and temperature values from our original set, and not the results for these values returned from the grids, in the rest of this work. We did not include corrections for the $\nu_{\rm max}$ scaling relation, because these are more difficult to obtain theoretically (Belkacem et al., 2011), and are probably negligible (Brogaard et al.,

2018). Note that Brogaard et al. (2018) found that using corrections by Rodrigues et al. (2017) delivers on average slightly smaller stellar properties than using Sharma & Stello (2016) due to differences in how these methods treat the solar surface effect. Since we used a wide range of bolometric corrections for various temperature perturbations, the method by Rodrigues et al. (2017) would be too computationally expensive, and we thus elected to use Sharma & Stello (2016), which may lead to differences of the order of $\sim 2\%$ in radius than if we had used Rodrigues et al. (2017) (White et al., 2011). We discuss the impact of this on our work in Section 4.3.5.

In order to obtain absolute magnitudes for our sample, we used $T_{\rm eff}$ and seismic radii, calculated using $\Delta\nu$ and $\nu_{\rm max}$ from the Y18 catalogue through Equation 4.11, to calculate the stellar luminosity as

$$L_* = 4\pi\sigma_{\rm sb} R^2 T_{\rm eff}^4 \,. \tag{4.12}$$

Here L_* is the luminosity of the star and $\sigma_{\rm sb}$ is the Stefan-Boltzmann constant. This was converted to a bolometric magnitude as in Casagrande & VandenBerg (2014),

$$M_{\rm bol} = -2.5 \log_{10}(L_*/L_\odot) + M_{\rm bol\odot} \,, \tag{4.13}$$

where $L_\odot$ is the solar luminosity, and we have adopted $M_{\rm bol\odot} = 4.75$ (Casagrande & VandenBerg, 2014, 2018a,b). We calculated the bolometric correction (BC) in the 2MASS K and *Gaia* G bands with the method described by Casagrande & VandenBerg (2014, 2018a,b) using $T_{\rm eff}$, [Fe/H] and $\log g$, and without accounting for extinction. Since we are using a distance-independent measure of luminosity to calculate an absolute magnitude, accounting for extinction in the BC would bias our results. Because our method requires tweaking our values for $T_{\rm eff}$, we recalculated the $\log g$ used to find the BC through the scaling relation in equation (4.10), as well as our values for $f_{\Delta\nu}$, for each different set of temperatures, and thus obtained a full set of bolometric corrections and corrections to the scaling relations for each

temperature perturbation. Our values of absolute magnitude were then given by

$$\hat{M}_\lambda = M_{\rm bol} - BC_\lambda \,, \tag{4.14}$$

where λ is the relevant band, $M_{\rm bol}$ is the bolometric luminosity and BC_λ is the bolometric correction in that band. Uncertainties on $\hat{M}_\lambda$ were propagated through from the uncertainties on seismic parameters and effective temperatures, including those on the solar seismic parameters. Uncertainties on the BCs were estimated using a Monte Carlo method with 5000 iterations for 1000 randomly selected stars from our sample. We found an uncertainty of 0.3 mag for all BCs in the G band. For the K band we found 0.05 mag for stars with a fractional temperature uncertainty of $< 2.5\%$, and 0.09 mag for those with larger fractional uncertainties on temperature. We discuss the systematic uncertainties on $f_{\Delta\nu}$ in Section 4.3.5.

4.3.3 Locating the Red Clump using hierarchical Bayesian modelling

In order to test systematics in asteroseismology and *Gaia* using the Red Clump (RC), we aimed to find the location and spread of the RC in absolute magnitude using both sets of data separately. To obtain these RC parameters, we fitted a model for the distribution of RC stars in 'true' absolute magnitude, either inferred from an observed absolute magnitude (asteroseismic) or inferred from apparent magnitude, parallax and extinction (astrometric).

We built a pair of Bayesian hierarchical models with latent parameters that allowed us to infer key values such as the distance and the true absolute magnitude from the data and the model. The latent parameters form a stepping stone between our population model, which is described by *hyperparameters*, and the observations. We used a latent parameter for each star to infer the 'true' distribution of the absolute magnitudes, while fitting our population level model to these inferred 'true' absolute magnitudes, instead of to the observations themselves. Many aspects of

our hierarchical models, especially those for the *Gaia* data, are similar to those used for the same purpose by H17 with some improvements.

To fit to the position and spread of RC stars while also isolating any outlier contaminants, we applied the mixture model (Hogg et al., 2010) utilised by H17. In a mixture model, instead of parameters being drawn from a single probability distribution, they are drawn from a linear combination of two separate probability distributions, modulated by a mixture-model weighting factor Q. This allows for the construction of more complex distributions (i.e. to account for outliers, as done here), or for the comparison between two competing distributions (see Chapter 5). For our mixture model we used two normal distributions: one for the inlier population of RC stars, with a mean $\mu_{\rm RC}$ and a standard deviation (spread) $\sigma_{\rm RC}$, and a broad outlier distribution centred in the same location ($\mu_{\rm RC}$) but with a spread of σ_o, which must always be larger than $\sigma_{\rm RC}$. The likelihood to obtain an absolute magnitude M_i given this mixture model is then

$$p(M_i|\theta_{\rm RC}) = Q\mathcal{N}(M_i|\mu_{\rm RC}, \sigma_{\rm RC}) + (1-Q)\mathcal{N}(M_i|\mu_{\rm RC}, \sigma_{\rm o})\,, \tag{4.15}$$

where M_i is the true absolute magnitude for a given datum i, $\theta_{RC} = \{\mu_{\rm RC}, \sigma_{\rm RC}, Q, \sigma_{\rm o}\}$ are the model hyperparameters (which inform the population of latent parameters) and $\mathcal{N}(x|\mu, \sigma)$ represents a normal distribution evaluated at x, with a mean μ and a spread σ.[8]

The asteroseismic model

For our asteroseismic model, we used a calculated measure of the absolute magnitude ($\hat{M}$) from asteroseismology, along with appropriate uncertainties ($\sigma_{\hat{M}}$) as our data. We used a latent variable model to infer the true value of the absolute magnitude. Given our data and the hyperparameters on our mixture model $\theta_{\rm RC}$, we can use Bayes' theorem to find the unnormalised posterior probability of our model:

[8]Note that the spread σ as listed in $\mathcal{N}(x|\mu, \sigma)$ is not a variance, but a standard deviation, since we are following the nomenclature used in `PyStan`.

$$p(\theta_{\rm RC}, M|\mathcal{D}) \propto p(\theta_{\rm RC}) \prod_{i=1}^{\rm N} p(\mathcal{D}_i|M_i) p(M_i|\theta_{\rm RC})\,. \tag{4.16}$$

Here, N is the number of points in our data set $\mathcal{D} = \{\hat{M}, \sigma_{\hat{M}}\}$, $p(\mathcal{D}_i|M_i)$ is our likelihood function, $p(\theta_{\rm RC})$ represents the priors on the hyperparameters, and $p(M_i|\theta_{\rm RC})$ is the probability to obtain our latent parameters (the true absolute magnitudes) given our hyperparameters.

The likelihood to obtain our data given our parameters is then

$$p(\mathcal{D}_i|M_i) = \mathcal{N}(\hat{M}_i|M_i, \sigma_{\hat{M}_i})\,, \tag{4.17}$$

where M_i is the true absolute magnitude. Here, M_i is a latent parameter that is drawn from from the likelihood function $p(M_i|\theta_{\rm RC})$ (Equation 4.15), to which our hyperparameters are fit. A PGM of the asteroseismic model is shown in Figure 4.9.

The astrometric model

Fitting the absolute magnitude for the *Gaia* DR2 sample required a more involved approach, since we wanted to work directly with parallax to treat its uncertainties appropriately (Luri et al., 2018). We used a set of three latent parameters, $\alpha_i = \{M_i, r_i, A_i\}$, where M_i is the absolute magnitude in a given band, r_i is the distance and A_i is the extinction in a given band. We also include two additional hyperparameters: $\varpi_{\rm zp}$, the parallax zero-point offset and L, the length scale of the exponentially decreasing space density prior on distance (Astraatmadja & Bailer-Jones, 2016a,b, 2017). This prior, which is necessary to treat negative parallax values, has already successfully been applied to *Gaia* DR2 data (Bailer-Jones et al., 2018) and its use is recommended for this purpose within the *Gaia* DR2 release papers (Luri et al., 2018).

Some extra care was also required in the treatment of parallax uncertainties for

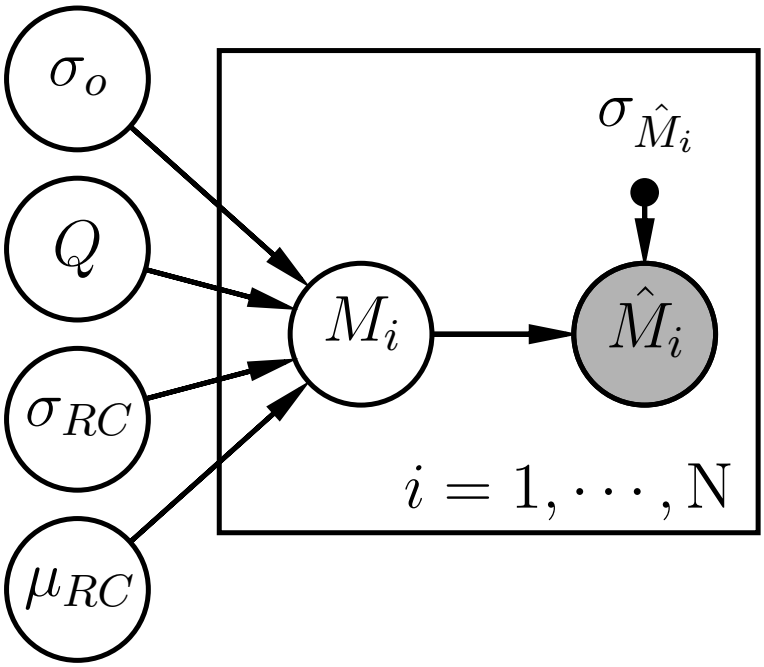

Figure 4.9: A probabilistic graphical model of the asteroseismic model, represented algebraically in Equation 4.16. Shaded circles indicate observed data, whereas solid black circles represent fixed parameters, such as the uncertainty on the observed data. The hyperparameters θ_{RC} can be seen on the left, and inform the set of latent parameters M_i, which in turn relate to the observed data $\hat{M}_i$ and $\sigma_{\hat{M}_i}$. N is the number of data points in our sample. The model structure in this figure is similar to our previous example of a line in Section 4.2, see Figures 4.2 and 4.4.

this sample. Lindegren et al. (2018) found parallaxes to be correlated on scales below 40°, with increasing strength at smaller separations, and quantified their covariance using quasar parallaxes. They found the positive covariance V_ϖ for these scales to be reasonably approximated by the fitted relation

$$V_\varpi(\theta) \simeq (285\,\mu\text{as}^2) \times \exp(-\theta/14^\circ)\,, \tag{4.18}$$

where θ is the angular separation between two targets in degrees. The fit corresponds to a RMS amplitude of $\sqrt{285\,\mu\text{as}^2} \approx 17\,\mu$as. This relation was recently applied by Zinn et al. (2019a), who found that the Lindegren et al. (2018) relation resulted in the best goodness-of-fit for their models of the parallax zero-point offset, over both a similar relation by Zinn et al. (2019a) based on TGAS data, and not including parallax covariances altogether.

We generated a covariance matrix $\underline{\Sigma}$ for our sample:

$$\Sigma_{ij} = V_\varpi(\theta_{ij}) + \delta_{ij}\sigma_{\hat{\varpi}_i}\sigma_{\hat{\varpi}_j}\,, \tag{4.19}$$

where θ_{ij} is the angular separation between stars i and j, and δ_{ij} is the Kronecker delta function.

Given these new additions, our set of data was $\mathcal{D} = \{\hat{\varpi}, \underline{\Sigma}, \hat{m}, \sigma_{\hat{m}}, \hat{A}\}$, where all symbols are as defined above and $\hat{A}$ is the band specific extinction. We can use Bayes' theorem, as before, to find the unnormalised posterior probability of our model as

$$p(\theta_{\text{RC}}, \varpi_{\text{zp}}, L, \alpha|\mathcal{D}) \propto p(\theta_{\text{RC}}, \varpi_{\text{zp}}, L, \alpha)\, p(\mathcal{D}|\theta_{\text{RC}}, \varpi_{\text{zp}}, L, \alpha)\,, \tag{4.20}$$

where $p(\mathcal{D}|\theta_{\text{RC}}, \varpi_{\text{zp}}, L, \alpha)$ is now our likelihood function and $p(\theta_{\text{RC}}, \varpi_{\text{zp}}, L, \alpha)$ represents the priors on our hyper- and latent parameters. Our likelihood function relates to two observables as,

$$p(\mathcal{D}|\theta_{\rm RC}, \varpi_{\rm zp}, L, \alpha) = p(\hat{\varpi}|r, \varpi_{\rm zp}, \underline{\Sigma}) \times p(\hat{m}|\alpha, \sigma_{\hat{m}})\,. \tag{4.21}$$

Note that the parallax only depends on the latent parameter for distance, r. Since parallax values are correlated, $p(\hat{\varpi}|r, \varpi_{\rm zp}, \underline{\Sigma})$ was evaluated for all data simultaneously, whereas $p(\hat{m}|\alpha, \sigma_{\hat{m}})$ was evaluated at every datum i. This means that our full posterior probability takes the form

$$\begin{aligned} &p(\theta_{\rm RC}, \varpi_{\rm zp}, L, \alpha|\mathcal{D}) \\ &\quad\propto p(\theta_{\rm RC}, \varpi_{\rm zp}, L)\, p(\hat{\varpi}|r, \varpi_{\rm zp}, \underline{\Sigma}) \;\times \prod_{i=1}^{\rm N} p(\hat{m}_i|\alpha_i, \sigma_{\hat{m}_i})\, p(\alpha_i|\theta_{\rm RC}, \varpi_{\rm zp}, L)\,, \end{aligned} \tag{4.22}$$

where the first term represents the priors on our hyperparameters, the second term is the likelihood to obtain our observed parallaxes, the third is the likelihood to obtain an observed magnitude, and the fourth gives the probability to obtain the latent parameters, given the hyperparameters.

The second component of Equation 4.22 is the probability of obtaining the observed parallax given our latent parameters and our covariance matrix. Since we treated our parallax uncertainties as correlated, we evaluated these probabilities for the full set using a multivariate normal distribution:

$$p(\hat{\varpi}|r, \varpi_{\rm zp}, \underline{\Sigma}) = \mathcal{N}(\hat{\varpi}|1/r + \varpi_{\rm zp}, \underline{\Sigma})\,, \tag{4.23}$$

where $1/r$ defines the *true* parallax. The latent parameters for the distance r_i were drawn from an exponentially decreasing space density prior (Bailer-Jones, 2015), which goes as

$$p(r_i|L) = \frac{1}{2L^3} r_i^2 \exp(-r_i/L)\,, \tag{4.24}$$

and thus depends on the length scale hyperparameter L. This prior has a mode at $2L$, beyond which it decreases exponentially.

The third component of Equation 4.22 is then

$$p(\hat{m}_i|\alpha_i, \sigma_{\hat{m}_i}) = \mathcal{N}(\hat{m}_i|m_i, \sigma_{\hat{m}_i}) \,, \tag{4.25}$$

where m_i is the *true* apparent magnitude, and is drawn from the relation

$$m_i = M_i + 5\log_{10}(r_i) - 5 + A_i \,. \tag{4.26}$$

Here, we have used the inferred true values for absolute magnitude, distance and extinction to calculate apparent magnitude. As for the seismic method, the true absolute magnitude M_i was drawn from the likelihood $p(M_i|\theta_{RC})$, as given in Equation 4.15. The final latent parameter A_i is given a prior as

$$p(A_i|\hat{A}_i) = \mathcal{N}(A_i|\hat{A}_i, 0.05) \,, \tag{4.27}$$

a normal distribution with a spread of 0.05 mag, where $\hat{A}_i$ is our observed value for the extinction (Green et al., 2018). A PGM of the astrometric model is shown in Figure 4.10.

Priors on the hyperparameters

The priors on the hyperparameters were, where possible, identical across both models. For the asteroseismic model, our priors took the form of

$$\begin{aligned}
\mu_{\rm RC} &\sim \mathcal{N}(\mu_{\rm H}, 1) \,, \\
\sigma_{\rm RC} &\sim \mathcal{N}(0, 1) \, \times \mathcal{U}(0, \infty) \,, \\
Q &\sim \mathcal{N}(1, 0.25) \, \times \mathcal{U}(0.5, 1) \,, \\
\sigma_{\rm o} &\sim \mathcal{N}(3\sigma_{\rm RC}, 2\sigma_{\rm RC}) \, \times \mathcal{U}(\sigma_{\rm RC}, \infty) \,,
\end{aligned} \tag{4.28}$$

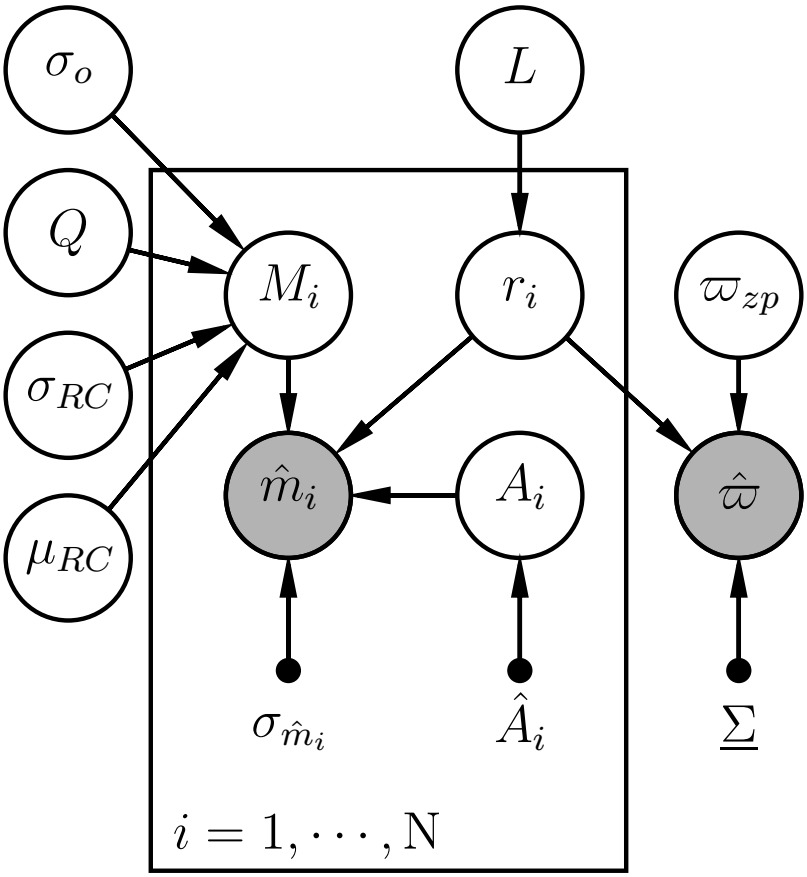

Figure 4.10: A probabilistic graphical model of the astrometric model, represented algebraically in Equation 4.22. Conventions are the same as for Figure 4.9. The full parallax covariance matrix is denoted as $\underline{\Sigma}$; it should be noted that the parallax likelihood is evaluated across the full set as a multivariate normal distribution.

where $\mu_{\rm H}$ is the absolute magnitude of the RC in the relevant passband, as reported by H17. Here, $\sigma_{\rm RC}$, Q and $\sigma_{\rm o}$ all have truncated priors, with limits indicated by the uniform prior on the right hand side of the equation. In order to evaluate the hierarchical mixture model in `PyStan`, $\sigma_{\rm o}$ has a prior expressed in units of $\sigma_{\rm RC}$ and must always be larger than 1 to ensure the two components of the mixture model do not switch roles. Q must fall within the range 0.5 to 1, because we expect an inlier-dominated sample. The spread $\sigma_{\rm RC}$ can not go below zero, and has no upper limit.

For the astrometric method, we introduced the two new parameters $\varpi_{\rm zp}$ and L, and applied a new prior to $\mu_{\rm RC}$ and $\sigma_{\rm RC}$, while the priors for the other hyperparam-

eters remained the same:

$$\begin{aligned}
\mu_{\rm RC} &\sim \mathcal{N}(\mu_{\rm RC,seis}, \sigma_{\mu_{\rm RC,seis}})\,, \\
\sigma_{\rm RC} &\sim \mathcal{N}(\sigma_{\rm RC,seis}, \sigma_{\sigma_{\rm RC,seis}}) \times \mathcal{U}(0, \infty)\,, \\
L &\sim \mathcal{U}(0.1, 4000)\,, \\
\varpi_{\rm zp} &\sim \mathcal{N}(0, 500)\,.
\end{aligned} \tag{4.29}$$

Here, $\mathcal{U}$ denotes a uniform distribution with the lower and upper limits as arguments, and the units of $\varpi_{\rm zp}$ and L are μas and kpc, respectively. The quantities $\mu_{\rm RC,seis}$ and $\sigma_{\rm RC,seis}$ are the medians of the posterior distributions on $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ from the asteroseismic model, and $\sigma_{\mu_{\rm RC,seis}}$ and $\sigma_{\sigma_{\rm RC,seis}}$ are the spreads on the posteriors, effectively allowing us to explore what value of the parallax-zero point offset, $\varpi_{\rm zp}$, recovers the results we see using asteroseismology.

Finally, for runs where we investigated the impact of literature values for $\varpi_{\rm zp}$ on our RC parameters, we set the priors on $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ to those used on our seismic run, and applied a prior on $\varpi_{\rm zp}$ as

$$\varpi_{\rm zp} \sim \mathcal{N}(\varpi_{\rm zp,lit}, \sigma_{\varpi_{\rm zp,lit}})\,. \tag{4.30}$$

Here, $\varpi_{\rm zp,lit}$ and $\sigma_{\varpi_{\rm zp,lit}}$ are values and uncertainties on said values from the literature.

We drew samples from the posterior distributions using `PyStan` (Carpenter et al., 2017; Van Hoey et al., 2013) with four chains and 5000 iterations, with half of the iterations used as burn-in. Appropriate convergence of our chains was evaluated using the Gelman-Rubin statistic ($\hat{R}$, Gelman & Rubin, 1992).

4.3.4 Results

Results from asteroseismology

To see how the absolute magnitude $\mu_{\rm RC}$ and spread $\sigma_{\rm RC}$ of the RC change given our input data, we applied two changes to calculations for seismic absolute magnitude. First, we perturbed the temperature by a value $\Delta T_{\rm eff}$ that ranged between -50 and 50 K, in steps of 10 K. Second, we propagated these temperatures, along with the original and unperturbed uncertainties on $T_{\rm eff}$, $\nu_{\rm max}$ and $\Delta\nu$, through the seismic scaling relations to find luminosity. We did this both with and without calibrations for the $\Delta\nu$ scaling relation obtained by the grid interpolation method by Sharma et al. (2016). The perturbed temperatures were also used in the grid interpolation required to obtain the correction (Sharma & Stello, 2016), and the corrections were thus recalculated for each change in temperature. We also calculated BCs for each set of temperatures, and recalculated a seismic $\log g$ given the perturbed temperatures for each calculation of the BCs (Casagrande & VandenBerg, 2014, 2018a,b). Seismic radii were calculated per Equation 4.11, which were in turn used to calculate luminosities and were combined with the BCs to compute our absolute magnitudes, resulting in 22 individual sets that differ in corrections to the seismic scaling relations and temperature scale, for both photometric bands.

Our results for our Y18 sample are shown in Tables 4.1 and 4.3 where we present the medians of the posterior distributions for our hyperparameters for the 2MASS K band and *Gaia* G band respectively, both with and without a correction to the $\Delta\nu$ scaling relation, for various changes in temperature scale. Uncertainties are given as the 1σ credible intervals. Where the posterior distributions are approximately Gaussian we quote a symmetric single uncertainty. The change of the posterior on the magnitude of the RC $\mu_{\rm RC}$ alone, given the input, can be seen in Figure 4.11.

For our APOKASC-2 temperature subsample of 1637 stars, we reran our models using the same methodology as before, simply substituting the temperatures and

temperature uncertainties reported in Pinsonneault et al. (2018) for those in Y18 for those stars, and making no other changes. Note that the change in temperature values carried through to the calculation of the bolometric corrections and corrections to the scaling relations for each run. The results of this are presented in Tables 4.2 and 4.4 for all hyperparameters, as with the run on the full sample. The change in the posteriors on the position of the RC is shown for this reduced sample in Figure 4.12.

	No Correction				Clump Corrected			
$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$
-50.0	-1.704 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.35^{+1.17}_{-1.01}$	-1.713 ± 0.002	0.034 ± 0.004	0.91 ± 0.01	$8.85^{+1.09}_{-0.93}$
-40.0	-1.709 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.33^{+1.22}_{-1.01}$	-1.718 ± 0.002	0.033 ± 0.004	0.91 ± 0.01	$9.11^{+1.12}_{-1.04}$
-30.0	-1.714 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.4^{+1.15}_{-1.04}$	-1.724 ± 0.002	0.033 ± 0.004	0.91 ± 0.01	$9.16^{+1.12}_{-0.96}$
-20.0	-1.719 ± 0.002	0.029 ± 0.003	0.92 ± 0.01	$10.55^{+1.15}_{-1.05}$	-1.73 ± 0.002	0.033 ± 0.004	0.91 ± 0.01	$9.22^{+1.05}_{-0.91}$
-10.0	-1.724 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.49^{+1.13}_{-1.03}$	-1.735 ± 0.002	0.033 ± 0.004	0.91 ± 0.01	$9.16^{+1.09}_{-0.98}$
0.0	-1.729 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.33^{+1.19}_{-1.01}$	-1.741 ± 0.002	0.033 ± 0.004	0.91 ± 0.01	$9.18^{+1.09}_{-0.92}$
10.0	-1.734 ± 0.002	0.029 ± 0.003	0.92 ± 0.01	$10.44^{+1.07}_{-0.97}$	-1.746 ± 0.002	0.032 ± 0.004	0.91 ± 0.01	$9.36^{+1.2}_{-1.05}$
20.0	-1.739 ± 0.002	0.03 ± 0.004	0.92 ± 0.01	$10.32^{+1.17}_{-1.02}$	-1.752 ± 0.002	0.033 ± 0.004	0.91 ± 0.01	$9.19^{+1.16}_{-1.01}$
30.0	-1.744 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.41^{+1.06}_{-0.99}$	-1.757 ± 0.002	0.032 ± 0.004	0.91 ± 0.01	$9.37^{+1.16}_{-1.02}$
40.0	-1.749 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.41^{+1.18}_{-1.02}$	-1.762 ± 0.002	0.032 ± 0.004	0.91 ± 0.01	$9.37^{+1.14}_{-0.97}$
50.0	-1.754 ± 0.002	0.03 ± 0.003	0.92 ± 0.01	$10.27^{+1.12}_{-1.01}$	-1.768 ± 0.002	0.032 ± 0.004	0.91 ± 0.01	$9.25^{+1.16}_{-1}$

Table 4.1: Medians of the posterior distributions for hyperparameters of our seismic model, for the 2MASS K band, for 5576 stars from the Y18 sample. Uncertainties are taken as the 1σ credible intervals, and are listed as a single value for cases where the posterior was approximately Gaussian. Values are listed for data that have been left uncorrected (No Correction) and data with an appropriate correction to the seismic scaling relations (Clump Corrected). $\Delta T_{\rm eff}$ is the global shift to our values of $T_{\rm eff}$, $\mu_{\rm RC}$ is the position of the RC in absolute magnitude, $\sigma_{\rm RC}$ is the spread of the RC in absolute magnitude, Q is the mixture model weighting factor (and the effective fraction of stars considered inliers), and σ_o is the spread of our outlier population, expressed in terms of $\sigma_{\rm RC}$.

	No Correction				Clump Corrected			
$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$
-50.0	-1.659 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.2^{+1.2}_{-1.09}$	-1.663 ± 0.003	0.031 ± 0.005	0.89 ± 0.02	$8.46^{+1.19}_{-1.06}$
-40.0	-1.664 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.14^{+1.18}_{-1.08}$	-1.669 ± 0.003	0.032 ± 0.005	0.89 ± 0.02	$8.4^{+1.16}_{-1.1}$
-30.0	-1.669 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.13^{+1.23}_{-1.11}$	-1.675 ± 0.003	0.031 ± 0.005	0.89 ± 0.02	$8.53^{+1.16}_{-1.06}$
-20.0	-1.674 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.15^{+1.26}_{-1.1}$	-1.681 ± 0.003	0.031 ± 0.005	0.89 ± 0.02	$8.43^{+1.24}_{-1.06}$
-10.0	-1.679 ± 0.003	0.03 ± 0.004	0.9 ± 0.02	$9.11^{+1.18}_{-1.09}$	-1.687 ± 0.003	0.032 ± 0.005	0.89 ± 0.02	$8.37^{+1.23}_{-1.11}$
0.0	-1.684 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.13^{+1.25}_{-1.1}$	-1.693 ± 0.003	0.031 ± 0.005	0.89 ± 0.02	$8.5^{+1.18}_{-1.08}$
10.0	-1.689 ± 0.003	0.03 ± 0.004	0.9 ± 0.02	$9.08^{+1.2}_{-1.08}$	-1.698 ± 0.003	0.032 ± 0.005	0.89 ± 0.02	$8.41^{+1.2}_{-1.06}$
20.0	-1.694 ± 0.003	0.03 ± 0.004	0.9 ± 0.02	$9.04^{+1.26}_{-1.06}$	-1.704 ± 0.003	0.032 ± 0.005	0.89 ± 0.02	$8.44^{+1.21}_{-1.08}$
30.0	-1.699 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.1^{+1.17}_{-1.07}$	-1.71 ± 0.003	0.033 ± 0.005	0.9 ± 0.02	$8.29^{+1.21}_{-1.05}$
40.0	-1.704 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.12^{+1.18}_{-1.1}$	-1.715 ± 0.003	0.032 ± 0.005	0.89 ± 0.02	$8.43^{+1.23}_{-1.07}$
50.0	-1.709 ± 0.003	0.029 ± 0.004	0.9 ± 0.02	$9.15^{+1.24}_{-1.09}$	-1.721 ± 0.003	0.032 ± 0.005	0.89 ± 0.02	$8.39^{+1.19}_{-1.07}$

Table 4.2: Same as Table 4.1, except for a subsample of stars from the APOKASC-2 (Pinsonneault et al., 2018) sample.

	No Correction				Clump Corrected			
$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$
-50.0	0.35 ± 0.003	0.181 ± 0.004	0.98 ± 0.01	$2.73^{+0.58}_{-0.42}$	0.34 ± 0.003	0.193 ± 0.004	0.99 ± 0.01	$2.77^{+0.66}_{-0.46}$
-40.0	0.336 ± 0.003	0.181 ± 0.004	0.98 ± 0.01	$2.73^{+0.56}_{-0.4}$	0.325 ± 0.003	0.192 ± 0.004	0.99 ± 0.01	$2.78^{+0.66}_{-0.45}$
-30.0	0.323 ± 0.003	0.18 ± 0.004	0.98 ± 0.01	$2.72^{+0.54}_{-0.4}$	0.311 ± 0.003	0.19 ± 0.004	0.99 ± 0.01	$2.79^{+0.58}_{-0.43}$
-20.0	0.309 ± 0.003	0.179 ± 0.004	0.98 ± 0.01	$2.72^{+0.54}_{-0.4}$	0.297 ± 0.003	0.188 ± 0.004	0.98 ± 0.01	$2.72^{+0.58}_{-0.41}$
-10.0	0.295 ± 0.003	0.178 ± 0.004	0.98 ± 0.01	$2.69^{+0.53}_{-0.38}$	0.282 ± 0.003	0.187 ± 0.004	0.98 ± 0.01	$2.71^{+0.58}_{-0.4}$
0.0	0.282 ± 0.003	0.177 ± 0.004	0.98 ± 0.01	$2.68^{+0.52}_{-0.38}$	0.268 ± 0.003	0.187 ± 0.004	0.98 ± 0.01	$2.73^{+0.58}_{-0.42}$
10.0	0.268 ± 0.003	0.177 ± 0.004	0.98 ± 0.01	$2.71^{+0.53}_{-0.39}$	0.254 ± 0.003	0.185 ± 0.004	0.98 ± 0.01	$2.7^{+0.58}_{-0.41}$
20.0	0.255 ± 0.003	0.176 ± 0.004	0.98 ± 0.01	$2.7^{+0.51}_{-0.37}$	0.24 ± 0.003	0.184 ± 0.004	0.98 ± 0.01	$2.71^{+0.56}_{-0.4}$
30.0	0.241 ± 0.003	0.175 ± 0.004	0.98 ± 0.01	$2.68^{+0.51}_{-0.36}$	0.226 ± 0.003	0.183 ± 0.004	0.98 ± 0.01	$2.7^{+0.55}_{-0.4}$
40.0	0.228 ± 0.003	0.174 ± 0.004	0.98 ± 0.01	$2.67^{+0.48}_{-0.36}$	0.213 ± 0.003	0.182 ± 0.004	0.98 ± 0.01	$2.7^{+0.52}_{-0.4}$
50.0	0.215 ± 0.003	0.173 ± 0.004	0.98 ± 0.01	$2.68^{+0.48}_{-0.36}$	0.199 ± 0.003	0.181 ± 0.004	0.98 ± 0.01	$2.69^{+0.53}_{-0.38}$

Table 4.3: Same as Table 4.1, except for the *Gaia G* band, for 5576 stars from the Y18 sample.

	No Correction				Clump Corrected			
$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$
-50.0	0.53 ± 0.004	0.118 ± 0.006	$0.96^{+0.02}_{-0.03}$	$3.19^{+0.65}_{-0.46}$	0.526 ± 0.004	0.128 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.29^{+0.81}_{-0.55}$
-40.0	0.516 ± 0.004	0.117 ± 0.005	$0.96^{+0.02}_{-0.03}$	$3.18^{+0.65}_{-0.46}$	0.51 ± 0.004	0.127 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.31^{+0.8}_{-0.55}$
-30.0	0.501 ± 0.004	0.116 ± 0.006	$0.96^{+0.02}_{-0.03}$	$3.19^{+0.62}_{-0.45}$	0.495 ± 0.004	0.127 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.29^{+0.76}_{-0.54}$
-20.0	0.486 ± 0.004	0.116 ± 0.006	$0.96^{+0.02}_{-0.03}$	$3.19^{+0.65}_{-0.47}$	0.479 ± 0.004	0.126 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.28^{+0.77}_{-0.54}$
-10.0	0.472 ± 0.004	0.115 ± 0.006	$0.96^{+0.02}_{-0.03}$	$3.19^{+0.63}_{-0.45}$	0.464 ± 0.004	0.126 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.29^{+0.74}_{-0.55}$
0.0	0.457 ± 0.004	0.114 ± 0.006	$0.96^{+0.02}_{-0.03}$	$3.2^{+0.64}_{-0.45}$	0.449 ± 0.004	0.125 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.27^{+0.79}_{-0.53}$
10.0	0.443 ± 0.004	0.113 ± 0.006	$0.95^{+0.02}_{-0.03}$	$3.17^{+0.63}_{-0.44}$	0.434 ± 0.004	0.124 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.25^{+0.8}_{-0.52}$
20.0	0.429 ± 0.004	0.113 ± 0.006	$0.96^{+0.02}_{-0.03}$	$3.21^{+0.62}_{-0.44}$	0.419 ± 0.004	0.124 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.25^{+0.73}_{-0.53}$
30.0	0.414 ± 0.004	0.112 ± 0.006	$0.95^{+0.02}_{-0.03}$	$3.18^{+0.61}_{-0.43}$	0.404 ± 0.004	0.123 ± 0.005	$0.97^{+0.01}_{-0.02}$	$3.25^{+0.72}_{-0.54}$
40.0	0.4 ± 0.004	0.112 ± 0.006	$0.95^{+0.02}_{-0.03}$	$3.19^{+0.59}_{-0.44}$	0.389 ± 0.004	0.122 ± 0.005	$0.97^{+0.02}_{-0.02}$	$3.24^{+0.72}_{-0.5}$
50.0	0.386 ± 0.004	0.111 ± 0.006	$0.95^{+0.02}_{-0.03}$	$3.2^{+0.57}_{-0.43}$	0.375 ± 0.004	0.122 ± 0.006	$0.97^{+0.01}_{-0.02}$	$3.25^{+0.71}_{-0.51}$

Table 4.4: Same as Table 4.1, except for the *Gaia G* band, for a subsample of stars from the APOKASC-2 (Pinsonneault et al., 2018) sample.

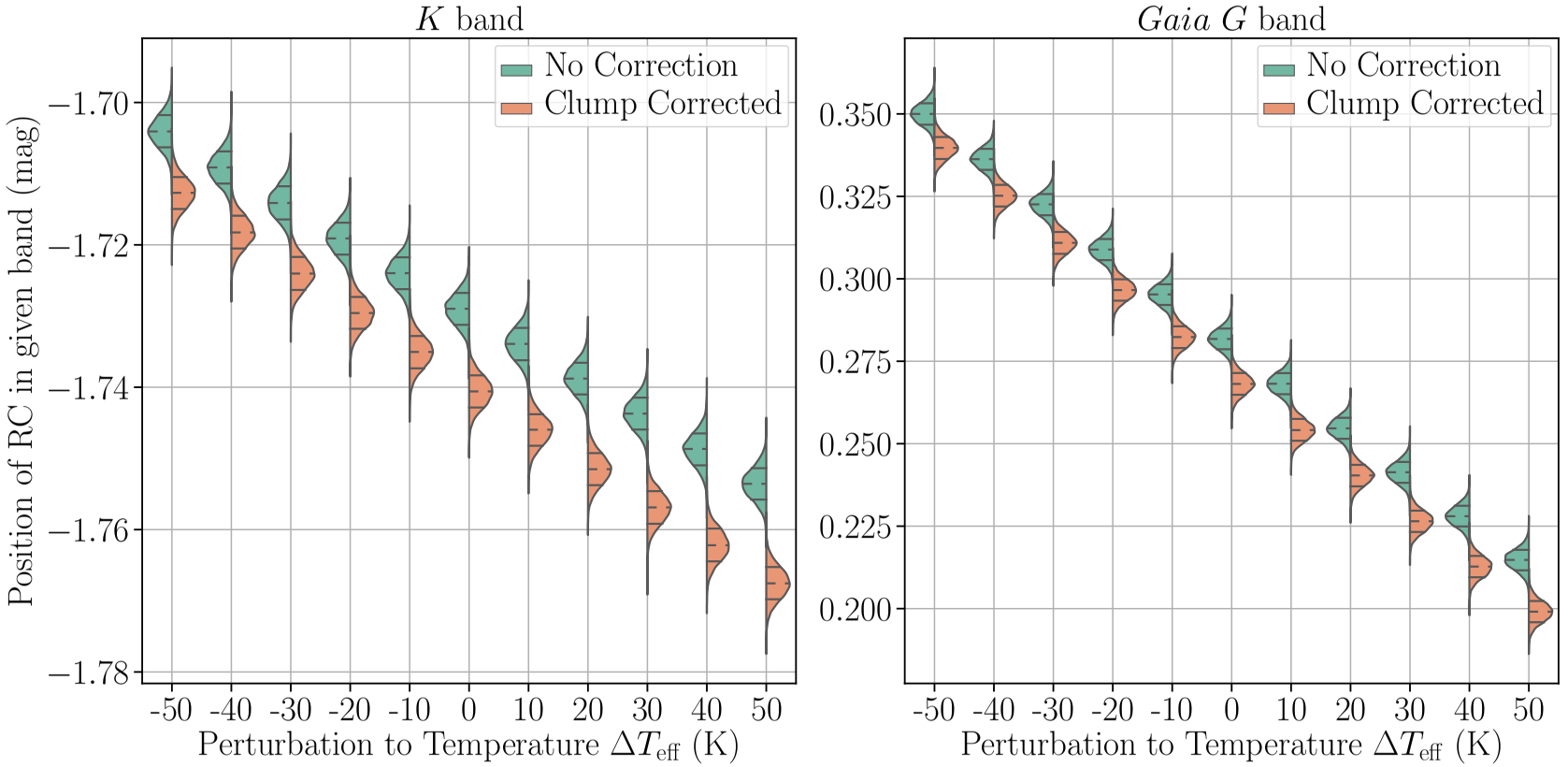

Figure 4.11: The posterior distributions on the position of the Red Clump in the 2MASS K band (left) and *Gaia* G band (right), as a function of overall perturbation to the temperature values $\Delta T_{\rm eff}$ using asteroseismology, both with (orange) and without (green) corrections to the $\Delta\nu$ scaling relation (Sharma et al., 2016). The dashed horizontal lines indicate the median on the posteriors, and the solid horizontal lines represent the 1σ credible intervals. The posteriors' magnitudes along the x-axis are indicative of power with arbitrary units, whereas their shape along the y-axis indicates the spread in the posterior result.

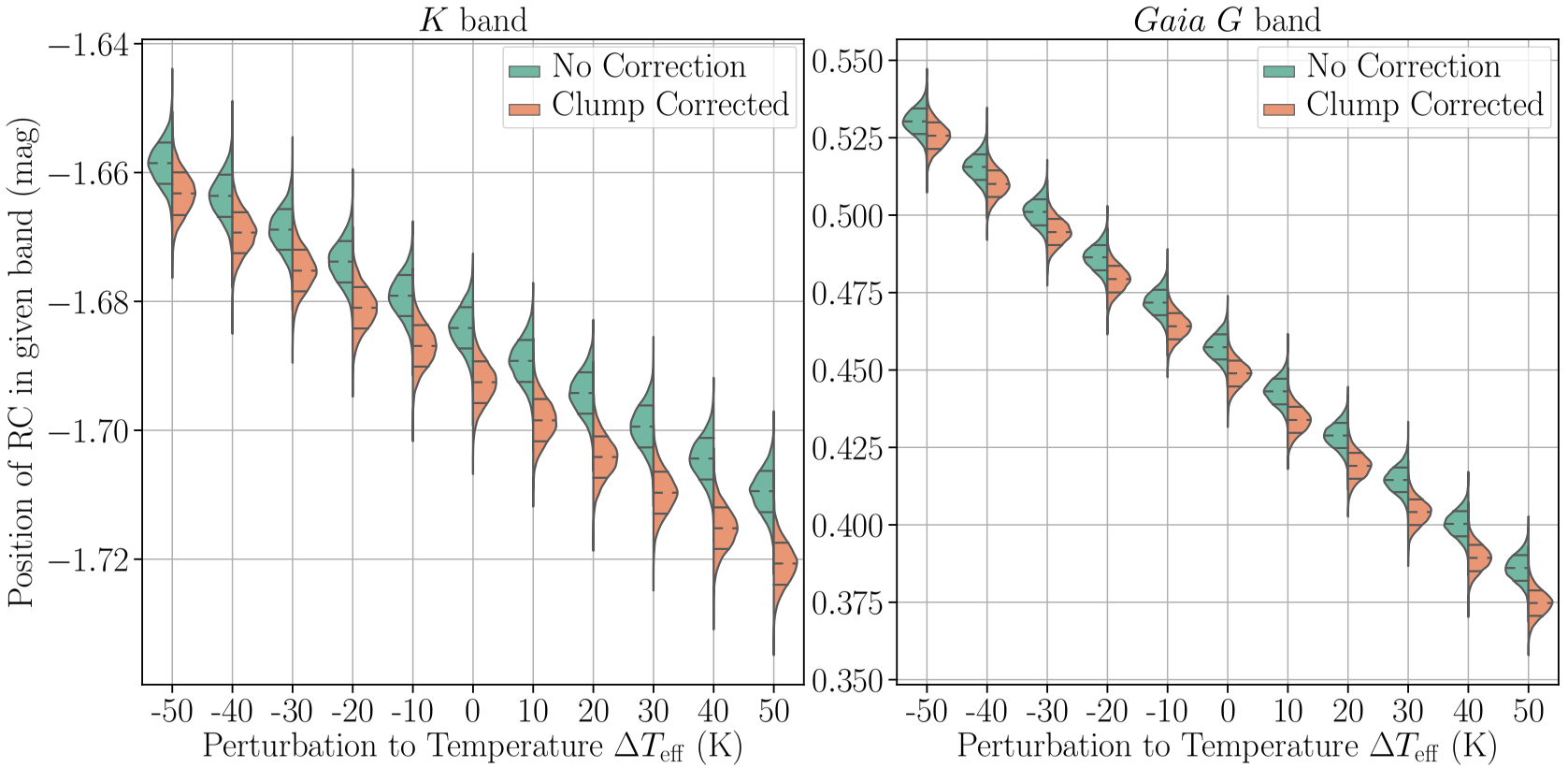

Figure 4.12: The same as Figure 4.11, except using only stars both in our sample and the APOKASC-2 (Pinsonneault et al., 2018) sample, replacing $T_{\rm eff}$ with those reported in APOKASC-2.

Results from *Gaia*

Given our results from asteroseismology, we wish to determine the parallax zero-point offset, $\varpi_{\rm zp}$, that recovers our values of the absolute magnitude and spread of the RC. Since $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ represent astrophysical observables that should be consistent across both data sets, we used a description of the posterior distributions from these parameters from our seismic model as a highly informative prior in our *Gaia* model. This yields the parallax offset required to recover the same magnitude and spread of the RC found using seismology. We passed in the seismic posteriors for $\Delta T_{\rm eff}$ being -50, 0, and $+50$ K from our runs on our full sample and the reduced APOKASC-2 sample, and thus ran our model for 6 different RC magnitudes and spreads in each band. Additionally, we used the median values of each latent parameter M_i from the application of our seismic model to our full sample, along with distance estimates by Bailer-Jones et al. (2018) and observed extinctions from Green et al. (2018), as initial guesses in our *Gaia* model for computational efficiency. No other values were changed on each run.

Following the relation presented in Equation 4.18 (Lindegren et al., 2018) we treated our parallax uncertainties as correlated as a function of position on the sky across the entire *Kepler* field, similarly to previous work by Zinn et al. (2019a). While the model equation presented by Lindegren et al. (2018) describes the covariance well for a wide range of separations, individual covariances oscillate around the model at separations below 1°, and the model no longer holds at all for separations below 0.125°. To ensure that our treatment of the parallax covariances was sensible, we ran our *Gaia* model on a reduced sample of 1000 stars, randomly selected from across the entire *Kepler* field to ensure sparsity. This reduced sample contained no angular separations in the range $< 0.125°$[9].

In Tables 4.5, 4.6, 4.7 and 4.8 we present the medians on the posterior distri-

[9]The data were shuffled using the `sklearn.utils.shuffle` function with a random seed of 24601.

butions of our hyperparameters for our *Gaia* model, given RC-corrected seismic positions and spreads for the RC at different temperature offsets $\Delta T_{\rm eff}$ for both the Y18 and APOKASC-2 samples. In Figure 4.13, we present the posterior distributions of $\varpi_{\rm zp}$ given the 6 values for the position of the RC used each in the K and G bands.

In order to probe the impact of literature values for $\varpi_{\rm zp}$ on an inference of our RC parameters, we reran our *Gaia* model for the K and *Gaia* G bands with a strongly informative prior on $\varpi_{\rm zp}$ (see Equation 4.30). We did this for the same reduced sample of 1000 stars from our Y18 sample. For all these runs, we applied the same priors used for $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ as in the asteroseismic runs (see Equation 4.28). We used the parallax zero-point offsets reported by Lindegren et al. (2018) ($-29\,\mu$as, with an assumed uncertainty of $1\,\mu$as), Zinn et al. (2019a) ($-52.8\,\mu$as with a total uncertainty of $3.4\,\mu$as), Riess et al. (2018) ($-46 \pm 13\,\mu$as), Sahlholdt & Silva Aguirre (2018) ($-35 \pm 16\,\mu$as) and Stassun & Torres (2018) ($-82 \pm 33\,\mu$as). Not all these zero-point offsets would be applicable to our sample due to differences in colour, magnitude, and position. We instead used them as representative of $\varpi_{\rm zp}$ in the literature to study their impact on our inferences only. In addition, we also ran with a prior of $0 \pm 1\,\mu$as in an attempt to recreate the H17 work (albeit accounting for parallax covariances), as well as a single run with no strongly informative priors on $\varpi_{\rm zp}$, $\mu_{\rm RC}$ or $\sigma_{\rm RC}$, thus finding our own measure of the zero-point offset.

In Tables 4.9 and 4.10 we present the medians and 1σ credible intervals on the posterior distributions for the hyperparameters of our *Gaia* model given the conditions stated above, as well as naming the source of the used parallax zero-point offset, and an expression of the prior applied to $\varpi_{\rm zp}$. Note that the inferred value of $\varpi_{\rm zp}$ may differ significantly within the uncertainties on any of the literature values used. In Figure 4.14 we present the medians and 1σ credible intervals on the posterior distributions for $\mu_{\rm RC}$ given our chosen values for $\varpi_{\rm zp}$, with the result from the 'uninformed' run shown with bold red error bars.

$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	L (pc)	$\varpi_{\rm zp}$ $(\mu$as)
-50.0	-1.71 ± 0.002	0.041 ± 0.003	$0.58^{+0.05}_{-0.05}$	$5.49^{+0.52}_{-0.47}$	$908.63^{+16.55}_{-15.89}$	$-24.09^{+12.84}_{-12.76}$
0.0	-1.737 ± 0.002	0.04 ± 0.003	$0.55^{+0.05}_{-0.03}$	$5.61^{+0.5}_{-0.47}$	$920.12^{+17.18}_{-16.61}$	$-19.5^{+12.4}_{-12.46}$
50.0	-1.764 ± 0.002	0.041 ± 0.004	$0.53^{+0.04}_{-0.02}$	$5.5^{+0.5}_{-0.48}$	$930.95^{+18.07}_{-16.83}$	$-14.81^{+12.57}_{-12.98}$

Table 4.5: Medians of the posterior distributions for hyperparameters of our *Gaia* model, for the 2MASS K band, for a randomly selected subsample of 1000 stars from the Y18 sample. Uncertainties are taken as the 1σ credible intervals, and are listed as a single value for cases where the posterior was approximately Gaussian. Priors were imposed on $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ corresponding to the results for these values using seismic Clump Corrected data in Table 4.1, for the temperature shifts shown in the $\Delta T_{\rm eff}$ column. L is the length scale of the exponentially decaying space density prior on distance (Bailer-Jones et al., 2018), and $\varpi_{\rm zp}$ is the parallax zero-point offset. All other symbols are the same as for Table 4.1.

$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	L (pc)	$\varpi_{\rm zp}$ $(\mu$as)
-50.0	-1.661 ± 0.003	0.04 ± 0.003	0.6 ± 0.05	$5.76^{+0.55}_{-0.5}$	$888^{+16.38}_{-15.78}$	$-33.53^{+12.93}_{-12.97}$
0.0	-1.689 ± 0.003	0.04 ± 0.004	0.59 ± 0.05	$5.66^{+0.53}_{-0.51}$	$899.36^{+16.72}_{-16.04}$	$-28.33^{+12.96}_{-12.92}$
50.0	-1.715 ± 0.003	0.041 ± 0.004	0.57 ± 0.05	$5.51^{+0.55}_{-0.49}$	$910.68^{+16.83}_{-16.4}$	$-23.47^{+13.25}_{-13.13}$

Table 4.6: Same as Table 4.5, except for priors imposed on $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ corresponding to the results for these values using seismic Clump Corrected data in Table 4.2 (the APOKASC-2 subsample), for the temperature shifts shown in the $\Delta T_{\rm eff}$ column.

$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	L (pc)	$\varpi_{\rm zp}$ $(\mu$as)
-50.0	0.346 ± 0.003	0.19 ± 0.003	$0.97^{+0.01}_{-0.02}$	$3.1^{+0.79}_{-0.7}$	$948.41^{+18.15}_{-17.96}$	$-9.96^{+13.1}_{-13.18}$
0.0	0.277 ± 0.003	0.188 ± 0.004	$0.95^{+0.02}_{-0.05}$	$2.64^{+0.77}_{-0.63}$	$978.9^{+18.35}_{-17.46}$	$1.14^{+12.8}_{-12.81}$
50.0	0.209 ± 0.003	0.184 ± 0.004	$0.74^{+0.12}_{-0.13}$	$1.71^{+0.36}_{-0.2}$	$1008.77^{+18.85}_{-18.48}$	$10.76^{+13.13}_{-13.21}$

Table 4.7: Same as Table 4.5, except for the *Gaia* G band, with priors imposed on $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ corresponding to the results for these values using seismic Clump Corrected data in Table 4.3, for the temperature shifts shown in the $\Delta T_{\rm eff}$ column.

$\Delta T_{\rm eff}$ (K)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	L (pc)	$\varpi_{\rm zp}$ $(\mu$as)
-50.0	0.527 ± 0.004	0.13 ± 0.005	$0.82^{+0.05}_{-0.07}$	$2.53^{+0.36}_{-0.3}$	$874.12^{+16.56}_{-16.1}$	$-39.02^{+12.98}_{-13.16}$
0.0	0.455 ± 0.004	0.129 ± 0.005	$0.79^{+0.07}_{-0.08}$	$2.42^{+0.36}_{-0.29}$	$903.23^{+16.8}_{-16.68}$	$-26.84^{+13.1}_{-12.97}$
50.0	0.385 ± 0.004	0.127 ± 0.005	$0.68^{+0.09}_{-0.1}$	$2.22^{+0.27}_{-0.22}$	$931.92^{+17.53}_{-17}$	$-14.94^{+12.58}_{-13.04}$

Table 4.8: Same as Table 4.5, except for the *Gaia* G band, with priors imposed on $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ corresponding to the results for these values using seismic Clump Corrected data in Table 4.4 (the APOKASC-2 subsample), for the temperature shifts shown in the $\Delta T_{\rm eff}$ column.

Source	$\varpi_{\rm zp}$ prior $(\mu$as)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ $(\sigma_{\rm RC})$	L (pc)	$\varpi_{\rm zp}$ $(\mu$as)
Lindegren+ 18	$\mathcal{N}(-29.0, 1.0)$	-1.638 ± 0.017	$0.075^{+0.016}_{-0.015}$	$0.78^{+0.09}_{-0.11}$	$3.28^{+0.64}_{-0.56}$	$888.56^{+25.41}_{-24.36}$	$-29.07^{+1}_{-0.99}$
Zinn+ 18	$\mathcal{N}(-52.8, 3.4)$	-1.631 ± 0.017	$0.074^{+0.016}_{-0.015}$	$0.77^{+0.09}_{-0.1}$	$3.3^{+0.65}_{-0.57}$	$885.76^{+24.4}_{-23.73}$	$-51.92^{+3.21}_{-3.21}$
Riess+ 18	$\mathcal{N}(-46.0, 13.0)$	-1.634 ± 0.017	$0.076^{+0.017}_{-0.015}$	$0.78^{+0.09}_{-0.11}$	$3.26^{+0.64}_{-0.57}$	$886.59^{+25}_{-24.12}$	$-42.22^{+9.33}_{-9.16}$
Sahlholdt & Silva Aguirre18	$\mathcal{N}(-35.0, 16.0)$	-1.634 ± 0.017	$0.073^{+0.016}_{-0.015}$	$0.77^{+0.09}_{-0.11}$	$3.33^{+0.69}_{-0.58}$	$887.37^{+24.06}_{-23.89}$	$-37^{+10.17}_{-10.42}$
Stassun & Torres 18	$\mathcal{N}(-82.0, 33.0)$	-1.632 ± 0.017	$0.072^{+0.017}_{-0.016}$	$0.76^{+0.09}_{-0.11}$	$3.36^{+0.64}_{-0.59}$	$885.77^{+24.33}_{-23.01}$	$-44.55^{+12.62}_{-12.59}$
Hawkins+ 17	$\mathcal{N}(0.0, 1.0)$	-1.648 ± 0.018	$0.075^{+0.017}_{-0.015}$	$0.78^{+0.09}_{-0.11}$	$3.31^{+0.64}_{-0.57}$	$893.39^{+24.6}_{-24}$	$-0.22^{+0.99}_{-1.01}$
Uninformed	$\mathcal{N}(0.0, 1000.0)$	-1.634 ± 0.018	$0.074^{+0.017}_{-0.015}$	$0.77^{+0.09}_{-0.11}$	$3.3^{+0.64}_{-0.58}$	$887.27^{+24.12}_{-23.82}$	$-38.38^{+13.83}_{-13.54}$

Table 4.9: Medians on the posterior distributions for hyperparameters on our *Gaia* model, for the 2MASS K band, for a randomly selected subsample of 1000 stars from the Y18 sample. Uncertainties are taken as the 1σ credible intervals, and are listed as single values for cases where the posterior was approximately Gaussian. Highly informative priors, shown in the '$\varpi_{\rm zp}$ prior' column, were imposed on $\varpi_{\rm zp}$ corresponding to estimates for this parameter from the literature, listed in bold print in the Source column. Additionally, we applied a custom prior to place $\varpi_{\rm zp}$ near zero in order to recreate conditions similar to the H17 work, and an extremely broad prior on $\varpi_{\rm zp}$ in order to find a value given no strong constraints on $\varpi_{\rm zp}$, $\mu_{\rm RC}$ or $\sigma_{\rm RC}$. $\mathcal{N}(\mu, \sigma)$ indicates a normal distribution with mean μ and standard deviation σ.

Source	$\varpi_{\rm zp}$ prior (μas)	$\mu_{\rm RC}$ (mag)	$\sigma_{\rm RC}$ (mag)	Q	$\sigma_{\rm o}$ ($\sigma_{\rm RC}$)	L (pc)	$\varpi_{\rm zp}$ (μas)
Lindegren+ 18	$\mathcal{N}(-29.0, 1.0)$	0.542 ± 0.016	$0.138^{+0.014}_{-0.018}$	$0.86^{+0.07}_{-0.12}$	$2.61^{+0.48}_{-0.34}$	$868.2^{+17.41}_{-17.09}$	$-29.06^{+0.98}_{-1.01}$
Zinn+ 18	$\mathcal{N}(-52.8, 3.4)$	0.548 ± 0.016	$0.139^{+0.014}_{-0.018}$	$0.86^{+0.07}_{-0.12}$	$2.62^{+0.49}_{-0.35}$	$865.44^{+16.95}_{-17.15}$	$-52.18^{+3.27}_{-3.31}$
Riess+ 18	$\mathcal{N}(-46.0, 13.0)$	0.545 ± 0.016	$0.14^{+0.013}_{-0.017}$	$0.87^{+0.07}_{-0.11}$	$2.62^{+0.48}_{-0.34}$	$867.13^{+17.23}_{-17.55}$	$-44.23^{+9.06}_{-9.32}$
Sahlholdt & Silva Aguirre 18	$\mathcal{N}(-35.0, 16.0)$	0.545 ± 0.016	$0.136^{+0.015}_{-0.021}$	$0.85^{+0.08}_{-0.14}$	$2.62^{+0.47}_{-0.34}$	$867.15^{+17.3}_{-17.05}$	$-39.29^{+9.86}_{-10.27}$
Stassun & Torres 18	$\mathcal{N}(-82.0, 33.0)$	0.546 ± 0.017	$0.138^{+0.014}_{-0.018}$	$0.86^{+0.07}_{-0.12}$	$2.61^{+0.46}_{-0.33}$	$866.11^{+17.76}_{-17.02}$	$-47.86^{+12.18}_{-12.51}$
Hawkins+ 17	$\mathcal{N}(0.0, 1.0)$	0.534 ± 0.015	$0.14^{+0.013}_{-0.018}$	$0.87^{+0.06}_{-0.12}$	$2.64^{+0.5}_{-0.35}$	$872.01^{+17.8}_{-17.38}$	$-0.23^{+1}_{-1.01}$
Uninformed	$\mathcal{N}(0.0, 1000.0)$	0.546 ± 0.016	$0.139^{+0.013}_{-0.019}$	$0.87^{+0.07}_{-0.13}$	$2.62^{+0.49}_{-0.34}$	$866.26^{+17.53}_{-16.86}$	$-42.66^{+13.48}_{-13.14}$

Table 4.10: Same as Table 4.9, except for the *Gaia G* band.

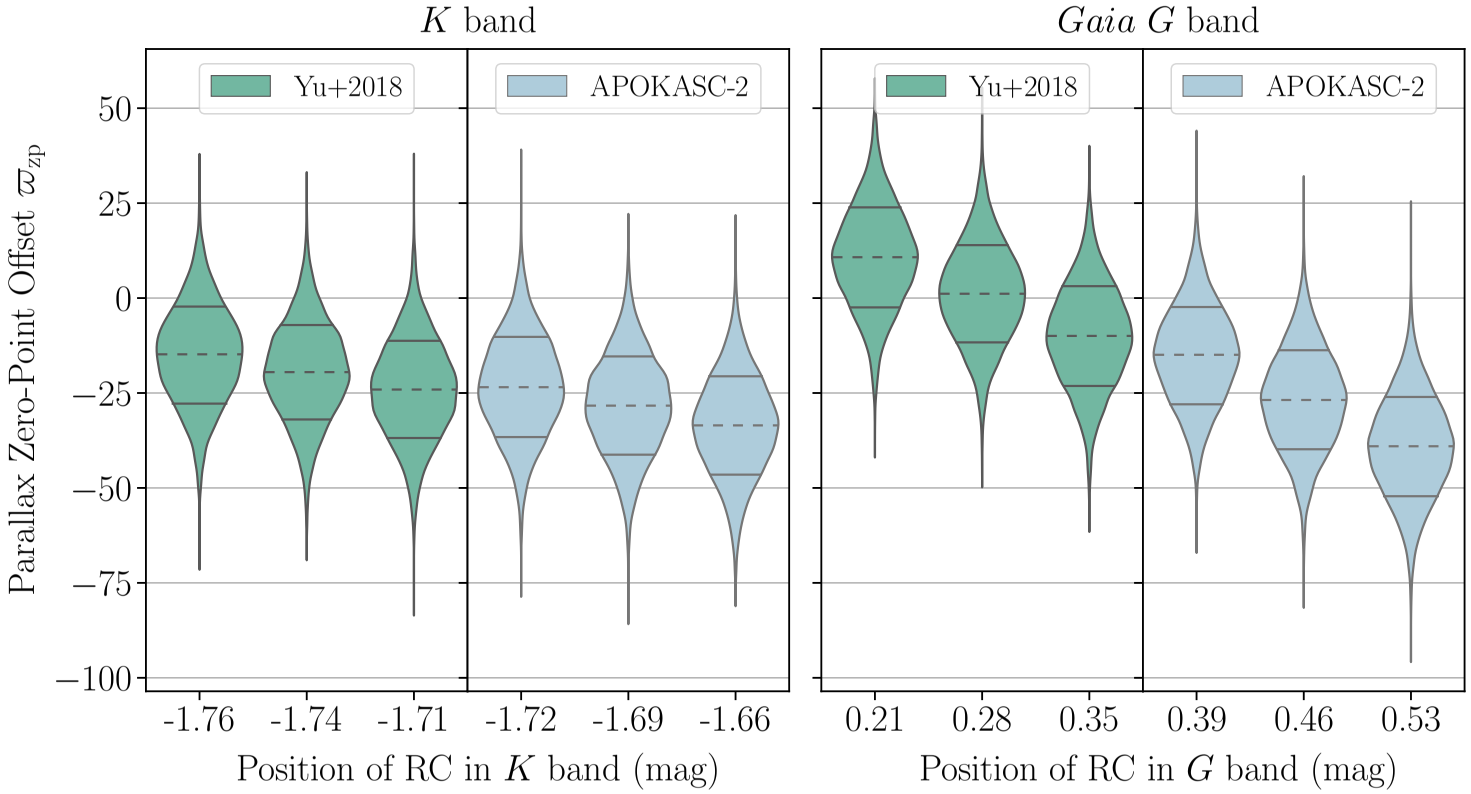

Figure 4.13: The posterior distributions on the parallax zero-point offset ϖ_{zp}, as a function of the absolute magnitude of the RC used to calibrate this value, for 1000 randomly selected stars across the *Kepler* field. The RC magnitudes on the x-axis correspond to those obtained from seismology for perturbations to the temperature values ΔT_{eff} of -50, 0, and $+50$ K, from runs on our full sample (Yu et al., 2018) and the APOKASC-2 sample (Pinsonneault et al., 2018). The dashed horizontal lines indicate the median on the posteriors, and the solid horizontal lines represent the 1σ credible intervals. The posteriors' magnitudes along the x-axis are indicative of power with arbitrary units, whereas their shape along the y-axis indicates the spread in the posterior result, and is reflected across the x-axis.

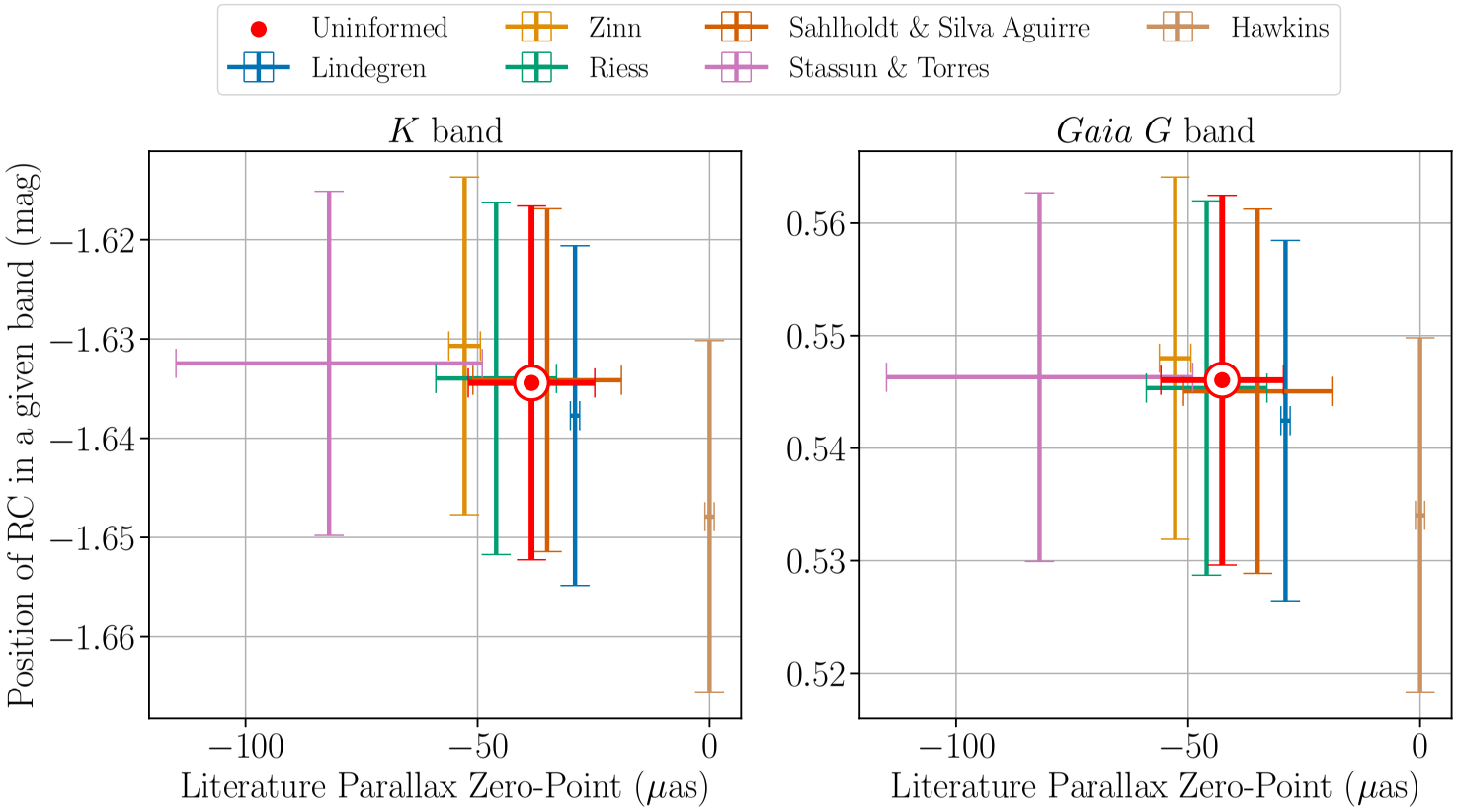

Figure 4.14: The the 1σ credible intervals for the posterior distributions on $\mu_{\rm RC}$, as a function of the value for $\varpi_{\rm zp}$ used as an informative prior on $\varpi_{\rm zp}$, for 1000 randomly selected stars across the *Kepler* field in both the 2MASS K and *Gaia* G bands. The error bars on the x-axis correspond to the formal uncertainties for literature values, or are otherwise specified in the text. the 'uninformed' value corresponds to a run of our *Gaia* model with no strong constraints on $\varpi_{\rm zp}$, and in this case the x-axis error bars correspond go the 1σ credible intervals on the inferred value for $\varpi_{\rm zp}$.

4.3.5 Discussion

Luminosity of the Red Clump

Figures 4.11 and 4.12 show the posteriors on the inferred absolute magnitude of the RC, $\mu_{\rm RC}$, for the K and *Gaia* G bands given changes to effective temperature and corrections to the scaling relations. There is a clear relation between the overall offset in $T_{\rm eff}$ and the inferred magnitude of the RC, where a change of about 20 K results in a difference of more than 1σ. The overall relation between the clump magnitude and temperature is expected, given the large impact of temperature on the calculations for absolute magnitude; luminosity calculated via the seismic scaling relations scales with temperature to a power of 4.5, and bolometric corrections calculated through the Casagrande & VandenBerg (2018b) method rely on both $T_{\rm eff}$ and $\log g$, which is calculated using $T_{\rm eff}$. The small uncertainties on $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ indicate the ability of hierarchical models to leverage a large number of individual uncertainties to fit to a population model, given that the uncertainties on our data for $T_{\rm eff}$ are well above the shifts in temperature we are applying.

We also see that the scaling relation corrections appear to be degenerate with a small temperature offset. A change of $\sim$ 20 K to the temperatures provides a similar clump magnitude as when applying a correction to the scaling relations. At higher temperatures, the difference in the magnitude of the RC between corrected and uncorrected scaling relations increases. This shows that the $T_{\rm eff}$ values have a significant impact on the $f_{\Delta\nu}$ obtained through the Sharma & Stello (2016) method, even at relatively small $T_{\rm eff}$ shifts.

The values for $\mu_{\rm RC}$ in both bands are fainter for the subset of stars using APOKASC-2 temperatures than those using temperatures from Mathur et al. (2017). This reflects the relation we already saw between $T_{\rm eff}$ and $\mu_{\rm RC}$ for the Y18 stars, since the stars in the APOKASC-2 subsample represent a population subset of lower-temperature stars, as well as having lower values for $T_{\rm eff}$ in the APOKASC-

2 catalogue itself. However, the fact that APOKASC-2 stars represent a lower-temperature population only accounts for a shift in a measured median absolute magnitude of ~ 0.028 mag in K and ~ 0.12 mag in G. The use of APOKASC-2 temperatures for the subset shifts the absolute magnitudes even fainter, by another ~ 0.028 mag and ~ 0.07 mag in K and G, respectively. At the precision afforded to us by hierarchical models, these shifts caused by the choice of temperatures become statistically significant.

Due to the nature of the K band minimizing the effects of metallicity on the RC spread, there is an extensive literature on the value of $\mu_{\rm RC}$ in K. It was found by Alves (2000) to be -1.62 ± 0.03 (with a consistent measurement by Udalski, 2000), but later placed at -1.54 ± 0.04 by Groenewegen (2008). A recent review by Girardi (2016) found a median literature value of -1.59 ± 0.04 mag, which was applied by Davies et al. (2017) to calibrate TGAS parallaxes. New work by Chen et al. (2017) has used RC stars identified using asteroseismology to find -1.626 ± 0.057 mag, and the precursor to our hierarchical Bayesian approach, H17, used TGAS parallaxes to find -1.61 ± 0.01 mag. Using the same method, H17 reported an absolute magnitude of 0.44 ± 0.01 mag in the *Gaia* G band.

Our RC magnitudes for both the K and *Gaia* G bands are much closer to those reported in literature when we used APOKASC-2 stars and temperatures alone. For the K band, we found values within 1σ of Chen et al. (2017) for $\Delta T_{\rm eff} \leq 20\,K$ when using corrections to the scaling relations, although our results are otherwise incompatible with the literature for K.However in the G band we found values for $\mu_{\rm RC}$ compatible with H17 when using APOKASC-2 stars for $\Delta T_{\rm eff}$ of 0 or $+10\,K$ both with and without corrections to the scaling relations. The disagreement found only in the K band could be due to our choice of bolometric corrections or corrections to the scaling relations, or due to H17's choice of extinction coefficient, which is twice as large as the coefficient we used in our *Gaia* models, and would bias the absolute magnitudes of their stars towards brighter values. Alternatively, it could be due

to H17 not accounting for known spatial correlations in parallax (Lindegren et al., 2016; Zinn et al., 2019a) or possible parallax zero-point offsets (Brown, 2018).

In Tables 4.9 and 4.10, we attempt to recreate the H17 work, albeit including parallax covariances, and find values for $\mu_{\rm RC}$ that are compatible with a temperature offset of $\Delta T_{\rm eff} < -50\,K$ for both photometric bands. Finally, allowing $\varpi_{\rm zp}$ to vary as a free parameter with loose prior constraints finds $\mu_{\rm RC} = -1.634 \pm 0.018$ mag in the K band and 0.546 ± 0.016 mag in the G band. These values imply that a shift to the temperature scales of $-50\,K$ or more is appropriate when using temperatures for seismology of the Red Clump.

Spread of the Red Clump

In principle, the spread of the RC, like its luminosity, is a property of a RC population and depends on the mass and metallicity of the sample (Girardi, 2016; Salaris & Girardi, 2002). Our hierarchical approach allows us to study the 'true' spread of the RC, by evaluating the uncertainties on individual measures of absolute magnitude.

As seen for the K band in Tables 4.1 and 4.2, the spread of the RC is consistent within 1σ for all perturbations of temperature, corrections to the scaling relations, and between both the Y18 and APOKASC-2 temperatures. This indicates that $\sigma_{\rm RC}$ is only weakly dependent on the choice of temperature scale, and that any effects of the APOKASC-2 sample only representing a small subset in metallicity are minimal for the K band.

The spread of the RC due to mass and metallicity is minimised in the 2MASS K band (Salaris & Girardi, 2002), which would lead us to expect a broader spread of the RC in the *Gaia* G band. We see this effect in Tables 4.3 and 4.4, where the reported spreads are ~ 4 to 6 times larger in magnitude. Surprisingly, we do not see the same consistency for the values of $\sigma_{\rm RC}$ for the G band, but instead find that the inferred value of $\sigma_{\rm RC}$ varies inversely with temperature beyond 1σ from $-50\,K$ to $50\,K$. This trend of $\sigma_{\rm RC}$ with $\Delta T_{\rm eff}$ is likely to be an effect of the bolometric correction, as we do not see a compatible trend in K. It should also be noted that

we would expect extinction to play a larger role in the G band, possibly contributing to this effect.

For the *Gaia* G band we also see that the value for $\sigma_{\rm RC}$ is lower for the APOKASC-2 sample than for the full Y18 sample. This reduction is likely because the APOKASC-2 sample draws temperatures from a uniform spectroscopy source (and thus temperature scale) whereas the Y18 temperatures come from a variety of sources, broadening the distribution of RC stars.

The similar hierarchical approach taken by H17 found a spread of 0.17 ± 0.02 mag in K and 0.20 ± 0.02 mag in G using TGAS parallaxes. The agreement within 1σ for the G band for the Y18 sample agrees with the inferred APOKASC-2 spread being an underestimate. The estimates found in our work for $\sigma_{\rm RC}$ in K are an order of magnitude lower. This is probably due to our sample size (increased from H17 by a factor of 5) and asteroseismology providing more precise measurements for these stars than TGAS (Davies et al., 2017), allowing the hierarchical method to more closely constrain the true underlying spread.

Tables 4.9 and 4.10 show the results of our attempt to recreate the H17 work, accounting for parallax covariances and including a parallax zero-point offset. Using *Gaia* parallaxes, we found a $\sigma_{\rm RC}$ in K that is larger than our value from seismology. The results presented in Tables 4.5 and 4.6, where the the seismic $\sigma_{\rm RC}$ in K has been applied as a prior on the *Gaia* model, show an inlier fraction Q that is lower than we would expect for this sample. This implies that *Gaia* DR2 is underestimating the uncertainties for stars considered 'outliers', and not including them in the inlier population.

For the G band, we found a value for $\sigma_{\rm RC}$ in agreement with our seismic value using APOKASC-2 temperatures. In this instance, as opposed to the results shown in Table 4.8 at similar $\sigma_{\rm RC}$, we find an inlier fraction Q in the expected range. This is probably due to the simultaneous inference of a more appropriate value for $\mu_{\rm RC}$, which is closer to values established in literature (H17). For this reason, the spreads

reported in Tables 4.4 and 4.10 are our best estimates for the 'true' spread of the RC in the G band.

With our measurement of $\sigma_{\rm RC} = 0.03\,$mag in the K band, we can use standard error propagation through Equation 4.26 (setting extinction to zero) to find that this spread yields a precision in distance of $\sim 1\%$ for our sample, subject to mass and metallicity. This is a factor of 5 improvement from the precision reported by H17. When using $\sigma_{\rm RC} = 0.14\,$mag for the G band we find a distance precision of $\sim 6\%$, in line with the findings by H17.

The *Gaia* parallax zero-point offset

The *Gaia* DR2 parallax zero-point offset, while small, can still have an effect on results, and is widely applied in studies using DR2 data (Luri et al., 2018; Bailer-Jones et al., 2018), with potentially far-reaching consequences (Shanks et al., 2019). The offset has been estimated through calibration with eclipsing binaries (Stassun & Torres, 2018), Cepheids (Riess et al., 2018), asteroseismology (Zinn et al., 2019a; Sahlholdt & Silva Aguirre, 2018), kinematics (Schönrich et al., 2019) and quasars (Lindegren et al., 2018).

In Tables 4.5, 4.6, 4.7 and 4.8 we present our inferred model parameters given our values for $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ found through asteroseismology at different temperature shifts $\Delta T_{\rm eff}$, effectively 'calibrating' *Gaia* DR2 to see what offset recovers a given set of RC parameters.

Figure 4.13 shows the posterior distributions for $\varpi_{\rm zp}$ given our seismic priors from different temperature shifts, where there is a clear trend of $\varpi_{\rm zp}$ with seismic $\mu_{\rm RC}$, and thus with temperature. This trend was also found in recent results by Khan et al. (2019), where a comparison of *Gaia* parallaxes and seismic distances obtained through the seismic scaling relations found that a temperature shift of $100\,K$ caused a shift in $\varpi_{\rm zp}$ of $10-15\,\mu$as for RC stars, although it should be noted that they found this effect largely reduced when using grid modelling techniques (Rodrigues et al., 2017).

It is also apparent in Figure 4.13 that the uncertainty on $\varpi_{\rm zp}$ is significant, and consistent for all model conditions, due to the parallax covariances presenting a systematic lower limit on parallax uncertainties for this sample. Given a $\mu_{\rm RC}$ in K closer to literature values, with the run corresponding to APOKASC-2 temperatures using $\Delta T_{\rm eff} = -50\,K$, we found a $\varpi_{\rm zp}$ within 1σ of the uncertainties on all literature values for $\varpi_{\rm zp}$ in the *Kepler* field discussed in this work. This is both an encouraging sign of a consistent $\varpi_{\rm zp}$ in the *Kepler* field, and further indication that seismology would be improved by reducing the temperature scale. For the *Gaia* G band, the run closest to the existing literature ($\Delta T_{\rm eff} = 0$) is consistent with all values for $\varpi_{\rm zp}$ besides Stassun & Torres (2018).

Given a selection of values for $\varpi_{\rm zp}$ reported in the literature, we applied informative priors on $\varpi_{\rm zp}$ in our *Gaia* model, and allowed $\mu_{\rm RC}$ and $\sigma_{\rm RC}$ to explore the parameter space freely. The results of this are shown in Tables 4.9 and 4.10, for the K and G bands respectively. The credible intervals for $\mu_{\rm RC}$ are shown in Figure 4.14. For both bands, we found that the choice of $\varpi_{\rm zp}$ from the literature had no impact beyond 1σ on either of the RC properties for any values used. When using a tightly constrained $\varpi_{\rm zp}$ of zero (in an attempt to recreate H17) we found the largest overall change. It is also interesting to note that for a prior corresponding to the Stassun & Torres (2018) value, the inferred value for $\varpi_{\rm zp}$ is reduced to lie closer to those found in other works for the *Kepler* field.

Finally, running the *Gaia* model with uninformative priors on both $\varpi_{\rm zp}$ and the RC parameters produced a parallax zero-point offset of $(-38 \pm 13)\,\mu$as in K and $(-42 \pm 13)\,\mu$as in G for the *Kepler* field. These values are consistent with one another and with the existing literature, and also agree with recent results by Khan et al. (2019) for RC stars in APOKASC-2. Given the uncertainties on the inferred values of $\varpi_{\rm zp}$, we see a fundamental uncertainty limit on *Gaia* parallaxes of $\sim 13\,\mu$as as a result of spatial covariances between parallaxes. Encouragingly, this implies that for our RC sample in the *Kepler* field, the choice of parallax-zero point offset does

not dramatically impact the inferred luminosities, given a proper treatment of the spatial parallax covariances. However, this may not generalize to populations more sparsely sampled in space, and in other magnitude ranges, given the known relation between the parallax zero-point offset, G band magnitude and colour (Zinn et al., 2019a; Lindegren et al., 2018).

Corrections to the seismic scaling relations

In Section 4.3.5, we have compared results with and without corrections to the $\Delta\nu$ seismic scaling relation, $f_{\Delta\nu}$, derived from Sharma & Stello (2016). It is known that stellar models do not not accurately reproduce the $\Delta\nu$ of the Sun (off by about 1%), due to the so-called surface effect (Christensen-Dalsgaard et al., 1988; White et al., 2011). Corrections to the scaling relation $f_{\Delta\nu}$ derived without accounting for the surface effect (i.e. Sharma & Stello, 2016) can produce radii that differ on the order of $\sim 2\%$ from methods that do (such as Rodrigues et al., 2017)[10]. As a check, we considered the impact that this may have on our inferred values for the RC magnitude.

To compare the calculated RC populations in the K and *Gaia* G bands, we obtained radii using $f_{\Delta\nu}$ obtained through Sharma & Stello (2016). We then used bolometric corrections for no temperature offset to calculate the absolute magnitudes using both those radii and those same radii reduced by both 1.6% and 2.4%. We found that a reduction on radius in the range $(2 \pm 0.4)\%$ resulted in a global shift toward brighter bolometric magnitudes by 44^{+9}_{-8} mmag.

In Tables 4.2 and 4.4 we report the absolute magnitude of the RC (for no temperature offset) in the APOKASC-2, $f_{\Delta\nu}$-corrected, sample of -1.69 mag in K and 0.45 mag in G. A shift of 0.04 mag applied to both bands is enough to reconcile our seismic results with those obtained through *Gaia* for both the K and G bands, as well as those from the literature. Note however that this is not the case when applied to the Y18 sample (see Tables 4.1 and 4.3), where this shift applied in both bands

[10] A small uncertainty in radius translates to a larger uncertainty in magnitude, as $L \propto R^2 {T_{\rm eff}}^4$.

would not be enough to reconcile the seismic results for the absolute magnitude of the RC with any measures both in this work or in the literature.

Calibrating *Gaia* and asteroseismology

Our initial aim with this work was to calibrate the *Gaia* parallax zero-point offset, $\varpi_{\rm zp}$, using asteroseismology. Given the large change in the absolute magnitude of the RC, $\mu_{\rm RC}$, with relatively small changes in temperature for our large RC population, and consequently the shift in inferred $\varpi_{\rm zp}$ given these values for $\mu_{\rm RC}$, it proved difficult to definitively calibrate *Gaia* parallaxes using seismology.

The reverse however, seems more possible. We found that the various parallax offsets reported in the literature, when used as informative priors on our *Gaia* model, all resulted in similar values for $\mu_{\rm RC}$ in both the 2MASS K and *Gaia* G bands (as shown in Tables 4.9 and 4.10), and inferred values for $\varpi_{\rm zp}$ that lie closer together for those literature values with large uncertainties (Stassun & Torres, 2018; Riess et al., 2018; Sahlholdt & Silva Aguirre, 2018). Imposing a prior for $\varpi_{\rm zp}$ to lie close to zero showed a departure beyond 1σ from the $\mu_{\rm RC}$ values found otherwise, indicating that $\varpi_{\rm zp}$ does have a measurable effect on the inferred RC luminosity. Finally, applying no strongly informative priors on the RC parameters nor $\varpi_{\rm zp}$ led to inferred values of $\mu_{\rm RC}$ and $\varpi_{\rm zp}$ being consistent with values in the literature, albeit with a large uncertainty of $\sim 13\,\mu$as on the parallax zero-point offset, implying a fundamental limit on the uncertainty on this offset given the spatial parallax covariances.

Given that the choice of parallax zero-point offset did not dramatically affect the inferred luminosity of the clump (see Tables 4.9 and 4.10 and Figure 4.14), we can reasonably use any value of $\varpi_{\rm zp}$ reported in the literature, including from this work, to attempt a calibration of seismology. Given the results for our runs on *Gaia* data with RC parameters constrained by seismology (Tables 4.6 and 4.8), we expect that $\mu_{\rm RC} = -1.634\,$mag in K and in $0.546\,$mag in G would be roughly consistent with a temperature offset $\Delta T_{\rm eff}$ between $\sim -100\,K$ and $\sim -70\,K$ for temperatures in the APOKASC-2 catalogue (which, as has been noted, are already

lower than those reported by Mathur et al., 2017, for the same stars). An offset of this size would fall within known systematic uncertainties on temperatures inferred from seismology (Slumstrup et al., 2019). However, it should be noted that this shift in temperature scale is degenerate with the scaling relations underestimating radii by $\sim 2\%$ compared to our estimates for radius using corrections by Sharma & Stello (2016), as discussed above 4.3.5.

In order to confirm these proposed shifts to temperature, we reran our asteroseismic model on our APOKASC-2 subsample for a range of temperature shifts extended down to $-110\,\mathrm{K}$ for both the K and *Gaia* G bands, with RC-corrected scaling relations. We found that when considering the K band, our calibration value for μ_{RC} from *Gaia* corresponds to within 1σ with a temperature shift of between -110 and $-70\,\mathrm{K}$. When considering the G band, the *Gaia* μ_{RC} corresponds to within 1σ for a shift between -70 and $-50\,\mathrm{K}$. Given that any calibrated correction to the temperature scale should be applied globally to the full APOKASC-2 subsample, we find that a temperature shift of $-70\,\mathrm{K}$ to the temperatures of our RC subsample of APOKASC-2 would produce seismic absolute magnitudes of the clump consistent with those found using *Gaia* DR2.

We only ran this test for the APOKASC-2 subsample, for which temperatures were all drawn from a uniform spectroscopic source. Since the temperatures for the full Y18 are not, claims about changes to temperature scales for this sample would be inappropriate.

The ability to make this inference reliably rests on our hierarchical treatment, as initially set out by H17, and treatment of the spatial correlations in parallax reported by Lindegren et al. (2018). As we improve our understanding of these correlations, our inferences using this and similar hierarchical models will improve. Similarly, it is known that population effects in age, metallicity and temperature, among others, have an effect on the inferred luminosity of the RC (Girardi, 2016). Our hierarchical model, can be further improved by accounting for these effects, as

well as by including parameters that check for consistent colours, as suggested by H17. As these hierarchical models improve in future work, so will our understanding of the RC, and our ability to calibrate asteroseismology.

4.4 Conclusions

Using two hierarchical models, based on the work by H17, we inferred the spread and position in absolute magnitude of a sample of 5576 Red Clump (RC) stars in the 2MASS K and *Gaia* G bands. We first did this using absolute magnitudes obtained through a completely distance-independent asteroseismic method, and probed systematics in asteroseismology by varying the temperatures of the sample, applying corrections to the scaling relations, and running our model on a subsample of stars with separate spectroscopic temperatures reported in APOKASC-2 (Pinsonneault et al., 2018). We then applied the results from seismology as strongly informative priors on the position and spread of the clump for our second hierarchical model. We applied this to *Gaia* DR2 data in order to see how the parallax zero-point varied, taking into account spatial correlations of parallaxes reported by Lindegren et al. (2018). We then applied strongly informative priors on the parallax zero-point in our *Gaia* model and allowed the RC parameters to roam more freely, to study the impact of published values for the zero-point offset on the RC. Finally, we performed a run of the *Gaia* model with no strongly informative priors on any parameters.

We leave the reader with the following conclusions:

1. By applying the H17 hierarchical model, with improvements to account for spatial correlations of parallaxes and to marginalize over the parallax zero-point offset ($\varpi_{\rm zp}$), we find a mean value for $\varpi_{\rm zp}$ in the *Kepler* field to be $-41 \pm 10\,\mu$as for our sample, consistent with all existing measures of $\varpi_{\rm zp}$ in the *Kepler* field. This offset results in a Red Clump magnitude of -1.634 ± 0.018 mag in K and 0.546 ± 0.016 mag in G for our sample.

2. Applying a hierarchical model to our sample of absolute magnitudes obtained

from asteroseismology, we find a spread of the RC in the 2MASS K band of ~ 0.03 mag independent of our changes made to the sample, an order of magnitude lower than the value reported previously using *Gaia* TGAS parallaxes in H17. This extremely small spread highlights the power of seismology and the potential of the RC in the K band as a standard candle. In the *Gaia* G band we find a spread of ~ 0.13 mag using APOKASC-2 temperatures, which is consistent with results found using *Gaia* TGAS parallaxes.

3. We find that a small global change in temperature ($\sim 10 - 20$ K) can affect the inferred absolute magnitude of the RC from seismology by more than 1σ, and is degenerate with the application of a correction $f_{\Delta\nu}$ to the seismic scaling relations.

4. We find values for the absolute magnitude of the RC from seismology to agree within 1σ with those inferred from *Gaia* DR2 parallaxes in both the K and G bands, only if a global temperature shift of ~ -70 K is applied to our RC subsample of APOKASC-2 stars. This shift is within expected systematic uncertainties on spectroscopic techniques. These differences are also degenerate with a shift in seismic radius of 2%, which is within the uncertainty imposed by choice of corrections to the scaling relations[11].

As shown in this research, the use of a hierarchical Bayesian mixture model for a population of RC stars continues to be an excellent tool for working with *Gaia* DR2 parallaxes. It allowed us to make the new additions of a parallax zero-point offset as a free parameter, and to treat spatial correlations between parallaxes. Using the hierarchical nature of the data (all RC stars being subject to the same physical constraints) and latent variables, we were able to craft a generative model that described the RC well enough to draw novel inferences about its spread, the *Gaia*

[11] Recent work by Zinn et al. (2019b) has suggested the latter is more likely the cause of this disagreement.

zero-point offset, and asteroseismic techniques.

The natural next step for this model is the same as it was for our basic example of a line: to make it more generative. First and foremost: our models here do not account for the spread on the RC imposed by metallicity and colour. Including those parameters as part of the modelling process is crucial in order to ensure that our inferences of the RC are not biased, and to constrain the exact luminosity and spread of the RC more precisely.

All analysis performed as part of this Chapter is stored in an online public repository[12].

[12] https://www.github.com/ojhall94/halletal2019

Chapter 5

Stellar Rotation and Gyrochronology with Asteroseismology

This chapter is adapted from an earlier draft of Hall et al. (in prep.), of which I am first author. The study has been submitted to Nature Astronomy pending an initial assessment by an editor. I performed the majority of the work, with exception of the development of stellar models of rotational evolution, which was done by Prof. Jennifer van Saders. Section 5.3.2 includes some text written by- or paraphrased from text written by- Prof. van Saders. The mixture model used to distinguish between different stellar models was co-developed with Dr. Guy Davies. In order to retain the integrity of this study as an independent piece of work, there is some introductory material repeated or expanded upon from Chapters 2 and 3.

5.1 Introduction

Gyrochronology is the study of a star's rotation period as a function of its age. Over the course of a star's main sequence (MS) lifetime, magnetic winds will cause it to lose angular momentum. The rate of loss depends on the depth of a star's convection zone (Barnes, 2010), which is a strong function of mass and temperature, and so gyrochronology relations are commonly described as a function of colour. Over the course of a star's MS evolution, its rate of rotation will rapidly settle on to a plane

in age-colour-rotation space (Barnes, 2007). As a result, knowing the rotation and colour of a star provides a promising way to estimate stellar age, to a precision of $\sim 10\%$ for stars most similar to the Sun (Meibom et al., 2015). This is not only useful for estimating what age a star is (e.g. for population studies, Claytor et al., 2020), but also for estimating what age a star *should appear as* in isolation (e.g. for locating merger products, Leiner et al., 2019).

Gyrochronology was first calibrated on young clusters (Barnes, 2003); populations with fixed age, but a range of colours and rotation rates measurable from spot modulation on the stellar surface. Long time-series observations with the *Kepler* mission (Borucki et al., 2010) first extended this to 2.5 Gyr. Calibrations have been done using the short time-series observations of the *K2* mission (Howell et al., 2014) for the M67 cluster at ~ 4 Gyr (Barnes et al., 2016; Gonzalez, 2016), although the accuracy of the measured rotation periods has been questioned (Esselstein et al., 2018).

Recently, *Kepler* observations of field stars (i.e. not in clusters), aged using asteroseismology (Silva Aguirre et al., 2015), found issues with gyrochronology when looking at rotation rates of stars older than the Sun for the first time. Stars beyond solar age appeared to experience a reduced rate of angular momentum loss compared to younger stars. No single gyrochronology relation that was a function of colour, age and rotation could reconcile observations of stars older and younger than the Sun (Angus et al., 2015; Nielsen et al., 2015; Davies et al., 2015).

Attempts have been made to explain this issue from both observational and theoretical points of view. First, van Saders et al. (2016) proposed the theory of *weakened magnetic braking*, where stars beyond a certain critical Rossby number[1] experience a change in the efficiency of rotational braking due to magnetic winds. The exact mechanism by which this happens is still subject to debate, but my be a result of a shift in magnetic field morphology (see e.g. Réville et al., 2015; Garraffo

[1]The Rossby number is defined as $Ro \equiv P/\tau_{\rm cz}$, where $\tau_{\rm cz}$ is the convective turnover timescale, and P is the rotation. The Rossby number scales inversely with stellar activity.

et al., 2016; Metcalfe et al., 2016; van Saders et al., 2016; See et al., 2019).

Second, In their large-scale survey of rotation from star spot modulation, McQuillan et al. (2014) pointed out a lack of old, slowly rotating stars older than the Sun, possibly introducing an observational bias. This apparent lack of old stars was confirmed using full magnetic braking models by (Matt et al., 2015). More recently, van Saders et al. (2019) examined the possibility that above a Rossby number of ~ 2, the detection of a rotation period from star spot modulation becomes less likely, creating a trend that could be misinterpreted as the weakened magnetic braking. Finally, Lorenzo-Oliveira et al. (2019) found no signatures of weakened magnetic braking for ages below 5.3 Gyr, when studying solar analogues.

Asteroseismology – the study of stellar oscillations – can contribute to resolving this problem. The high-quality time series data from the *Kepler* mission have allowed for precise measures of radii and masses for main sequence stars (Chaplin et al., 2013), and stellar age by studying individual oscillation frequencies and comparing these to models (Silva Aguirre et al., 2015, 2017; Creevey et al., 2017; Serenelli et al., 2017). But most interesting for the current state of the field is the ability to obtain independent measures of stellar rotation, by measuring modes of oscillation split by a star's spin. An asteroseismic measure of rotation does not require detection of magnetic surface activity, unlike conventional techniques. While spot modulation needs significant activity to be measurable, asteroseismology can measure rotation for quiescent stars. Asteroseismology can therefore both probe a broader age range and stars with Rossby numbers above a proposed critical threshold where surface detection can no longer be measured.

While measurements of spot modulation will probe the rotation period of star spots on the stellar surface, asteroseismology instead measures rotation in the region where asteroseismic pressure modes (p-modes) are most sensitive (see e.g. Davies et al., 2015). In main sequence stars, this is heavily weighted towards the near-surface of the star (see Lund et al., 2014, for a comprehensive theoretical justifica-

tion), and empirical evidence has shown that measures of rotation from asteroseismology do not diverge from those measured from spot modulation or spectroscopy in a statistically significant way for stellar ensembles (Nielsen et al., 2015; Benomar et al., 2015). However it should be noted that for individual stars the effects of latitudinal differential rotation may cause meaningful differences in rotation measurements (as observed in HD 173701, Karoff et al., 2018).

Asteroseismology is the ideal tool for distinguishing between whether weakened braking occurs, or whether the paucity of old, slowly rotating stars is merely detection bias. In this work, we extracted oscillation frequencies and new internal rotation estimates for 94 of the highest signal-to-noise stars observed by *Kepler*, referred to as the 'Kages' (Silva Aguirre et al., 2015; Davies et al., 2016) and LEGACY (Silva Aguirre et al., 2017; Lund et al., 2017) samples. Using the self-consistent age estimates reported in Silva Aguirre et al. (2015, 2017), we made comparisons to models of stellar rotational evolution for different magnetic braking scenarios (van Saders et al., 2019). Through comparisons to measures of rotation from star spot modulation, we used our new sample to revisit the study of differences between surface and seismic rotation measurements, as in Nielsen et al. (2015) and Benomar et al. (2015). Finally, we have drawn conclusions on the consequences of these new data for gyrochronology, and detail the next steps required to make gyrochronology more robust for all MS stars.

This Chapter is laid out as follows: Section 5.2 details and justifies our data set. Section 5.3 details the asteroseismic fitting techniques and the models used to study our new rotation measurements. We present our results in Section 5.4 and discuss them in context of similar work in Section 5.5. In Section 5.6 we present our conclusions and recommendations for future work.

5.2 Data

5.2.1 Asteroseismic Data

Our aim is to obtain asteroseismic rotation rates of main sequence stars to high precision and accuracy. To do so robustly requires detection of multiples of both the dipole ($\ell = 1$) and quadrupole ($\ell = 2$) oscillation modes, of which the latter have significantly lower signal-to-noise. In order to ensure quadrupole modes are measurable, we used a sample of the highest signal-to-noise targets observed with *Kepler*; combining the 'Kages' (Silva Aguirre et al., 2015; Davies et al., 2016) and LEGACY (Silva Aguirre et al., 2017; Lund et al., 2017) catalogues[2]. In cases where both catalogues contained the same target, we used the stellar parameters reported in LEGACY. These two samples together contain 95 stars, all of which have been subject to an extensive asteroseismic analysis with reported age estimates from comparisons to stellar models through multiple pipelines.

Neither the LEGACY nor Kages mode extraction papers report inclination angles or asteroseismic rotation, instead focusing on robust inference of mode frequencies for detailed comparisons to models. To extract asteroseismic rotation, we need to repeat the extraction of mode frequencies from these data. As we are using the same *Kepler* data, we want to avoid explicitly using the full posterior information from the LEGACY and Kages papers above as our priors. Instead, we can use their posterior data as first guesses for our parameters.

For our asteroseismic power-spectrum data we use the unweighted power spectra from the KASOC pipeline (Handberg & Lund, 2014)[3]. We do not apply any additional treatment to these data. For 16 Cyg A & B (KIC 12069424 and KIC 12069449) we use the KEPSEISMIC light curves (García et al., 2011)[4], which have

[2]For these catalogues, Davies et al. (2016) and Lund et al. (2017) cover the mode extraction through frequency fitting, and Silva Aguirre et al. (2015, 2017) cover the modelling using the mode frequencies to obtain stellar parameters.

[3]Obtainable from the KASOC webpage.

[4]Obtainable from MAST.

significantly better signal-to-noise for these two stars.

For our sample, we used the asteroseismic ages obtained by `BASTA` (BAyesian STellar Algorithm, Silva Aguirre et al., 2015) in the Kages and LEGACY catalogues. These ages have been obtained by comparisons of measured oscillation properties to stellar models (i.e. detailed modelling), accounting for an expanded range of metallicities. `BASTA` is thoroughly compared to four other seismic modelling techniques in Silva Aguirre et al. (2017). While uncertainties found through `BASTA` are typically higher than for other techniques, only `BASTA` and `ASTFIT` (Aarhus STellar Evolution Code Christensen-Dalsgaard, 2008) recover the radius, mass and age of the Sun, when applied to solar data. Although the uncertainties on `ASTFIT` ages are overall lower, they are not reported in Silva Aguirre et al. (2015) for the Kages sample. In order to maintain a stellar age sample that is internally consistent between both catalogues (and to err on the side of caution with larger uncertainties), we use age results from `BASTA` for both the Kages and LEGACY samples.

For our stellar masses we use asteroseismic model masses obtained by `BASTA` reported in Kages and LEGACY, in order to maintain internal consistency with the age measurements. We note that age and mass posteriors from `BASTA` are correlated, but choose not to account for the unpublished correlations in this work. As described in the catalogue papers, for Kages stars atmospheric properties ($T_{\rm eff}$ and [Fe/H]) were measured through high-resolution spectroscopy reported in Huber et al. (2013a). For LEGACY stars, atmospheric properties were similarly taken from Buchhave & Latham (2015) for most stars in the catalogue, and complemented by other values from the literature for the remaining stars (see Table 3 of Silva Aguirre et al., 2017).

Stellar classification

Between the Kages and LEGACY catalogues, we have a stellar sample spanning surface gravities of $3.8\,\mathrm{dex} < \log(\mathrm{g}) < 4.6\,\mathrm{dex}$ and effective temperatures of $5000\,\mathrm{K} < T_{\rm eff} < 6700\,\mathrm{K}$. For the purposes of studying gyrochronology, we wanted to be able to single out stars for which the loss of angular momentum via magnetized winds is

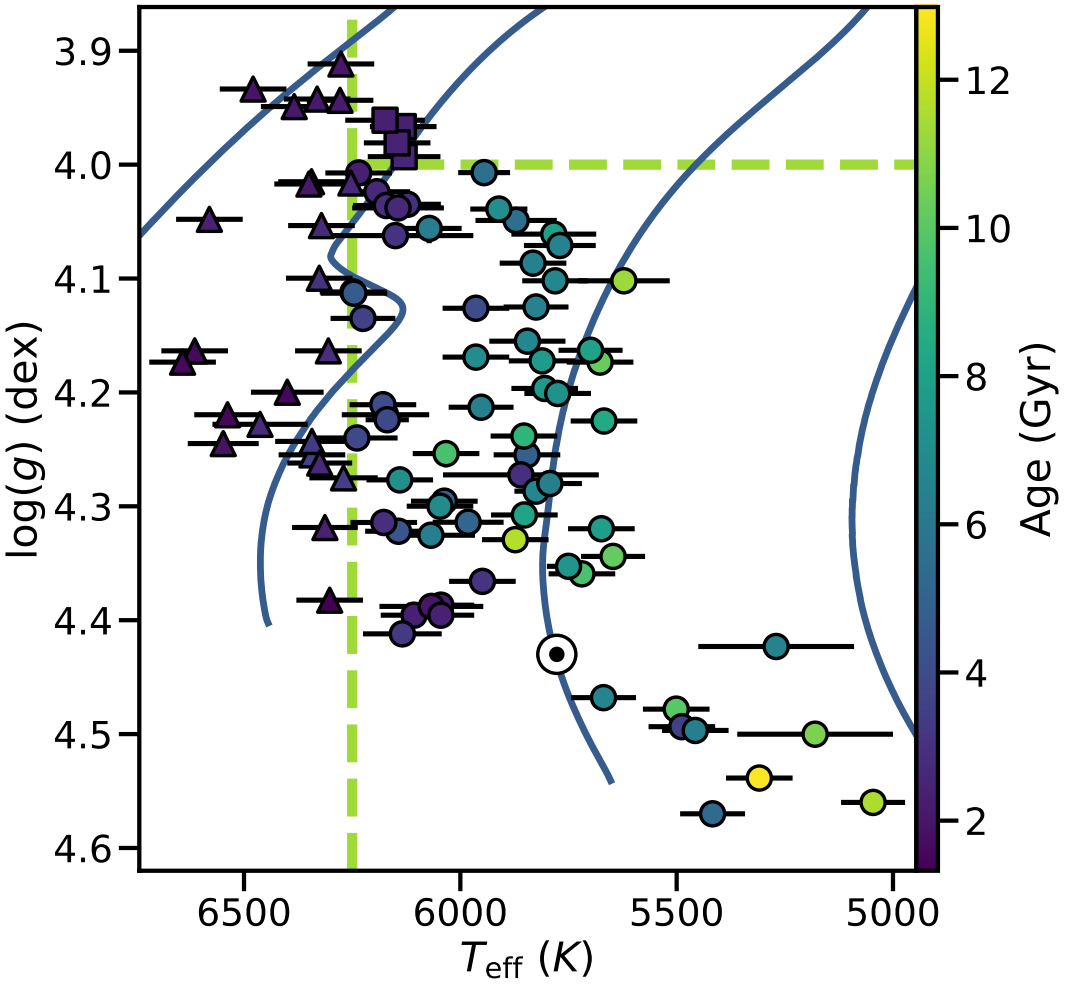

Figure 5.1: Our sample of stars from the LEGACY and Kages catalogues (Silva Aguirre et al., 2015, 2017) coloured by asteroseismic age. All properties shown are as reported in LEGACY and Kages. The dashed lines indicate our classification boundaries: main sequence (circles, $T_{\rm eff} < 6250\,{\rm K}$, $\log g < 4\,{\rm dex}$), sub-giants (squares, $T_{\rm eff} < 6250\,{\rm K}$, $\log g > 4\,{\rm dex}$), and 'hot' stars (triangles, $T_{\rm eff} > 6250\,{\rm K}$). The Sun is denoted by the '$\odot$' symbol, for clarity, and has an age of 4.6Gyr. The solid lines are evolutionary tracks generated using MESA (Paxton et al., 2018), for a metallicity of $Z = 0.01493$ and helium content of $Y = 0.26588$. Left to right, they represent masses of 1.5, 1.25, 1 and $0.75\,M_{\odot}$.

the dominant source of rotational evolution. This ruled out stars with $T_{\rm eff} > 6250\,{\rm K}$, the so called 'Kraft Break' (Kraft, 1967; van Saders & Pinsonneault, 2013; Angus et al., 2015). The Kraft Break roughly separates stars with relatively large convective envelopes (at lower temperatures) from those with thin and tenuous convective envelopes (at higher temperatures), at which point the efficiency of magnetic winds as angular momentum transport is reduced. We also wanted to avoid stars that have started to evolve up onto the sub-giant branch, where the outer envelope begins to expand. The interior angular momentum transport associated with these structural changes is uncertain, making them challenging targets for gyrochronology. While there are no identified mixed dipole modes[5] in the 95 selected stars from the Kages and LEGACY catalogues, indicating an evolved structure, we can still apply a conservative classification of stars with $\log g < 4.0\,{\rm dex}$, the same selection as applied by García et al. (2014). Our conclusions are largely insensitive to these categorical assignments, and they are intended solely to help explore results category-by-category.

Of our full sample of 95 targets, 67 stars fall in the range $\log g > 4.0\,{\rm dex}$, $T_{\rm eff} < 6250\,{\rm K}$, which we class as belonging to the main sequence (hereafter 'MS stars'). 4 targets in our sample have $\log g < 4.0\,{\rm dex}$, $T_{\rm eff} < 6250\,{\rm K}$ which we class as being sub-giants (hereafter 'SG stars'). The remaining 24 stars in our sample have $T_{\rm eff} > 6250\,{\rm K}$, which we class as being hot MS stars (hereafter 'Hot stars'). Our final sample, as classified, can be seen in Figure 5.1.

5.3 Method

5.3.1 Asteroseismic Model

In order to extract signatures of stellar rotation from the asteroseismic mode frequencies, we built a model that simultaneously treats the convective background

[5]In evolved stars, gravity modes propagating in the radiative interior will interfere with non-radial p-modes, changing their structure and creating so-called 'mixed modes' (see e.g. Scuflaire, 1974; Kjeldsen et al., 1995; Bedding et al., 2010).

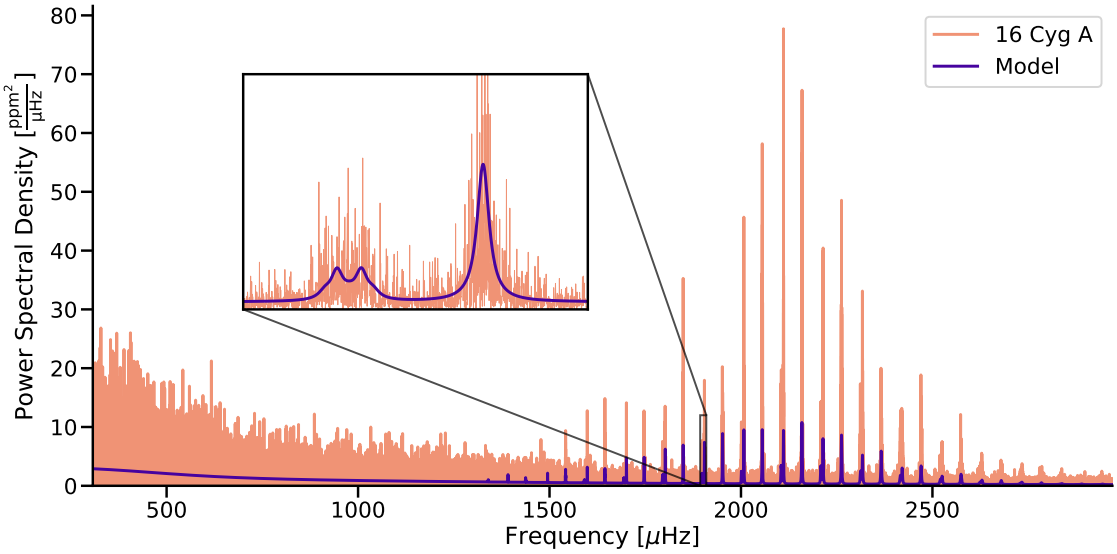

Figure 5.2: A power spectrum of four years of *Kepler* observations of 16 Cyg A (KIC 12069424). Plotted over the top is the model resulting from the fit to the data described in this work. The model implements both the mode frequencies, seen on the right hand side of the plot, and the convective background, the effects of which are seen on the left. We have cropped out low frequencies for clarity. *Inset*: A zoom in on a radial (right) and quadrupole (left) ($\ell = 0, 2$) pair of modes. The quadrupole mode is split into five components by the star's rotation, displaying two distinct peaks. The height and spacing of the mode components is a function of the star's rotational splitting (0.56 μHz, equivalent to $P_{\rm rot} = 20.5$ days) and angle of inclination (45°).

($B(\nu)$), the oscillations ($O(\nu)$), and the white noise (W) in the style of e.g. Davies et al. (2015). Our data, observed with *Kepler*, are also subject to the apodization (see Section 3.2.6) of signals in the frequency-power spectrum (Chaplin et al., 2011). The apodization in power is given by

$$\eta^2(\nu) = \mathrm{sinc}^2 \left(\frac{\pi}{2} \frac{\nu}{\nu_{\rm nyq}} \right) , \tag{5.1}$$

where $\nu_{\rm nyq}$ is the Nyquist frequency for the *Kepler* short cadence, which was treated as a free parameter in our model. Apodization only affects signals with characteristic timescales, meaning that it does not affect the white noise level, only the oscillations and convective background components. Given the above, our comprehensive model for the frequency-power spectrum is

$$M(\nu) = W + \eta^2(\nu)[O(\nu) + B(\nu)] \, . \tag{5.2}$$

Convective Background

To model the convective background we used three Harvey components (Harvey, 1985), which express the background in power as Lorentzian-like functions centred on zero frequency. The Harvey components take the form

$$H(\nu, a, b, x) = \frac{4a^2/b}{1 + (2\pi b\nu)^x}\,, \tag{5.3}$$

where a and b are the free parameters in our model, and x is fixed. The three Harvey components together form our background function as

$$B(\nu) = H(\nu, a, b, x = 4) + H(\nu, c, d, x = 4) + H(\nu, j, k, x = 2)\,, \tag{5.4}$$

where we have differentiated the parameters for the different components (Kallinger et al., 2014).

Modes of Oscillation

Modes of oscillation appear in the power spectrum as Lorentzian peaks (Chaplin & Basu, 2017). Due to stellar rotation, each mode with an angular degree of $\ell > 0$ is split into its $(2\ell + 1)$ Lorentzians components, labelled by m, the azimuthal order. For all $\ell = (0, 1, 2)$ modes identified in Davies et al. (2016) and Lund et al. (2017) we add a (set of) Lorentzian(s) to our model, building a composite model representing all visible modes. The construction of our oscillation model takes the form

$$O(\nu) = \sum_n \sum_\ell \sum_{m=-\ell}^{\ell} \frac{H_{n,\ell,m}}{1 + \frac{4}{\Gamma_{n,\ell}^2}(\nu - \nu_{n,\ell,m})^2}\,, \tag{5.5}$$

where n is the radial order of a mode (i.e. the overtone number of the oscillation), $H_{n,\ell,m}$ is the height of the mode, $\Gamma_{n,\ell}$ is the linewidth of the mode (approximated to be equal for all split components at a single n and ℓ) and $\nu_{n,\ell,m}$ is the frequency of the mode. The range of n differs per star depending on how many radial orders were reported in LEGACY or Kages, and the range of ℓ depends on how many angular

degrees were reported for the corresponding radial order.

Mode Frequencies and Rotational Splitting

The mode frequencies of main sequence stars are described by the well known asymptotic expression (see e.g. Tassoul, 1980; Vrard et al., 2016). The asymptotic expression defines the locations of the modes as regularly spaced, with structured deviation around $\nu_{\rm max}$, the frequency of maximum oscillation amplitude (see also Section 2.3.3). Expanded to include rotational splitting and azimuthal order, the asymptotic expression takes the form

$$\nu_{n,\ell,m} = \Delta\nu \left(n + \epsilon + \delta\nu_{0,\ell} + \frac{\alpha}{2}(n - \frac{\nu_{\rm max}}{\Delta\nu} + \epsilon)^2 \right) + m\nu_s \,, \tag{5.6}$$

where $\Delta\nu$ is the large frequency separation, n is the radial order, ϵ is a a phase offset (to treat the breaking-down of this relation at low frequencies), $\delta\nu_{0,\ell}$ is the small frequency separation between two oscillations of different ℓ at the same radial order n, α describes the curvature of the spacing around $\nu_{\rm max}$, and ν_s is the rotational splitting in μHz. Note that here we have expressed the small separation $\delta\nu_{0,\ell}$ as a fraction of $\Delta\nu$. In order to improve the computational efficiency of this analysis, we fixed $\Delta\nu$ to the values reported in LEGACY and Kages.

Instead of calculating mode frequencies directly from Equation 5.6 for the model, we instead treated the individual frequencies as latent parameters, drawn from Equation 5.6 (see also Chapter 4). The parameters $\nu_{n,\ell,m}$ were allowed to vary within an uncertainty σ_ℓ, which has a single value for each angular degree and also varied as a free parameter. This allowed us to marginalise out small shifts in frequency due to sudden changes in the stellar structure (called 'glitches'). The mode frequency latent parameters were drawn from a normal distribution using Equation 5.6 as a mean function, as

$$\nu_{n,\ell,m} \sim \mathcal{N}(\nu_{n,\ell,m}, \sigma_\ell) \,, \tag{5.7}$$

where the expression $\nu_{n,\ell,m}$ on the right hand side represents the contents of Equation 5.6. As is the case in previous Chapters, $\mathcal{N}$ represents a normal distribution of the form $\mathcal{N}(\mu, \sigma)$, where μ is the mean and σ is the standard deviation, and the symbol '$\sim$' indicates that the parameters on the left hand side of the equation (the frequencies of the modes in our model) are drawn from the probability function on the right hand side.

Mode Linewidth

The linewidths of asteroseismic p modes vary roughly as a function of mode frequency, and do so slowly relative to $\Delta\nu$. This can be expressed as an empirical relation, as seen in e.g. Davies et al. (2014); Appourchaux et al. (2016), and for the LEGACY sample in Lund et al. (2017). However, this relation has six free parameters, none of which are directly relevant to this work. Instead of fitting this relation, we chose to employ a more flexible Gaussian Process (GP, Rasmussen & Williams, 2006) to act as a prior on the linewidths. This can be considered as modelling the linewidths as correlated measurements. As our GP covariance kernel we used a Squared Exponential Kernel to capture the slight periodicity of linewidth with frequency, as

$$K_{i,j} = \rho^2 \exp\left[-\frac{(n_{\mathrm{f},i} - n_{\mathrm{f},j})^2}{2L^2}\right], \tag{5.8}$$

where n_f is the fractional radial order of a given mode. For each star, the overtone numbers n were rescaled to be between 0 (for the lowest n) and 1 (for the highest n)[6]. This approximation was used to describe the change in linewidth as a function of frequency without depending on the exact frequencies of the modes. $K_{i,j}$ represents an element of the covariance matrix $\underline{K}$, describing the covariance between two values of linewidth at different fractional radial orders. The GP kernel has two hyperparameters: ρ, which determines the spread of the kernel, and L, which de-

[6]Overtones of $\ell = 2$ modes were increased by 1 to ensure this approximation applied to all modes.

termines the length scale in terms of $n_{\rm f}$. The length scale was significantly larger than the large frequency separation ($\Delta\nu$) in all cases, and so we considered the use of fractional radial orders a valid approximation in this model.

A linear function was used for the mean of the GP, as

$$\mu = m \times n_{\rm f} + c\,, \tag{5.9}$$

where m and c are the slope and intercept of the line. The linewidth latent parameters were then drawn from the multivariate probability distribution

$$\Gamma \sim \mathcal{N}(\mu, \underline{K})\,, \tag{5.10}$$

where Γ represents the linewidths of all the modes in the model. The parameters m, c and ρ were marginalised over, whereas L was fixed to a pre-determined value.

Mode Heights and Angle of Inclination

The height in power of each mode, $H_{n,\ell,m}$, varies not only as a function of distance in frequency from $\nu_{\rm max}$, but also due to observation conditions, such as inclination angle and passband. In our model, we treated $H_{n,\ell,m}$ as a deterministic parameter, as

$$H_{n,\ell,m} = \varepsilon_{\ell,m}(i) \times \frac{2(A_{n,\ell})^2}{\pi\Gamma_{n,\ell}}\,, \tag{5.11}$$

where $\varepsilon_{\ell,m}(i)$ modulates the height as a function of inclination angle (i) (see below), and $A_{n,\ell}$ and $\Gamma_{n,\ell}$ are the mode amplitude and linewidth respectively for a given radial order and angular degree. Instead of modelling and modulating height directly, we instead sampled in amplitude and linewidth. In order to mitigate the correlations between height and linewidth in our sampling process (Toutain & Appourchaux, 1994).

As done above for the mode frequencies and linewidths, the mode amplitudes

$A_{n,\ell}$ were also treated as latent parameters drawn a probability distribution governed by hyperparameters. We used a Gaussian function $G(\nu)$, centred on $\nu_{\rm max}$, so that

$$G(\nu) = A \times \exp\left[-\frac{(\nu - \nu_{\rm max})^2}{2w^2}\right], \tag{5.12}$$

where A is the amplitude at $\nu_{\rm max}$, and w is the width of the Gaussian function, both free parameters in our model. The parameter $\nu_{\rm max}$ in this equation is the same as that used in Equation 5.6. The mode amplitude latent parameters were then drawn from the probability distribution

$$A_{n,\ell} \sim \mathcal{N}(G(\nu_{n,\ell}) \times V_\ell, \sigma_A), \tag{5.13}$$

where V_ℓ is a free parameter for the mode visibility of different angular degrees, which should be consistent for all *Kepler* observations. The mode visibility for V_0 is fixed at 1, and $V_{1,2}$ are treated as free parameters. The parameter σ_A, the uncertainty on the distribution, is the same for all amplitudes regardless of angular degree.

The angle of inclination of the star with respect to earth changes the net perturbation by a given mode when integrated across the stellar disc, changing the amplitudes of modes of different azimuthal orders. This is a geometric problem, and is expressed by $\varepsilon_{\ell,m}(i)$, which takes the form (Gizon & Solanki, 2003)

$$\varepsilon_{\ell,m}(i) = \frac{(\ell - |m|)!}{(\ell + |m|)!}\left[P_\ell^{|m|}(\cos(i))\right]^2, \tag{5.14}$$

where $P_\ell^{|m|}$ are associated Legendre functions. For the first three angular degrees, they take the form (Handberg & Campante, 2011)

$$
\begin{aligned}
\varepsilon_{0,0}(i) &= 1\,, \\
\varepsilon_{1,0}(i) &= \cos^2(i)\,, \\
\varepsilon_{1,\pm1}(i) &= \frac{1}{2}\sin^2(i)\,, \\
\varepsilon_{2,0}(i) &= \frac{1}{4}(3\cos^2(i)-1)^2\,, \\
\varepsilon_{2,\pm1}(i) &= \frac{3}{8}\sin^2(2i)\,, \\
\varepsilon_{2,\pm2}(i) &= \frac{3}{8}\sin^4(i)\,,
\end{aligned}
\tag{5.15}
$$

where the sum of available components for a single ℓ are normalized to one, to maintain the equipartition of energy.

Likelihood Function

If data have Gaussian distributed noise in the time domain, it will appear in the frequency domain distributed in power as a χ^2 distribution with two degrees of freedom (χ^2_2 hereafter, Woodard, 1984). The noise properties of χ^2_2 distributed data are multiplicative, and require a specific treatment when fitting to these data. As our frequency bins are independent, we used the likelihood function described in Anderson et al. (1990),

$$
\ln p(\mathcal{P}|M) = \sum_{j=0}^{N-1}\left[\ln[M_j(\nu)] + \frac{\mathcal{P}_j}{M_j(\nu)}\right]\,, \tag{5.16}
$$

where $\mathcal{P}$ is the power spectral density (and thus our data), and $M(\nu)$ represents our model. The subscript j denotes an individual datum, for a total of N data. We have omitted the dependence of the model $M(\nu)$ on our parameters, for clarity. This equation is functionally equivalent to the evaluation of a gamma distribution of the form $\gamma(\mathcal{P}|1,\beta)$, where $\beta = 1/M(\nu)$, which is the implementation we used in our sampling process.

Fitting the background

Fitting the convective background, apodization and white noise component must be done for the full range of the power spectrum in order to be accurately constrained. However fitting a single model to the full range of frequencies is computationally inefficient when we are mainly interested in the modes of oscillation, which occupy a relatively small range of frequencies.

In order to speed up this process, we first fit the background independently to a subset of our data for each star. We created this subset by removing all frequencies within in a range $0.1 \times \Delta\nu$ below and above the minimum and maximum mode frequencies reported in LEGACY and Kages. For KIC 3427720 we also removed frequencies in the range $90\,\mu\text{Hz} < \nu < 400\,\mu\text{Hz}$, where there were large peaks not of asteroseismic origin, skewing the background fit.

For each star we fit the model function (see Eq. 5.2), as

$$M_B(\nu) = W + \eta^2(\nu) B(\nu)\,, \tag{5.17}$$

where $B(\nu)$ is the same background model described in Equation 5.4. The parameter components of our background fit are then

$$\phi_B = \{\log(a), \log(b), \log(c), \log(d), \log(j), \log(k), W, \nu_{\text{nyq}}\}\,,$$

where we sampled the parameters of the Harvey components in log space.

We fit our model to the background data using `PyStan` (Carpenter et al., 2017; Van Hoey et al., 2013), run for 10,000 iterations on each star[7].

Obtaining First Guesses and Prior Values

In order to utilise some of the prior measurements of our targets without using them as hard constraints on our parameters, some of our model equations were fit to

[7]These `PyStan` runs were initiated with a random seed of 11, as were all other random processes in this Chapter.

LEGACY and Kages data to obtain first guesses and mean values on hyperparameter priors.

For first guesses for parameters in the asymptotic expression, we fit Equation 5.6, *not* including the rotational component $m\nu_s$, to the $\ell = (0, 1, 2)$ mode frequencies reported in LEGACY and Kages for each star, using their reported uncertainties. This yielded estimates of $\hat{\epsilon}$, $\widehat{\delta\nu_{01}}$, $\widehat{\delta\nu_{02}}$ and $\hat{\alpha}$, where the hat symbol '$\widehat{\ }$' indicates a prior value (e.g. $\hat{\nu}_{\rm max}$ is taken from LEGACY or Kages). While not precise, as we did not mitigate any perturbations due to glitches (sudden changes in interior sound speed, which the latent parameter treatment takes care of in the full model), these rough results act as functional first guesses and prior mean values. The relation was fit to each star using PyMC3 (Salvatier et al., 2016) using 5000 iterations on 4 chains.

To obtain first guesses for the parameters used to set the GP prior on linewidth, we fit a GP constructed as in Equation 5.10 to the linewidths of the $\ell = 0$ modes reported in LEGACY. Linewidths were not reported for the other angular degrees in LEGACY, but the estimates may be generalised to other ℓ, as linewidth is a strong function of frequency. The relation was fit to each star using PyMC3 using 2500 iterations on 4 chains.

Fitting the LEGACY linewidths yielded rough estimates of $\hat{m}$, $\hat{c}$, $\hat{\rho}$ and L for each star. As is noted in Equation 5.10, L was fixed to this fit value when fitting our full model to our data. For stars in Kages, for which no linewidths were reported, we instead fixed these prior values to $\hat{m} = 1$, $\hat{c} = 0.5$, $\hat{\rho} = 0.1$, and the length scale to $L = 0.3$. These values were chosen to reflect those found for the LEGACY stars.

Finally, we obtained prior values for the Gaussian function describing the distribution of mode amplitudes around $\nu_{\rm max}$. For the amplitude, we used the mode amplitude of the highest peak in the spectrum, which was typically at or near $\nu_{\rm max}$. For the mode width we used the empirical function (Lund et al., 2017)

$$\hat{w} \approx 0.25 \times \nu_{\rm max}\,, \tag{5.18}$$

where $\hat{w}$ is the first guess for the width. For the mode visibilities, we use $\hat{V}_1 = 1.2$ and $\hat{V}_2 = 0.7$, which roughly reflect the results for these parameters reported in the LEGACY catalogue.

Priors on our Hyperparameters

Given our first guesses and measured prior values, we define the prior probabilities of the hyperparameters on which our model depends. For mode frequencies, these are

$$\begin{aligned}
\nu_{\rm max} &\sim \mathcal{N}(\hat{\nu}_{\rm max}, 10)\,, \\
\epsilon &\sim \mathcal{N}(\hat{\epsilon}, 1)\,, \\
\alpha &\sim \ln\mathcal{N}(\ln(\hat{\alpha}), 0.01)\,, \\
\delta\nu_{01} &\sim \ln\mathcal{N}(\ln(\widehat{\delta\nu_{01}}), 0.1)\,, \\
\delta\nu_{02} &\sim \ln\mathcal{N}(\ln(\widehat{\delta\nu_{02}}), 0.1)\,, \\
\sigma_{0,1,2} &\sim \mathcal{C}_{1/2}(\beta = 2)\,,
\end{aligned} \tag{5.19}$$

where $\ln\mathcal{N}$ represents a log-Normal distribution and $\mathcal{C}_{1/2}$ represents a half-Cauchy distribution[8], and other symbols are as described above. All three hyperparameters $\sigma_{0,1,2}$ describing the uncertainty on the latent parameters of different angular degree were subject to the same prior.

For the mode linewidths, our hyperparameter priors took the form

[8]The half-Cauchy distribution ensures the standard deviations do not inflate to large numbers, and is generally well-behaved close to zero in the case of stars with little deviation from Eq. 5.6 (Gelman, 2006).

$$
\begin{aligned}
m &\sim \mathcal{N}(\hat{m}, 1)\,, \\
c &\sim \mathcal{N}(\hat{c}, 1)\,, \\
\rho &\sim \ln\mathcal{N}(\ln(\hat{\rho}), 0.1)\,,
\end{aligned}
\tag{5.20}
$$

where the conventions are the same as above. For our mode amplitudes, they took the form

$$
\begin{aligned}
w &\sim \ln\mathcal{N}(\ln(\hat{w}), 10)\,, \\
A &\sim \ln\mathcal{N}(\ln(\hat{A}), 1)\,, \\
V_1 &\sim \ln\mathcal{N}(\ln(\hat{V}_1), 0.1)\,, \\
V_2 &\sim \ln\mathcal{N}(\ln(\hat{V}_2), 0.1)\,, \\
\sigma_A &\sim \mathcal{C}_{1/2}(\beta = 1)\,.
\end{aligned}
\tag{5.21}
$$

As the convective background had already been fit to our data excluding the region where the modes are present, the results from that fit could be used as extremely informative priors on our fit to the region containing the modes, where there is little information present to constrain the background. To do so, we modelled the background parameters ϕ_B in our full model as being drawn from a multivariate normal distribution as

$$
\phi_B \sim \mathcal{N}(\hat{\phi}_B, \underline{\Sigma}_{\hat{\phi}_B})\,,
\tag{5.22}
$$

where $\hat{\phi}_B$ are the median values of our posterior distributions from our prior background fit, and $\underline{\Sigma}_{\hat{\phi}_B}$ is the full covariance matrix of all the posterior distributions from our prior background fit, taking into account the correlations between the different Harvey components.

Finally, but most importantly, we defined the priors on the rotational parameters:

the mode splitting (ν_s), and the inclination angle (i). In order to give these an appropriate treatment, we made two reparametrizations. First, we sampled the projected rotational splitting, $\nu_{\rm s}\sin(i)$, which is more efficiently sampled due to the strong correlations between i and $\nu_{\rm s}$ (Ballot et al., 2006, 2008). A prior was applied over this as

$$\nu_s \sin(i) \sim \ln\mathcal{N}(\ln(0.75), 0.75)\,, \tag{5.23}$$

where conventions are as above. This subjective prior was chosen to reflect that most stars will have a solar-like rotation, with a long tail to allow for the fastest rotators. Second, we sampled in $\cos(i)$, and gave it a prior of

$$\begin{aligned} \cos(i) &\sim \mathcal{U}(0,1)\,, \text{which is equivalent to stating} \\ p(i) &= \sin(i)\,, \end{aligned} \tag{5.24}$$

where the $\mathcal{U}(0,1)$ indicates a uniform prior between 0 and 1. Using a uniform prior on $\cos(i)$ allowed us to account for the geometric effect that stars with a large inclination angle with respect to us are more common (Chaplin & Basu, 2017).

Fitting procedure

Using our prior information and model described above, we fit Equation 5.2 to our data $\mathcal{P}$, using the likelihood function described in Section 5.3.1.

In order to speed up the fitting process, we only applied our model to the region of the power spectrum that contains visible modes of oscillation. We created this subset by removing all frequencies outside a range $0.25 \times \Delta\nu$ below and above the minimum and maximum mode frequencies reported in LEGACY and Kages. This region overlaps minimally with the data used to fit for our prior information of the convective background (see Section 5.3.1), by design.

To improve computational efficiency, we reduced the number of oscillation modes

being fit in five targets. For 16 Cyg A & B, KIC 7970740 and KIC 8478994, we excluded any modes with a Bayes Factor ($\ln(K)$) of less than 6, as reported in LEGACY (for details, see Davies et al., 2016; Lund et al., 2017; Kass & Raftery, 1995). For KIC 8478994, which is reported without a value for $\ln(K)$ in Kages, we only included modes of an overtone number that contained a detection for all of $\ell = (0, 1, 2)$, retaining 5 sets of higher signal-to-noise overtones. We do not expect this reduced scope to bias our results, although they may reduce the precision on our measured rotation rates.

We fit our model to our power spectrum data using PyMC3, using 2500 iterations each on 4 chains. We assessed convergence using the Gelman-Rubin statistic ($\hat{R}$, Gelman & Rubin, 1992) and the number of effective samples ($n_{\rm eff}$), and flagged targets for which one or both of these exceeded expected values for the stellar rotational parameters.

An example of our model fit to an asteroseismic power spectrum of 16 Cyg A is shown in Figure 5.2.

5.3.2 Distinguishing between Gyrochronology Models

Stellar Models

In order to evaluate the implications of our seismic ensemble for gyrochronology, we compared our data to two models of rotational evolution adapted from those described in van Saders et al. (2019): a 'standard' model, which assumes a traditional angular momentum transport through magnetically driven stellar winds (Skumanich, 1972; Kawaler, 1988), and a weakened magnetic braking (hereafter WMB) model, which is identical to the standard model in its input physics, except for the condition that angular momentum loss ceases above a critical Rossby number of $Ro_{\rm crit} = 1.97$.

The braking models used in this work have several free parameters: the overall normalization (f_k); A disk locking timescale ($T_{\rm disk}$) and period ($P_{\rm disk}$) that set the

initial conditions for rotation; the critical angular velocity that marks the transition from saturated to unsaturated spin down ($\omega_{\rm crit}$), and the critical Rossby number, above which stars conserve angular momentum ($Ro_{\rm crit}$), mentioned above. $T_{\rm disk}$, $P_{\rm disk}$ and $\omega_{\rm crit}$ are calibrated to match the behaviour in young open clusters, but have little impact on the rotational evolution beyond ~ 1 Gyr in solar-mass stars. Both f_k and $Ro_{\rm crit}$ affect the late-time evolution. Both models adopt $\omega_{\rm crit} = 3.4 \times 10^{-5}\,{\rm s}^{-1}$, $P_{\rm disk} = 8.1$, $T_{\rm disk} = 0.28$ and $f_k = 6.6$. In the weakened magnetic braking model, $Ro_{\rm crit} = 1.97$. For further details, see van Saders & Pinsonneault (2013); van Saders et al. (2016, 2019).

Instead of comparing asteroseismic measurements of age and rotation directly to rotational models of individual stars, we instead compared them to synthetic populations based on rotational models, similar to those presented in van Saders et al. (2019), with some adaptations. To construct a synthetic population of rotating stars, van Saders et al. (2019) applied the rotational braking laws described above to a galactic population model. Their approach used a TRILEGAL (Girardi et al., 2012) Milky-Way simulation, tuned for the *Kepler* field and using the standard population values from Girardi et al. (2015). The TRILEGAL simulations were then matched in bins to observations of stellar parameters of the *Kepler* field (Mathur et al., 2017), with the aim of replicating the *Kepler* selection effect.

We used the same approach here, with two changes. First, we updated the distribution of stellar parameters matched to the TRILEGAL populations. Instead of Mathur et al. (2017), we used the Berger et al. (2020) temperatures and luminosities, and 2MASS K-band magnitudes as our stellar parameter distributions for the *Kepler* field, which are more accurate following the use of *Gaia* DR2 parallaxes (Gaia Collaboration et al., 2018). The TRILEGAL simulation was matched to the Berger et al. (2020) sample using a nearest-neighbours approach (as opposed to binning), based on the density of stars on the HR-diagram.

Second, we made further changes to account for possible binarity in the sample.

If the first step is performed blindly, blended binaries in the Berger et al. (2020) sample cause an overestimation of the number of old stars. In order to overcome this, we:

1. blended the TRILEGAL stars with binary companions drawn from a flat mass-ratio distribution, using the binary fraction of Raghavan et al. (2010),

2. recalculated the 'observed' luminosities and magnitudes assuming that each binary pair was blended, and

3. shifted these stars' temperatures following the g-K relation presented in Berger et al. (2020).

This new distribution was used for the nearest-neighbour matching. Once drawn we dropped the binary companion and used the true TRILEGAL properties of those stars. For stars with $M < 0.4\,M_{\odot}$, binary contributions were ignored. Every binary was assumed to result in a blend, regardless of separation. This results in slightly more young stars than reality, because young, blended binary systems contaminate regions of the HR-diagram where one expects to find old stars, and the number of blends is overestimated by assuming every binary system is a blend.

Our asteroseismic sample of stars with short cadence observations are subject to additional selection functions not included in the creation of the model populations above. We did *not* explicitly account for these asteroseismic selection functions in our model, by design. Both the standard and WMB models contain stars with the same fundamental parameters (mass, radius, effective temperature, metallicity) but a different period based on the choice of rotational evolution prescription. Applying an asteroseismic selection function that depends on these fundamental parameters (such as that described in Chaplin et al., 2011) would apply an identical prior to both models, therefore providing no net effect on our posterior distribution. Additionally, we expected any seismic selection function to be relatively flat (and therefore unin-

formative) on a star-by-star scale, on which we run our model analysis (as opposed to a population scale, see below).

Bayesian Mixture Models

In order to determine whether weakened magnetic braking occurs on the main sequence, we compared our sample of seismic age and rotation, along with temperature and mass, to the two stellar population models of the *Kepler* field adapted from van Saders et al. (2019), discussed above. We evaluated both stellar models in a Bayesian framework, with the rationale that we wanted to determine which of the two models (standard or WMB) is most likely to reproduce our observed data. Each model sample contained temperature ($T_{\rm eff}$), mass (M), age (t), metallicity ([Fe/H]) and rotation (P) information.

In order to draw probabilistic inference about the models, we built a five-dimensional Kernel Density Estimate (KDE) of both models using the `statsmodels` package. We used a bin width (setting the resolution of the KDE) of $0.02\,M_{\odot}$ in mass, 10 K in $T_{\rm eff}$, 0.01 in $\ln(t)$, 0.01 dex in [Fe/H] and 0.01 in $\ln(P)$. Note that we treated age and rotation in log space, where the posterior estimates from asteroseismology more closely resemble normal distributions. This approach translates the population models to a probability distribution we can use in a Bayesian framework.

We evaluated our data against both models simultaneously by treating the data as being drawn from a mixture of both model KDEs. In this mixture model structure, the two KDEs were modulated by a weighting factor, $P_{\rm s}$. In the limit $P_{\rm s} \to 1$, the data are most likely drawn from the standard model. In the limit $P_{\rm s} \to 0$, the data are most likely drawn from the WMB model.

The posterior probability of obtaining $P_{\rm s}$ and additional parameters θ given our data $\mathcal{D}$ is $p(P_{\rm s}, \theta|\mathcal{D})$. Using Bayes equation (see Eq. 1.1), we can express this as:

$$p(P_{\rm s}, \theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)\ p(\theta|P_{\rm s}, \kappa_{\rm s}, \kappa_{\rm WMB})\ p(P_{\rm s})\,, \tag{5.25}$$

where $\kappa_{\rm s}$ and $\kappa_{\rm WMB}$ are the KDE functions for the standard and weakened magnetic braking models respectively, and θ here are parameters, $\theta = \{M, T_{\rm eff}, \ln(t), [{\rm Fe/H}], \ln(P)\}$. As done above for the asteroseismic model, the parameters θ are latent parameters. Using this approach allowed our model to properly take into account the observational uncertainties on our data.

The second component of Equation 5.25 describes the probability of obtaining our latent parameters θ given our KDEs and the mixture model weighting parameter $P_{\rm s}$, and is described by the mixture model

$$p(\theta|P_{\rm s}, \kappa_{\rm s}, \kappa_{\rm WMB}) = P_{\rm s} \times \kappa_{\rm s}(\theta) + (1 - P_{\rm s}) \times \kappa_{\rm WMB}(\theta)\,, \tag{5.26}$$

where all parameters are as described above. This probability function describes a distribution that is a mixture of both KDEs. While the KDEs are constant, $P_{\rm s}$ is a free parameter, and so the shape of this distribution can vary. The latent parameters θ are drawn from this distribution, and therefore from some combination of the two stellar models.

The first component in Equation 5.25 describes the likelihood of obtaining the parameters θ given our data and their observational uncertainty. It takes the form

$$p(\mathcal{D}|\theta) = \mathcal{N}(\mathcal{D}|\theta, \sigma_{\mathcal{D}})\,, \tag{5.27}$$

a normal distribution evaluating the latent parameters θ against the observations, with observational uncertainty $\sigma_{\mathcal{D}}$. This approach means that in each parameter space (such as age), the age drawn from the stellar model mixture is entered into the likelihood equation with our static observations. The value of this equation (and thus the likelihood) will increase if θ is closer to the observations, and the mixture model will be modulated in a manner that maximises this probability, inferring whether one stellar model is more likely to produce our data than the other.

The final term, $p(P_{\rm s})$, represents the prior on the mixture model weight, which

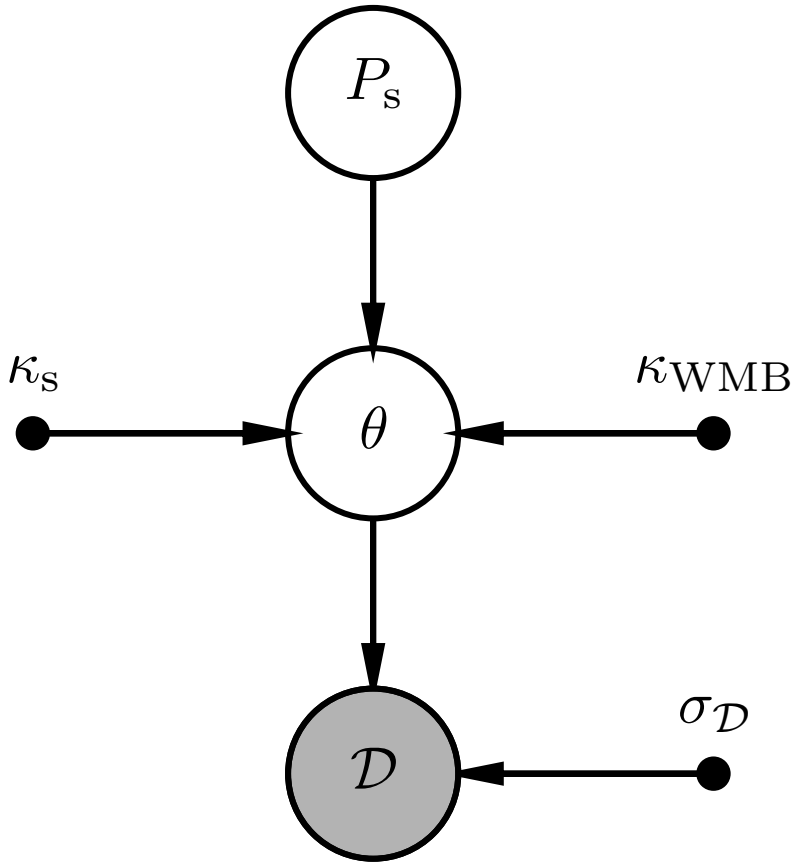

Figure 5.3: A probabilistic graphical model represented algebraically in Equation 5.25. Shaded circles indicate observed data, and solid black circles represent other fixed information, such as the KDEs and observational uncertainties. The remaining circles represent parameters. The symbols $\mathcal{D}$, θ and $\sigma_{\mathcal{D}}$ all represent sets of parameters or data. Here, $\kappa_{\rm s}$ and $\kappa_{\rm WMB}$ represent the KDE's of standard and weakened magnetic braking model populations respectively. $P_{\rm s}$ is the mixture model weighting factor. The latent parameters θ, our observations $\mathcal{D}$ and their uncertainties $\sigma_{\mathcal{D}}$ include temperature ($T_{\rm eff}$), mass (M), log-age ($\ln(t)$), metallicity ([Fe/H]) and log-rotation ($\ln(P)$). This model is *hierarchical*, as all the latent parameters are drawn from the common probability distribution set by $P_{\rm s}$ and described in Equation 5.26.

is uniform between 0 and 1. A visual representation of our model is shown in Figure 5.3.

Typically, this model would evaluate all stars in our sample against the stellar models simultaneously for a single posterior estimate of $P_{\rm s}$. At 95 stars, in 5 parameter spaces, this totals 476 free parameters to marginalise over. This is not an issue for Hamiltonian Monte Carlo (HMC, Betancourt & Girolami, 2013), however the use of Pythonic KDE functions, over which a probabilistic gradient can not

be measured, reduces HMCs effectiveness. Alternative Markov Chain Monte Carlo techniques (MCMC, Foreman-Mackey et al., 2013) can more efficiently sample the KDE functions, but can not treat the large number of hierarchical parameters. To overcome this, we fit our model to each star to obtain an independent individual posterior distribution for $P_{\rm s}$, and multiplied these afterwards to obtain a combined posterior. This comes with the benefit of easily allowing us to calculate the combined posterior for different stellar classifications (see Section 5.2.1), at the expense of marginalising for single value of $P_{\rm s}$ directly. As we are not interested in an exact value for $P_{\rm s}$, but rather where the posterior probability density of this parameter is highest, we found this trade-off justifiable.

Fitting procedure

For the sake of efficiency, the parameter spaces of the stellar models were reduced before calculating the KDEs. These cuts were made in M, $T_{\rm eff}$, $\ln(t)$ and [Fe/H], removing any stars in the models that fall more than $3 \times \sigma_{\mathcal{D}}$ outside the observations. Our observables M, t and P have asymmetric uncertainties from the Bayesian asteroseismic analysis. In order to err on the side of caution, we used the larger uncertainties as $\sigma_{\mathcal{D}}$ in each parameter space. We do not expect this to bias our results, as the posteriors are approximately Gaussian in our chosen parametrizations, and we employed a latent variable structure.

KIC 6278762 was excluded from this stellar model analysis, because its age fell more than 3σ outside of the highest age in the stellar models[9], and KICs 7106245 and 8760414 were excluded for the same reason due to low metallicities (-0.99 and -0.92 respectively).

We fit our model Equation 5.25 using `emcee` (Foreman-Mackey et al., 2013), using 32 walkers for a total of 7500 samples per walker, of which the first 2500 were discarded as a burn-in.

[9]This is actually a metallicity issue, as the oldest stars have metallicities outside the range of the rotational model grids.

After fitting, we took a normalised histogram of each posterior estimate, using 100 bins. With each bin approximating the value of the posterior function for $P_{\rm s}$, we were able to multiply the array of 100 bins for all stars, resulting in an approximate combined posterior estimate for $P_{\rm s}$ with a 100 bin resolution.

5.4 Results

5.4.1 Asteroseismology

We fit our asteroseismic model, described in Section 5.3.1, to all 95 stars in the LEGACY and Kages catalogues. The only star which could not converge on a single result was KIC 8478994, which had some of the lowest signal-to-noise modes. The results for the remaining 94 stars are shown in Table 5.1 alongside relevant data reported in Kages and LEGACY. We report the projected splitting, inclination (transformed from $\cos(i)$) and rotation period (transformed from ν_s). The summary statistics are taken as the median of the posterior distribution, with uncertainties being the $15.9^{\rm th}$ and $84.1^{\rm st}$ percentiles. To calculate parameters with no direct posterior samples (e.g. rotation), the full posterior samples were transformed before taking the summary statistics.

In order to be careful with our conclusions, we flagged any sub-optimal conditions of the final fit. Of 94 stars, we flagged 5 for which the convergence metric, $\hat{R}$, was greater than 1.01 (low concern) and 2 stars over 1.1 (medium concern) on rotational parameters, to avoid non-converged results. It should be noted that for the 2 stars with an $\hat{R}$ over 1.1, neither had values high enough to cause enough concern to exclude them entirely. We also performed by-eye checks of the sampled chains on all hyperparameters and of the best-fit model compared to both the raw and smoothed asteroseismic data. We found no issues in the by-eye investigation of the 94 stars. KIC 8478994 was excluded both for a poor fit as well as high $\hat{R}$ on rotational parameters[10]. KICs 6603624, 8760414 and 8938364 are reported in Table 5.1, but

[10]Bearing in mind that it is difficult to verify a fit by-eye on low-signal to noise spectra.

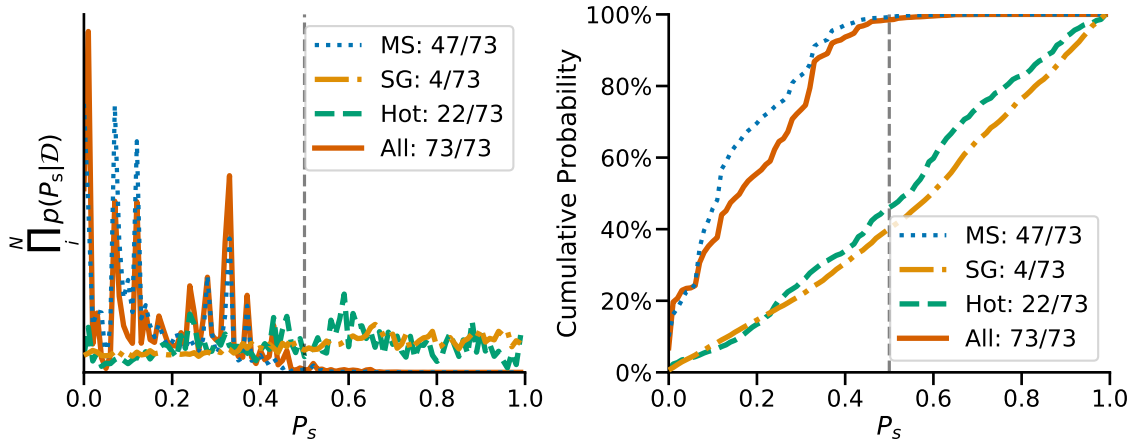

Figure 5.4: Posterior estimates of the mixture model parameter P_s, broken down by stellar classification. A value of P_s close to zero indicates that the data are more consistent with a rotational evolution that includes weakened magnetic braking, whereas a value close to 1 indicates that the data are more consistent with a standard rotational evolution scenario. *Left:* the product of histograms of the posterior estimates for 73 stars, or subsets thereof for different stellar classifications. *Right:* the cumulative posterior probability for 73 stars, or subsets thereof for different stellar types.

were excluded from the gyrochronology analysis due to strong disagreement with literature values, leaving a sample of 91 stars (see Section 5.5.1). Finally, we flagged any stars with fewer than 1000 effective samples of ν_s.

5.4.2 Gyrochronology

We fit our mixture model, described in Section 5.3.2, to a sub-sample of 89 stars, where we excluded KICs 6278762 and 7106245 as they had age and metallicity respectively that fell far outside the functional range of the stellar models. To ensure our results were robust, we also excluded any stars with $n_{\rm eff} < 1000$ for rotational splitting, and those with $\hat{R} > 1.1$. The remaining sample of 73 stars contained 4 sub-giants, 22 'hot' stars, and 47 main sequence stars.

The full joint posterior, obtained by multiplying the individual posteriors for P_s, as well as said individual posteriors are shown in Figure 5.4. The left-hand side of Figure 5.4 ($P_s < 0.5$) holds 98.4% of the total joint probability, when considering the ensemble of 73 stars. When only considering the 47 stars classified as main sequence, this rises to 99.2%. The stellar sample used for this analysis, alongside those that

were excluded, is shown in Figure 5.5, set against both the standard and WMB models in period-temperature space. Stars excluded from the model analysis are shown without uncertainties. Stars with metallicities outside the functional range of the van Saders et al. (2019) models ($|[\mathrm{Fe/H}]| > 0.4$) are marked. All 5 of these stars are metal-poor, and have metallicities between -0.99 and -0.44.

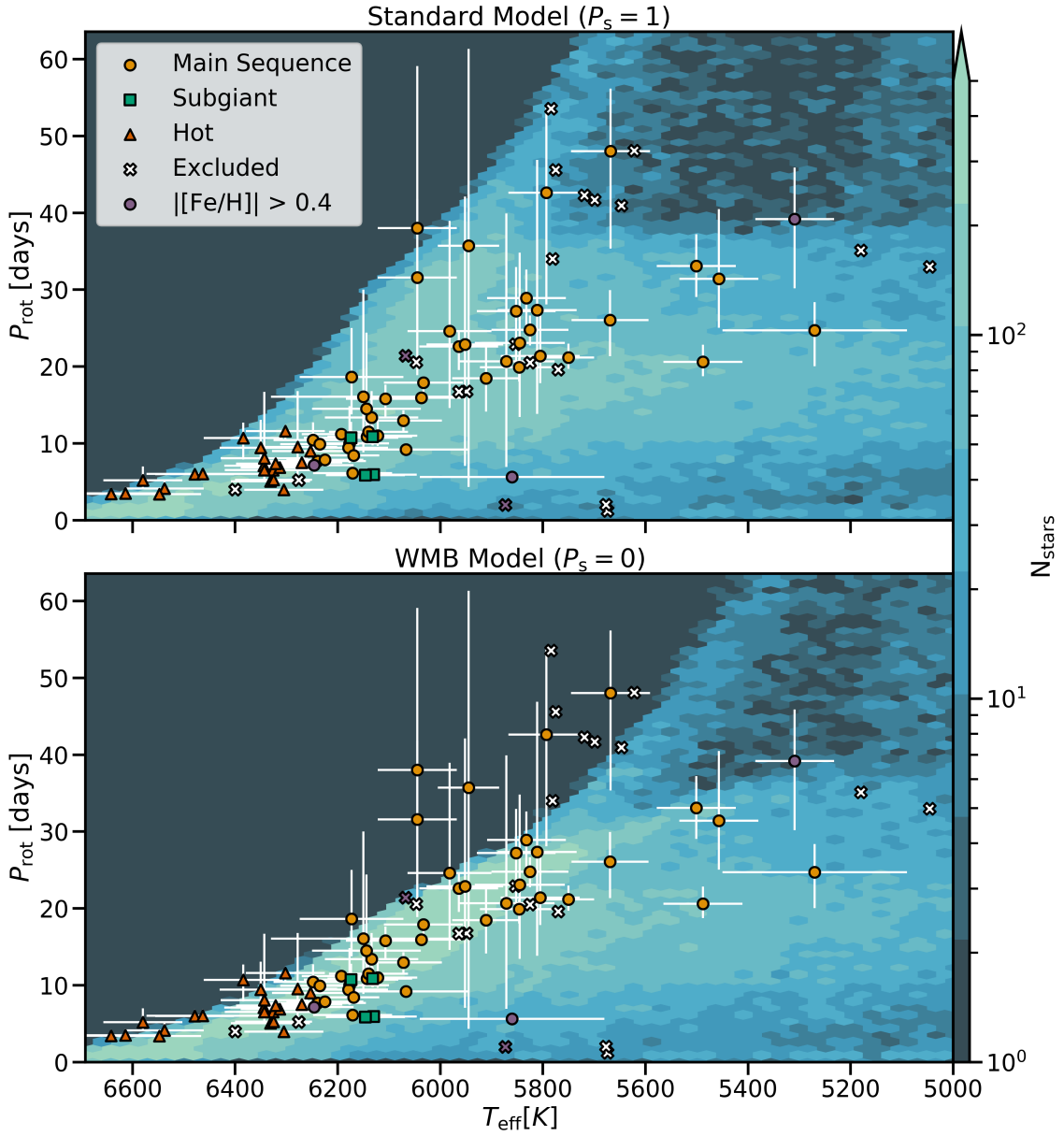

Figure 5.5: Stars for which rotation was measured in this work, plotted over the van Saders et al. (2019) modelled stellar populations based on standard (top) and weakened magnetic braking (WMB, bottom) models of stellar evolution. The 23 Stars that were excluded from the gyrochronology analysis are marked with crosses and without uncertainties (see text). Circles, squares and triangles denote main sequence, sub-giant, and 'hot' stars respectively. Stars with a metallicity of $|[Fe/H]| > 0.4$ are coloured differently (purple), to indicate that they fall outside the functional range of the van Saders et al. (2019) stellar models. Note that only period-temperature space is shown here, but when evaluating between the two model prescriptions we also considered mass, age and metallicity.

Table 5.1: Parameters for the 94 stars for which seismic rotation rates were obtained in this work. Temperature ($T_{\rm eff}$), age, mass, metallicity ([Fe/H]) and surface gravity ($\log(g)$) are adopted from the LEGACY (L, Lund et al., 2017; Silva Aguirre et al., 2017) and Kages (K, Silva Aguirre et al., 2015; Davies et al., 2016) catalogues, as listed in the Source column. Projected splitting ($\nu_{\rm s}\sin(i)$), inclination angle (i) and asteroseismic rotation ($P_{\rm rot}$) are from this work. Uncertainties were taken using the 15.9$^{\rm th}$ and 84.1$^{\rm st}$ percentiles of posterior distributions on the parameters, which are frequently asymmetrical in linear space. Reported values are the median of the posteriors. For parameters with no direct posterior samples (e.g. rotation) the full posterior samples were transformed before taking the summary statistics. The stellar type denotes whether a star is roughly classified as belonging to the main sequence (MS), Sub-Giants (SG) or 'hot' stars (H) (see text). The flags indicate the following: 0; no issues, used in the gyrochronology analysis. 1; has either a number of effective samples $n_{\rm eff} < 1000$ for the asteroseismic splitting, or Gelman-Rubin convergence metric of $\hat{R} > 1.1$, indicating that rotation measurements for these stars are less robust than those with a flag of 0. 2; was found to strongly disagree with multiple literature values, excluded from the gyrochronology analysis. 3; fell outside the model range of the stellar models, and were therefore not used in the gyrochronology analysis. Table is continued on the next page.

KIC	$T_{\rm eff}$ [K]	Age [Gyr]	Mass [$M_\odot$]	[Fe/H] [dex]	$\log(g)$ [dex]	$\nu_{\rm s}\sin(i)$ [μHz]	i [°]	$P_{\rm rot}$ [days]	Flag	Type	Source
1435467	6326±77	$3.02^{+0.50}_{-0.35}$	$1.32^{+0.03}_{-0.05}$	0.01±0.10	$4.100^{+0.009}_{-0.009}$	$1.58^{+0.10}_{-0.09}$	$63.4^{+10.2}_{-6.6}$	$6.5^{+0.8}_{-0.6}$	0	H	L
2837475	6614±77	$1.63^{+0.11}_{-0.18}$	$1.43^{+0.02}_{-0.02}$	0.01±0.10	$4.163^{+0.007}_{-0.007}$	$3.12^{+0.08}_{-0.08}$	$70.7^{+6.0}_{-4.4}$	$3.5^{+0.2}_{-0.2}$	0	H	L
3425851	6343±85	$3.32^{+0.85}_{-0.64}$	$1.18^{+0.05}_{-0.05}$	-0.04±0.10	$4.243^{+0.008}_{-0.008}$	$1.17^{+0.48}_{-0.62}$	$60.9^{+20.1}_{-22.7}$	$8.1^{+8.6}_{-2.7}$	0	H	K
3427720	6045±77	$2.23^{+0.24}_{-0.24}$	$1.11^{+0.02}_{-0.01}$	-0.06±0.10	$4.387^{+0.005}_{-0.004}$	$0.30^{+0.06}_{-0.06}$	$56.4^{+22.9}_{-23.4}$	$31.6^{+10.2}_{-11.8}$	0	MS	L
3456181	6384±77	$2.09^{+0.13}_{-0.13}$	$1.50^{+0.03}_{-0.02}$	-0.15±0.10	$3.949^{+0.008}_{-0.009}$	$0.92^{+0.08}_{-0.08}$	$58.2^{+20.4}_{-17.7}$	$10.7^{+2.0}_{-2.8}$	0	H	L
3544595	5669±75	$6.63^{+0.62}_{-0.57}$	$0.90^{+0.01}_{-0.01}$	-0.18±0.10	$4.468^{+0.003}_{-0.003}$	$0.40^{+0.04}_{-0.04}$	$66.0^{+15.6}_{-13.9}$	$26.1^{+3.9}_{-4.7}$	0	MS	K
3632418	6193±77	$2.63^{+0.18}_{-0.18}$	$1.41^{+0.02}_{-0.02}$	-0.12±0.10	$4.024^{+0.008}_{-0.008}$	$0.98^{+0.03}_{-0.03}$	$72.3^{+10.0}_{-7.4}$	$11.2^{+0.6}_{-0.7}$	0	MS	L
3656476	5668±77	$8.37^{+1.72}_{-1.57}$	$1.04^{+0.05}_{-0.04}$	0.25±0.10	$4.225^{+0.008}_{-0.010}$	$0.21^{+0.02}_{-0.02}$	$62.4^{+18.9}_{-20.5}$	$48.0^{+8.1}_{-12.7}$	0	MS	L
3735871	6107±77	$2.35^{+1.04}_{-0.85}$	$1.09^{+0.04}_{-0.04}$	-0.04±0.10	$4.396^{+0.007}_{-0.007}$	$0.69^{+0.05}_{-0.05}$	$70.4^{+13.4}_{-15.1}$	$15.8^{+1.8}_{-2.5}$	0	MS	L

Continued on next page

Table 5.1: *Continued from previous page.*

KIC	$T_{\rm eff}$ [K]	Age [Gyr]	Mass [$M_\odot$]	[Fe/H] [dex]	$\log(g)$ [dex]	$\nu_{\rm s}\sin(i)$ [μHz]	i [°]	$P_{\rm rot}$ [days]	Flag	Type	Source
4141376	6134±91	$3.27^{+0.59}_{-0.64}$	$1.02^{+0.02}_{-0.03}$	-0.24±0.10	$4.412^{+0.004}_{-0.003}$	$0.76^{+0.13}_{-0.13}$	$64.0^{+17.5}_{-16.7}$	$13.4^{+3.4}_{-3.0}$	0	MS	K
4143755	5622±106	$11.27^{+1.50}_{-1.35}$	$0.92^{+0.02}_{-0.03}$	-0.40±0.11	$4.102^{+0.002}_{-0.001}$	$0.18^{+0.08}_{-0.05}$	$45.8^{+30.6}_{-27.3}$	$48.1^{+27.4}_{-32.7}$	1	MS	K
4349452	6270±79	$3.45^{+0.81}_{-0.72}$	$1.16^{+0.04}_{-0.05}$	-0.04±0.10	$4.275^{+0.008}_{-0.007}$	$1.50^{+0.09}_{-0.09}$	$79.7^{+7.1}_{-10.0}$	$7.5^{+0.5}_{-0.6}$	0	H	K
4914423	5845±88	$6.67^{+0.69}_{-0.62}$	$1.10^{+0.02}_{-0.03}$	0.07±0.11	$4.155^{+0.004}_{-0.004}$	$0.42^{+0.15}_{-0.14}$	$61.3^{+19.8}_{-30.5}$	$23.1^{+11.7}_{-9.6}$	0	MS	K
4914923	5805±77	$7.57^{+1.66}_{-1.79}$	$1.06^{+0.06}_{-0.05}$	0.08±0.10	$4.197^{+0.008}_{-0.010}$	$0.39^{+0.03}_{-0.03}$	$46.6^{+13.3}_{-8.1}$	$21.4^{+5.4}_{-3.5}$	0	MS	L
5094751	5952±75	$6.35^{+1.05}_{-1.05}$	$1.07^{+0.04}_{-0.04}$	-0.08±0.10	$4.213^{+0.007}_{-0.008}$	$0.39^{+0.27}_{-0.16}$	$51.5^{+26.7}_{-31.1}$	$22.9^{+19.3}_{-15.8}$	0	MS	K
5184732	5846±77	$4.85^{+1.57}_{-0.88}$	$1.15^{+0.04}_{-0.06}$	0.36±0.10	$4.255^{+0.010}_{-0.008}$	$0.55^{+0.02}_{-0.02}$	$71.3^{+11.2}_{-10.8}$	$19.9^{+1.3}_{-1.9}$	0	MS	L
5773345	6130±84	$2.55^{+0.26}_{-0.24}$	$1.47^{+0.03}_{-0.03}$	0.21±0.09	$3.993^{+0.008}_{-0.007}$	$1.08^{+0.08}_{-0.08}$	$33.7^{+2.8}_{-2.5}$	$5.9^{+0.7}_{-0.5}$	0	SG	L
5866724	6169±50	$3.89^{+0.59}_{-0.48}$	$1.20^{+0.03}_{-0.03}$	0.09±0.08	$4.224^{+0.007}_{-0.005}$	$1.34^{+0.07}_{-0.08}$	$80.9^{+6.3}_{-9.4}$	$8.4^{+0.5}_{-0.5}$	0	MS	K
5950854	5853±77	$8.93^{+1.12}_{-1.15}$	$0.97^{+0.03}_{-0.03}$	-0.23±0.10	$4.238^{+0.007}_{-0.007}$	$0.29^{+0.64}_{-0.12}$	$27.6^{+46.0}_{-15.2}$	$22.9^{+34.2}_{-19.7}$	1	MS	L
6106415	6037±77	$5.03^{+1.28}_{-1.12}$	$1.07^{+0.05}_{-0.04}$	-0.04±0.10	$4.295^{+0.009}_{-0.009}$	$0.69^{+0.02}_{-0.02}$	$73.1^{+8.4}_{-6.3}$	$16.0^{+0.7}_{-0.8}$	0	MS	L
6116048	6033±77	$9.58^{+2.16}_{-1.90}$	$0.94^{+0.05}_{-0.05}$	-0.23±0.10	$4.254^{+0.009}_{-0.012}$	$0.63^{+0.02}_{-0.02}$	$76.3^{+9.0}_{-10.2}$	$17.9^{+0.8}_{-1.2}$	0	MS	L
6196457	5871±94	$5.52^{+0.51}_{-0.48}$	$1.21^{+0.02}_{-0.03}$	0.17±0.11	$4.049^{+0.005}_{-0.004}$	$0.43^{+0.28}_{-0.19}$	$52.5^{+26.0}_{-29.1}$	$20.7^{+19.2}_{-13.7}$	0	MS	K
6225718	6313±76	$2.41^{+0.53}_{-0.43}$	$1.16^{+0.03}_{-0.03}$	-0.07±0.10	$4.319^{+0.005}_{-0.007}$	$0.81^{+0.03}_{-0.03}$	$29.1^{+2.1}_{-1.8}$	$6.9^{+0.6}_{-0.5}$	0	H	L
6278762	5046±74	$11.54^{+0.99}_{-0.94}$	$0.74^{+0.01}_{-0.01}$	-0.37±0.09	$4.560^{+0.002}_{-0.003}$	$0.30^{+0.09}_{-0.09}$	$62.2^{+19.0}_{-29.5}$	$33.0^{+13.6}_{-13.5}$	1, 3	MS	K
6508366	6331±77	$2.06^{+0.13}_{-0.14}$	$1.53^{+0.03}_{-0.02}$	-0.05±0.10	$3.942^{+0.005}_{-0.007}$	$2.28^{+0.04}_{-0.04}$	$87.0^{+2.1}_{-3.2}$	$5.1^{+0.1}_{-0.1}$	0	H	L
6521045	5825±75	$6.50^{+0.46}_{-0.56}$	$1.11^{+0.02}_{-0.02}$	0.02±0.10	$4.125^{+0.004}_{-0.004}$	$0.45^{+0.03}_{-0.02}$	$75.7^{+9.7}_{-11.2}$	$24.8^{+1.9}_{-2.0}$	0	MS	K

Continued on next page

Table 5.1: *Continued from previous page.*

KIC	$T_{\rm eff}$ [K]	Age [Gyr]	Mass [$M_\odot$]	[Fe/H] [dex]	$\log(g)$ [dex]	$\nu_{\rm s}\sin(i)$ [μHz]	i [°]	$P_{\rm rot}$ [days]	Flag	Type	Source
6603624	5674±77	$7.82^{+0.94}_{-0.86}$	$1.01^{+0.03}_{-0.02}$	0.28±0.10	$4.320^{+0.004}_{-0.005}$	$1.13^{+0.13}_{-0.13}$	$6.9^{+0.8}_{-0.8}$	$1.2^{+0.0}_{-0.0}$	2	MS	L
6679371	6479±77	$1.95^{+0.18}_{-0.16}$	$1.53^{+0.04}_{-0.02}$	0.01±0.10	$3.934^{+0.007}_{-0.008}$	$1.90^{+0.05}_{-0.06}$	$82.1^{+5.5}_{-7.2}$	$6.0^{+0.2}_{-0.2}$	0	H	L
6933899	5832±77	$6.34^{+0.72}_{-0.62}$	$1.13^{+0.03}_{-0.03}$	-0.01±0.10	$4.087^{+0.008}_{-0.007}$	$0.36^{+0.02}_{-0.02}$	$64.3^{+16.1}_{-14.0}$	$28.9^{+3.7}_{-4.8}$	0	MS	L
7103006	6344±77	$2.47^{+0.22}_{-0.24}$	$1.42^{+0.04}_{-0.02}$	0.02±0.10	$4.015^{+0.007}_{-0.007}$	$1.36^{+0.08}_{-0.09}$	$56.8^{+14.8}_{-8.8}$	$7.1^{+1.3}_{-1.0}$	0	H	L
7106245	6068±102	$6.27^{+1.06}_{-1.06}$	$0.92^{+0.02}_{-0.04}$	-0.99±0.19	$4.325^{+0.007}_{-0.007}$	$0.32^{+0.16}_{-0.10}$	$33.4^{+38.4}_{-14.7}$	$21.4^{+23.8}_{-13.2}$	1, 3	MS	L
7206837	6305±77	$2.90^{+0.42}_{-0.30}$	$1.30^{+0.03}_{-0.03}$	0.10±0.10	$4.163^{+0.008}_{-0.007}$	$1.53^{+0.12}_{-0.12}$	$31.7^{+3.2}_{-2.8}$	$4.0^{+0.6}_{-0.4}$	0	H	L
7296438	5775±77	$7.23^{+1.49}_{-1.77}$	$1.08^{+0.06}_{-0.05}$	0.19±0.10	$4.201^{+0.009}_{-0.010}$	$0.20^{+0.06}_{-0.06}$	$50.5^{+28.2}_{-31.1}$	$45.6^{+23.4}_{-29.0}$	1	MS	L
7510397	6171±77	$2.82^{+0.14}_{-0.16}$	$1.37^{+0.02}_{-0.02}$	-0.21±0.10	$4.036^{+0.007}_{-0.004}$	$0.64^{+0.06}_{-0.06}$	$19.9^{+2.0}_{-2.0}$	$6.1^{+0.7}_{-0.6}$	0	MS	L
7670943	6463±110	$2.78^{+0.62}_{-0.51}$	$1.24^{+0.04}_{-0.05}$	0.09±0.11	$4.228^{+0.008}_{-0.008}$	$1.83^{+0.14}_{-0.14}$	$75.7^{+9.8}_{-11.5}$	$6.0^{+0.6}_{-0.6}$	0	H	K
7680114	5811±77	$7.68^{+1.45}_{-1.28}$	$1.06^{+0.04}_{-0.05}$	0.05±0.10	$4.172^{+0.008}_{-0.010}$	$0.26^{+0.05}_{-0.04}$	$37.0^{+34.3}_{-16.3}$	$27.3^{+19.6}_{-13.5}$	0	MS	L
7771282	6248±77	$3.24^{+0.35}_{-0.32}$	$1.29^{+0.03}_{-0.03}$	-0.02±0.10	$4.112^{+0.007}_{-0.007}$	$1.01^{+0.14}_{-0.18}$	$69.5^{+14.6}_{-17.8}$	$10.4^{+2.3}_{-1.7}$	0	MS	L
7871531	5501±77	$9.96^{+1.93}_{-1.77}$	$0.83^{+0.03}_{-0.02}$	-0.26±0.10	$4.478^{+0.007}_{-0.005}$	$0.33^{+0.03}_{-0.03}$	$71.3^{+12.0}_{-13.2}$	$33.1^{+4.1}_{-4.1}$	0	MS	L
7940546	6235±77	$2.33^{+0.08}_{-0.08}$	$1.40^{+0.03}_{-0.01}$	-0.20±0.10	$4.007^{+0.003}_{-0.001}$	$1.14^{+0.03}_{-0.03}$	$78.9^{+7.4}_{-7.9}$	$9.9^{+0.3}_{-0.4}$	0	MS	L
7970740	5309±77	$12.98^{+1.36}_{-2.00}$	$0.73^{+0.03}_{-0.01}$	-0.54±0.10	$4.539^{+0.004}_{-0.005}$	$0.26^{+0.03}_{-0.02}$	$60.7^{+17.4}_{-15.3}$	$39.2^{+6.7}_{-9.0}$	0	MS	L
8006161	5488±77	$3.59^{+1.53}_{-1.45}$	$0.98^{+0.03}_{-0.03}$	0.34±0.10	$4.494^{+0.007}_{-0.007}$	$0.34^{+0.02}_{-0.02}$	$37.0^{+4.1}_{-3.4}$	$20.6^{+2.2}_{-1.8}$	0	MS	L
8077137	6072±75	$6.23^{+0.56}_{-1.23}$	$1.12^{+0.04}_{-0.05}$	-0.09±0.10	$4.056^{+0.010}_{-0.013}$	$0.84^{+0.06}_{-0.07}$	$72.8^{+11.3}_{-12.6}$	$13.0^{+1.3}_{-1.5}$	0	MS	K
8150065	6173±101	$3.83^{+0.99}_{-0.67}$	$1.19^{+0.04}_{-0.05}$	-0.13±0.15	$4.220^{+0.008}_{-0.008}$	$0.54^{+0.11}_{-0.13}$	$64.0^{+18.0}_{-21.1}$	$18.6^{+6.4}_{-4.9}$	0	MS	L

Continued on next page

Table 5.1: *Continued from previous page.*

KIC	$T_{\rm eff}$ [K]	Age [Gyr]	Mass [$M_\odot$]	[Fe/H] [dex]	$\log(g)$ [dex]	$\nu_{\rm s}\sin(i)$ [μHz]	i [°]	$P_{\rm rot}$ [days]	Flag	Type	Source
8179536	6343±77	$3.54^{+1.04}_{-0.81}$	$1.16^{+0.05}_{-0.06}$	-0.03±0.10	$4.255^{+0.010}_{-0.010}$	$1.46^{+0.10}_{-0.09}$	$55.7^{+13.3}_{-7.4}$	$6.5^{+1.2}_{-0.8}$	0	H	L
8228742	6122±77	$2.89^{+0.16}_{-0.18}$	$1.38^{+0.02}_{-0.02}$	-0.08±0.10	$4.035^{+0.007}_{-0.005}$	$0.64^{+0.04}_{-0.04}$	$37.9^{+6.2}_{-4.0}$	$11.0^{+2.0}_{-1.3}$	0	MS	L
8292840	6239±94	$3.85^{+0.81}_{-0.75}$	$1.15^{+0.05}_{-0.05}$	-0.14±0.10	$4.240^{+0.008}_{-0.008}$	$1.45^{+0.07}_{-0.07}$	$76.2^{+8.9}_{-9.1}$	$7.7^{+0.5}_{-0.5}$	0	MS	K
8349582	5699±74	$8.03^{+0.80}_{-0.70}$	$1.07^{+0.02}_{-0.02}$	0.30±0.10	$4.163^{+0.004}_{-0.003}$	$0.23^{+0.07}_{-0.06}$	$59.6^{+20.7}_{-24.1}$	$41.7^{+19.6}_{-14.9}$	1	MS	K
8379927	6067±120	$1.99^{+0.85}_{-0.75}$	$1.12^{+0.04}_{-0.04}$	-0.10±0.15	$4.388^{+0.008}_{-0.007}$	$1.12^{+0.02}_{-0.02}$	$63.3^{+2.5}_{-2.3}$	$9.2^{+0.3}_{-0.2}$	0	MS	L
8394589	6143±77	$4.45^{+0.94}_{-0.83}$	$1.04^{+0.04}_{-0.03}$	-0.29±0.10	$4.322^{+0.008}_{-0.008}$	$1.01^{+0.03}_{-0.03}$	$71.1^{+7.9}_{-5.9}$	$10.9^{+0.6}_{-0.6}$	0	MS	L
8424992	5719±77	$9.61^{+1.92}_{-1.74}$	$0.92^{+0.04}_{-0.04}$	-0.12±0.10	$4.359^{+0.007}_{-0.007}$	$0.22^{+0.06}_{-0.06}$	$59.1^{+21.3}_{-30.3}$	$42.3^{+19.4}_{-17.7}$	1	MS	L
8494142	6144±106	$2.62^{+0.26}_{-0.24}$	$1.42^{+0.03}_{-0.02}$	0.13±0.10	$4.038^{+0.005}_{-0.005}$	$0.67^{+0.22}_{-0.28}$	$62.8^{+18.5}_{-23.8}$	$14.5^{+9.9}_{-4.5}$	0	MS	K
8554498	5945±60	$5.60^{+0.45}_{-0.42}$	$1.20^{+0.02}_{-0.03}$	0.17±0.05	$4.007^{+0.003}_{-0.003}$	$0.25^{+0.21}_{-0.09}$	$48.1^{+30.0}_{-36.6}$	$35.7^{+25.6}_{-31.4}$	0	MS	K
8694723	6246±77	$4.69^{+0.48}_{-0.51}$	$1.14^{+0.02}_{-0.02}$	-0.42±0.10	$4.113^{+0.007}_{-0.009}$	$0.92^{+0.05}_{-0.05}$	$34.7^{+3.4}_{-2.7}$	$7.2^{+0.8}_{-0.6}$	0	MS	L
8760414	5873±77	$11.66^{+1.28}_{-1.61}$	$0.81^{+0.03}_{-0.02}$	-0.92±0.10	$4.329^{+0.006}_{-0.005}$	$0.69^{+0.26}_{-0.42}$	$7.6^{+2.3}_{-1.7}$	$2.0^{+4.5}_{-0.4}$	2, 3	MS	L
8866102	6325±75	$2.60^{+0.56}_{-0.53}$	$1.23^{+0.04}_{-0.04}$	0.01±0.10	$4.262^{+0.008}_{-0.007}$	$2.15^{+0.04}_{-0.04}$	$78.2^{+6.6}_{-4.7}$	$5.3^{+0.2}_{-0.2}$	0	H	K
8938364	5677±77	$10.25^{+0.56}_{-0.65}$	$0.99^{+0.01}_{-0.01}$	-0.13±0.10	$4.173^{+0.007}_{-0.002}$	$0.61^{+0.27}_{-0.50}$	$8.7^{+57.5}_{-2.3}$	$2.0^{+87.0}_{-0.2}$	2	MS	L
9025370	5270±180	$6.55^{+1.26}_{-1.13}$	$0.97^{+0.03}_{-0.03}$	-0.12±0.18	$4.423^{+0.007}_{-0.004}$	$0.43^{+0.04}_{-0.04}$	$67.5^{+15.2}_{-19.1}$	$24.7^{+3.7}_{-4.7}$	0	MS	L
9098294	5852±77	$8.08^{+0.99}_{-0.73}$	$0.97^{+0.02}_{-0.03}$	-0.18±0.10	$4.308^{+0.005}_{-0.007}$	$0.36^{+0.04}_{-0.04}$	$58.2^{+21.0}_{-16.3}$	$27.2^{+5.7}_{-7.0}$	0	MS	L
9139151	6302±77	$1.32^{+0.94}_{-0.75}$	$1.18^{+0.04}_{-0.05}$	0.10±0.10	$4.382^{+0.008}_{-0.008}$	$0.95^{+0.04}_{-0.04}$	$73.5^{+11.0}_{-11.0}$	$11.6^{+0.8}_{-1.1}$	0	H	L
9139163	6400±84	$1.60^{+0.22}_{-0.22}$	$1.40^{+0.03}_{-0.02}$	0.15±0.09	$4.200^{+0.009}_{-0.008}$	$1.59^{+0.07}_{-0.08}$	$33.5^{+3.0}_{-3.0}$	$4.0^{+0.3}_{-0.3}$	1	H	L

Continued on next page

Table 5.1: *Continued from previous page.*

KIC	$T_{\rm eff}$ [K]	Age [Gyr]	Mass [$M_\odot$]	[Fe/H] [dex]	$\log(g)$ [dex]	$\nu_{\rm s}\sin(i)$ [μHz]	i [°]	$P_{\rm rot}$ [days]	Flag	Type	Source
9206432	6538±77	$1.53^{+0.21}_{-0.30}$	$1.38^{+0.04}_{-0.02}$	0.16±0.10	$4.220^{+0.005}_{-0.007}$	$1.55^{+0.17}_{-0.20}$	$34.3^{+5.7}_{-4.2}$	$4.1^{+1.0}_{-0.5}$	0	H	L
9353712	6278±77	$2.15^{+0.11}_{-0.13}$	$1.51^{+0.03}_{-0.02}$	-0.05±0.10	$3.943^{+0.007}_{-0.005}$	$0.75^{+0.16}_{-0.17}$	$37.6^{+29.1}_{-12.6}$	$9.5^{+7.3}_{-3.9}$	0	H	L
9410862	6047±77	$6.93^{+1.49}_{-1.33}$	$0.97^{+0.05}_{-0.04}$	-0.31±0.10	$4.300^{+0.009}_{-0.008}$	$0.41^{+0.09}_{-0.08}$	$46.4^{+28.3}_{-16.7}$	$20.6^{+9.8}_{-8.1}$	1	MS	L
9414417	6253±75	$2.65^{+0.16}_{-0.18}$	$1.40^{+0.02}_{-0.03}$	-0.13±0.10	$4.016^{+0.005}_{-0.005}$	$1.09^{+0.05}_{-0.05}$	$58.1^{+9.7}_{-6.9}$	$9.0^{+1.1}_{-0.9}$	0	H	L
9592705	6174±92	$2.33^{+0.18}_{-0.16}$	$1.51^{+0.03}_{-0.02}$	0.22±0.10	$3.961^{+0.003}_{-0.004}$	$0.93^{+0.09}_{-0.09}$	$59.3^{+17.8}_{-12.1}$	$10.7^{+2.1}_{-1.9}$	0	SG	K
9812850	6321±77	$2.71^{+0.46}_{-0.35}$	$1.37^{+0.04}_{-0.05}$	-0.07±0.10	$4.053^{+0.008}_{-0.009}$	$1.54^{+0.09}_{-0.08}$	$81.0^{+6.4}_{-10.0}$	$7.4^{+0.4}_{-0.5}$	0	H	L
9955598	5457±77	$6.29^{+1.95}_{-1.84}$	$0.90^{+0.04}_{-0.03}$	0.05±0.10	$4.497^{+0.007}_{-0.005}$	$0.29^{+0.04}_{-0.04}$	$53.4^{+20.8}_{-11.9}$	$31.4^{+9.1}_{-6.4}$	0	MS	L
9965715	5860±180	$2.92^{+0.86}_{-0.75}$	$1.21^{+0.04}_{-0.05}$	-0.44±0.18	$4.272^{+0.008}_{-0.009}$	$1.75^{+0.05}_{-0.05}$	$58.3^{+3.5}_{-3.2}$	$5.6^{+0.3}_{-0.3}$	0	MS	L
10068307	6132±77	$2.36^{+0.08}_{-0.10}$	$1.47^{+0.01}_{-0.02}$	-0.23±0.10	$3.967^{+0.004}_{-0.004}$	$0.71^{+0.03}_{-0.03}$	$41.7^{+6.0}_{-4.1}$	$10.9^{+1.5}_{-1.1}$	0	SG	L
10079226	5949±77	$3.06^{+0.70}_{-0.65}$	$1.12^{+0.02}_{-0.03}$	0.11±0.10	$4.366^{+0.005}_{-0.005}$	$0.65^{+0.07}_{-0.08}$	$75.1^{+10.6}_{-23.4}$	$16.8^{+2.2}_{-3.0}$	1	MS	L
10162436	6146±77	$2.46^{+0.10}_{-0.11}$	$1.45^{+0.02}_{-0.01}$	-0.16±0.10	$3.981^{+0.005}_{-0.005}$	$0.84^{+0.05}_{-0.05}$	$25.5^{+2.1}_{-1.8}$	$5.9^{+0.6}_{-0.4}$	0	SG	L
10454113	6177±77	$2.89^{+0.56}_{-0.53}$	$1.17^{+0.02}_{-0.03}$	-0.07±0.10	$4.314^{+0.005}_{-0.005}$	$0.76^{+0.07}_{-0.07}$	$40.9^{+26.9}_{-9.6}$	$10.0^{+4.8}_{-2.5}$	0	MS	L
10514430	5784±98	$7.84^{+0.40}_{-0.91}$	$1.06^{+0.04}_{-0.02}$	-0.11±0.11	$4.061^{+0.004}_{-0.004}$	$0.18^{+0.05}_{-0.05}$	$57.2^{+23.5}_{-32.0}$	$53.6^{+23.6}_{-27.2}$	1	MS	K
10516096	5964±77	$7.01^{+1.33}_{-1.45}$	$1.06^{+0.05}_{-0.06}$	-0.11±0.10	$4.169^{+0.010}_{-0.011}$	$0.48^{+0.03}_{-0.03}$	$71.8^{+12.5}_{-16.2}$	$22.6^{+1.9}_{-3.1}$	0	MS	L
10586004	5770±83	$6.43^{+0.64}_{-0.61}$	$1.18^{+0.02}_{-0.03}$	0.29±0.10	$4.071^{+0.005}_{-0.005}$	$0.48^{+0.16}_{-0.17}$	$59.3^{+20.6}_{-22.4}$	$19.6^{+11.1}_{-6.5}$	1	MS	K
10644253	6045±77	$2.39^{+1.12}_{-0.96}$	$1.10^{+0.04}_{-0.04}$	0.06±0.10	$4.396^{+0.007}_{-0.008}$	$0.24^{+0.08}_{-0.08}$	$55.6^{+24.0}_{-30.3}$	$38.0^{+21.1}_{-19.1}$	0	MS	L
10666592	6350±80	$2.11^{+0.29}_{-0.24}$	$1.50^{+0.04}_{-0.04}$	0.26±0.08	$4.017^{+0.009}_{-0.007}$	$0.91^{+0.11}_{-0.11}$	$47.2^{+27.2}_{-14.8}$	$9.5^{+3.6}_{-3.2}$	0	H	K

Continued on next page

Table 5.1: *Continued from previous page.*

KIC	$T_{\rm eff}$ [K]	Age [Gyr]	Mass [$M_\odot$]	[Fe/H] [dex]	$\log(g)$ [dex]	$\nu_{\rm s}\sin(i)$ [μHz]	i [°]	$P_{\rm rot}$ [days]	Flag	Type	Source
10730618	6150±180	$3.05^{+0.46}_{-0.29}$	$1.34^{+0.04}_{-0.05}$	-0.11±0.18	$4.062^{+0.008}_{-0.007}$	$0.56^{+0.23}_{-0.25}$	$55.6^{+23.7}_{-26.2}$	$16.1^{+13.9}_{-7.2}$	0	MS	L
10963065	6140±77	$7.15^{+1.92}_{-1.61}$	$0.99^{+0.06}_{-0.06}$	-0.19±0.10	$4.277^{+0.011}_{-0.011}$	$0.67^{+0.03}_{-0.03}$	$41.8^{+4.7}_{-3.6}$	$11.5^{+1.3}_{-1.0}$	0	MS	L
11081729	6548±82	$1.88^{+0.59}_{-0.42}$	$1.30^{+0.04}_{-0.05}$	0.11±0.10	$4.245^{+0.010}_{-0.009}$	$3.35^{+0.10}_{-0.10}$	$82.9^{+4.9}_{-5.9}$	$3.4^{+0.1}_{-0.1}$	0	H	L
11133306	5982±82	$5.14^{+0.86}_{-0.88}$	$1.06^{+0.04}_{-0.03}$	-0.02±0.10	$4.314^{+0.007}_{-0.004}$	$0.39^{+0.14}_{-0.13}$	$58.0^{+21.9}_{-24.1}$	$24.6^{+14.3}_{-10.0}$	0	MS	K
11253226	6642±77	$1.60^{+0.06}_{-0.13}$	$1.41^{+0.02}_{-0.01}$	-0.08±0.10	$4.173^{+0.005}_{-0.004}$	$2.55^{+0.11}_{-0.12}$	$49.3^{+6.3}_{-4.4}$	$3.4^{+0.4}_{-0.3}$	0	H	L
11295426	5793±74	$6.31^{+0.32}_{-0.34}$	$1.07^{+0.01}_{-0.02}$	0.12±0.07	$4.280^{+0.003}_{-0.003}$	$0.22^{+0.03}_{-0.03}$	$55.6^{+22.9}_{-20.9}$	$42.6^{+11.5}_{-14.5}$	0	MS	K
11401755	5911±66	$7.10^{+0.61}_{-0.59}$	$1.06^{+0.03}_{-0.02}$	-0.20±0.06	$4.039^{+0.004}_{-0.004}$	$0.55^{+0.09}_{-0.10}$	$64.6^{+17.4}_{-19.5}$	$18.5^{+4.7}_{-4.3}$	0	MS	K
11772920	5180±180	$10.67^{+2.73}_{-2.97}$	$0.83^{+0.04}_{-0.04}$	-0.09±0.18	$4.500^{+0.005}_{-0.008}$	$0.31^{+0.03}_{-0.04}$	$70.8^{+12.6}_{-14.3}$	$35.1^{+5.3}_{-4.7}$	1	MS	L
11807274	6225±75	$3.59^{+0.78}_{-0.45}$	$1.24^{+0.04}_{-0.04}$	0.00±0.08	$4.135^{+0.009}_{-0.007}$	$1.42^{+0.07}_{-0.07}$	$76.8^{+8.8}_{-9.4}$	$7.9^{+0.5}_{-0.5}$	0	MS	K
11853905	5781±76	$6.71^{+0.77}_{-0.67}$	$1.12^{+0.02}_{-0.03}$	0.09±0.10	$4.102^{+0.004}_{-0.005}$	$0.27^{+0.09}_{-0.09}$	$54.7^{+24.5}_{-28.8}$	$34.0^{+18.9}_{-17.2}$	1	MS	K
11904151	5647±74	$10.23^{+0.83}_{-0.67}$	$0.92^{+0.01}_{-0.02}$	-0.15±0.10	$4.344^{+0.003}_{-0.003}$	$0.24^{+0.06}_{-0.07}$	$64.9^{+18.1}_{-28.4}$	$40.9^{+16.5}_{-12.8}$	1	MS	K
12009504	6179±77	$3.97^{+0.57}_{-0.43}$	$1.17^{+0.02}_{-0.04}$	-0.08±0.10	$4.211^{+0.007}_{-0.005}$	$1.16^{+0.03}_{-0.04}$	$71.3^{+8.0}_{-5.7}$	$9.4^{+0.5}_{-0.5}$	0	MS	L
12069127	6276±77	$2.01^{+0.11}_{-0.13}$	$1.57^{+0.03}_{-0.02}$	0.08±0.10	$3.912^{+0.005}_{-0.004}$	$0.54^{+1.26}_{-0.33}$	$16.6^{+54.2}_{-6.3}$	$5.2^{+39.9}_{-3.9}$	1	H	L
12069424	5825±50	$6.67^{+0.81}_{-0.77}$	$1.05^{+0.02}_{-0.02}$	0.10±0.03	$4.287^{+0.007}_{-0.007}$	$0.40^{+0.01}_{-0.01}$	$44.7^{+6.2}_{-2.9}$	$20.5^{+2.0}_{-1.1}$	1	MS	L
12069449	5750±50	$7.39^{+0.89}_{-0.91}$	$0.99^{+0.02}_{-0.02}$	0.05±0.02	$4.353^{+0.007}_{-0.005}$	$0.31^{+0.01}_{-0.01}$	$34.0^{+3.0}_{-2.4}$	$21.2^{+1.8}_{-1.5}$	0	MS	L
12258514	5964±77	$4.05^{+0.18}_{-0.16}$	$1.26^{+0.01}_{-0.01}$	0.00±0.10	$4.126^{+0.004}_{-0.003}$	$0.39^{+0.04}_{-0.03}$	$35.0^{+22.9}_{-6.6}$	$16.7^{+10.0}_{-3.6}$	1	MS	L
12317678	6580±77	$2.46^{+0.22}_{-0.18}$	$1.34^{+0.04}_{-0.01}$	-0.28±0.10	$4.048^{+0.008}_{-0.009}$	$1.27^{+0.13}_{-0.14}$	$35.3^{+10.1}_{-5.2}$	$5.2^{+1.8}_{-0.9}$	0	H	L

5.5 Discussion

5.5.1 Verifying asteroseismic results

Priors on rotational parameters

In our Bayesian analysis, we have placed weakly informative priors on our sampled rotational parameters, $\nu_{\rm s}\sin(i)$ and $\cos(i)$ (see Section 5.3.1). The prior is especially important for the angle of inclination, which is hardest to infer from the data. We are able to validate the robustness of our asteroseismic results by confirming that their posterior distributions are data-dominated, and not prior-dominated. We can do so by comparing the 68% credible regions of the posterior estimates of $\nu_{\rm s}\sin(i)$ and i against the 68% credible regions of their priors.

A comparison between prior and posterior is shown for 94 stars in $\nu_{\rm s}\sin(i)$, i and $P_{\rm rot}$ in Figure 5.6, arranged by age. In this figure, results with means and errorbars that closely resemble the prior distribution can be interpreted as prior-dominated (i.e. poorly informed by the data). The projected splitting, $\nu_{\rm s}\sin(i)$, is overall well constrained, with only one star appearing prior dominated. This is expected, as the projected splitting is what we observe on the star before decoupling inclination and rotation. The angle of inclination i, sampled as $\cos(i)$, more closely follows the prior distribution in most cases. This is not a cause for concern, as our chosen prior accurately reflects the geometries of our observed stars, but it does highlight the key source of uncertainty in our rotation measurements. Combining the two, the rotation period $P_{\rm rot}$ has no stars directly corresponding to the effective prior on period, and globally follows a trend with increasing age, as we would expect from gyrochronology. The three outliers with fast rotation at late ages (KICs 6603624, 8760414 and 8938364) are likely to be faulty measurements in this work based on comparisons to the literature (see below).

The measurements of rotation rate as presented in this work are a product of

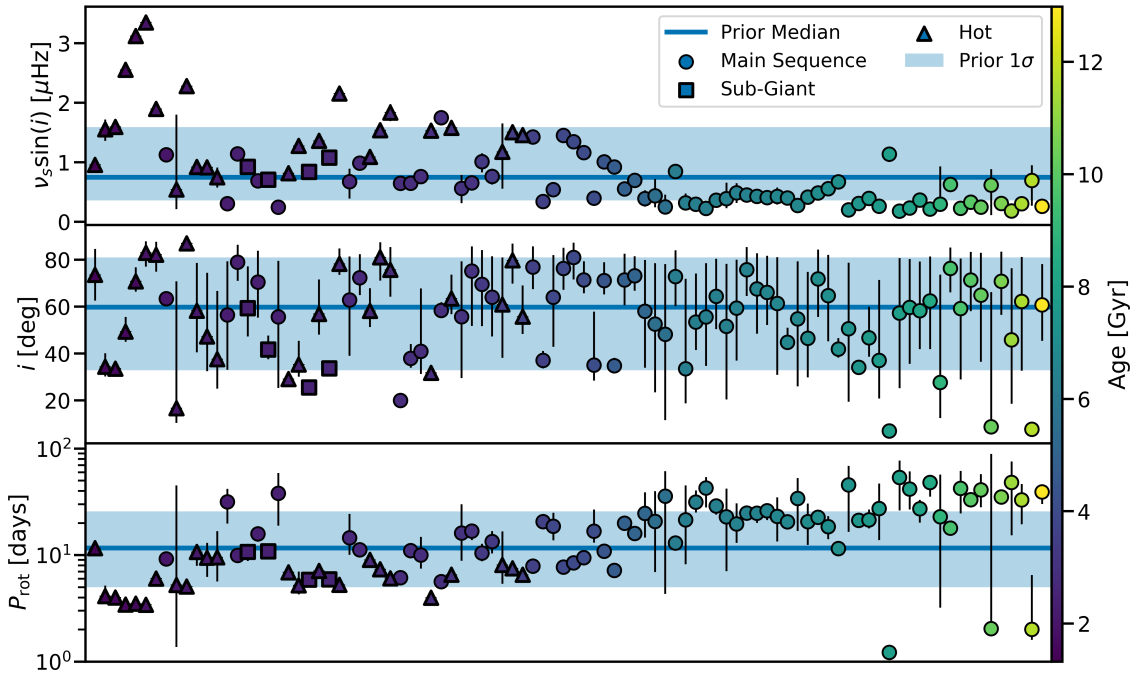

Figure 5.6: Comparisons between posterior estimates of rotational parameters (data points and error bars) against priors on said parameters (shaded regions). If error bars can not be seen they are smaller than the points. In both cases, the extent of the error bars and shaded regions indicate the 68% credible interval (1σ) of the posterior and prior distributions respectively. The solid lines indicate the median of the prior distributions. In this figure, results with means and errorbars that closely resemble the prior distribution can be interpreted as prior-dominated (i.e. poorly informed by the data). All stars are sorted and coloured by age. In the case of inclination angle i and rotation period $P_{\rm rot}$, the displayed priors are transformed from the priors imposed on the sampled parameters from which their posteriors were derived.

our Bayesian sampling of both projected splitting and angle of inclination. As seen in Figure 5.6, while there are instances of i or $\nu_{\rm s}\sin(i)$ closely resembling the prior (and therefore being prior-dominated), there are no cases of this when looking at the resulting period measurements, as they will have been informed by at least one strongly data-driven parameter (judging from Figure 5.6, commonly $\nu_{\rm s}\sin(i)$). From this, we can conclude that our ensemble of asteroseismic rotation rates is not strongly dominated by the priors imposed on projected splitting and inclination in our Bayesian analysis.

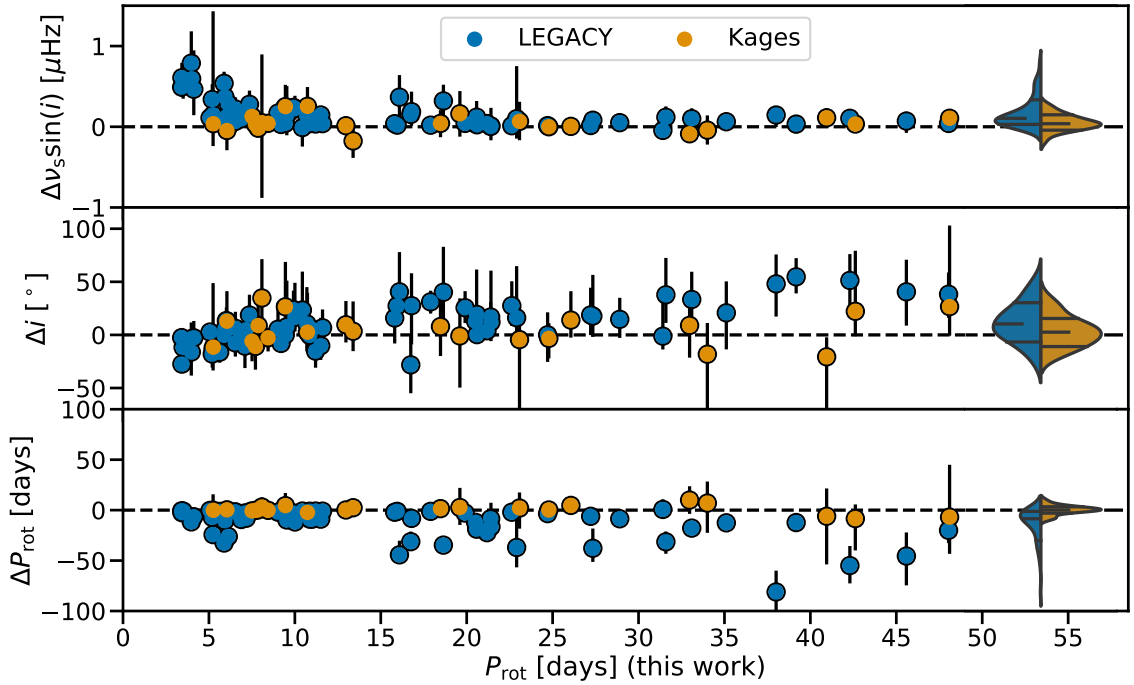

Figure 5.7: Comparisons between posterior estimates of rotational parameters from this work, LEGACY and Kages (Davies et al., 2016; Lund et al., 2017, private communication). Residuals are plotted against stellar rotation obtained in this work. The Δ indicates the literature value subtracted from the value obtained in this work (i.e. stars above the zero-line have higher values in this work). The violin plots on the right show the distribution of the residuals of LEGACY and Kages targets. The dashed and solid lines in the violin plots indicate the median and 16$^{\text{th}}$ and 84$^{\text{th}}$ percentiles, respectively. Ten stars have been excluded from this comparison: KICs 5094751, 6196457, 8349582, 8494142, 8554498, 105114430 and 11133306 all have extremely low rotation periods in Kages, with high uncertainties. Conversely, KICs 6603624, 8760414 and 8938364 have extremely high rotation periods in LEGACY with low uncertainties.

Comparisons to previous studies

In order to validate our results, we compared our rotational parameters to those obtained in the literature, as well as those resulting from the work presented in LEGACY and Kages, which were unreported and received through private communication.

Comparisons with LEGACY and Kages are shown in Figure 5.7 for projected splitting, inclination angle, and rotation period. In all three cases we show the values obtained in this work subtracted from external values. On the right of Figure 5.7, we show the distribution of the residuals, including the 16$^{\text{th}}$ and 84$^{\text{th}}$ percentiles.

The projected splitting, which is typically most closely constrained, is generally in good agreement with both studies, however LEGACY finds consistently lower $\nu_{\rm s}\sin(i)$ for the fastest rotators, deviating from our work by over 1σ. While this looks severe compared to the rest of the sample, at these rotation rates the splitting will lie around $5\,\mu$Hz, so the fractional difference remains relatively low. Neither LEGACY nor Kages used an isotropic prior for the inclination angle in their analyses, instead opting for a uniform prior. As posterior estimates of inclination angle are only loosely data-driven, the introduction of an isotropic prior should result in our analysis reporting globally higher inclination angles. This effect is seen in the comparisons of both inclination angles and rotation rates, where we find overall lower rotation rates compared to LEGACY for stars at very similar $\nu_{\rm s}\sin(i)$.

A number of stars are excluded from this comparison and compared individually as extreme outliers: KICs 5094751, 6196457, 8349582, 8494142, 8554498, 105114430 and 11133306 all have fast rotation rates (< 5 days) in Kages, compared to a broader spread of rotation rates in this work. At similar values for $\nu_{\rm s}\sin(i)$, Kages found much lower inclination angles with highly asymmetrical uncertainties, which are discriminated against in our work through our prior. Based on a comparison between the summary statistics of these stars, we conclude that the results found in this work have better marginalised over inclination angle, improving our measure of rotation.

Conversely, KICs 6603624, 8760414 and 8938364 have extremely slow rotation periods in LEGACY, but extremely *fast* (< 3 days) in this work. KICs 8760414 and 8938364 are excluded from the gyrochronology analysis below based on their number of effective samples and by-eye investigations. This exclusion is further justified by this comparison; these two stars have ages greater than 10 Gyr, making their fast rotation rates highly unlikely under any model of rotational evolution. Both these stars have low inclination angles ($< 10^\circ$), at which point the power in the split components of the seismic modes is so low that it becomes difficult to probe the measure of splitting. The posterior estimate of P for KIC 6603624 is well-resolved

in our analysis, but with an age of 7.8 Gyr, its measured rotation of 1.2 days is also highly unlikely under any model of rotational evolution. The LEGACY estimate of rotation is similarly extreme at 378 days. As this star also has a low inclination angle, we choose to discount its rotation at this stage as being wrongly constrained.

Finally, we look at KICs 8424992 and 10644253, which are included in Figure 5.7 but have $\Delta P_{\rm rot} < -50$. Both these stars are found to have low inclination angles in LEGACY, but larger inclination angles in this work. KIC 8424992 is already excluded from the gyrochronology analysis below based on its number of effective samples, and we choose to trust our rotation measurement for 10644253 over that found in LEGACY, which is anomalously high (199 days) for its age (2.4 Gyr).

We also compared our asteroseismic estimates of stellar rotation with similar studies in the literature, shown in Figure 5.8. These included: Davies et al. (2015), a study of the binary solar analogues 16 Cyg A & B; Nielsen et al. (2015), which our catalogue shares 5 stars with; and Benomar et al. (2018), an asteroseismic study of differential rotation with which our catalogue shares 40 targets. For the latter, we used their reported splitting value a_1, which represents the rotational splitting in the case of uniform latitudinal rotation. Unlike the comparison in Figure 5.7, these published values provide a clearer picture of the systematic uncertainty in asteroseismic rotation measurements.

Overall, Figure 5.8 shows no strong disagreements between our asteroseismic measurements for stellar rotation and those from the literature. The scatter of the residuals lies cleanly around the zero line, with a mean and spread of $-0.01^{+3.24}_{-1.82}$ days. The apparent increase in uncertainty for slow rotating stars is a product of period differences in Figure 5.8 not being shown on a fractional scale, and more slowly rotating stars being more difficult to constrain using asteroseismology. The agreement within 1σ here is encouraging, indicating that these independent Bayesian analyses are finding appropriate uncertainties on these rotation measurements.

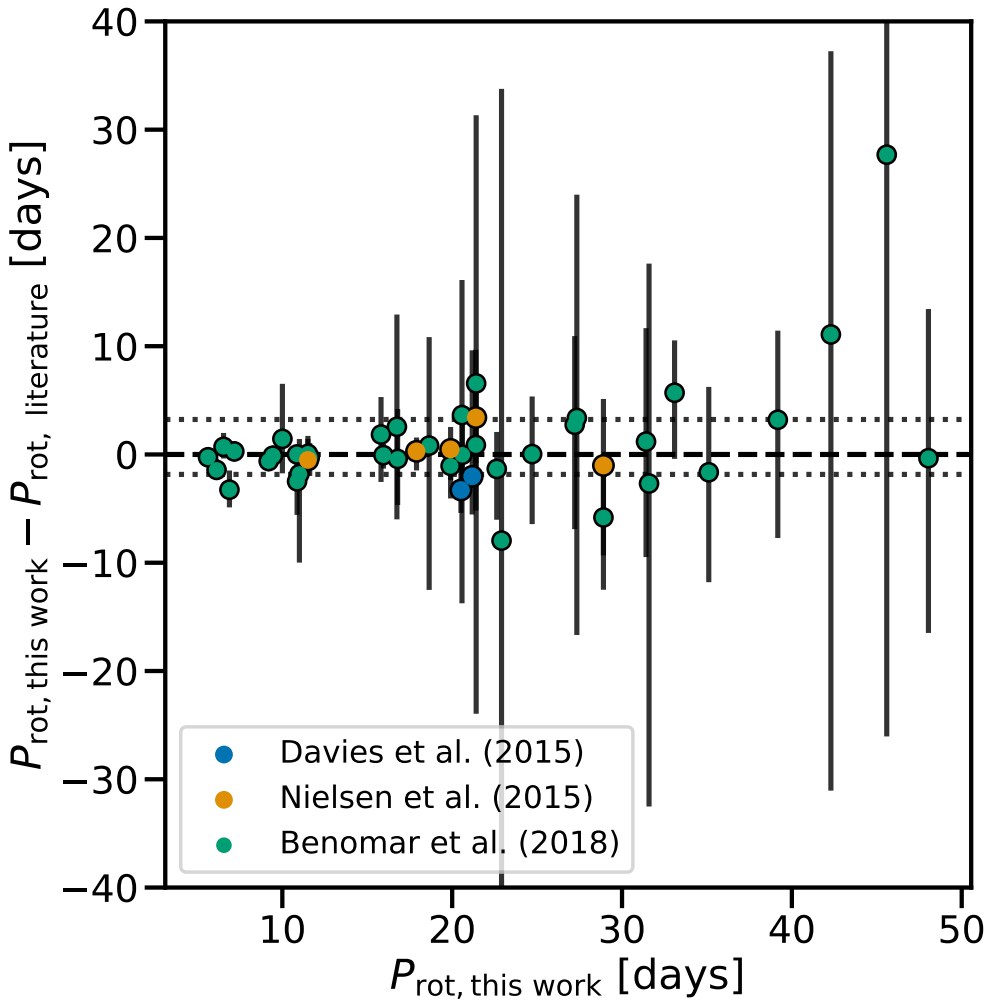

Figure 5.8: Comparisons between posterior estimates of asteroseismic rotation period from this work. Literature sources are: Davies et al. (2015) (16 Cyg A & B), Nielsen et al. (2015) (5 stars) and Benomar et al. (2018) (40 stars). We used the reported parameter a_1 from Benomar et al. (2018), which represents the rotational splitting in the case of uniform latitudinal rotation in their model.

Three stars were excluded from the comparison in Figure 5.8: KICs 6603624, 8760414 and 8938364. These are the same targets found to be outliers in a comparison to the LEGACY and Kages measurements (see above), with anomalously fast rotation rates and low inclination angles. The comparison with Benomar et al. (2018) further reinforces the decision to exclude them from the gyrochronology analysis, and to doubt the rotation measurements for these three stars presented in this work.

It is of note that 16 Cyg A, as reported by Davies et al. (2015) was found to be

rotating slightly faster in our analysis (deviating within 2σ), despite the fit being performed on the same data. We found an inclination angle that is slightly lower than Davies et al. (2015) for 16 Cyg A but a similar projected splitting, which would explain finding a lower value of rotation.

Internal vs external rotation

Different methods of rotation measurement probe different regions of stars. Asteroseismology of main sequence stars probes internal rotation in the near surface layers, where the observed p-modes are most sensitive (Lund et al., 2014; Davies et al., 2015). The more traditional technique of star-spot modulation instead probes the rotation rates of star spots on the surface. A distinct difference between rotation rates obtained through these different techniques would hold information about differential rotation (both latitudinal and radial) of near-surface layers, such as those we observe in the Sun (Beck, 2000).

Nielsen et al. (2015) performed a comparative analysis of 5 Sun-like stars observed with *Kepler* for which both star-spot modulation and asteroseismic rotation could be measured. They found no statistically significant difference between the two techniques.

Similarly, Benomar et al. (2015) performed an analysis on a larger sample of 22 stars. In this work they not only considered rotation rates from spots, but also spectroscopic measures of the projected surface rotation $\mathrm{v}\sin(i)$, which they found to be more reliable. Comparisons with asteroseismic observations again showed no significant differential radial rotation.

With our expanded sample of asteroseismic rotation we can perform a similar analysis, to both validate our sample and probe radial differential rotation. Figure 5.9 shows a comparison between spectroscopic $\mathrm{v}\sin(i)$ measurements as reported in LEGACY and Kages (left) and Benomar et al. (2018) (right). In these cases the asteroseismic $\mathrm{v}\sin(i)$ has been calculated using our measure of asteroseismic $\nu_s \sin(i)$ and the known asteroseismic radii. Three stars (KICs 6603624, 8760414

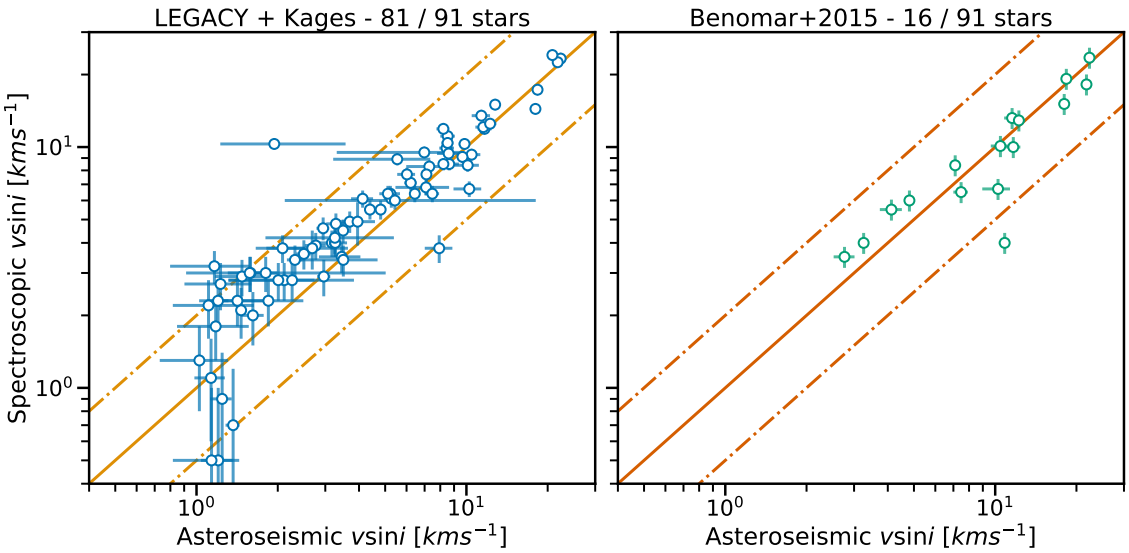

Figure 5.9: Comparisons between asteroseismic and spectroscopic measures of projected surface rotation ($v \sin(i)$). All asteroseismic (x-axis) values are from this work, all spectroscopic (y-axis) values are from the literature. *Left*: comparisons to 81 stars values reported in LEGACY and Kages. *Right*: comparisons to 16 stars observed by Benomar et al. (2015). Asteroseismic values are transformed from projected splitting ($\nu_s \sin(i)$) using the asteroseismic radius measurements presented in LEGACY and Kages. The solid lines indicate the 1:1 line, while the dash-dotted lines represent 2:1 and 1:2 lines.

and 8938364) have been excluded from this figure due to strong disagreements of measured rotation rates with the literature (see above).

For the LEGACY and Kages measurements, we find no strong deviation from the 1:1 line except at very low velocities, which is likely due to biases inherent to spectroscopic line broadening measurements (Doyle et al., 2014). For the Benomar et al. (2015) sample stars lie a lot closer to the 1:1 line.

Overall, there appears to be a global offset where spectroscopic measurements of projected rotation appear faster than asteroseismic measures. Based on the LEGACY and Kages $v \sin(i)$ values, this offset is roughly 18% (i.e. spectroscopic projected rotation rates are higher than asteroseismic rates). This offset is much smaller ($\sim$ 5%) for the Benomar et al. (2015) sample, albeit for far fewer stars. These offsets are within the typical disagreement between spectroscopic methods, based on comparisons of projected rotation measurements for red giant stars (see Figure 2, Doyle, Davies, Smalley, Chaplin & Elsworth, tay), especially at $< 5\,\mathrm{km\,s^{-1}}$.

We also repeated the Nielsen et al. (2015) comparison of asteroseismic rotation rates to surface rotation rates, using surface rotation measurements from Nielsen et al. (2013) and García et al. (2014). Figure 5.10 shows comparisons for 48 stars with rotation rates from both techniques, including 4 from the original Nielsen et al. (2015) sample. Rotation from spot-modulation are subject to measuring multiples of the true rotation rate, unlike an asteroseismic analysis (McQuillan et al., 2014), and therefore some stars may appear at the 2:1 and 1:2 lines on the Figure.

We repeated the analysis in Nielsen et al. (2015) by fitting a line of the form $P_{\rm surf} = m \times P_{\rm seis}$. For the purposes of this fit we used the larger of the asymmetrical uncertainties on the seismic rotation from this work. In order to avoid biasing the fit due to outliers, we only included stars below the 1.8:1 line.

Our fit found a value of $m = 0.96 \pm 0.03$, showing a close agreement ($< 2\sigma$) between asteroseismic and surface rotation rates on a population level. Of the 40 stars that were part of this analysis, 21 had a median value of $P_{\rm s} < 0.5$ in the gyrochronology analysis. We repeated the model fit for these stars only, and found a value of $m = 0.96 \pm 0.04$. As we find no statistically significant deviation, we conclude that our asteroseismic ensemble is in agreement with surface rotation measurements of these stars, and can therefore be used to draw inferences about gyrochronology.

The overall agreement within the ensemble between the photometric measures of surface rotation and the asteroseismic rotation rates is in line with previous studies of Sun-like stars (Gizon et al., 2013; Chaplin et al., 2013; Nielsen et al., 2015; Benomar et al., 2015), where different measures of rotation appear indistinguishable. Individual deviations from this agreement, as appears to be the case for a number of stars that lie distinctly off the 1:1 line, may warrant a more in-depth analysis, which we leave to future work.

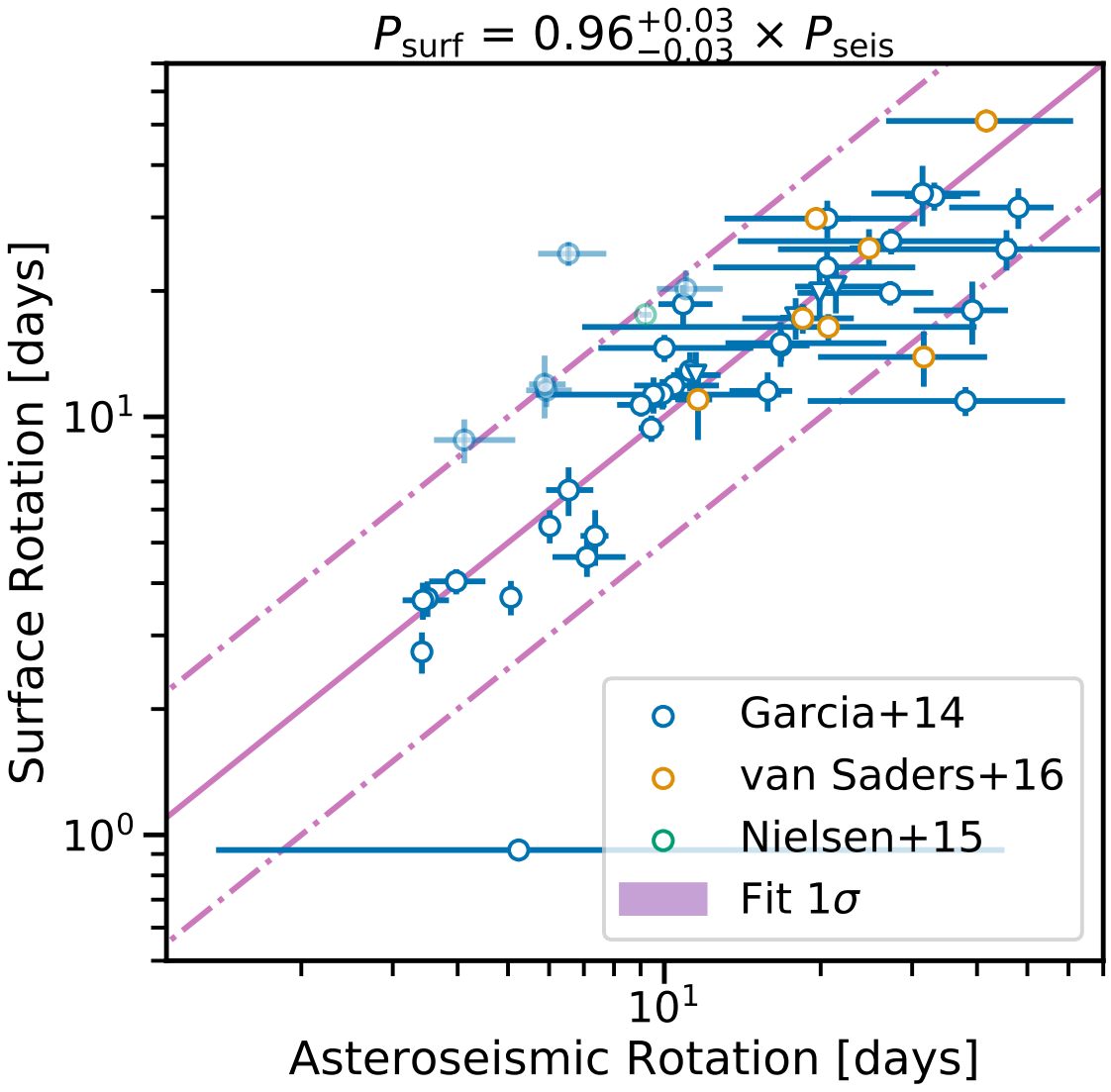

Figure 5.10: Comparisons between asteroseismic and photometric measures of stellar rotation, for 48 stars. Literature values were taken, in order of priority, from: García et al. (2014) and Nielsen et al. (2013) (i.e. if a value was reported in Nielsen et al. (2013) and also in García et al. (2014), the latter was used). Stars used in van Saders et al. (2016) are highlighted. The four triangles represent stars included in the original Nielsen et al. (2015) work. The solid line indicates 1:1 line, while the dash-dotted lines represent 2:1 and 1:2 lines. The shaded region around the 1:1 line is the 1σ credible interval of the fit to the data, the result of which is shown as the title of the figure. Stars that were not included in the fitting process are transparent.

5.5.2 Verifying consequences for gyrochronology

Rotational models for different stellar types

Of the 73 stars for which posterior probability distributions of P_s were obtained, 4 were sub-giants, 22 were 'hot' stars, and 47 were main sequence stars. In theory, different stellar types should hold different diagnostic information about the standard or WMB models. Our 'hot' stars should hold little to no information, as they lie above the Kraft break, where traditional gyrochronology relations used to build the stellar models break down (specifically, they do not lose angular momentum with time). Similarly, the rotational evolution of sub-giant stars is dominated by envelope expansion, complicating the interpretation of rotation.

We constructed joint posterior distributions for P_s for the three different stellar types. The 4 sub-giant stars together weakly preferred the standard model. The 22 'hot' stars on the other hand only did not strongly prefer one model over the other, which is as expected. The van Saders et al. (2019) models make simplifications that are less appropriate for hot stars, and thus the models are less trustworthy above the Kraft break.

Finally, the 47 MS stars were most strongly in favour of weakened magnetic braking and dominate the full joint posterior. Even when excluding both sub-giants and 'hot' stars from our sample and considering MS stars alone, our asteroseismic sample still lies in favour of a model of weakened magnetic braking at a critical Rossby number of 1.97.

Limits of our stellar models

The models presented in van Saders et al. (2019) were constructed for metallicities of $-0.4 < [\mathrm{Fe/H}] < 0.4$ dex, in steps of 0.1 dex. Our sample of 91 stars contained 4 stars with metallicities below -0.4 dex, and are shown as shaded symbols in Figure 5.5. Of these, 3 were included in our final sample of 73 stars used to evaluate our stellar models: KICs 7970740, 8684723 and KIC 9965715. All three are classed as

MS stars, with metallicities of -0.54 ± 0.10, -0.42 ± 0.10 and -0.44 ± 0.18 dex respectively, placing them within 2σ of the metallicity limits of our stellar models. KICs 8684723 and 9965715 strongly agrees with the WMB model, whereas KIC 7970740 weakly prefers the standard model. Excluding these stars was not found to significantly alter the joint posterior distribution shown in Figure 5.4.

Recently, Amard & Matt (2020) compared different rotational evolution models (van Saders & Pinsonneault, 2013; Matt et al., 2015, of which we use the former in this work) while studying the effect metallicity has on rotation. They found that metal-rich stars spin down significantly more effectively than metal poor stars (and vice-versa). The population of 73 stars used in our stellar model comparisons is roughly centred on a [Fe/H] of 0 dex, with a spread of 0.16 dex, with no stars significantly above or below ± 0.4 dex (as discussed above). While differences in stellar rotational evolution as a function of metallicity in this region are still somewhat pronounced, they are much less so than for more metal-rich or poor stars (see Figure 2 of Amard & Matt, 2020). However for stars with [Fe/H] < 0, the Matt et al. (2015) model prescription sees stars spin down more slowly than the models used in this work (i.e. they will rotate faster at later ages). It therefore is important to note that this work only explores the presence of weakened magnetic braking for a single braking law. A comparison the the Matt et al. (2015) braking model using this work's asteroseismic ensemble of rotation will be explored in a future paper.

When constructing KDEs from our stellar model samples (see Section 5.3.2), we selected fixed resolutions (or bin widths) for the KDEs. In mass, the bin width of $0.02\,M_{\odot}$, was larger than the uncertainties of 24 of 73 stars used to construct our joint posterior. For the stars with the smallest uncertainties, this significantly limits the size of the KDE being evaluated (as subsections of the full stellar models are used to evaluate individual stars, for computational efficiency). In order to confirm that these stars do not significantly affect the ensemble's preference towards the WMB model, we recalculated the joint posterior distribution for $P_{\rm s}$, excluding stars

with an uncertainty on mass smaller than the KDE bin-width While the 24 stars with small uncertainties do favour the WMB model, they do so very weakly, whereas the remaining 49 stars with larger uncertainties strongly favour the WMB model. Their removal from the total joint posterior probability does not significantly alter it from the distribution shown in Figure 5.4.

Systematic uncertainties from asteroseismology

In our model analysis, we used asteroseismic mass and age obtained using `BASTA`, as reported in LEGACY and Kages. Asteroseismic properties obtained through stellar models can be subject to systematic errors arising from differences in input physics and choice of stellar models not included in the reported statistical uncertainties. A quantification of these different systematic effects can be found in Section 4 of Silva Aguirre et al. (2015), based on the Kages catalogue. Combining their reported median systematic uncertainties due to input physics results in median uncertainties of 20% (up from 14%) on age and 5% (up from 3%) on mass for the Kages sample. For LEGACY, the median uncertainties are 18% (up from 10%) and 5.6% (up from 4%) for age and mass respectively.

In order to check the effect of systematic seismic uncertainties on our results for gyrochronology, we re-ran our model analysis described in Section 5.3.2 after inflating uncertainties on mass and age. We increased uncertainties by the fractional difference between the `BASTA` statistical uncertainties and the median full statistical and systematic uncertainties described above. For example, a LEGACY star with a mass of $2.0 \pm 0.5\,M_{\odot}$ would have its uncertainty inflated by 1.6% of its mass, to $0.53\,M_{\odot}$.

The results we found by repeating our model analysis with inflated uncertainties on asteroseismic mass and age closely replicated those found in our initial analysis. Specifically: subgiants preferred the standard model, and MS stars strongly favoured the WMB model. The 'hot' stars in this case slightly preferred the WMB model more when compared to the unaltered ensemble. The full joint posterior distribution

for 73 stars still strongly lies in favour of the WMB model even with the inflated uncertainties, including when considering only stars on the main sequence, when excluding stars outside the models' metallicity range, and when excluding stars with low uncertainties on mass (see Section 5.5.2).

To further test the limits of this analysis, we reran our mixture model fit, this time only shifting the asteroseismic ages younger by the systematic uncertainty, and retaining the statistical uncertainty (i.e. the ages of LEGACY and Kages stars were reduced by 8% and 6% respectively). In this scenario where all asteroseismic ages are overestimates, 'true' fast rotators at young ages would have been mistaken for fast rotators at old ages, suggesting the presence of weakened magnetic braking where none existed. Despite the shift in age, the results from this mixture model fit closely matched those found for the unaltered ensemble, finding 95.8% of the total posterior probability to lie below $P_{\rm s} = 0.5$ (as opposed to 98.4% for the unaltered ensemble). While the preference for the WMD model is slightly reduced, we conclude that the use of asteroseismic age is valid for the gyrochronology analyses presented in this work.

Binaries and Planet Hosts

For stars to be good probes of existing gyrochronology relations, their rotational evolution must occur in isolation. If a star interacts with a close binary companion (through tides, or a merger) the natural angular momentum loss can be disturbed, causing gyrochronology to mispredict ages (Leiner et al., 2019; Fleming et al., 2019). Between LEGACY and Kages, we have 8 known binaries.

First, KIC 8379927, KIC 7510397, KIC 10454113 and KIC 9025370 are spectroscopic binaries. This does not affect the asteroseismic analysis, but may affect their rotational evolution. Of these, KICs 8379927, 7510397 and 9025370 were included in the gyrochronology analysis. None of them preferred one model strongly over the other, with all three finding flat posteriors for $P_{\rm s}$.

Second, the binary pairs of KIC 9139151 & 9139163 and 16 Cyg A & B are

individually observed binary components with wide orbital separations, so we do not expect their binarity to have affected their rotational evolution.

While we chose not to account for star-planet tidal interactions in this work, we note that this may also disturb the natural stellar rotational evolution when tidal forces are at play (Maxted et al., 2015; Gallet & Delorme, 2019; Benbakoura et al., 2019), although this has been disputed by observations of asteroseismic planet hosts (Ceillier et al., 2016).

Evidence for Weakened Magnetic Braking in the literature

Weakened magnetic braking was first proposed by van Saders et al. (2016), in response to stars with spot rotation rates faster than expected from gyrochronology at their asteroseismic ages. This discrepancy was also indicated at around the same time in other studies of asteroseismic ages of main sequence stars (Nielsen et al., 2015; Angus et al., 2015; Davies et al., 2015). The theory of weakened magnetic braking has been both reinforced by recent studies (Metcalfe & Egeland, 2019), as well as disputed (Lorenzo-Oliveira et al., 2019), at least at the critical Rossby number originally proposed.

As a sanity check that the evidence in favour of weakened magnetic braking presented in this work is independently robust, we recreated the joint posterior probability for $P_{\rm s}$, this time looking only at MS stars, and excluding any of the 22 stars that were included in the original van Saders et al. (2016) sample. Of the 47 MS stars used to distinguish between stellar models, 16 were included in van Saders et al. (2016). Both the 16 van Saders et al. (2016) stars as well as the remaining MS stars favoured the WMB model in our analysis, although the van Saders et al. (2016) stars did so more strongly, not including any stars that favoured the standard model. Specifically, when considering only the van Saders et al. (2016) stars, 96.3% of the total joint probability lay below $P_{\rm s} = 0.5$, compared to 91.7% when considering only those stars *not* included in the van Saders et al. (2016) study. While the stars

initially used to propose the WMB model are still those most strongly in favour of it, it is encouraging that the inclusion of older, more quiescent stars and the use of asteroseismic rotation rates finds the same conclusion.

Finally, we comment on the results found by Lorenzo-Oliveira et al. (2019). Using spot rotation measurements of 14 solar twins, compared to grids of stellar models, they found marginal evidence favouring a standard rotational evolution. If weakened magnetic braking were to take place, it would most likely be outside the range of their sample, at $Ro_{\rm crit} \geq 2.29$, compared to the value of 1.97, used in this work. Our sample does not overlap with theirs, so we can not directly compare to their results. However our work does find statistical agreement between our asteroseismic rotation rates and ages, and a model of weakened magnetic braking at a critical Rossby number of 1.97. While a comparison to models of different Rossby numbers is outside the scope of this study, the LEGACY and Kages rotation sample may facilitate such studies in future work.

5.6 Conclusions

We fit an asteroseismic model to 95 stars in the Kages (Silva Aguirre et al., 2015; Davies et al., 2016) and LEGACY (Silva Aguirre et al., 2017; Lund et al., 2017) catalogues, representing the highest signal-to-noise main sequence stars observed in the *Kepler* short cadence. We simultaneously fit for oscillation frequencies, convective background, and stellar rotation in a hierarchical model. We obtained asteroseismic rotation rates for 91 stars, validated against previously published and unpublished asteroseismic rotation measurements, making our new rotation catalogue the largest self-consistent sample of asteroseismic rotation in main sequence stars to date.

Recent work by van Saders et al. (2019) proposed two different models of stellar rotational evolution: a 'standard' (or smooth) evolution model, and a weakened magnetic braking (WMB) model, where at a certain point in a star's main sequence lifetime it ceases to lose angular momentum. By describing both models as KDEs, we

evaluated them in a hierarchical Bayesian mixture model against our observations, modulated by a mixture model weighting parameter P_s. This evaluation was done in five dimensions, namely: asteroseismic mass, age and rotation, and spectroscopic effective temperature and metallicity. If P_s was close to 0 a given star was more likely to be drawn from the WMB model, and vice-versa. Of our sample of 91 stars in which we measured asteroseismic rotation, 73 stars were suitable for this analysis.

We leave the reader with the following conclusions:

1. We found that our ensemble was more likely to be drawn from the WMB model than the standard model, for stars evolved under a van Saders & Pinsonneault (2013) braking law. In other words, our asteroseismic observations strongly favour a model of weakened magnetic braking first proposed in van Saders et al. (2016), at a critical Rossby number of $Ro_{\rm crit} = 1.97$. This work expands upon the analysis done in van Saders et al. (2019), by including quiescent stars older than the Sun. This conclusion was found to be robust against choice of ensemble members and potentially underestimated asteroseismic systematic uncertainties. A comparison to other braking laws (e.g. Matt et al., 2015) using our asteroseismic ensemble will take place in future work.

2. We compared our new asteroseismic rotation rates with surface rotation measures from spot modulation. Our findings replicate those in similar research by Nielsen et al. (2015) and Benomar et al. (2015), in the sense that we find no statistically significant difference between seismic and spot modulation measures of stellar rotation for our ensemble. However we note that while on a population level there is no difference between the two measures of rotation, there are a number of individual stars that lie significantly away from the 1:1 line which may warrant further analysis.

3. We present new techniques for asteroseismic 'peak-bagging' frequency analyses, specifically: expressing well known empirical asteroseismic relations in

a hierarchical modelling structure to account for effects not included in our model (such as glitches), and replacing uninformative empirical relations with Gaussian Process priors.

4. The new asteroseismic rotation catalogue presented in this work will act as an entry point for more detailed studies outside the scope of this work, such as comparisons between different braking laws and individually modelled values of stellar rotation at different critical Rossby numbers.

Finally, in the near future our asteroseismic rotation catalogue will be further complemented with improved surface rotation and atmospheric parameters from spectroscopic surveys such as LAMOST (Deng et al., 2012), 4MOST (de Jong et al., 2014) and WEAVE (Dalton et al., 2014), and new asteroseismic measurements of age from the *K2* and *TESS* missions. With these surveys, large scale ensemble asteroseismology will continue to increase the possibilities for understanding gyrochronology.

All analysis performed as part of this work is stored in an online public repository[11].

[11] https://github.com/ojhall94/malatium

Chapter 6

Conclusions and Future Prospects

In this thesis, I have presented two studies showing how asteroseismology can be combined with Bayesian statistical techniques to make new inferences of astrophysics. In particular I focused on asteroseismology of solar-like oscillators: main sequence, sub-giant and red giant stars with convective outer envelopes that oscillate in the same manner as the Sun. By studying changes on the surfaces of these stars, asteroseismology provides a window into their internal processes and fundamental properties.

In Chapter 2, we saw how asteroseismology of solar-like oscillators using data from the CoRoT, *Kepler*, *K2* and *TESS* missions has brought new insights into stars throughout the Milky Way. Due to the turbulent motion of convective outer layers of these stars, standing waves are formed inside the stellar cavity. These standing waves can be described using spherical harmonics, and produce modes of oscillation regularly spaced in frequency. This pattern of oscillations has two well-observed global properties: $\Delta\nu$, the spacing between consecutive overtones, which is related to the sound travel time inside a star; and $\nu_{\rm max}$, the frequency at which oscillation amplitudes are highest, which is related to the surface-gravity and temperature of a star.

If we convert a time series of a solar-like oscillator to the frequency-domain, its modes of oscillation will appear as regularly spaced peaks in a comb-like struc-

ture around $\nu_{\rm max}$. In Chapter 3, we saw how we can make this conversion using the Fourier Transform, or more specifically through techniques that approximate a Fourier Transform, such as the Lomb-Scargle periodogram. Using the correct normalisations based on Parseval's Theorem, modes of oscillation will appear in the periodogram at predictable amplitudes.

As discussed in Chapter 3, the open-source Python package 'Lightkurve' hosts tools that can calculate a rudimentary $\Delta\nu$ and $\nu_{\rm max}$, given a well-resolved periodogram of a solar-like oscillator. These tools use an autocorrelation function, detecting an excess of power in the frequency-domain to locate $\nu_{\rm max}$. Using the regularly spaced nature of the 'comb' of modes, $\Delta\nu$ can also be found. Along with a value for effective temperature, $\nu_{\rm max}$ and $\Delta\nu$ can provide a value for mass, radius and surface gravity for a solar-like oscillator, using the seismic scaling relations presented in Chapter 2, subject to the assumption that these relations apply equally to all solar-like oscillators.

In Chapter 4, I introduced the concept of hierarchical Bayesian modelling; a statistical framework where multiple so-called 'latent' parameters all draw shared information from a distribution described by 'hyperparameters'. This technique is particularly useful for measuring a common property of a large ensemble of stars.

We fit two hierarchical models to an ensemble of 5576 Red Clump (RC) stars; evolved low-mass red giant stars that burn helium in their cores. These stars are expected to have very similar core masses, and therefore luminosities. As such, the stars in our ensemble could be modelled as being drawn from the same underlying distribution, characterised by hyperparameters. This approach had two aims: to use the common luminosity of this sample to calibrate distances to these stars measured by the *Gaia* mission, and to assess the assumptions used in the calculation of stellar radius using asteroseismic scaling relations.

Applying a hierarchical model to measurements of distance from *Gaia* and mag-

nitudes from 2MASS, we found that *Gaia* Data Release 2 (DR2) parallaxes in the *Kepler* field were subject to a zero-point offset of $-41 \pm 10\,\mu$as. This is in agreement with similar studies in the literature. Applying this offset to the *Gaia* distances, we inferred that our sample of RC stars had a mean absolute magnitude of -1.634 ± 0.018 mag in the 2MASS K band, and 0.546 ± 0.016 mag in the *Gaia* G band.

We repeated the analysis using asteroseismic data. Instead of calculating luminosities for the stars in our ensemble using *Gaia* parallaxes, we calculated their asteroseismic radius through the seismic scaling relations. Combined with temperature, stellar radius provides a value for luminosity that is distance independent. The absolute magnitudes of the RC stars inferred through asteroseismology alone did not agree with those found using *Gaia* and 2MASS data. We found that by making a global shift of -70 K in the temperature values used in the asteroseismic analysis, the asteroseismic absolute magnitude of the RC would agree with the *Gaia* estimate. However this same effect could be reproduced by keeping temperature constant and changing asteroseismic stellar radius on the order of 2%. While small, this change in radius is within the typical systematic uncertainty imposed by the choice of model-motivated corrections to the asteroseismic scaling relations (i.e. the use of other available models not considered in this work would have found the two estimates of the RC absolute magnitude to agree).

While the absolute magnitude of the RC changed depending on choice of input data and model prescriptions, the *spread* of the distribution of RC absolute magnitudes was more consistent. Using the asteroseismic estimates for the RC absolute magnitudes (which had an overall lower uncertainty), we found the spread of the RC to be 0.03 mag in the K band, and 0.13 mag in the G band. When using the RC as a standard candle, this spread in the K band implies a 1% precision on distance, a factor of 5 improvement from the previously most precise measurement of this property. For the G band this is 6%, in line with previous studies.

In Chapter 5 I presented a more detailed asteroseismic analysis of main sequence stars. By modelling modes of oscillation, a granulation background, and the effects of stellar rotation simultaneously, we measured asteroseismic rotation rates for 91 stars. This sample contains some of the highest signal-to-noise main sequence stars observed in the *Kepler* field and represents the largest self-consistent catalogue of asteroseismic rotation for such stars to date. As part of this so-called 'peak-bagging' analysis, we introduced new techniques into the fitting process. Specifically: expressing the asymptotic relation (presented in Chapter 2) as a hierarchical model and expressing the relationship between oscillation mode linewidth and frequency as a Gaussian Process. These approaches meant that our model could more easily account for sources of uncertainty that we did not explicitly include, such as sudden changes in the structure of the stellar interior (causing so-called 'glitches').

Conventionally, stellar rotation is measured by observing brightness modulations as star spots rotate in and out of view on the stellar surface. This technique is applicable to a vast number of stars, but is more difficult for older, more quiescent stars. Asteroseismology instead probes stellar rotation by observing how rotation affects modes of oscillation. This has no strong dependency on age, allowing for measurements of rotation in stars that are otherwise inaccessible. Comparing rotation rates from asteroseismology and those from spot modulation, we found no statistically significant difference between the two, in line with previous studies on smaller samples.

As stars age along the main sequence, they lose angular momentum (and therefore slow down) in a manner that is predictable for young stars, but unpredictable for old stars (the study of which is called *gyrochronology*). Understanding this relation would allow astronomers to estimate stellar age for old stars using a measurement of the star's rotation. In order to help understand the age-rotation relation, we used our stellar ensemble to distinguish between two population models of stellar rota-

tional evolution: a 'standard' (i.e. smooth) model, where the angular momentum loss mechanism is consistent throughout a main sequence lifetime; and a 'weakened magnetic braking' (WMB) model, where a star ceases to lose angular momentum once it has evolved to a point where its Rossby number (an expression of the rotation relative to the convective turnover timescale) is equal to a critical value of $Ro_{\rm crit} = 1.97$, meaning that old stars will rotate faster than they would under the standard scenario.

We evaluated our data against both models in a hierarchical Bayesian mixture model, calculating the posterior probability that each star was drawn from either the standard or the WMB model. Overall, we found that our ensemble strongly agreed with the WMB model (i.e. a scenario where angular momentum loss ceases at a certain stage in the main sequence evolution), in the sense that $> 98\%$ of our posterior probability for the full ensemble lay in favour of the WMB model over the standard model. We found this result to be robust against various checks, such as comparing different sub-sets of the ensemble, and inflating asteroseismic uncertainties on age, mass and radius to account for systematic uncertainties in choice of stellar model prescriptions.

Future Prospects

The two studies presented in Chapters 4 and 5 of this thesis are a 'first-foot-in-the-door' look at how asteroseismology can be used to provide meaningful insights into the adjacent fields of standard candles and gyrochronology. With this, I mean to say that both studies present methods that can be immediately expanded upon to improve the precision and scope of the inference obtained from asteroseismic ensembles of red giant and main sequence stars.

The study of core-helium burning Red Clump stars presented in Chapter 4 did not explicitly account for effects of temperature (or colour) and metallicity on the

luminosity and spread of the Clump, which are both factors that impact stellar luminosity. Our work was recently expanded upon by Chan & Bovy (2020), who incorporated a relationship between the RC absolute magnitude, temperature, metallicity and abundance of alpha elements, applied to an all-sky sample using the APOGEE (Majewski et al., 2017) survey. While our measurements for the mean RC absolute magnitude are consistent, they find an overall larger spread of 0.09 mag in the K band, implying that the precision achieved through asteroseismology in this thesis was an overestimate.

We are currently expanding upon both works, by improving the Hall et al. (2019) model presented in this thesis. The novelty of this new approach is that instead of using an empirical relation to describe the relationship between stellar properties, we instead model the absolute magnitude, metallicity, colour, and alpha abundance of our stellar ensemble as being drawn from a single multivariate distribution. This approach also marginalises over the correlations between these various properties. Accounting for the correlations between these stellar properties will further improve the use of the RC as a standard candle, allowing for a precision higher than if one only had a measurement of a star's magnitude. The closer we get to a model that fully describes the relationship between the RC luminosity and fundamental stellar properties, the closer we get to really understanding its use as a standard candle.

Similarly, the entry of large ensembles of asteroseismic rotation rates into the study of stellar rotational evolution is a relatively new one. The work presented in this thesis only distinguished between a model of smooth rotational evolution, and one with a very specific critical Rossby number of $Ro_{\rm crit} = 1.97$. This is far from a comprehensive analysis of the field of gyrochronology. Recent works have implied the presence of a much larger critical Rossby number ($Ro_{\rm crit} > 2.29$, Lorenzo-Oliveira et al., 2019), and that metallicity has a much stronger impact on angular momentum loss in rotational evolution models not considered in this thesis (Matt et al., 2015;

Amard & Matt, 2020).

In a similar fashion to the Red Clump study, this study of stellar rotation can be improved by considering multiple models of rotational evolution, and marginalising over the value of the critical Rossby number at which angular momentum loss ceases. While computationally intensive, this could be achieved by individually modelling the stars in this ensemble under different model prescriptions of stellar rotational evolution, including different critical Rossby numbers. A more comprehensive study of model comparisons is a natural next step for this asteroseismic ensemble, and will provide a more complete overview of the state of the field of rotational evolution on the main sequence.

The *Kepler* and CoRoT missions started an era of large-data asteroseismology, providing distance-independent measures of fundamental stellar properties and stars' internal physics. However while *Kepler* provided high-precision data, it was for a single patch on the sky, subject to local biases, and size-limited. At time of writing, this is set to change significantly. The now complete *K2* mission, following on from *Kepler*, has made multiple observations around the ecliptic, containing multitudes of solar-like oscillators yet to be analysed. More recently, the all-sky *TESS* mission is about to complete its nominal 2-year run, with a confirmed extended mission to follow.

The increased sample size of asteroseismic observations will have a significant impact on the intersection of asteroseismology and adjacent fields. While Chapter 5 presented a study of 91 main sequence stars with asteroseismic age and rotation, the short-cadence observations of *TESS* and *K2* will provide a likely additional ~ 350 main sequence stars for which asteroseismic age can be determined. Along with measures of stellar rotation from surface modulation, this will increase the number of stars with asteroseismic ages with which gyrochronology can be calibrated by a factor of 4. For Red Clump stars, for which oscillations can be measured using a

longer 30 minute cadence, *TESS*'s Full Frame Images are set to provide detections of asteroseismic oscillations in hundreds of thousands of evolved stars, many belonging to the Clump. This increase in detections of Clump stars will bring asteroseismic studies of the Red Clump in line with spectroscopic surveys, and allow us to study the properties of the Red Clump throughout the entire sky.

In this thesis, we have discussed open-source tools such as Lightkurve, as well as contemporary Bayesian statistical techniques such as hierarchical latent variable models, mixture models, and Gaussian Processes. At time of writing, new statistical techniques are continuing to bring new insights into the field of astrophysics. Neural networks, for example, can provide relationships between asteroseismology and fundamental stellar parameters without having to formalise said relationship empirically (see e.g. Ness et al., 2015). Using Gaussian Processes, new asteroseismic models of $\nu_{\rm max}$ and $\Delta\nu$ can be fit directly to the time-domain, making asteroseismology possible even in low signal-to-noise cases with short observations (see e.g. Farr et al., 2018; Foreman-Mackey et al., 2017). As asteroseismic analyses of solar-like oscillators become more accessible, distance-independent fundamental parameters of these stars become a foundation on which adjacent fields of study begin their analyses. The synergy between observational asteroseismology and astrophysics is rapidly growing into one of the most exciting aspects of modern astronomy, bolstered by expansive space-based missions and Bayesian statistics.

Bibliography

Aerts C., Christensen-Dalsgaard J., Kurtz D. W., 2010, Asteroseismology, Astronomy and Astrophysics Library. Springer Science+Business Media B.V., Heidelberg, Germany

Alves D. R., 2000, The Astrophysical Journal, 539, 732

Amard L., Matt S. P., 2020, The Astrophysical Journal, 889, 108

Anderson E. R., Duvall T. L., Jefferies S. M., 1990, The Astrophysical Journal, 364, 699

Angus R., Aigrain S., Foreman-Mackey D., McQuillan A., 2015, Monthly Notices of the Royal Astronomical Society, 450, 1787

Appourchaux T., 2004, Astronomy & Astrophysics, 428, 1039

Appourchaux T., 2014, A crash course on data analysis in asteroseismology

Appourchaux T., Corbard T., 2019, Astronomy and Astrophysics, 624, A106

Appourchaux T., Gizon L., Rabello-Soares M.-C., 1998, Astronomy and Astrophysics Supplement Series, 132, 107

Appourchaux T., et al., 2010, Astronomy and Astrophysics Review, 18, 197

Appourchaux T., et al., 2016, Astronomy & Astrophysics, 595, C2

Appourchaux T., Boumier P., Leibacher J. W., Corbard T., 2018, Astronomy and Astrophysics, 617, A108

Arenou F., et al., 2018, Astronomy & Astrophysics, 616, A17

Astraatmadja T. L., Bailer-Jones C. A. L., 2016a, The Astrophysical Journal, 832, 137

Astraatmadja T. L., Bailer-Jones C. A. L., 2016b, The Astrophysical Journal, 833, 119

Astraatmadja T. L., Bailer-Jones C., 2017, American Astronomical Society Meeting Abstracts, p. 134.06

Astropy Collaboration et al., 2013, Astronomy and Astrophysics, 558, A33

Astropy Collaboration et al., 2018, The Astronomical Journal, 156, 123

Baglin A., Michel E., Auvergne M., Team C., 2006, Proceedings of SOHO 18/GONG 2006/HELAS I, Beyond the spherical Sun, 624, 34

Bahng J., Schwarzschild M., 1963, The Astrophysical Journal, 137, 901

Bailer-Jones C. A. L., 2015, Publications of the Astronomical Society of the Pacific, 127, 994

Bailer-Jones C. a. L., Rybizki J., Fouesneau M., Mantelet G., Andrae R., 2018, The Astronomical Journal, 156, 58

Ball W. H., Gizon L., 2017, arXiv:1702.02570 [astro-ph]

Ballot J., Turck-Chièze S., García R. A., 2004, Astronomy and Astrophysics, 423, 1051

Ballot J., Garcia R. A., Lambert P., 2006, Monthly Notices of the Royal Astronomical Society, 369, 1281

Ballot J., Appourchaux T., Toutain T., Guittet M., 2008, Astronomy & Astrophysics, 486, 867

Barnes S. A., 2003, The Astrophysical Journal, 586, 464

Barnes S. A., 2007, The Astrophysical Journal, 669, 1167

Barnes S. A., 2010, The Astrophysical Journal, 722, 222

Barnes S. A., Weingrill J., Fritzewski D., Strassmeier K. G., Platais I., 2016, The Astrophysical Journal, 823, 16

Basu S., et al., 1997, Monthly Notices of the Royal Astronomical Society, 292, 243

Basu S., Chaplin W. J., Elsworth Y., New R., Serenelli A. M., 2009, The Astrophysical Journal, 699, 1403

Batalha N. M., 2014, Proceedings of the National Academy of Sciences of the United States of America, 111, 12647

Batalha N. M., et al., 2011, The Astrophysical Journal, 729, 27

Bayes M., Price M., 1763, Philosophical Transactions of the Royal Society of London Series I, 53, 370

Beck J. G., 2000, Solar Physics, 191, 47

Beck P. G., 2013, PhD thesis

Beck P. G., De Ridder J., Aerts C., Kallinger T., Hekker S., García R. A., Mosser B., Davies G. R., 2012, Astronomische Nachrichten, 333, 967

Bedding T. R., 2014, Solar-like oscillations: An observational perspective. p. 60

Bedding T. R., et al., 2010, The Astrophysical Journal Letters, 713, L176

Bedding T. R., et al., 2011, Nature, 471, 608

Belkacem K., Goupil M. J., Dupret M. A., Samadi R., Baudin F., Noels A., Mosser B., 2011, Astronomy and Astrophysics, 530, A142

Belkacem K., Samadi R., Mosser B., Goupil M.-J., Ludwig H.-G., 2013, Progress in Physics of the Sun and Stars: A New Era in Helio- and Asteroseismology, 479, 61

Bell K. J., 2020, in American Astronomical Society Meeting Abstracts. American Astronomical Society Meeting Abstracts. p. 106.06

Bellinger E. P., 2019, Monthly Notices of the Royal Astronomical Society, 486, 4612

Bellinger E. P., 2020, Monthly Notices of the Royal Astronomical Society: Letters, 492, L50

Bellinger E. P., Hekker S., Angelou G. C., Stokholm A., Basu S., 2019, Astronomy & Astrophysics, 622, A130

Benbakoura M., Réville V., Brun A. S., Le Poncin-Lafitte C., Mathis S., 2019, The Astrophysical Journal, 621, A124

Benomar O., Takata M., Shibahashi H., Ceillier T., García R. A., 2015, Monthly Notices of the Royal Astronomical Society, 452, 2654

Benomar O., et al., 2018, Science, 361, 1231

Berger T. A., Huber D., van Saders J. L., Gaidos E., Tayar J., Kraus A. L., 2020, arXiv:2001.07737 [astro-ph]

Betancourt M. J., Girolami M., 2013, arXiv e-prints, p. arXiv:1312.0906

Borucki W. J., Koch D., Team K. S., 2010, AAS/Division for Planetary Sciences Meeting Abstracts #42, 42, 47.03

Bouchy F., Carrier F., 2001, Astronomy & Astrophysics, 374, L5

Brogaard K., et al., 2018, Monthly Notices of the Royal Astronomical Society, 476, 3729

Brown A. G. A., 2018, Astrometry and Astrophysics in the Gaia Sky, 330, 13

Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, The Astrophysical Journal, 368, 599

Buchhave L. A., Latham D. W., 2015, The Astrophysical Journal, 808, 187

Buzasi D. L., 2000, Experiment Design and Data Reduction for Seismology: Lessons Learned from WIRE. p. 9

Buzasi D. L., et al., 2005, The Astrophysical Journal, 619, 1072

Cannon R. D., 1970, Monthly Notices of the Royal Astronomical Society, 150, 111

Carpenter B., et al., 2017, Journal of Statistical Software, 76, 1

Casagrande L., VandenBerg D. A., 2014, Monthly Notices of the Royal Astronomical Society, 444, 392

Casagrande L., VandenBerg D. A., 2018a, Astrophysics Source Code Library, p. ascl:1805.022

Casagrande L., VandenBerg D. A., 2018b, Monthly Notices of the Royal Astronomical Society, 475, 5023

Casagrande L., et al., 2014, The Astrophysical Journal, 787, 110

Casagrande L., et al., 2016, Monthly Notices of the Royal Astronomical Society, 455, 987

Ceillier T., et al., 2016, Monthly Notices of the Royal Astronomical Society, 456, 119

Chan V. C., Bovy J., 2020, Monthly Notices of the Royal Astronomical Society, 493, 4367

Chaplin W. J., Basu S., 2008, Solar Physics, 251, 53

Chaplin W. J., Basu S., 2014, Space Science Reviews, 186, 437

Chaplin W. J., Basu 2017, Asteroseismic Data Analysis: Foundations and Techniques, 1st edn. Princeton University Press, Princeton, New Jersey

Chaplin W. J., Miglio A., 2013, Annual Review of Astronomy and Astrophysics, 51, 353

Chaplin W. J., et al., 1999, The Monthly Notices of the Royal Astronomical Society, 308, 405

Chaplin W. J., Houdek G., Appourchaux T., Elsworth Y., New R., Toutain T., 2008, Astronomy & Astrophysics, 485, 813

Chaplin W. J., et al., 2010, The Astrophysical Journal, 713, L169

Chaplin W. J., et al., 2011, Science, 332, 213

Chaplin W. J., et al., 2013, The Astrophysical Journal, 766, 101

Chaplin W. J., Elsworth Y., Davies G. R., Campante T. L., Handberg R., Miglio A., Basu S., 2014, Monthly Notices of the Royal Astronomical Society, 445, 946

Chaplin W. J., et al., 2015, Publications of the Astronomical Society of the Pacific, 127, 1038

Chaplin W. J., et al., 2020, Nature Astronomy

Chen Y. Q., Casagrande L., Zhao G., Bovy J., Silva Aguirre V., Zhao J. K., Jia Y. P., 2017, The Astrophysical Journal, 840, 77

Christensen-Dalsgaard J., 2002, Reviews of Modern Physics, 74, 1073

Christensen-Dalsgaard J., 2008, Astrophysics and Space Science, 316, 13

Christensen-Dalsgaard J., Duvall Jr. T. L., Gough D. O., Harvey J. W., Rhodes Jr. E. J., 1985, Nature, 315, 378

Christensen-Dalsgaard J., Dappen W., Lebreton Y., 1988, Nature, 336, 634

Christensen-Dalsgaard J., et al., 2010, The Astrophysical Journal, 713, L164

Claverie A., Isaak G. R., McLeod C. P., van der Raay H. B., Cortes T. R., 1979, Nature, 282, 591

Claytor Z. R., van Saders J. L., Santos Â. R. G., García R. A., Mathur S., Tayar J., Pinsonneault M. H., Shetrone M., 2020, The Astrophysical Journal, 888, 43

Corsaro E., De Ridder J., 2015, in European Physical Journal Web of Conferences. p. 06019 (arXiv:1509.08311), doi:10.1051/epjconf/201510106019

Couvidat S., Turck-Chieze S., Gracia R. A., Corbard T., 2003a, The Search for the Solar Core Dynamics with SoHO/GOLF. p. 276

Couvidat S., Turck-Chièze S., Kosovichev A. G., 2003b, The Astrophysical Journal, 599, 1434

Cowling T. G., 1941, Monthly Notices of the Royal Astronomical Society, 101, 367

Cox J. P., 1980, Research supported by the National Science Foundation Princeton, NJ, Princeton University Press, 1980. 393 p.

Creevey O. L., et al., 2017, Astronomy and Astrophysics, 601, A67

Dalton G., et al., 2014, Project overview and update on WEAVE: the next generation wide-field spectroscopy facility for the William Herschel Telescope. p. 91470L, doi:10.1117/12.2055132

Davies G. R., Miglio A., 2016, Astronomische Nachrichten, 337, 774

Davies G. R., Chaplin W. J., Elsworth Y., Hale S. J., 2014, Monthly Notices of the Royal Astronomical Society, 441, 3009

Davies G. R., et al., 2015, Monthly Notices of the Royal Astronomical Society, 446, 2959

Davies G. R., et al., 2016, Monthly Notices of the Royal Astronomical Society, 456, 2183

Davies G. R., et al., 2017, Astronomy and Astrophysics, 598, L4

Deheuvels S., et al., 2012, The Astrophysical Journal, 756, 19

Deheuvels S., et al., 2014, Astronomy and Astrophysics, 564, A27

Deng L.-C., et al., 2012, Research in Astronomy and Astrophysics, 12, 735

Deubner F.-L., 1975, Astronomy and Astrophysics, 44, 371

Deubner F.-L., Gough D., 1984, Annual Review of Astronomy and Astrophysics, 22, 593

Domingo V., Fleck B., Poland A. I., 1995, Solar Physics, 162, 1

Doyle A. P., Davies G. R., Smalley B., Chaplin W. J., Elsworth Y., 2014, Monthly Notices of the Royal Astronomical Society, 444, 3592

Dziembowski W., 1977, Acta Astronomica, 27, 203

Eisenstein D. J., et al., 2011, The Astronomical Journal, 142, 72

Elsworth Y., Howe R., Isaak G. R., McLeod C. P., Miller B. A., New R., Wheeler S. J., Gough D. O., 1995, Nature, 376, 669

Elsworth Y., Hekker S., Basu S., Davies G. R., 2017, Monthly Notices of the Royal Astronomical Society, 466, 3344

Esselstein R., Aigrain S., Vanderburg A., Smith J. C., Meibom S., Van Saders J., Mathieu R., 2018, The Astrophysical Journal, 859, 167

Farr W. M., et al., 2018, The Astrophysical Journal Letters, 865, L20

Fleming D. P., Barnes R., Davenport J. R. A., Luger R., 2019, The Astrophysical Journal, 881, 88

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publications of the Astronomical Society of the Pacific, 125, 306

Foreman-Mackey D., Agol E., Ambikasaran S., Angus R., 2017, The Astronomical Journal, 154, 220

Fossat E., et al., 2017, Astronomy and Astrophysics, 604, A40

Gai N., Basu S., Chaplin W. J., Elsworth Y., 2011, The Astrophysical Journal, 730, 63

Gaia Collaboration et al., 2016, Astronomy & Astrophysics, 595, A2

Gaia Collaboration et al., 2018, Astronomy & Astrophysics, 616, A1

Gallet F., Delorme P., 2019, Astronomy & Astrophysics, 626, A120

Garcia R. A., Ballot J., 2019, Living Reviews in Solar Physics, 16, 4

García R. A., Turck-Chièze S., Jiménez-Reyes S. J., Ballot J., Pallé P. L., Eff-Darwich A., Mathur S., Provost J., 2007, Science, 316, 1591

García R. A., Mathur S., Ballot J., Eff-Darwich A., Jiménez-Reyes S. J., Korzennik S. G., 2008, Solar Physics, 251, 119

García R. A., et al., 2011, Monthly Notices of the Royal Astronomical Society, 414, L6

García R. A., et al., 2014, Astronomy and Astrophysics, 572, A34

Garraffo C., Drake J. J., Cohen O., 2016, Astronomy & Astrophysics, 595, A110

Gaulme P., et al., 2016, The Astrophysical Journal, 833, L13

Gehan C., Mosser B., Michel E., 2017, in Reylé C., Di Matteo P., Herpin F., Lagadec E., Lançon A., Meliani Z., Royer F., eds, SF2A-2017: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. p. Di (arXiv:1710.05647)

Gelman A., 2006, Bayesian Analysis, 1, 515

Gelman A., Rubin D. B., 1992, Statistical Science, 7, 457

Girardi L., 1999, Monthly Notices of the Royal Astronomical Society, 308, 818

Girardi L., 2016, Annual Review of Astronomy and Astrophysics, 54, 95

Girardi L., et al., 2012, Astrophysics and Space Science Proceedings, 26, 165

Girardi L., Marigo P., Bressan A., Rosenfield P., 2013, The Astrophysical Journal, 777, 142

Girardi L., Barbieri M., Miglio A., Bossini D., Bressan A., Marigo P., Rodrigues T. S., 2015, in Asteroseismology of Stellar Populations in the Milky Way. p. 125 (arXiv:1409.2276), doi:10.1007/978-3-319-10993-0_14

Gizon L., Solanki S. K., 2003, The Astrophysical Journal, 589, 1009

Gizon L., et al., 2013, Proceedings of the National Academy of Science, 110, 13267

Goldreich P., Keeley D. A., 1977a, The Astrophysical Journal, 211, 934

Goldreich P., Keeley D. A., 1977b, The Astrophysical Journal, 212, 243

Goldreich P., Kumar P., 1988, The Astrophysical Journal, 326, 462

Gonzalez G., 2016, Monthly Notices of the Royal Astronomical Society, 463, 3513

Goode P. R., Thompson M. J., 1992, The Astrophysical Journal, 395, 307

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010, 5, 65

Gough D. O., 1983, in Shaver P. A., Kunth D., Kjar K., eds, Primordial Helium. pp 117–136

Gough D. O., Thompson M. J., 1990, Monthly Notices of the Royal Astronomical Society, 242, 25

Gough D. O., et al., 1996, Science, 272, 1296

Grec G., Fossat E., Pomerantz M., 1980, Nature, 288, 541

Green G. M., et al., 2018, Monthly Notices of the Royal Astronomical Society, 478, 651

Grenander U., 1959, Probability and Statistics: The Harald Cramr Volume. Alqvist & Wiksell

Groenewegen M. a. T., 2008, Astronomy and Astrophysics, 488, 935

Hacking P., et al., 1999, The Wide-Field Infrared Explore (WIRE) Mission. p. 409

Hale S. J., Howe R., Chaplin W. J., Davies G. R., Elsworth Y. P., 2016, Solar Physics, 291, 1

Hall O. J., et al., 2019, Monthly Notices of the Royal Astronomical Society, 486, 3569

Handberg R., Campante T. L., 2011, Astronomy and Astrophysics, 527, A56

Handberg R., Lund M. N., 2014, Monthly Notices of the Royal Astronomical Society, 445, 2698

Handler G., 2009, Monthly Notices of the Royal Astronomical Society, 398, 1339

Handler G., 2013, Planets, Stars and Stellar Systems. Volume 4: Stellar Structure and Evolution, p. 207

Hansen C. J., Cox J. P., van Horn H. M., 1977, The Astrophysical Journal, 217, 151

Hart A. B., 1954, Monthly Notices of the Royal Astronomical Society, 114, 17

Hart A. B., 1956, Monthly Notices of the Royal Astronomical Society, 116, 38

Harvey J., 1985, Future Missions in Solar, Heliospheric & Space Plasma Physics, 235, 199

Hawkins K., Leistedt B., Bovy J., Hogg D. W., 2017, Monthly Notices of the Royal Astronomical Society, 471, 722

Hekker S., et al., 2011, Monthly Notices of the Royal Astronomical Society, 414, 2594

Hoffman M. D., Gelman A., 2014, Journal of Machine Learning Research, 15, 1593

Hogg D. W., Bovy J., Lang D., 2010, arXiv e-prints, p. arXiv:1008.4686

Houdek G., Balmforth N. J., Christensen-Dalsgaard J., Gough D. O., 1999, Astronomy and Astrophysics, 351, 582

Howell S. B., et al., 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Huber D., Stello D., Bedding T. R., Chaplin W. J., Arentoft T., Quirion P. O., Kjeldsen H., 2009, Communications in Asteroseismology, 160, 74

Huber D., et al., 2011a, The Astrophysical Journal, 731, 94

Huber D., et al., 2011b, The Astrophysical Journal, 743, 143

Huber D., et al., 2013a, Science, 342, 331

Huber D., et al., 2013b, The Astrophysical Journal, 767, 127

Huber D., et al., 2014, The Astrophysical Journal Supplement Series, 211, 2

Huber D., et al., 2017, The Astrophysical Journal, 844, 102

Huber D., et al., 2019, The Astronomical Journal, 157, 245

Jofré E., Petrucci R., Saffe C., Saker L., de la Villarmois E. A., Chavero C., Gómez M., Mauas P. J. D., 2015, Astronomy and Astrophysics, 574, A50

Kallinger T., et al., 2014, Astronomy & Astrophysics, 570, A41

Karoff C., et al., 2018, The Astrophysical Journal, 852, 46

Kass R. E., Raftery A. E., 1995, Journal of the American Statistical Association, 90, 773

Kawaler S. D., 1988, The Astrophysical Journal, 333, 236

Kennedy J. R., GONG Team 1994, GONG, A Global Network of Automated Solar Telescopes. p. 188

Khan S., Hall O. J., Miglio A., Davies G. R., Mosser B., Girardi L., Montalbán J., 2018, The Astrophysical Journal, 859, 156

Khan S., et al., 2019, Astronomy and Astrophysics, 628, A35

Kiefer R., Roth M., 2018, The Astrophysical Journal, 854, 74

Kjeldsen H., Bedding T. R., 1995, Astronomy and Astrophysics, 293, 87

Kjeldsen H., Frandsen S., 1992, Publications of the Astronomical Society of the Pacific, 104, 413

Kjeldsen H., Bedding T. R., Viskum M., Frandsen S., 1995, The Astronomical Journal, 109, 1313

Kjeldsen H., et al., 2003, The Astronomical Journal, 126, 1483

Kraft R. P., 1967, The Astrophysical Journal, 150, 551

Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1

Lebreton Y., Goupil M. J., Montalbán J., 2014a, in EAS Publications Series. pp 99–176 (arXiv:1410.5336), doi:10.1051/eas/1465004

Lebreton Y., Goupil M. J., Montalbán J., 2014b, in EAS Publications Series. pp 177–223 (arXiv:1410.5337), doi:10.1051/eas/1465005

Ledoux P., 1951, The Astrophysical Journal, 114, 373

Lefebvre S., García R. A., Jiménez-Reyes S. J., Turck-Chièze S., Mathur S., 2008, Astronomy & Astrophysics, 490, 1143

Leibacher J. W., Stein R. F., 1971, Astrophysical Letters, 7, 191

Leighton R. B., Noyes R. W., Simon G. W., 1962, The Astrophysical Journal, 135, 474

Leiner E., Mathieu R. D., Vanderburg A., Gosnell N. M., Smith J. C., 2019, The Astrophysical Journal, 881, 47

Lightkurve Collaboration et al., 2018, Lightkurve: Kepler and TESS time series analysis in Python (ascl:1812.013)

Lindegren L., et al., 2016, Astronomy and Astrophysics, 595, A4

Lindegren L., et al., 2018, Astronomy & Astrophysics, 616, A2

Lomb N. R., 1976, Astrophysics and Space Science, 39, 447

Lopes I., Turck-Chieze S., 1994, Astronomy and Astrophysics, 290, 845

Lorenzo-Oliveira D., et al., 2019, Monthly Notices of the Royal Astronomical Society, 485, L68

Lund M. N., Miesch M. S., Christensen-Dalsgaard J., 2014, The Astrophysical Journal, 790, 121

Lund M. N., et al., 2016a, Publications of the Astronomical Society of the Pacific, 128, 124204

Lund M. N., et al., 2016b, Monthly Notices of the Royal Astronomical Society, 463, 2600

Lund M. N., et al., 2017, The Astrophysical Journal, 835, 172

Luri X., et al., 2018, Astronomy & Astrophysics, 616, A9

Maíz Apellániz J., Weiler M., 2018, Astronomy and Astrophysics, 619, A180

Majewski S. R., et al., 2017, The Astronomical Journal, 154, 94

Martić M., et al., 1999, Astronomy & Astrophysics, 351, 993

Mathur S., García R. A., Huber D., Regulo C., Stello D., Beck P. G., Houmani K., Salabert D., 2016, The Astrophysical Journal, 827, 50

Mathur S., et al., 2017, The Astrophysical Journal Supplement Series, 229, 30

Matt S. P., Brun A. S., Baraffe I., Bouvier J., Chabrier G., 2015, The Astrophysical Journal, 799, L23

Maxted P. F. L., Serenelli A. M., Southworth J., 2015, Astronomy and Astrophysics, 577, A90

Mazumdar A., et al., 2014, The Astrophysical Journal, 782, 18

McQuillan A., Mazeh T., Aigrain S., 2014, The Astrophysical Journal Supplement Series, 211, 24

Meibom S., Barnes S. A., Platais I., Gilliland R. L., Latham D. W., Mathieu R. D., 2015, Nature, 517, 589

Metcalfe T. S., Egeland R., 2019, The Astrophysical Journal, 871, 39

Metcalfe T. S., Mathur S., Doğan G., Woitaszek M., 2012, First Results from the Asteroseismic Modeling Portal. p. 213

Metcalfe T. S., et al., 2014, The Astrophysical Journal Supplement Series, 214, 27

Metcalfe T. S., Egeland R., van Saders J., 2016, The Astrophysical Journal Letters, 826, L2

Miglio A., et al., 2009, Astronomy and Astrophysics, 503, L21

Miglio A., et al., 2010, Astronomy and Astrophysics, 520, L6

Miglio A., et al., 2012, Monthly Notices of the Royal Astronomical Society, 419, 2077

Miglio A., et al., 2013, Monthly Notices of the Royal Astronomical Society, 429, 423

Miglio A., et al., 2016, Monthly Notices of the Royal Astronomical Society, 461, 760

Mosser B., Appourchaux T., 2009, Astronomy and Astrophysics, 508, 877

Mosser B., et al., 2010, Astronomy and Astrophysics, 517, A22

Mosser B., et al., 2011, Astronomy and Astrophysics, 532, A86

Mosser B., et al., 2012a, Astronomy and Astrophysics, 537, A30

Mosser B., et al., 2012b, Astronomy and Astrophysics, 540, A143

Mosser B., et al., 2012c, Astronomy and Astrophysics, 548, A10

Mosser B., Vrard M., Belkacem K., Deheuvels S., Goupil M. J., 2015, Astronomy and Astrophysics, 584, A50

Murphy S. J., Shibahashi H., Kurtz D. W., 2013, Monthly Notices of the Royal Astronomical Society, 430, 2986

Ness M., Hogg D. W., Rix H.-W., Ho A. Y. Q., Zasowski G., 2015, The Astrophysical Journal, 808, 16

Nielsen M. B., Gizon L., Schunker H., Karoff C., 2013, Astronomy and Astrophysics, 557, L10

Nielsen M. B., Schunker H., Gizon L., Ball W. H., 2015, Astronomy and Astrophysics, 582, A10

Paxton B., Bildsten L., Dotter A., Herwig F., Lesaffre P., Timmes F., 2011, The Astrophysical Journal Supplement Series, 192, 3

Paxton B., et al., 2013, The Astrophysical Journal Supplement Series, 208, 4

Paxton B., et al., 2015, The Astrophysical Journal Supplement Series, 220, 15

Paxton B., et al., 2018, The Astrophysical Journal Supplement Series, 234, 34

Pereira F., et al., 2019, The Monthly Notices of the Royal Astronomical Society, 489, 5764

Pinsonneault M. H., et al., 2014, The Astrophysical Journal Supplement Series, 215, 19

Pinsonneault M. H., et al., 2018, The Astrophysical Journal Supplement Series, 239, 32

Planck Collaboration et al., 2018, arXiv e-prints, p. arXiv:1807.06209

Plaskett H. H., 1916, The Astrophysical Journal, 43, 145

Prša A., et al., 2016, The Astronomical Journal, 152, 41

Raghavan D., et al., 2010, The Astrophysical Journal Supplement Series, 190, 1

Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning. Adaptive Computation and Machine Learning, MIT Press, Cambridge, Mass

Rauer H., et al., 2014, Experimental Astronomy, 38, 249

Rendle B. M., et al., 2019a, Monthly Notices of the Royal Astronomical Society, 484, 771

Rendle B. M., et al., 2019b, Monthly Notices of the Royal Astronomical Society, 490, 4465

Réville V., Brun A. S., Matt S., Strugarek A., Pinto R., 2015, The Astrophysical Journal, 798, 116

Ricker G. R., et al., 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Riess A. G., et al., 2018, The Astrophysical Journal, 861, 126

Rodrigues T. S., et al., 2017, Monthly Notices of the Royal Astronomical Society, 467, 1433

Roxburgh I. W., 2009, Astronomy and Astrophysics, 506, 435

Roxburgh I. W., Vorontsov S. V., 2006, Monthly Notices of the Royal Astronomical Society, 369, 1491

Sahlholdt C. L., Silva Aguirre V., 2018, The Monthly Notices of the Royal Astronomical Society, 481, L125

Salaris M., Girardi L., 2002, Monthly Notices of the Royal Astronomical Society, 337, 332

Salvatier J., Wiecki T. V., Fonnesbeck C., 2016, PeerJ Computer Science, 2, e55

Scargle J. D., 1982, The Astrophysical Journal, 263, 835

Schönrich R., McMillan P., Eyer L., 2019, The Monthly Notices of the Royal Astronomical Society, 487, 3568

Schunker H., Schou J., Gaulme P., Gizon L., 2018, Solar Physics, 293, 95

Schuster A., 1898, Terrestrial Magnetism (Journal of Geophysical Research), 3, 13

Scuflaire R., 1974, Astronomy and Astrophysics, 36, 107

See V., et al., 2019, The Astrophysical Journal, 886, 120

Serenelli A., et al., 2017, The Astrophysical Journal Supplement Series, 233, 23

Shanks T., Hogarth L. M., Metcalfe N., 2019, The Monthly Notices of the Royal Astronomical Society, 484, L64

Sharma S., Stello D., 2016, Asfgrid: Asteroseismic parameters for a star (ascl:1603.009)

Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, The Astrophysical Journal, 822, 15

Silva Aguirre V., et al., 2013, The Astrophysical Journal, 769, 141

Silva Aguirre V., et al., 2015, Monthly Notices of the Royal Astronomical Society, 452, 2127

Silva Aguirre V., et al., 2017, The Astrophysical Journal, 835, 173

Silva Aguirre V., et al., 2020, The Astrophysical Journal, 889, L34

Skrutskie M. F., et al., 2006, The Astronomical Journal, 131, 1163

Skumanich A., 1972, The Astrophysical Journal, 171, 565

Slumstrup D., Grundahl F., Silva Aguirre V., Brogaard K., 2019, Astronomy & Astrophysics, 622, A111

Stassun K. G., Torres G., 2018, The Astrophysical Journal, 862, 61

Stello D., Bruntt H., Preston H., Buzasi D., 2008, The Astrophysical Journal Letters, 674, L53

Stello D., Chaplin W. J., Basu S., Elsworth Y., Bedding T. R., 2009a, Monthly Notices of the Royal Astronomical Society: Letters, 400, L80

Stello D., et al., 2009b, The Astrophysical Journal, 700, 1589

Stello D., et al., 2013, Period Spacings of Most Red Giants Observed by Kepler. p. 167

Stello D., et al., 2015, The Astrophysical Journal, 809, L3

Sweigart A. V., Greggio L., Renzini A., 1990, The Astrophysical Journal, 364, 527

Tassoul M., 1980, The Astrophysical Journal Supplement Series, 43, 469

Thomas A. E. L., et al., 2019, Monthly Notices of the Royal Astronomical Society, 485, 3857

Thompson M. J., et al., 1996, Science, 272, 1300

Toutain T., Appourchaux T., 1994, Astronomy and Astrophysics, 289, 649

Turck-Chièze S., et al., 1997, Solar Physics, 175, 247

Turck-Chièze S., et al., 2001, The Astrophysical Journal Letters, 555, L69

Turck-Chièze S., et al., 2004, The Astrophysical Journal, 604, 455

Udalski A., 2000, The Astrophysical Journal, 531, L25

Ulrich R. K., 1970, The Astrophysical Journal, 162, 993

Unno W., Osaki Y., Ando H., Saio H., Shibahashi H., 1989, Nonradial oscillations of stars, Tokyo: University of Tokyo Press, 1989, 2nd ed.

Van Eylen V., et al., 2014, The Astrophysical Journal, 782, 14

Van Eylen V., Agentoft C., Lundkvist M. S., Kjeldsen H., Owen J. E., Fulton B. J., Petigura E., Snellen I., 2018, Monthly Notices of the Royal Astronomical Society, 479, 4786

Van Hoey S., van der Kwast J., Nopens I., Seuntjens P., 2013, in EGU General Assembly Conference Abstracts. EGU General Assembly Conference Abstracts. pp EGU2013–10059

Vandakurov Y. V., 1967, Astronomicheskii Zhurnal, 44, 786

VanderPlas J. T., 2018, The Astrophysical Journal Supplement Series, 236, 16

VanderPlas J. T., Ivezić Ž., 2015, The Astrophysical Journal, 812, 18

VanderPlas J., Connolly A. J., Ivezic Z., Gray A., 2012, in Proceedings of Conference on Intelligent Data Understanding (CIDU. pp 47–54 (arXiv:1411.5039), doi:10.1109/CIDU.2012.6382200

Verner G. A., Roxburgh I. W., 2011, arXiv e-prints, p. arXiv:1104.0631

Viani L. S., Basu S., Chaplin W. J., Davies G. R., Elsworth Y., 2017, The Astrophysical Journal, 843, 11

Viani L. S., Basu S., Corsaro E., Ball W. H., Chaplin W. J., 2019, The Astrophysical Journal, 879, 33

Vorontsov S. V., Baturin V. A., Pamiatnykh A. A., 1991, Nature, 349, 49

Vrard M., Mosser B., Samadi R., 2016, Astronomy and Astrophysics, 588, A87

White T. R., Bedding T. R., Stello D., Christensen-Dalsgaard J., Huber D., Kjeldsen H., 2011, The Astrophysical Journal, 743, 161

White T. R., et al., 2013, Monthly Notices of the Royal Astronomical Society, 433, 1262

Woodard M. F., 1984, PhD thesis

Yu J., Huber D., Bedding T. R., Stello D., Hon M., Murphy S. J., Khanna S., 2018, The Astrophysical Journal Supplement Series, 236, 42

Yu J., Bedding T. R., Stello D., Huber D., Compton D. L., Gizon L., Hekker S., 2020, Monthly Notices of the Royal Astronomical Society, 493, 1388

Zhao G., Zhao Y.-H., Chu Y.-Q., Jing Y.-P., Deng L.-C., 2012, Research in Astronomy and Astrophysics, 12, 723

Zinn J. C., Pinsonneault M. H., Huber D., Stello D., 2019a, The Astrophysical Journal, 878, 136

Zinn J. C., Pinsonneault M. H., Huber D., Stello D., Stassun K., Serenelli A., 2019b, ApJ, 885, 166

de Jong R. S., et al., 2014, 4MOST: 4-metre Multi-Object Spectroscopic Telescope. p. 91470M, doi:10.1117/12.2055826

van Saders J. L., Pinsonneault M. H., 2013, The Astrophysical Journal, 776, 67

van Saders J. L., Ceillier T., Metcalfe T. S., Silva Aguirre V., Pinsonneault M. H., García R. A., Mathur S., Davies G. R., 2016, Nature, 529, 181

van Saders J. L., Pinsonneault M. H., Barbieri M., 2019, The Astrophysical Journal, 872, 128

van der Walt S., Colbert S. C., Varoquaux G., 2011, Computing in Science and Engineering, 13, 22

Publications and Contributions

Publications

First Author Publications:

Hall, O. J., Davies, G. R., Elsworth, Y. P. and 9 co-authors
Testing asteroseismology with Gaia DR2: Hierarchical models of the Red Clump
Monthly Notices of the Royal Astronomical Society, 2019
Summary: Constrained the luminosity of the Red Clump and the *Gaia* DR2 parallax zero-point offset simultaneously using hierarchical latent variable models.
doi:10.1093/mnras/stz1092, arXiv:1904.07919

Contributing Author Publications:

Khan, S., **Hall, O. J.**, Miglio, A., Davies, G. R., Mosser, B., Girardi, L., Montalbán, J.
The Red-giant Branch Bump Revisited: Constraints on Envelope Overshooting in a Wide Range of Masses and Metallicities
The Astrophysical Journal, 2018
Contribution: Used Mixture Models to constrain the position of the Red Giant Branch Bump.
doi:10.3847/1538-4357/aabf90, arXiv:1804.06669

Bugnet, L., García, R. A., Davies, G. R., Mathur, S., Corsaro, E., **Hall, O. J.**, Rendle, B. M.
FliPer: A global measure of power density to estimate surface gravities of main-sequence solar-like stars and red giants
Astronomy & Astrophysics, 2018
Contribution: Helped develop the FliPer metric and its machine learning implementation.
doi:0.1051/0004-6361/201833106, arXiv:1809.05105

Silva Aguirre, V., Stello, D., Stokholm, A. and 75 coauthors including **Hall, O. J.**
Detection and characterisation of oscillating red giants: first results from the TESS satellite
The Astrophysical Journal, 2019
Contribution: Obtained fundamental seismic parameters for stellar sample.

doi:10.3847/2041-8213/ab6443, arXiv:1912.07604

Chaplin, W., Serenelli, A. M., Miglio, A. and 82 coauthors including **Hall, O. J.**
Age dating of an early Milky Way merger via asteroseismology of the naked-eye star νIndi
Nature Astronomy, 2020
Contribution: Advised on systematic uncertainties in spectroscopic methods.
doi:10.1038/s41550-019-0975-9, arXiv:2001.04653

Bugnet, L., García, R. A., Mathur, S., Davies, G. R., **Hall, O. J.**, Lund, M. N., Rendle, B. M.
FliPer$_{Class}$: In search of solar-like pulsators among TESS targets
arXiv e-prints, 2019
Contribution: Aided with interpretation of systematic uncertainties on effective temperature.
doi:10.1051/0004-6361/201834780, arXiv:1902.09854

Huber, D., Chaplin, W. J., Chontos, A and 139 co-authors including **Hall, O. J.**
A Hot Saturn Orbiting An Oscillating Late Subgiant Discovered by TESS
arXiv e-prints, 2019
Contribution: Checked proper use and interpretation of *Gaia* parallaxes.
doi:10.3847/1538-3881/ab1488, arXiv:1901.01643

Davies, G. R., Lund, M. N., Miglio, A., Elsworth, Y. P. and 13 co-authors including **Hall, O. J.**
Using red clump stars to correct the Gaia *DR1 parallaxes*
Astronomy & Astrophysics, 2017
Contribution: Verified results found by lead authors.
doi:10.1051/0004-6361/201630066, arXiv:1701.02506

Software Publications

Lightkurve Collaboration, Cardoso, J. V. d. M., Hedges, C., Gully-Santiago, M., Saunders, N., Cody, A-M., Barclay, T., **Hall, O. J.**, Sagear, S., Turtelboom, E., Zhang, J., Tzanidakis, A., Mighell, K., Coughlin, J., Bell, K., Berta-Thompson, Z., Williams, P., Dotson, J., Barentsen, G.
Lightkurve: Kepler and TESS time series analysis in Python
Astrophysics Source Code Library, 2018
Contribution: Led development of the '`periodogram`' and '`seismology`' modules.
ascl:1812.013

White Papers

Khullar, G., Kholer, S., Konchady, T. and 32 co-authors including **Hall, O. J.**
Astrobites as a Community-led Model for Education, Science Communication, and Accessibility in Astrophysics
arXiv e-prints, 2019
arXiv:1907.09496

Presentations

Asteroseismology & Rotational Evolution: Bayesian Inference in Stellar Astrophysics
Invited Talk, CSH Symposium, 2020
Centre for Space and Habitability, Switzerland

Asteroseismic Follow-Up of *CHEOPS* Target Hosts
Invited Talk, ESA Research Fellow Jamboree, 2020
European Space Research and Technology Centre, The Netherlands

Asteroseismology & Applied Statistics
Group Seminar, 2019
University of Exeter, UK
Available Online

Accessible Asteroseismology with Lightkurve
Invited Talk, TASC5/KASC12, 2019
Massachusetts Institute of Technology, USA
Available Online

Testing asteroseismology with *Gaia* DR2: Hierarchical Models & the Red Clump
Birmingham-Warwick Science Meet-Up, 2018
University of Warwick, UK
Available Online

Testing Asteroseismology with *Gaia* DR2: Luminosity of the Red Clump
TASC4/KASC11, 2018
Aarhus University, Denmark
Available Online

Estimating *TESS* backgrounds with mixture models – Update
T'DA 2, 2017
University of Aarhus, Denmark

Estimating *TESS* backgrounds with mixture models
T'DA 1, 2016
University of Birmingham, UK

Contributions to Open-Source Code

Lightkurve

I led the development of the `periodogram` and `seismology` modules of Lightkurve, which facilitate asteroseismic frequency-domain analysis of observations from *Kepler*, *K2* and *TESS* (see Chapter 3). As part of this, I spent 3 weeks with the development team at NASA Ames, USA. My core contributions are through two major pull requests, but I have remained an active contributor through additional smaller pull requests and by engaging in ongoing discussions about changes to the package.

- Pull Request: Enhancing lightcurve.periodogram()
- Pull Request: Add asteroseismic parameter estimator for solar-like oscillators

eleanor

The Python package `eleanor` enables users to download, detrend and analyse data from *TESS* Full Frame Images (FFIs). A visualisation feature in `eleanor` allows users to plot a pixel-by-pixel time-series for a cut-out of a FFI. I contributed a pull request that gave the option to plot a pixel-by-pixel amplitude periodogram instead.

- Pull Request: Add option to visualize pixel_by_pixel amplitude spectrum

TESS Data for Asteroseismology (T'DA)

The TASOC (*TESS* Asteroseismic Science Operations Centre) T'DA code is an ongoing collaboration to provide an open-source data pipeline for asteroseismology with the *TESS* mission, specifically its Full Frame Images (FFIs). The T'DA code is still in development at time of writing, and aims to provide basic calibrations of raw data, extract flux and perform instrumental corrections, and classify the observed object using machine learning. As part of the T'DA collaboration, I have

made contributions to the `photometry` and `corrections` components of the code. For `photometry`, I performed a thorough comparison of different FFI background estimation methods. For the `corrections`, I helped implement the stellar ensemble correction method into the code.

- Pull Request: Writing Unit Tests for Ensemble.py
- Pull Request: Ensemble algorithm and code update for corrected data release
- GitHub Repository: SkyBackground

Carbon Footprint

Between travel and use of computational resources, astronomy is a carbon-intensive field. As a retrospective of this 3.5 year Ph.D. I have attempted to estimate the carbon footprint of my research.

Over the course of this Ph.D. 28 individual economy-class flights were taken for research purposes. Together, these are responsible for a carbon footprint of 10.77 tonnes of CO_2[1]. Of these, 6 transatlantic flights were responsible for 63% of emissions from flights.

Research done for this Ph.D made use of the University of Birmingham's BlueBEAR high performance computing cluster. My BlueBEAR energy use totalled 790.5 kWh, which translates to a carbon footprint of 0.2 tonnes of CO_2[2]. This assumes the perfect efficiency of each CPU and does not include other computational overhead (e.g. motherboards, RAM, cooling etc.).

Combined, the carbon footprint of this PhD is **10.97 tonnes of CO_2.** This is equivalent to driving my 2017 Vauxhall Corsa around the circumference of the Earth 2.28 times.

This carbon footprint does not include taxis to and from airports, local transportation on location, domestic or international train travel, printed paper, the purchase or running of personal computing equipment, the cost of obtaining ground- and space based data, the individual carbon footprints of my research collaborators, streaming YouTube music for 8 hours a day or the amount of coffee I consumed.

[1] Calculated using https://www.carbonfootprint.com/. The additional impact high-altitude flights have on global warming is included, and accounts for a factor of 1.891.

[2] Based on the UK Government's most recent Greenhouse gas reporting factors.